

高压釜泄漏声音的高频高阶空间交互识别算法研究

李衍志, 郭丽敏, 张维国, 古健, 宗井彬, 张凯, 刘君

(中国恩菲工程技术有限公司, 北京 100038)

摘要: 高压釜是湿法冶金领域常用的重要设备, 存在危险气体泄漏的风险; 同时, 泄漏会导致高压釜内压不稳, 严重时甚至引起爆炸, 威胁设备和生产安全; 针对高压釜泄漏导致的危险和威胁, 提出了一种高压釜泄漏声音的高频高阶空间交互识别算法, 用于监测高压釜泄漏发生时的声音, 从而及时发现泄漏并采取相应的措施, 确保设备和生产安全; 该算法首先通过高通滤波器消除低频噪声对于识别结果的干扰, 然后利用递归门控卷积块实现高频分量在高阶空间的交互, 最后使用全卷积层识别高压釜泄漏的声音; 实验结果表明, 所提算法具有较好的高压釜泄漏识别效果, 平均置信度达到 0.93, 以 0.65 作为置信度阈值时, 识别准确率可达到 99.5%; 在处理速度上, 算法能够每秒识别 60 个 5 秒长的音频文件, 满足实时性的需求。

关键词: 湿法冶金; 高压釜; 泄漏声音识别; 递归门控卷积

High Frequency and High-Order Spatial Interaction Recognition Algorithm for Autoclave Leaking Voice

LI Yanzhi, GUO Limin, ZHANG Weiguo, GU Jian, ZONG Jingbing, ZHANG Kai, LIU Jun

(China Enfi Engineering Corporation, Beijing 100038, China)

Abstract: An autoclave is important equipment commonly applied in the field of hydrometallurgy, which has a risk of hazardous gas leaks. Additionally, this leak leads to the pressure unstable in the autoclave, potentially causing explosions that threaten both the equipment and production safety. To address this issue, a high frequency high-order spatial interaction recognition algorithm based on the sound of high-pressure autoclave leaks is proposed. This algorithm is used to monitor the sounds associated with autoclave leaks, thus timely detect leaks, and make a decision to ensure equipment and production safety. Firstly, the low-frequency noise interference of the recognition results is eliminated through a high-pass filter. Then, a recursive gated convolutional block is used to achieve the interaction of high-frequency components in high-order spatial dimensions. Finally, a fully convolutional layer is utilized to recognize the sound of autoclave leaks. Experimental results demonstrate that the proposed algorithm achieves good recognition results for the leakage of the autoclave, with an average confidence level of 0.93. The recognition accuracy can reach up to 99.5%, with a confidence threshold of 0.65. During the processing speed, the algorithm can recognize 60 audio files per second, with each file lasting 5 seconds, meeting its real-time requirement.

Keywords: hydrometallurgy; autoclave; leak voice recognition; recursive gated convolution

0 引言

高压釜是有色金属湿法冶金常用的关键设备^[1-2], 存在危险气体泄漏的风险。高压釜泄漏造成的安全事故也屡见不鲜, 2020年9月25日下午合肥循环经济示范园, 发生高压釜泄漏事件造成2人死亡3人重伤的结果^[3]。2022年4月27日林芝市巴宜区建筑构件产业园, 发生高压釜爆炸事故造成2人死亡6人受伤^[4]。湿法冶金过程中高压釜内常有大量高温高压的流体和有毒有害气体, 如硫酸蒸汽(H_2SO_4), 氢气(H_2), 二氧化硫(SO_2)和氯化氢(HCl)等, 其中 H_2SO_4 蒸汽具有较强的腐蚀性, H_2 密度低还原性极强, 是易燃易爆气体^[5], SO_2 和 HCl 都同时具备酸性和腐蚀性^[6]。由于上述有毒有害气体一旦发生泄漏会造成严

重的安全事故^[7], 因此泄漏检测对于保证现场财产和人员安全具有极其重大的意义。

文献[8]中分析了高压装置的泄漏声音的频率特性, 指出影响泄漏声音识别的主要影响成分 is 高频成分。在湿法冶金过程中, 高压釜内多为高温高压液体, 釜体泄漏后会有气相介质溢出, 发生泄漏时气相介质在高压的作用下冲出釜体从而发出泄漏的声音。这一过程受到高压作用的影响, 导致声音信号中存在丰富的高频成分。高频成分的存在是泄漏声音的显著特征, 而且这些高频成分携带了关于泄漏事件的关键信息。这为通过声音传感器采集声音来识别高压釜泄漏提供了基础。

泄漏声音识别相关方法可以分为基于传统算法的识别

收稿日期: 2023-08-30; 修回日期: 2023-10-12。

基金项目: 国家重点研发计划(2022YFB3304901)。

作者简介: 李衍志(1994-), 男, 硕士。

通讯作者: 郭丽敏(1982-), 女, 博士。

引用格式: 李衍志, 郭丽敏, 张维国, 等. 高压釜泄漏声音的高频高阶空间交互识别算法研究[J]. 计算机测量与控制, 2024, 32(10): 169-174.

