

基于改进 MFCC-OCSVM 和贝叶斯优化 BiGRU 的 GIS 异常工况声纹识别算法

庄小亮¹, 李乾坤¹, 刘紫罡², 张禄亮², 季天瑶², 张长虹³

(1. 南方电网超高压输电公司, 广州 510663; 2. 华南理工大学电力学院, 广州 510640; 3. 南方电网超高压输电公司电力科研院, 广州 510663)

摘要: 为了准确识别气体绝缘开关柜(gas insulated switchgear, GIS)设备的异常工况, 提出了一种基于加权梅尔频率谱系数单类支持向量机(Mel frequency cestrum coefficient-one class support vector machine, MFCC-OCSVM)和贝叶斯优化的门控循环单元(bidirectional gate recurrent unit, BiGRU)声纹识别算法。首先, 利用基于F统计量的MFCC对声纹数据进行加权特征提取, 突出重要特征并减弱噪声的影响, 然后利用OCSVM对加权后的特征进行异常检测并去除异常值, 提高数据质量。为解决样本不平衡问题, 采用合成少数类过采样技术(synthetic minority over-sampling technique, SMOTE)进行声纹样本的均衡。最后, 应用基于贝叶斯优化的BiGRU模型进行声纹识别。以某气体绝缘全封闭组合电器(gas insulated switchgear, GIS)为例, 采集了20类不同工况下操纵机构的声音样本, 与多种经典分类模型进行对比。结果显示, 所提算法取得的最高平均识别准确率达到92.8%, 相比于自适应增强、朴素贝叶斯和线性判别分析算法分别提升了30.1%、14.7%和11.5%。通过消融实验进一步评估和验证了所提算法各个流程对声纹识别的实际效果和性能影响, 研究成果可为GIS设备异常工况的声纹识别提供高效技术路线。

关键词: GIS设备; 梅尔频谱倒谱系数; 单类支持向量机; 双向门控循环单元; 声纹识别; 贝叶斯优化

Voiceprint Recognition Algorithm for GIS Anomalous Conditions Based on Improved MFCC-OCSVM and Bayesian Optimized BiGRU

ZHUANG Xiaoliang¹, LI Qiankun¹, LIU Zigang², ZHANG Luliang², JI Tianyao², ZHANG Changhong³

(1. EHV Power Transmission Company, CSG, Guangzhou 510663, China; 2. School of Electric Power Engineering, South China University of Technology, Guangzhou 510640, China; 3. Electric Power Research Institute of EHV Power Transmission Company, CSG, Guangzhou 510663, China)

Abstract: To accurately identify abnormal conditions in gas insulated switchgear (GIS) equipment, a voiceprint recognition algorithm is proposed based on weighted Mel frequency cestrum coefficient-one class support vector machine (MFCC-OCSVM) and Bayesian optimized bidirectional gate recurrent unit (BiGRU). Firstly, weighted extractions of voiceprint data are performed using MFCC based on the F-statistic, highlighting important features and reducing the influence of noise. Subsequently, OCSVM is utilized to detect anomalies and remove anomalous values from the weighted features to improve data quality. To address the issue of sample imbalance, synthetic minority over-sampling technique (SMOTE) is employed to balance voiceprint samples. Finally, voiceprint recognition is carried out using a BiGRU model based on Bayesian optimization. Taking a certain GIS equipment as an example, sound samples from 20 different operating conditions are collected and compared with various classical classification models. The results demonstrate that the proposed algorithm achieves the highest average recognition accuracy of 92.8%, resulting in improvements of 30.1%, 14.7% and 11.5% compared to adaptive boosting, Naïve Bayes, and linear discriminate analysis, respectively. Ablation study further assesses and validates the practical effects and performance impacts of each process in the proposed algorithm. Research results provide an efficient technical approach for voiceprint recognition of anomalous conditions in GIS.

