

基于 Mel-CNN 模型的磨煤机声纹识别研究

朱斌¹, 李锲^{1†}, 曹宏², 常达³

(1. 国家能源集团永州发电有限公司, 湖南 永州 425000; 2 北京华控智加科技有限公司, 北京 100080)

3. 国网淄博供电公司, 山东 淄博 255022)

摘要:磨煤机是燃煤电站的重要辅助设备之一, 它的运行状况直接关系到机组的安全稳定运行。因此, 提出了一种基于 Mel-CNN 模型的磨煤机轴承声纹识别方法, 以有效检测磨煤机的运行状态。首先利用噪声采集平台模拟了各种轴承故障作用下的噪声信号, 提取 Mel 谱特征, 然后将磨煤机轴承各状态下产生的 Mel 谱图输入到 CNN 模型中, 该模型对各状态下生成的 Mel 谱图进行训练、学习和特征提取。最终实验结果表明, Mel-CNN 模型在训练时间、模型大小指标等方面优于其他模型, 在轴承故障声纹识别方面具有较大优势。

关键词:磨煤机; Mel 谱图; 声纹识别; CNN; 轴承故障

中图分类号:TP39

文献标识码:A

Research on Voiceprint Recognition of Coal Mill Based on Mel-CNN Model

ZHU Bin¹, LI Qie^{1†}, CAO Hong², CHANG Da³

(1. CHN Energy Group Yongzhou Power Generation Co., Ltd., Yongzhou, Hunan 425000, China;

2. Beijing Huakong Zhijiao Technology Co., Ltd., Beijing 100080, China;

3. State Grid Zibo Power Supply Company, Zibo, Shandong 255022, China)

Abstract: Coal mill is one of the important auxiliary equipment of coal-fired power station, and its running condition is directly related to the safe and stable operation of the unit. A method based on Mel-CNN model is proposed to identify the bearing vowels of coal mill, so as to detect the running state of coal mill effectively. Firstly, the noise signal under the action of various bearing failure is simulated by the noise acquisition platform, and the Mel spectrum features are extracted. Then, the Mel spectrum generated under each state of the mill bearing is input into the CNN model, which trains, learns and extracts the Mel spectrum generated under each state. The final experimental results show that the Mel-CNN model is superior to other models in terms of training time and model size, and has great advantages in bearing fault voicing recognition.

Key words: coal mill; Mel spectrum; voiceprint recognition; CNN; bearing failure

磨煤机作为燃煤机组的核心辅助设备, 其运行状态对机组的安全稳定运行至关重要, 一旦磨煤机出现故障, 锅炉的燃料供应将无法保障, 进而导致锅炉与汽轮机的输出功率失衡。在火力发电和工

业生产中, 磨煤机的重要性不言而喻, 其轴承的健康状况直接决定了设备的运行效率和稳定性。因此, 开展磨煤机的有效故障预警与诊断, 确保其正常运行, 具有极其重要的意义^[1-2]。

收稿日期: 2025-03-22

基金项目: 国网山东省电力公司科技项目资助(52060324000C)

作者简介: 朱斌(1985—), 男, 湖南辰溪人, 本科, 工程师, 研究方向: 火力发电厂脱硫脱硝设备运维。

† 通信联系人, E-mail: qielli@aliyun.com

目前,国内外学者对声纹识别研究取得了诸多进展,文献[3]提出了一种结合 Mel-CNN 和电流频谱的轴承外圈故障声纹检测方法,通过 CNN 重新提取 Mel 声纹特征,获取高维抽象特征,并利用声纹融合电流神经网络模型实现故障诊断。文献[4]开发了一种基于卷积时域音频分离网络(Conv-TASNet)的变压器音频降噪识别网络,利用该网络去除环境噪声,并结合卷积神经网络(CNN)进行声纹识别。文献[5]提出了一种基于加权梅尔频率倒谱系数(WMFCC)和单类支持向量机(OCS-VM)的声纹识别算法,并结合贝叶斯优化对门控循环单元(GRU)网络进行优化,通过消融实验验证了各流程对声纹识别性能的影响。文献[6]则提出了一种基于多粒度知识特征和 Transformer 网络的电力变压器故障声纹辨识方法,通过卷积自编码器实现特征降维和模型压缩,从而缩短训练时间并提高故障辨识效率。文献[7-8]采用 MFCC 特征结合深度学习进行声纹识别,还有研究探讨了 MFCC 特征在声纹识别中的应用前景。还有一种基于改进鲸鱼优化算法(IWOA)和 Transformer 的磨煤机故障预警模型,通过优化超参数和自适应阈值实现故障预警,验证了其在实际工程中的应用价值。这些研究表明,磨煤机故障诊断技术正朝着多方法融合、智能化和实时化的方向发展^[9]。