方法和基于深度学习的识别方法。传统方法主要根据声音信号的特定的一些时域和频域特征进行分析,需要调整设置每一步分析中的阈值,场景的适应性差,现场使用操作复杂也会受到周围环境噪声的影响^[9]。基于深度学习的声音识别算法^[10]将声音信号转化带有时频域特征的特征图后,通过神经网络来对特征图进行识别,但是没有考虑到声音的高频分量在高阶空间的交互,因此识别精度相对较低。

针对上述问题,本文提出了高压釜泄漏声音的高频高阶空间交互识别算法,其具体改进策略是对高压釜声音数据进行高通滤波后将其转化为声谱图,并通过递归门控卷积实现声谱信息的高阶空间交互,最后通过实验验证了该方法的可行性与优越性^[11-13]。

1 泄漏声音识别算法相关研究

1.1 传统泄漏声音识别算法

高压釜泄漏声音识别任务,旨在将采集到的声音数据分类为有泄漏和无泄漏两种。传统的识别方法主要是根据对原始声音信号或声音信号阵列的某些特征在时域和频域上的分析结果进行判断。例如,李明霞^[9]在 2008 年提出通过传统的特征提取技术消除声音信号的噪声后,将声音信号转化为梅尔谱,并根据梅尔谱特征建立隐马尔可夫模型,最后根据模型输出判定是否发生了泄漏。华科等人^[14]则提出了一种利用短时傅里叶变化,来检测管道泄漏的方法。短时傅里叶变换是一种将信号分解成时间窗内的频率成分的方法。对于时域信号 $x(t)$,其短时傅里叶变换如公式(1)所示:

$$x(t, f) = \int_{-\infty}^{\infty} x(\tau) \cdot \omega(t - \tau) e^{-j2\pi f\tau} d\tau \quad (1)$$

式中, $x(t, f)$ 是在时间 t 和频率 f 上的短时傅里叶变化结果,是一个复数,其中包含了信号在不同时间和频率上的频谱信息,是一个矩阵,通常由实部和虚部组成。 $\omega(t - \tau)$ 是窗口函数,通常使用汉宁窗或高斯窗,用于将原始信号 $x(t)$ 切分成小段。是频率, τ 是窗口函数 $\omega(t - \tau)$ 的时间偏移,表示在时域上移动窗口的位置,以便在不同的时间片段内进行傅里叶变换。 $e^{-j2\pi f\tau}$ 是复指数项,用于执行傅里叶变换,它将信号从时域转换到频域。将信号从时域转换到频域之后,在分析某些特定频率上的信号功率谱和强度,再与一个经验阈值对比就能得出高压釜是否发生了泄漏。

孙静等人^[15]则提出用小波分解消除噪声后,通过分帧加窗的方式来检测音频中的泄漏信号。小波分解的过程中首先将原始信号 $x[n]$ 进行小波分解,得到小波系数 w_i 和尺度参数 a_i ,如公式(2)所示:

$$w_i[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot \psi\left(\frac{k-n}{a_i}\right) \quad (2)$$

式中, $w_i[n]$ 是分解后的小波系数,表示在不同尺度和时间点 n 上的小波分量。 ψ 表示小波基函数。 k 代表时间域中离散时间步长。分解后需要对小波系数进行阈值处理,并以此来抑制噪声信号,阈值处理过程可用公式(3)所示:

$$w'_i[n] = \begin{cases} W_i[n] - T, & \text{if } W_i[n] > T \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

式中, T 是阈值,用于控制哪些小波系数被抑制。最后对处理后的小波系数进行逆变换,以重建噪声消除后的声音信号,重建过程可以用公式(4)表示:

$$X'[n] = \sum_i \sum_n w'_i[n] \cdot \psi\left(\frac{k-n}{a_i}\right) \quad (4)$$

式中, $X'[n]$ 表示经过小波逆变换重建的声音信号。其余符号的含义与公式(2)和公式(3)中对应符号的含义一致。

总的来说,传统算法特征提取能力弱,对于特征提取过程中的阈值依赖大。而高压釜发生泄漏时,其声音信号中各频率成分的分布取决于高压釜本身的结构、高压釜内压力的大小、高压釜内的温度、泄漏位置的形态等多种不可控因素,导致泄漏声音特征提取算法的适用性极差,现场工况和场景变化时需要实时调整不同的阈值。