Key words: gas insulated switchgear (GIS) equipment; Mel frequency cestrum coefficient (MFCC); one-class support vector machine (OCSVM); bidirectional gate recurrent unit (BiGRU); voiceprint recognition; Bayesian optimization

0 引言

气体绝缘开关设备(gas insulated switchgear, GIS)在现代电力系统中扮演着至关重要的角色,用于实现电能的分配、控制和保护^[1]。然而, GIS设备在长期运行过程中会面临各种异常工况和潜在的故障风险,这可能对电力系统的稳定性和可靠性造成严重影响^[2-7]。因此,必须对GIS设备开展在线监测和故障诊断,以确保其正常运行并减少潜在的停电风险^[8]。

传统的GIS设备监测方法通常基于接触式传感器,这些传感器可以提供有关设备温度、湿度和电流等物理参数的信息。然而,这种监测方法存在一些局限性,包括安装和维护的复杂性以及仅提供有限信息的问题^[9]。为了克服这些限制,非接触式监测方式如图像、声音、红外的方式应运而生^[10-12],其中声音方式成本最低,尤其适用于GIS等机械振动类设备。

声纹识别技术最早是在声音领域得到广泛应用,用于识别和验证个体的身份。声纹识别通过分析人的语音特征,如声调、音调、语速等,来识别和验证一个人的身份。目前,提取声音信号特征的方法主要分为线性预测倒谱系数(linear prediction cepstrum coefficient, LPCC)和梅尔频率倒谱系数(Mel frequency cepstrum coefficient, MFCC)等^[13-16]。其中, MFCC因其出色的提取性能已经成为声纹识别的主流选择。近年来,声纹识别技术逐渐被应用于电力设备的监测和诊断,尤其在特征提取和工况识别等方面已经取得了显著的成果。文献[17]将梅尔-格拉姆角差场时频谱图与ConvNeXt-T网络相结合,在变压器铁芯松动故障识别中达到了99.5%的准确率。文献[18]利用MFCC与两级SVM相结合,有效地检测出陌生样本,以提高在多噪声环境下识别的准确率。文献[19]采用主成分分析算法对GIS局部放电信号声纹特征进行降维优化,随后构建了SVM分类器对优化特征进行识别,提高了GIS绝缘缺陷的诊断准确性。

值得注意的是, MFCC最初是基于人声感知提

出的,而GIS设备操纵机构的振动声与人声有较大差异,且在采集中受到环境噪声的干扰。因此,直接将MFCC应用于GIS声纹工况识别准确率较低,需要进一步提取特征和降低噪声干扰。此外,在实际情况下每类GIS工况数据分布存在不均衡。某些工况数据可能比其他工况更常见,导致模型识别罕见工况时性能下降,因此须进行数据平衡处理。

为解决上述问题,本文采用了一系列方法来改进GIS异常工况声纹识别,为其提供新思路。首先,基于F统计量对MFCC特征进行了进一步加权提取,这有助于突出重要特征并减弱噪声的影响。接着,利用单类支持向量机(one-class support vector machine, OCSVM)算法对加权后的特征进行异常检测并去除异常值,从而提高了数据的质量。为了解决样本不平衡的问题,采用了合成少数类过采样技术(synthetic minority over-sampling technique, SMOTE)来均衡样本。最后,应用了基于贝叶斯优化的双向门控循环单元(bidirectional gate recurrent unit, BiGRU)模型来进行声纹识别,从而获得更高的准确率。本文以ZF23-126型GIS设备断路器为对象采集了20类不同的工况声音,与多种经典的分类模型进行对比以验证本文算法的优越性。

1 GIS异常工况声纹识别算法

1.1 声纹识别流程

GIS异常工况声纹识别流程包括特征提取和算法识别两个核心部分,特征提取旨在提取最具代表性的信息,算法识别则通过算法来确定声纹数据所属的具体类别。针对GIS声纹中噪声干扰及数据不平衡问题,本文提出了一种改进的GIS异常工况声纹识别算法,包括特征加权、噪声去除、数据平衡等,以提高识别准确性,如图1所示。

具体来说,首先对采集到的GIS声纹信号进行预处理操作,然后基于F统计量提取加权的MFCC特征,接着用OCSVM进行异常值检测,然后采用SMOTE进行样本均衡操作,最后基于贝叶斯优化的BiGRU模型进行异常工况的声纹识别。

基金项目:国家自然科学基金资助项目(52077081);中国南方电网有限责任公司科技项目(CGYKJXM20220069)。

Foundation item: Supported by the National Science Foundation of China(52077081);the Science and Technology Project of China Southern Power Grid Co., Ltd. (CGYKJXM20220069).