针对以上问题,以磨煤机轴承为研究对象,以噪声数据为研究基础,提出基于 Mel-CNN 模型的磨煤机轴承故障诊断方法,完成了轴承故障类型的识别。

1 声音信号处理

1.1 频谱图

频谱图是声信号分析与处理中的关键特征频谱图,能够反映信号在不同时间点的频率能量分布,全面建立声音信号的时域与频域关系,并最大限度地提取声音特征信息,从而为后续声纹特征的提取和学习提供便利。声纹谱图的绘制过程包括分帧、加窗和离散傅里叶变换。首先,截取一段磨煤机轴承的 2 s 噪声信号,并对其进行分帧处理。如果帧长过长,可能会降低特征量的准确性;而帧长过短,则可能导致无法提取有用的特征量。鉴于磨煤机轴承噪声比人声更为稳定,可以适当增加帧长,以确保信号特性的完整性。

在本文中,选择帧长为 $4096/n = 64$ ms,并设置重叠率为 50%,以实现帧间的平滑过渡。此外,若直接对分帧后的数据进行离散傅里叶变换,可能会出现频谱泄漏现象。因此,每个帧都需要先进行加窗处理。为了使信号的两端更加平滑,减少信号失真,本文选择了汉明窗。

$$\bar{\omega}(n) = \begin{cases} 0.54 - 0.46\cos[2\pi n/(N-1)], & 0 \leq n \leq N-1 \\ 0, & \text{else} \end{cases} \quad (1)$$

最后,对每帧数据加窗后进行离散傅里叶变换,以下公式为求得时频矩阵并绘制频谱图:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{N}kn}, \quad k = 0, 1, 2, \dots, N-1 \quad (2)$$

频谱图的横坐标表示分帧后的帧数(时间),纵坐标表示频率,颜色表示此时这个频率的大小,即功率谱密度。

1.2 梅尔频谱图

经过上述预处理得到的频谱图能够反映磨煤机轴承运行时的时域与频域关系。尽管频谱图在时间维度上被压缩为 62 列,但频率维度仍包含 2049 行,这会对后续 CNN 的训练和识别效率产生较大影响。因此,需要对频率维度进行降维处理,本文考虑采用 Mel 滤波器来实现这一目标。

Mel 滤波器依据人耳的感知特性,对线性频带进行非线性转换,增强低频部分的权重,同时降低高频部分的权重,从而突出有效频带的作用。Mel 频率与实际频率的转换关系如下:

$$f_{\text{mel}}(f) = 2595 \cdot \lg\left(1 + \frac{f}{700}\right) \quad (3)$$

式中, $f_{\text{mel}}(f)$ 为 mel 标度下的频率, f 为实际频率。

Mel 滤波器组由 M 个三角形滤波器构成($M=40$),每个滤波器的中心频率为 $f(m)$,且在 Mel 频率尺度上,相邻滤波器之间的间隔是均匀的。

滤波器组的传递函数为:

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

其中, $f(m)$ 定义为:

$$f(m) = \left(\frac{N}{f_s}\right) F_{\text{mel}}^{-1} \left[F_{\text{mel}}(f_l) + m \frac{F_{\text{mel}}(f_h) - F_{\text{mel}}(f_l)}{M+1} \right]$$

(5)