1.2 深度学习泄漏声音识别算法

数字信号的主要来源包括声音、文字及图像 3 种。深度学习技术在图像领域率先得到了发展,直到人们意识到声谱图也是一种图像的时候深度学习技术逐步进入到声音识别领域^[16]。基于深度学习的泄漏声音识别算法,可以分为基于卷积神经网络 CNN 的识别算法^[17]、基于残差网络 ResNet^[18]的识别算法和基于 Inception^[19]的识别算法。

相对于传统算法,基于 CNN 的声音识别算法,使用卷积神经网络来提取声谱图的特征。CNN 通过卷积层,可以有效地捕捉输入数据中的局部特征。这些卷积核(滤波器)可以学习识别不同的图像特征,如边缘、纹理和形状。此外, CNN 端到端训练方法允许模型自动优化权重参数,无需手工设计特征提取器。这种自动特征学习使得识别精度有了较大的提升^[20]。卷积神经网络层数越深,特征提取能力越强,但同时根据链式求导法则,网络越深梯度越不稳定。在神经网络中,由于网络通常由多个层组成,每一层都包含激活函数和权重,因此梯度的传播和计算需要使用链式求导法则,求导过程如公式(5)所示:

$$\frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dh} \cdot \frac{dh}{dx} \quad (5)$$

式中, $\frac{df}{dg}$ 、 $\frac{dg}{dh}$ 、 $\frac{dh}{dx}$ 分别表示某一层的求导结果。公式(5)只是仅有 3 层的神经网络的求导结果。实际中网络层数,往往会达到几百层甚至上千层。所以当链式求导法则中的梯度相乘时,如果某些梯度值小于 1,且这种情况在多个层中发生,那么梯度会指数级地变小,导致网络无法学习。与梯度消失相反,有时梯度值可能非常大,导致权重更新过大,破坏网络的稳定性。因此在反向传播过程中,梯度值可能变得非常小或非常大,这可能导致训练不稳定性 and 梯度消失或者爆炸问题。

为了解决声音识别任务中梯度不稳定的问题,基于 ResNet 的声音识别算法走入了大众的视野。ResNet 通过引入残差连接解决了深度神经网络中的梯度消失和梯度爆炸问题,薛珊等人^[21]将卷积神经网络优化为 ResNet 网络结构,用于低空无人机声音的识别,有效解决了卷积神经网络梯度消失和梯度爆炸的问题。这种结构的优势在于能够

在非常深的网络结构下进行训练, 加深网络层数可以提升精度, 而不会出现准确率突然下降的情况。Inception 则是同时使用了不同尺度卷积核的卷积神经网络, 通过多尺度并行处理的做法可以在保证模型精度的同时提高训练速度。刘杰等人^[22]通过 Inception 网络来解决管道泄漏声音识别中深度神经网络搭建的问题。Inception 网络结构在保证模型深度的同时, 通过不同层的网络结构来观察不同层次、不同尺度上的特征, 从而适用于大型数据集的训练。

现有基于深度学习的泄漏声音识别算法, 主要是面向如何搭建更深一级的网络增强神经网络的特征提取能力, 但在高压釜内高压的作用下, 发生泄漏时其声音通常以高频为主, 且随着高压釜内压强的变化, 声音的频率和幅值随着改变, 导致现有算法识别效果不佳。

2 高压釜泄漏声音识别算法设计

2.1 高压釜泄漏声音识别方法概述

高压釜泄漏声音识别神经网络结构如图 1 所示。神经网络的输入为高频分量的声谱图, 输出为神经网络预测的类别与置信度。在图 1 中, 识别神经网络前几个立方体模块为本文新提出的 gnBlock 模块, 其主要用于对输入的声谱图进行特征提取, 实现声音信号中高频分量的高阶空间交互。Conv 则表示卷积层, 对 gnBlock 输出的张量进行进一步的特征提取, 降低张量维度为全连接层进行声音的分类识别做准备。FullConv 表示全连接层模块, 该部分由两个全连接层实现, 第一个全连接层用于进一步降低数据维度输出 1×10 大小的张量, 第二个全连接层用于识别采集到的声音是属于发生了泄漏的类别 y_1 , 还是属于未发生泄漏的类别 y_2 。输出中的 c_1 和 c_2 则表示分类类别的置信度。置信度是泄漏声音判别结果识别的分类准确性的依据。泄漏声音识别任务中, 如果 c_1 和 c_2 都小于一定的阈值, 可以认为识别失败, 反之, 判定为识别成功, 并取置信度大的预测类别, 作为声音的识别结果。