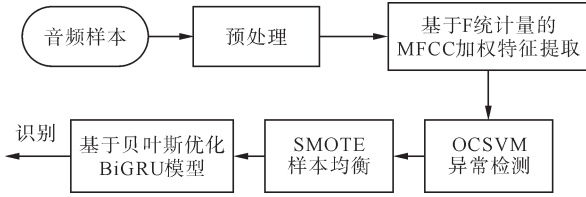


图1 声纹识别流程

Fig. 1 Flow chart of voiceprint recognition

1.2 预处理操作

在进行特征提取和算法识别之前通常需要采用一些预处理方法，如分帧和加窗。

为了保持信号分帧的连续性，在相邻两帧信号之间通常会引入一定的重叠。这种帧间重叠关系可以表示为：

$$f = (N - S_{\text{frame}}) / L_{\text{frame}} + 1 \quad (1)$$

式中： f 为帧数； S_{frame} 为帧移； L_{frame} 为每帧长度； N 为输入语音信号长度。

MFCC的特征提取通常会以20~30 ms的音频帧信号作为样本，这更符合语音信号能量快速变化的特点。针对GIS设备声音信号能量变化较为平稳的特点，本文选择 S_{frame} 为40 ms， L_{frame} 为20 ms，每个输入样本持续400 ms，帧之间重叠率为40%。

在分帧后，通常会将每个帧乘以一个窗函数（本文采用汉明窗），以减少信号在窗口边界处的不连续性，避免频谱泄漏和振铃效应等问题。汉明窗的计算公式如式(2)所示。

$$h(n) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right), & N \leq n \leq N-1 \\ 0, & \text{其他} \end{cases} \quad (2)$$

式中： N 为汉明窗口的长度； n 为第 n 个采样点。

1.3 基于F统计量的MFCC加权特征向量提取

MFCC是用于将输入信号从实际频率域转换为Mel频率域的倒谱参数。其基本思想是提取短帧的线性频谱，使用三角滤波器组将线性频谱映射到Mel频谱，然后进行对数变换和离散余弦变换^[20]。

线性频谱是通过快速傅立叶变换(fast Fourier transform, FFT)获得的。Mel频率与实际频率之间的关系可以近似表示如式(3)所示^[21]。

$$M(f) = 25951 \lg\left(1 + \frac{f}{700}\right) \quad (3)$$

式中： M 为Mel频率函数； f 为音频实际频率。

根据奈奎斯特采样定律，滤波器的最大频率应设置为采样频率的一半。中心频率与Mel滤波器之

间的关系可以表示如式(4)所示^[19]。

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

式中 $f(k)$ 为线性频谱中第 k 个滤波器的中心频率。

本文使用了52个三角滤波器进行特征提取。

最后，使用对数变换和离散余弦变换来计算倒谱系数^[19]。

$$c_n = \sum_{j=1}^{26} \lg(S_j) \cos\left((j-0.5) \frac{\pi n}{26}\right) \quad (5)$$

式中： S_j 为第 j 个Mel三角滤波器的输出； c_n 为分帧信号的第 n 维Mel倒谱系数。在特征提取过程后获得了每帧的26维Mel倒谱系数。

通常情况下，大部分分量对声纹识别贡献较大，因为它们包含了关于声音的关键信息。然而，也存在一些分量，其贡献相对较小，甚至可能对识别率产生负面影响，因为它们可能包含噪声或不相关的信息。对于GIS设备声纹信号，由于其特殊性质和噪声干扰，某些MFCC分量可能会受到其他更大的干扰，从而降低识别的准确性。为了克服这个问题，本文引入了基于F比的加权优化方法。

F比是一种统计量，可以用于衡量不同特征分量之间的差异和重要性。F比的数值越大表示特征在不同组(或类别)之间的方差差异较大，这意味着该特征对于区分不同组的能力较强，或者说对于任务的贡献较大。对于第 n 维MFCC分量的F比公式可以表示为^[22]：

$$F = \left[F_{\text{组内}} \right] / \left[F_{\text{组间}} \right] \quad (6)$$

$$F_{\text{组内}} = \frac{1}{N} \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^{n_i} (u(i, j) - u(i))^2 \quad (7)$$

$$F_{\text{组间}} = \frac{1}{N} \sum_{i=1}^N (u(i) - u)^2 \quad (8)$$

式中： F 为F比； $F_{\text{组内}}$ 和 $F_{\text{组间}}$ 分别为组内和组间方差； N 为样本总数量； n_i 为第 i 个样本中的帧数； $u(i, j)$ 为第 i 个样本中的第 j 个维度上的MFCC向量； $u(i)$ 为第 i 个样本所有维度上的MFCC向量的平均值； u 为所有样本的MFCC均值。