式中, $f(h)$, $f(l)$ 为滤波器频率的上下限, f_s 为变压器声纹采样的采样频率($f_s = 64 \text{ kHz}$), N 为短时傅里叶变换时的帧长。

2 基于卷积神经网络的声纹识别

2.1 CNN

卷积神经网络在图像识别方面具有突出的能力,它由多组卷积层、池化层和全连接层组成,通常用于对图像中的像素值进行分类,其基本结构如图 1 所示。

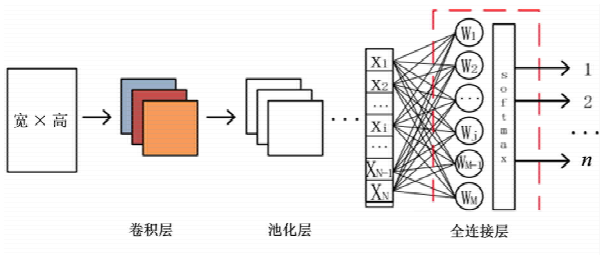


图 1 CNN 网络结构

二维卷积神经网络的计算公式为:

$$y_{xy} = f\left(\sum_i \sum_j \bar{w}_{ij} v_{(x+i)(y+j)} + b\right)$$

(6)

式中, y_{xy} 表示原始图位置的特征值(x, y), w_{ij} 表示卷积核的值, $v_{(x+i)(y+j)}$ 表示该位置的输入, ($x+i, y+j$)为偏差值。

将上面绘制的 Mel 谱图作为 CNN 的输入,输入图像的像素为($62 \times 40 \times 3$),其中,宽度 62 表示声音信号分帧后的时间分量,高度 40 表示 Mel 频率分量,深度 3 表示颜色通道,由于生成的光谱图是彩色的,所以有三个 RGB 通道。将磨煤机轴承各状态下产生的 Mel 谱图输入到 CNN 网络中,该网络对各状态下生成的 Mel 谱图进行训练、学习和特征提取,最终学习到能够识别磨煤机轴承各种故障的频谱图,从而完成基于卷积神经网络的磨煤机轴承故障识别。

2.2 数据采集与分析

噪声采集平台主要由电机、发电机和声传感器组成,发电机后面有一个可变负载,将声传感器的采样频率设为 48 kHz,频率响应置于 35 Hz~20 kHz 范围内,电机的额定转速固定为 1445 r/min。

按照国家标准 GB/T 3260-2007《无线电收发设备用电子元件和材料. 磨煤机用轴承》规定,磨煤机磨辊轴承端隙标准应为 0.2~0.4 mm。因此本文采用线切割加工技术,对轴承的内圈和外圈加工 0.5 mm 宽和 3 mm 深的沟槽,对球加工 0.5 mm 宽和 2 mm 深的沟槽,模拟磨煤机轴承故障裂纹。在本次实验中,收集了 5 组轴承噪声信号,每组噪声信号包含一个故障,共有正常轴承、内圈裂纹、外圈裂纹、内外圈裂纹和滚珠裂纹五种故障,轴承故障的安装位置和故障位置如图 2 所示。

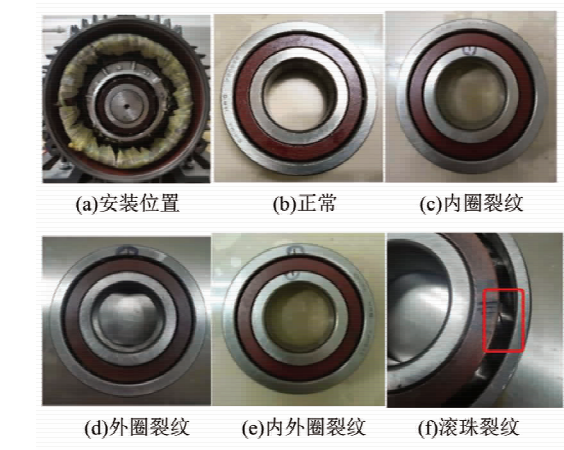


图 2 磨煤机轴承安装位置及五种故障类型

将采集到的 5 种磨煤机轴承类型的噪声数据切割成 1 s 音频文件,共获得 23498 个音频文件作为自制数据集。将不同预载下的声音信号分批转换为 Mel 谱图,并将这些图像分成 8 : 2 的训练集和验证集。