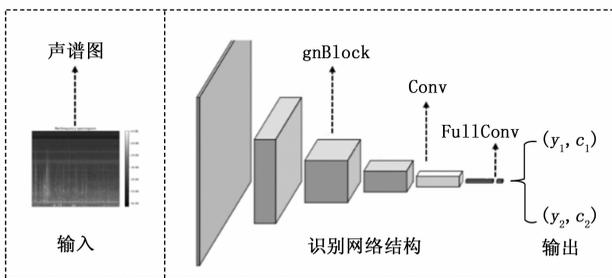


图 1 高压釜泄漏声音识别神经网络结构图

2.2 高频声音信号声谱图特征提取

高压釜正常工作无泄漏声音时, 声音传感器采集到的声音信号频谱如图 2 所示。图 2 中横坐标是声音信号的频率, 纵轴表示声音强度。由于高压釜周围还存在其他机械设备工作, 因此不可避免地存在相关的环境噪声。环境噪声的性质和频谱分布通常受到多种因素的影响, 如附近设备的类型和状态、周围环境的温度和湿度等。由于这些变

化, 不同场景下的环境噪声频谱分布可能会有所不同。在从环境噪声中准确提取高压釜泄漏的声音信号时, 采用高通滤波技术的主要目的是去除或削弱低频噪声, 从而突出高频成分, 这是因为高压釜泄漏声音通常表现为高频信号。高通滤波器允许通过频率高于某个特定阈值的信号, 而抑制低频信息的干扰。这种方式有效地消除了环境中的低频噪声, 使得声音信号更加清晰。

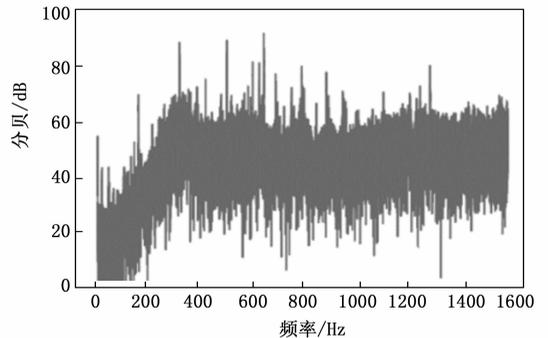


图 2 环境噪声频谱图

从图 2 中可以看出环境噪声的频率基本集中在低频段。经过对多个不同的高压釜周围环境噪声的采集和分析得出, 环境噪声基本频率基本都低于 1 000 Hz, 而泄漏声音频率通常都处于 1 000 Hz 以上。高通滤波器是数字信号处理中常用的滤波方法, 能够有效抑制声音传感器采集到的低频声音信号, 其频率响应函数如公式 (6) 所示:

$$|H(\omega)|^2 = \frac{1}{1 + \left(\frac{\omega}{\omega_c}\right)^{-2n}} \quad (6)$$

式中, ω 为频率, ω_c 为截止频率, n 为巴特沃斯滤波器的阶数, 阶数越高对于低频信息的抑制作用越好。为了避免出现漏检, 通过对泄漏声音的频率进行统计后发现其范围比较广, 实际应用中将 ω_c 定为 700 Hz。图 2 所示的环境噪声高通滤波后的频谱如图 3 所示, 包含泄漏信号的声音数据, 高通滤波后的频谱如图 4 所示。从图 2 与图 3 的对比中可以看出 700 Hz 以下的低频分量得到了有效地抑制。从图 3 与图 4 的对比中可以看出高频段中, 无泄漏状态信号的幅值明显低于有泄漏信号的幅值, 这也是做声音识别的基础。

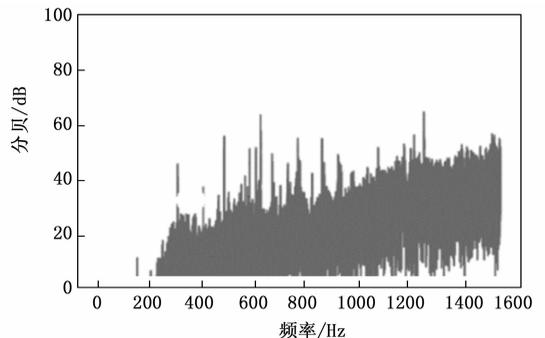


图 3 环境噪声高通滤波后的频谱

经过高通滤波之后, 低频段的声音信号得到了抑制, 只保留了高频部分的信息。现实场景中, 泄漏的声音同时具有

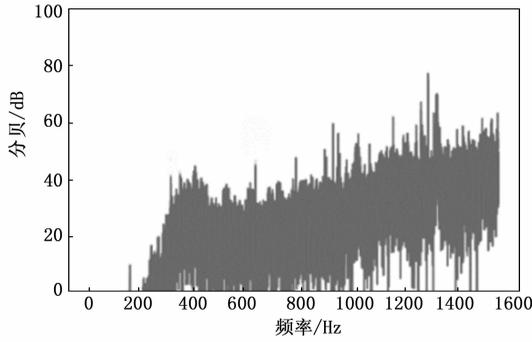


图 4 泄漏声音高通滤波后的频谱

时域和频域特性，声谱图同时具备了声音信号的时域特性和频域特性，是声音识别中常用的特征提取方式^[23]。除此之外也有研究者提出人耳对于高频声音的分辨程度低，据此提出了梅尔谱阶梯式的对高频分量进行处理^[24]。高压釜泄漏中主要任务就是对高频分量的特征进行提取和学习，选择原始的声谱图作为声音识别的特征图更符合实际情况。高压釜泄漏声音样本的声谱如图 5 所示，颜色表示声音强度的分贝值。从图 5 中可以看出，高通滤波后生成的有泄漏样本声谱图中，高频分量的分贝值大于低频分量的分贝值。