接着，用F比加权对原有的MFCC特征进行处理，高F比值特征将更加突出GIS设备声纹中的重

要特征,低F比值可以减少噪声和干扰,从而提高声纹特征的有效性和准确性。

1.4 基于OCSVM的异常检测

为了进一步增强加权MFCC特征的可分性,提高模型的性能和稳定性,通过引入OCSVM进行异常序列检测和去除异常点,从而净化数据集,提高声纹识别的准确性。

OCSVM是一种通过正样本来解决异常检测分类问题的算法^[23]。它的训练策略是将正样本映射到与核函数对应的特征空间中,并在数据和原点之间形成一个超平面,使得该超平面与原点之间的距离最大。假设这个超平面为:

$$\omega \cdot \Phi(x) - \rho = 0 \quad (9)$$

式中: $\omega = [\omega_1, \omega_2, \dots, \omega_d]$ 为法向量,决定超平面的方向,其中 d 为特征空间的维数; ρ 为单类超平面截距; $\Phi(x)$ 为法向映射函数, x 为输入的原始数据点。

优化目标是在正确分类的基础上最大化超平面与原点之间的距离。为防止参数求解出现耦合问题,可将问题描述为:

$$\begin{cases} \min_{w \in F, \xi \in R', \rho \in R} \frac{1}{2} \|\omega\|^2 + \frac{1}{v\ell} \sum_i \xi_i - \rho \\ \text{s.t. } (\omega \cdot \Phi(x_i)) \geq \rho - \xi_i, \xi_i \geq 0 \end{cases} \quad (10)$$

式中: ξ_i 为松弛变量; $v \in (0, 1)$ 用于调整松弛度,可以让一些边缘值位于超平面和原点之间; ℓ 为训练样本数量; F 为高维特征空间; R, R' 分别为超平面的截距和约束条件下的偏移量调整。

该优化目标的对偶问题为:

$$\begin{cases} \min_a \frac{1}{2} \sum_{ij} a_i a_j k(x_i, x_j) \\ \text{s.t. } 0 \leq a_i \leq \frac{1}{v\ell}, \sum_i a_i = 1 \end{cases} \quad (11)$$

式中: $k(x_i, x_j)$ 为高斯核函数, x_i, x_j 分别为输入的第 i, j 个样本数据点; a_i, a_j 为对应的拉格朗日乘子。

利用二次编程求解器求解对偶问题,可以得到支持向量 $\{x_i: i \in [\ell], \alpha_i > 0\}$, 进而得到预测函数 $f_{\text{OCSVM}}(x)$ 如式(12)所示。

$$f_{\text{OCSVM}}(x) = \sum_i \alpha_i k(x_i, x) - \rho \quad (12)$$

1.5 基于SMOTE的不平衡数据处理

在去除异常噪声数据后,声纹数据还存在样本

不平衡的问题,即某些类别的样本数量远远少于其他类别。这会导致模型在少数类别上的性能下降,因为它更容易识别常见的类别而忽略罕见的类别。

SMOTE算法基于重新采样,其核心思想是通过过度采样而不是复制来生成合成样本^[24]。该算法是通过在特征空间中操作少数类样本来生成新样本,从而提高少数类样本的识别率,削弱分类器对多数类样本的偏差效应。

图2为SMOTE过采样示意图。其中坐标轴 f_1 和 f_2 代表数据集中的两个特征维度。

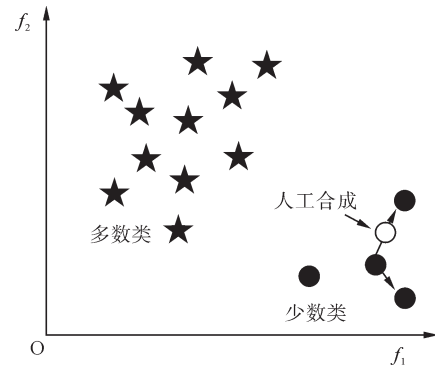


图2 SMOTE过采样示意图

Fig. 2 Flow chart of SMOTE oversampling

SMOTE过采样的基本步骤为:

- 1) 对于每个少数类别样本,计算其 k 个最近邻的样本(通常通过欧氏距离来计算距离, $k=5$);
- 2) 从少数类别样本中随机选择一个样本作为目标样本。从目标样本的 k 个最近邻中随机选择一个样本,记为邻居样本;
- 3) 随机选择一个值(通常在 $[0, 1]$ 区间内均匀分布)作为采样比例,记为 R_{alpha} ;
- 4) 使用以下公式生成合成样本: 合成样本 = 目标样本 + $R_{\text{alpha}} \times (\text{邻居样本} - \text{目标样本})$;
- 5) 重复上述步骤,生成足够数量的合成样本,以使得少数类别的样本数量达到所需的平衡水平。

1.6 基于贝叶斯优化的BiGRU识别模型

GRU是LSTM的一个变种,具有更简单的结构和更好的收敛性。GRU包括一个更新门和一个重置门^[25]。

对于给定的输入序列 $\mathbf{x} = [x_1, x_2, \dots, x_t]$, 标准的循环神经网络通过 t 次迭代来计算隐藏层的状态向量序列 $\mathbf{h}_t = [h_1, h_2, \dots, h_t]$ 和输出向量序列 $\mathbf{y}_t = [y_1, y_2, \dots, y_t]$ 。

$$h_t = \sigma_h(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (13)$$

$$y_t = \sigma_y(W_{hy}h_t + b_y) \quad (14)$$

式中: W_{xh} 为输入层与隐藏层层间权重矩阵; W_{hh} 为隐藏层层间权重矩阵; W_{hy} 为隐藏层与输出层层间权重矩阵; b_h 和 b_y 分别为隐藏层和输出层的偏置向量; σ_h 和 σ_y 分别为隐藏层和输出层激活函数。

GRU 的更新过程如式 (15) 所示^[24]。

$$\begin{cases} r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \\ z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \\ \tilde{h}_t = \tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \\ h_t = (1 - z_t) \odot \tilde{h}_t + z_t \odot h_{t-1} \end{cases} \quad (15)$$

式中: $\sigma(\cdot)$ 为 sigmoid 函数; $\tanh(\cdot)$ 为双曲正切函数。 U_r 和 U_z 分别为重置门和更新门前一个隐藏状态向量 h_{t-1} 的权重矩阵; $\tilde{h}_t$ 为候选隐藏状态, h_t 为最终隐藏状态; $\odot$ 为逐元素乘法; 向量 r_t 和 z_t 分别为重置门和更新门向量; b_r 和 b_z 分别为重置门和更新门的偏置向量; W_r 、 W_z 、 W_h 分别为重置门和更新门和隐藏层权重矩阵。

BiGRU 由前向 GRU 和后向 GRU 组成, 可以从两个角度获取 GIS 的音频信息。在每个时间步输入提供了两个相反方向的 GRU, 输出由这两个单向 GRU 决定。

对 BiGRU 模型选择合适的超参数对识别结果至关重要。通常有几种常见的超参数调整方法, 包括手动调整法、网格搜索法、随机搜索法和贝叶斯优化法。手动调整法涉及人工选择, 通常耗时长且难以找到最佳参数组合。网格搜索和随机搜索方法在测试超参数组合时未能充分利用之前的搜索结果, 因此可能效率较低。

贝叶斯优化则提供了一种高效且准确的选择方法^[26-27]。其核心涉及两个关键部分: 概率代理模型和采集函数。高斯过程是常见的概率代理模型。最常用的采集函数是置信区间最大化(probability of improvement, PI), 其核心思想在于根据当前的概率代理模型计算不同参数配置下目标函数性能提升的概率, 并选择具有最高概率提升的参数配置。

$$P_1(x) = P(f(x) \geq f(x^*) + \varepsilon) \quad (16)$$

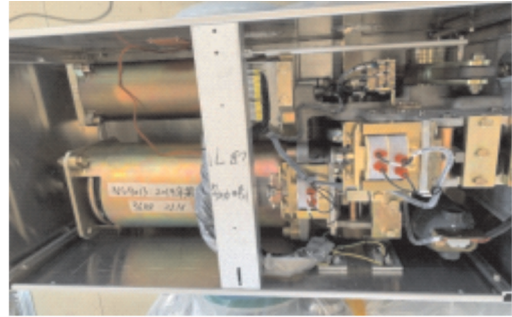
式中: $P_1(x)$ 为参数配置 x 下的采样点 PI 值; $P(\cdot)$ 为概率函数; $f(x)$ 为目标函数在参数配置 x 下的性能; x^* 为当前已知的最佳参数配置, 通常是已经观测到的目标函数的最佳值对应的参数。 ε 为一个小正数,