2.3 网络结构

根据样本集的大小和 Mel 谱的特点,设计合适的 CNN 来处理磨煤机轴承故障噪声的声纹特征,CNN 的结构如表 1 所示。

表 1 网络结构参数		
层	输出形状	参数
Rescaling_1	(400,600,3)	0
Conv2D	(400,600,16)	448
Max_pooling2D	(200,300,16)	0
Conv2D_1	(200,300,32)	4640
Max_pooling2D_1	(100,150,32)	0
Conv2D_2	(100,150,64)	18496
Max_pooling2D_2	(50,75,64)	0

2.4 Mel-CNN 模型参数设置

在网络训练过程中,超参数设置如下:损失函数

为稀疏分类交叉熵,mini-batch 大小和 epoch 数分别为 32 和 100,优化器为 adam,学习率为0.0001,ReLU 作为激活函数。结果表明,基于 Mel-CNN 的识别模型在轴承故障诊断中是可行的。

2.5 训练结果

从训练结果图(图 3)中可以看出,在训练集和验证集上都取得了良好的效果,经过 60 次迭代,准确率达到 99.91%,损失接近于 0。因此,该网络的训练效果非常好。

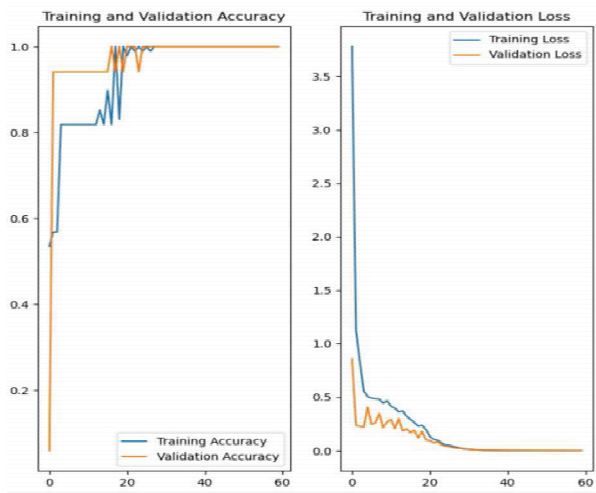


图 3 训练结果

3 实验分析

为了说明 Mel 谱语音特征较强的表征能力,首先比较了 Mel-CNN 模型和 MFCC-CNN 模型的诊断结果,如图 4 所示。

从图 4 可以看出,当网络结构和超参数设置一致时,基于 Mel 谱特征的磨煤机轴承故障声纹识别模型不仅收敛速度更快,精度更高,而且稳定性更强,以上结论验证了基于 Mel-CNN 的识别模型在轴承故障诊断中的可行性。

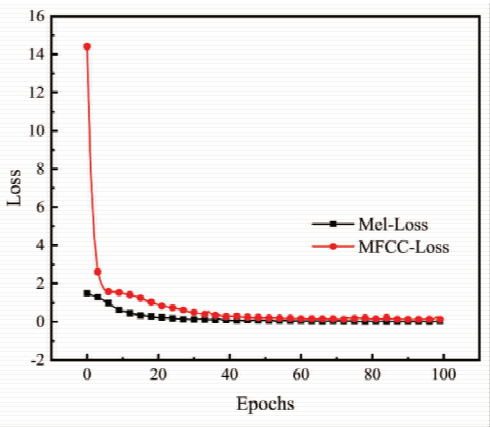


图 4 不同声纹特征识别模型对磨煤机轴承故障识别的准确性

为了进一步说明本文 Mel-CNN 模型的优势,选择了 SVM、RNN、GRU、MobileNet、DenseNet121、VGG16、ResNet50 以及 ACDIN 模型进行对比。模型输入均为本文生成的 Mel 谱样本集,模型测试集的准确率和损失值对比结果如图 5 所示。