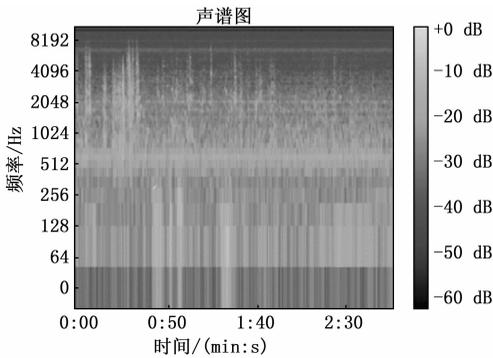


图 5 有泄漏样本声谱图

2.3 高压釜泄漏声音识别网络结构设计与优化

高压釜泄漏声音识别神经网络结构如图 1 所示。首先，声音信号经过预处理和特征提取步骤，得到了声谱图特征，这些特征主要包含了高频分量的信息。接下来，通过本文提出的 gnBlock 模块对高频分量信息进行高频特征提取，实现声音信号高频特征在高阶空间的交互。最后，通过全连接层进行分类预测，输出二分类的结果，即是否存在高压釜泄漏声音。本文提出的门控卷积块 gnBlock 与递归门控卷积的结构如图 6 所示。

门控卷积块 gnBlock 首先通过递归门控卷积对输入的张量 tensor 进行高阶空间交互特征提取，再通过 BN 层进行数据的归一化，解决梯度不稳定的问题避免出现梯度消失和梯度爆炸的问题，最后通过激活层激活后进行池化操作和卷积操作，进一步地对特征进行提取后输出提取到的特征给下一级。门控类似于循环神经网络 (RNN) 中的门控单元，它有助于模型更好地捕捉序列数据中的长期依赖关系，能实现声谱图中高频分量的交互。BN 层的作用是使数据的

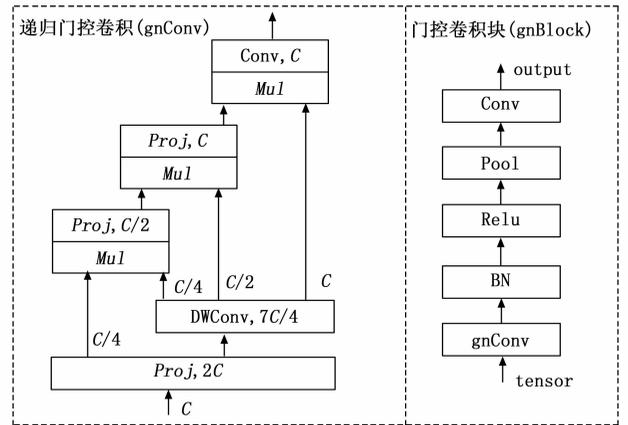


图 6 门控卷积块与递归门控卷积结构图

分布更加稳定，有助于加速模型的训练。

相对于 CNN、ResNet、Inception 这 3 种结构门控卷积块，gnBlock 除了通过 BN 层解决了梯度问题外，还通过递归门控卷积实现了声谱信息的高阶空间交互。神经网络结构中通过张量来表示声音信号的特征信息，实现张量信息在高阶空间的交互，其本质就是实现张量信息在高阶通道上的交互^[25]。一个张量同时具有宽度信息和通道信息，宽度信息指张量的宽和高，通道信息作为维度信息。因此实现滤除低频分量的声谱图在高阶通道上的交互，就实现了声音信号的高频高阶空间交互。

在泄漏过程中高压釜内压力变化后声音频率会发生变化，高频信号在高阶空间上的交互相对于在低阶空间上更容易捕获到这种变化的特征。本文尝试通过递归门控卷积在高阶空间上来学习这种变化的特征，增强神经网络的特征提取能力。如图 6 所示， C 表示高频声谱特征张量的通道数， $Proj$ 表示通过卷积层将张量的通道数扩大一倍， Mul 表示两个通道上的特征值相乘。高频声音信号在高阶空间的交互过程是：1) 将输入张量的通道数首先扩大一倍；2) 将扩大后的张量按照通道数进行划分，划分为 $1/4C$ 和 $7/4C$ 通道数的两个张量；3) 使用深度可分离卷积对 $7/4C$ 通道数的张量进行特征提取；4) 将提取后的特征在通道上划分为 $1/4C$ 、 $1/2C$ 、 C 共 3 个张量，然后两个 $1/4C$ 通道数的张量相乘后通过 $Proj$ 操作扩大其通道数；5) 按照图 6 所示的结构分别对 $1/2C$ 和 C 大小的张量相乘，并通过 $Proj$ 操作实现通道的变化，最后通过卷积层输出提取的高阶交互后的高频声音特征。