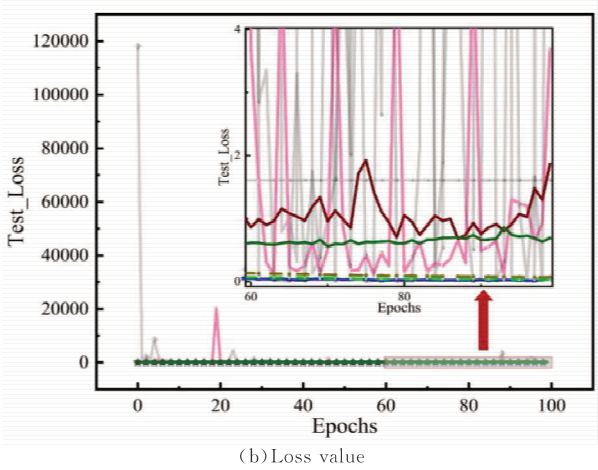
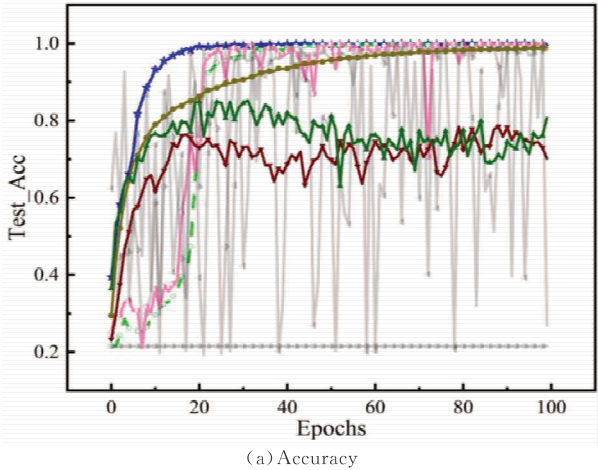
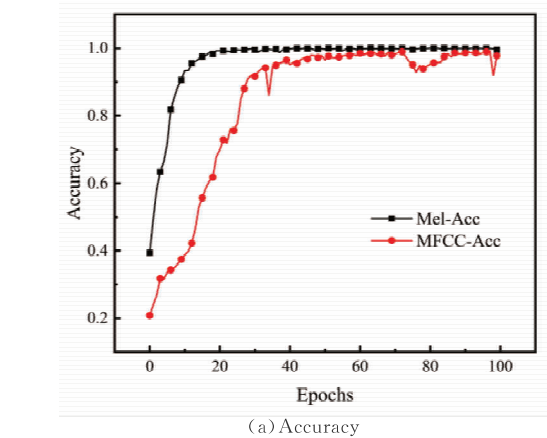


图 5 不同模型识别结果



(a) Accuracy

从图 5(a)所示的模型精度对比可以看出,灰线表示的 DenseNet121、VGG16 和 ACDIN 无法识别磨煤机轴承故障,其中 VGG16 网络模型精度仅为 20%左右,且保持不变,DenseNet121 和 ACDIN 模型不收敛,精度曲线呈现振荡趋势。结合图 5(a)的测试精度和图 5(b)的损失值曲线,本文提出的 Mel-CNN 模型和其他四种模型从第 25 次迭代开始基本稳定,MobileNet 模型自第 50 次迭代以来基本保持稳定,ResNet50 模型的稳定性不如 Mel-CNN 和 MobileNet,但最终测试集的准确率接近 99%。GRU 模型的测试精度稳步增长并最终保持稳定,但低于 Mel-CNN 模型,SVM 和 RNN 模型的精度水平通常低于其他模型,因此这些模型的稳定性也较差。为了全面分析网络模型在对照组中对磨煤机轴承故障的识别性能,对上述模型的相关参数进行比较和分析。详细结果如表 2 所示。

表 2 不同模型比较结果

模型	训练时间	准确率	Loss	模型大小
Mel-CNN	6.8 s	99.57%	0.042	8.92 MB
SVM	38 s	72.18%	0.672	7.81 MB
RNN	23 s	83.56%	0.326	21.47 MB
GRU	9 s	86.44%	0.187	11.32 MB
MobileNet	42 s	99.75%	0.037	32.45 MB
DenseNet121	156 s	98.45%	54.23	80.13 MB
VGG16	26 s	21.58%	1.654	157 MB
ResNet50	75 s	99.02%	3.782	208 MB
ACDIN	32 s	27.71%	16.11	5.8 MB

由表 2 可知,本文设计的 Mel-CNN 模型的识别精度明显优于 GRU、RNN 和 SVM 模型。此外,在识别准确率超过 99%的 Mel-CNN、MobileNet、ResNet50 三种模型中,Mel-CNN 模型在训练时间、模型大小指标等方面具有较大优势,进一步证明了 Mel-CNN 模型在轴承故障声纹识别方面的优势。

为了分析上述模型对五种磨煤机轴承故障类型的识别效果,引入混淆矩阵来表示训练模型对测试集中各种故障样本的识别精度,结果如图 6 所示。

由图 6 绘制的结果可以看出,4 种模型对正常磨煤机轴承类型的识别效果最好,准确率至少达到 99.7%。从图 6(a)可以看出,Mel-CNN 模型对内

外圈裂纹故障的识别效果与其他四种类型相比较差,准确率仅达到 98.17%,且主要是将内外圈裂纹误判为正常类型。从图 6(b)可以看出,MobileNet 模型对 5 种故障类型的平均识别精度较高,但对于滚珠裂纹和外圈裂纹两种故障类型,存在多达 3 种误判类型。从图 6(c)中可以看出,ResNet50 模型对滚珠裂纹轴承类型的识别效果最差,识别准确率仅为 95.92%,且滚珠裂纹多被误判为外圈裂纹类型。图 6(d)的结果表明,SVM 模型对内外圈裂纹、内圈裂纹和外圈裂纹三种磨煤机轴承类型的识别效果相对较差。

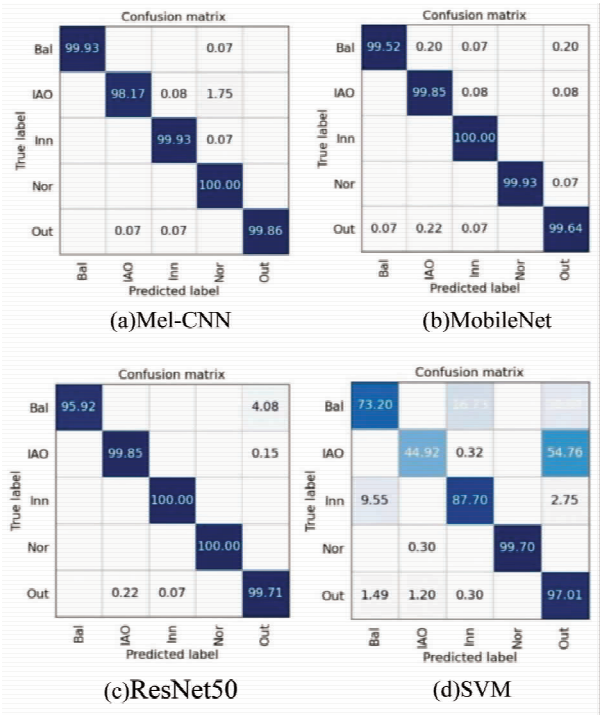


图 6 混淆矩阵

4 结 论

提出了一种基于 Mel-CNN 模型的磨煤机轴承声纹识别方法,以有效检测磨煤机的运行状态。首先利用噪声采集平台模拟了各种轴承故障作用下的噪声信号,提取 Mel 谱特征,然后将磨煤机轴承各状态下产生的 Mel 谱图输入到 CNN 模型中,该模型对各状态下生成的 Mel 谱图进行训练、学习和特征提取。最终实验结果表明,本文 Mel-CNN 模型在训练时间、模型大小指标等方面优于其他模型,在轴承故障声纹识别方面具有较大优势。

(下转第 200 页)