2.4 高压釜泄漏声音识别损失函数设计

高压釜泄漏声音识别任务中，损失函数用于指导模型优化。通过最小化损失函数，模型的参数被调整，以使模型在训练数据上的性能不断提升。这有助于模型更好地拟合训练数据，提高语音识别的准确性。高压釜泄漏声音识别中采用 CTC (Connectionist Temporal Classification) 损失函数，CTC 可以与深度神经网络结合使用，这些网络可以自动地学习声学特征与标签之间的映射关系。这使得模型可以在大规模数据集上进行端到端训练，提高了语音识

别的准确性。其表达式如公式 (7) 所示:

$$L(y, \hat{y}) = - \sum_{t=1}^T \log P(\hat{y}_t | y) \quad (7)$$

式中, $L(y, \hat{y})$ 为模型预测的标签序列 \hat{y} 与真实标签序列 y 之间的损失; T 为输入语音信号的时间步数, 也可以看作是声学特征序列的长度, 高压釜泄漏声音识别项目中为 5 s。 $P(\hat{y}_t | y)$ 是在给定真实标签序列的条件下, 在时间步 t 预测的标签为 \hat{y}_t 的概率, 这个概率是神经网络模型直接输出的。

3 高压釜泄漏声音识别实验

3.1 实验环境

实验的场地为中国恩菲偃师研发基地, 实验设备为湿法冶金所用的高压釜。高压釜在实验过程中发挥着关键作用, 通过其高压和高温条件, 有助于研究和开发新的材料和工艺。这些实验涉及了金属冶炼、材料合成及其他重要领域的研究。然而, 在实验过程中, 高压釜并不是唯一工作的设备。这些设备包括搅拌器、冷却系统、加热元件等, 它们各自发出特定的声音, 构成了实验室中的背景音。偃师研发基地, 高压釜容积为 2 000 L, 直径为 1 400 mm, 为科学研究提供了广泛的实验空间。在实验过程中, 高压釜能够稳定达到 230 °C 的高温, 相应的压力维持在 2. 8 MPa 左右, 创造了理想的高温高压实验条件。与此同时湿法冶金过程中涉及的多种工艺, 如反应、分离、过滤、清洗和干燥等, 也与实际生产过程一致。因此偃师研发基地的场景包括各种设备和工艺都是与实际工况一致的。然而, 值得注意的是, 虽然高压釜的内部温度和压力在实验和实际生产中保持一致, 但容积方面存在一些微小差距。在工厂中, 高压釜的容积主要为 2 000 L, 但也有一部分釜体容积较大, 达到 5 000 L, 或较小, 为 1 000 L。这种差异反映了不同生产场合的需求多样性。总的来说, 偃师研发基地的实验工况与实际工业生产场景高度一致, 因此偃师研发基地的测试效果在很大程度上能代表所提方法在实际工况下的检测的效果。

实验平台为一台强大的 64 位 Windows 系统, 具备出色的计算性能和内存容量。该系统搭载了高性能的 Intel (R) Xeon (R) Silver 4210R CPU @2.40 GHz 处理器, 系统内存大小达到了 56 GB, 为实验提供了充足的内存资源, 有助于处理大规模的数据集或运行内存密集型的应用程序。此外, 系统还配备了一块强大的 GPU, 即 NVIDIA GeForce RTX 3080Ti, 拥有 24 GB 的显存。

3.2 数据集介绍

实验过程中涉及有泄漏的声音数据和无泄漏的声音数据两类。由于泄漏是偶发事件, 因此现场实地采集到的有泄漏的真实数据不够充分。为了扩充深度学习所需的样本数据, 在现场长期使用高压釜的专家的指导下, 通过模拟的方式在不损害设备的前提下扩充了有泄漏的声音数据。

高压釜泄漏声音模拟主要包括升高高压釜内压到正常工作水平后对一些阀门做有限程度的扭动, 以及对一些结构类似的储气罐加压充气后释放其中的气体两种方式。结合测试

现场实际条件, 声音传感器的工作距离为 1.5~1.8 m。通过声音传感器采集到有泄漏无泄漏数据之后以 5 s 为时间单位分类分段将声音数据存储为 wav 文件。所有的无泄漏声音样本都是在现场实地采集得到, 最终共获得 20 000 个有泄漏样本和 20 000 个无泄漏样本, 均来自不同的场景和条件, 其中 30% 的样本用于测试, 70% 的样本用于训练。

3.3 评价指标

声音识别任务中常用评价指标是精度 (Precision), 其计算方法如公式 (8) 所示:

$$precision = \frac{TP}{TP + FP} \quad (8)$$

式中, TP 为判别正确的泄漏声音片段数量, FP 为判别错误的无泄漏片段样本数量。

3.4 泄漏声音识别实验参数设置及结果分析

训练和测试过程中, 每个音频文件的时间长度都设置为 5 s, 巴特沃斯高通滤波器的截止频率设置为 700 Hz, 巴特沃斯高通滤波器的阶数设置为 10, 高频分量的声谱图分辨率为 320×320 个像素。训练过程中学习率设置为 0.000 1。如上文所述所提的算法会输出两个类别预测结果及其对应的置信度值。通过对测试数据的统计分析, 本文所提算法平均置信度达到 0.93, 90% 以上的数据置信度在 0.8 以上。现场泄漏事件发生的概率通常相当低, 但另一方面, 周围的机械噪声可能会引入干扰, 增加了误报的风险。因此, 为了确保准确性和可靠性, 必须采取适当的措施来降低误报的可能性。在这种情况下, 综合考虑算法输出和现场情况, 将置信度阈值设置为 0.65。将阈值设置为 0.65, 确保了对预测的严格验证。具体来说, 只有当算法的输出中至少有一个结果的置信度高于这个阈值时, 才会认为预测是有效的。这种策略有助于避免过于敏感, 减少误报的概率。在预测被确定为有效的情况下, 进一步采用了一种智能策略来选择最终的识别结果。具体而言, 会比较算法输出的两个结果的置信度值, 并选择置信度较高的结果作为最终的识别结果。这种方式充分考虑了算法输出的不确定性, 确保了最终结果的可信度。

为了验证本文提出的算法在高压釜泄漏声音识别任务中的优越性, 进行了对比实验分析。所有参与对比的算法与本文算法输入都是 320320 个像素大小的声谱图, 参与对比的算法有基于 CNN 的算法^[11]、基于 ResNet 的算法^[21]、基于 Inception 的算法^[22]、基于 SVM 的算法^[23], 实验结果如表 1 所示, 表中速度列是指每秒识别时长为 5 s 的声音文件的个数。

表 1 各算法声音识别精度对比

| 算法 | 精度/% | 速度/(个/秒) |
|-----------|------|----------|
| CNN | 96 | 50 |
| ResNet | 97.3 | 45 |
| Inception | 97.6 | 58 |
| SVM | 95 | 20 |
| 本文算法 | 99.5 | 60 |

从表 1 中可以看出, 基于 SVM 的算法其识别精度最低且处理速度最慢, 其原因在于 SVM 算法的本质是将声谱图提升到更高维度的空间, 并在高维度空间区分两类数据, 数据越复杂涉及的场景越多就会用到越高的维度和空间, 处理速度和精度也因此受到限制。基于 ResNet 和基于 Inception 的算法解决了梯度弥散的问题, 把网络做得更深特征提取能力更强, 所以这两个算法的精度相对基于 CNN 的算法更高。本文的高压釜泄漏声音的高频高阶空间交互识别算法虽然在处理速度与基于 Inception 的算法基本持平, 但是精度上高于基于 Inception 算法的 1.9%, 算法精度的提升也印证了本文所提算法的有效性。通过对数据集中识别结果不准确的数据进行分析之后发现部分泄漏声音轻微的数据容易被识别成无泄漏数据。从表 1 中也可以看出, 本文算法能够在高精度的同时保持较快的处理速度, 为工业安全和环境监测提供了可靠的解决方案。

4 结束语

本文提出了一种高压釜泄漏声音的高频高阶空间交互识别算法, 并将其应用于中国恩菲偃师研发基地, 由对现场应用情况和实验数据的统计分析, 可以得出如下结论:

1) 所提算法对声音传感器采集到的湿法冶金常用设备高压釜泄漏声音信号进行实时识别, 能够帮助相关工作人员及时发现和处理高压釜异常泄漏, 保障生产安全;

2) 通过新提出的 gnBlock 模块能够有效地实现声音的高频信号在高阶空间交互, 增强神经网络的特征提取能力, 且识别精度达到 99.5%, 满足现场的实际需求;

3) 用来测试的数据集中模拟泄漏的数据也能被准确地识别, 说明了本文所提算法也适用于其他压力容器的泄漏检测。

参考文献:

- [1] 魏焕起. 对有色金属湿法冶金中高压釜设计的探讨 [J]. 有色设备, 1995 (3): 26-30.
- [2] 皮关华. 关于高压釜结垢原理分析及控制措施 [J]. 有色设备, 2018, 185 (6): 64-68.
- [3] 大皖新闻. 合肥循环经济示范园发生泄漏起火事故致 2 死 3 伤涉事企业道歉 [EB/OL]. (2023-06-27)[2023-11-24]. <https://baijiahao.baidu.com/s?id=1678871795671136755&wfr=spider&for=pc>.
- [4] 特种设备安全. 林芝市安元建材有限公司“4·27”蒸压釜爆炸事故调查报告 [EB/OL]. (2023-06-27)[2023-11-25]. <https://www.163.com/dy/article/I5C8T2M105149PS1.html>.
- [5] 龚皓, 李德前. 氢气的制取与性质实验的系列化设计 [J]. 中国现代教育装备, 2022 (22): 79-80.
- [6] 别如山, 樊庆铎. 喷消石灰脱除垃圾焚烧炉尾气中 HC_1 和 SO_2 的试验研究 [J]. 环境污染治理技术与设备, 2005, 6 (12): 69-72.
- [7] 郭丽敏, 张维国, 古健, 等. 基于机器人巡检的矿冶领域四足机器人开发研究 [J]. 有色设备, 2022, 36 (6): 30-35.
- [8] XU T, GENG M, LIANG C. Long distance large diameter

- heating pipeline leakage detection based on acoustic pressure sensor [C] //2020 Chinese Automation Congress (CAC). IEEE, 2020: 2289-2294.
- [9] 李明霞. 基于声音识别的新型输油泵阀门泄漏检测原型系统 [EB/OL]. (2023-06-28)[2023-11-25]. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201401&filename=1014124998.nh>.
- [10] 陈秋菊, 徐建国. 优化正交匹配追踪和短时谱估计用于声音识别 [J]. 计算机工程与应用, 2020, 56 (7): 162-169.
- [11] 俞颂华, 王汝凉. 基于 CNN 与 MFCC 的城市场景声音识别 [J]. 广西师范学院学报 (自然科学版), 2019, 36 (1): 50-56.
- [12] AZARANG A, HANSEN J, KEHTARNAVAZ N. Combining data augmentations for CNN-based voice command recognition [C] //2019 12th International Conference on Human System Interaction (HSI). IEEE, 2019: 17-21.
- [13] SUPARATPINYO S, SOONTHORNPHISAJ N. Smart voice recognition based on deep learning for depression diagnosis [J]. Artificial Life and Robotics, 2023: 1-11.
- [14] 华科, 叶昊, 王桂增, 等. 基于声波的输气管道泄漏检测与定位技术 [J]. 华中科技大学学报 (自然科学版), 2009, 37 (s1): 181-183.
- [15] 孙静, 聂士明, 史宝军. 基于小波和经验模态分解的雾气泄漏声音端点检测算法 [J]. 河北工业大学学报, 2022, 51 (6): 10-18.
- [16] 李永, 范雪, 杨鸿波. 声谱图在汉语普通话声调识别中的应用 [J]. 信息通信, 2017, 175 (7): 89-92.
- [17] YUAN X W. Design of voice recognition acoustic compression system based on neural network [J]. Wireless Personal Communications, 2022, 127 (3): 2121-2139.
- [18] WU Z, SHEN C, VAN DEN HL A. Wider or deeper: revisiting the resnet model for visual recognition [J]. Pattern Recognition, 2019, 90: 119-133.
- [19] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Venice, 2016: 2818-2826.
- [20] 王若平, 房宇, 陈达亮, 等. 基于小波包 FBank 图谱与 CNN 的车内异响识别研究 [J]. 重庆理工大学学报 (自然科学), 2020, 34 (7): 1-9.
- [21] 薛珊, 卫立炜, 顾宸瑜, 等. 基于改进残差网络的低空无人机声音识别方法 [J]. 长春理工大学学报 (自然科学版), 2022, 45 (4): 100-107.
- [22] 刘杰, 朱正伟. 基于稀疏轻量卷积神经网络的管道泄漏检测 [J]. 电子测量技术, 2022, 45 (19): 131-135.
- [23] 孙继平, 余星辰, 王云泉. 基于声谱图和 SVM 的煤矿瓦斯和煤尘爆炸识别方法 [J]. 煤炭科学技术, 2023, 51 (2): 366-376.
- [24] 赵扬青, 彭智才, 蒋雨涵, 等. 音频的梅尔频率倒谱系数特征抽取过程 [J]. 信息技术与信息化, 2023 (1): 104-111.
- [25] RAO Y, ZHAO W, TANG Y, et al. Hornet: efficient high-order spatial interactions with recursive gated convolutions [J]. Advances in Neural Information Processing Systems, 2022, 35: 10353-10366.