通常用于控制性能提升敏感度。

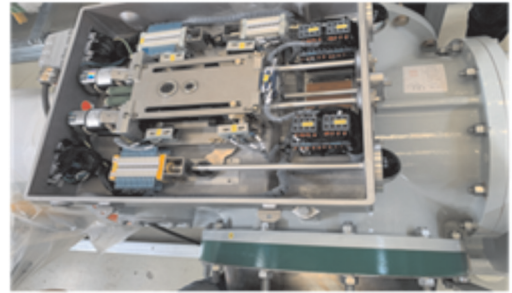
2 应用案例与实验细节

2.1 实验描述

以某 GIS 设备制造厂生产的 ZF23-126 型 GIS 设备为研究对象分别模拟其断路器、隔离开关操作系统中常见的机械故障, 采集操动机构中通过声音可以体现出来的异常和缺陷, 如图 3 所示。



(a) 断路器



(b) 隔离开关

图 3 实验模型

Fig. 3 Test model

首先, 针对断路器的储能模块进行故障模拟, 在 GIS 设备外壳中放置了 3 个麦克风, 距离地面 1.5 m, 距离设备外壳 0.9 m。储能电机的电压分别调整为 $65\%U_n$ 、 U_n 和 $120\%U_n$, U_n 为储能电机的额定电压, 模拟了电压过低和过高的故障情况, 每种状态进行了多次操作采集声音信号。

为了确保采集的数据更加可靠, 在每次操作之间通过调整实验条件来确保数据的多样性和全面性。例如, 对于断路器转速过低工况, 通过调节节流阀的幅度来控制液压油的流速, 节流阀调节得更紧, 转速会相应更低; 对于隔离开关分合闸不到位的工况, 通过调节弹簧的阻力来模拟传动机构的卡涩, 弹簧阻力越大, 卡涩问题就越严重; 对于储能电机空转故障的工况, 通过拧松接线端固定螺栓来

模拟不同程度的故障。这些调节方式不仅增加了数据的多样性,也更加符合实际应用场景中的需求。

最后根据GIS设备的不同工况及缺陷,建立20种不同运行状态下的音频样本库,如表1所示。

表1 GIS设备的各类工况

Tab. 1 Various working conditions of GIS equipment

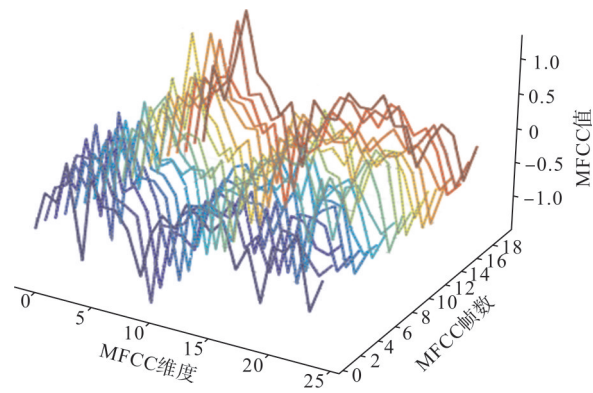
GIS 设备	工况类别	编号
断路器	储能电机空转	0
	储能电机电压过低	1
	储能电机额定电压	2
	储能电机电压过高	3
	断路器速度过低	4
弹簧机构	断路器速度正常	5
	储能电机空转	6
	储能电机电压过低	7
	储能电机额定电压	8
	储能电机电压过高	9
隔离开关	机构拒分拒合	10
	储能电机异常	11
	储能电机电压过低	12
	储能电机额定电压	13
	储能电机电压过高	14
电动机	储能电机电压过低	15
	储能电机额定电压	16
	储能电机电压过高	17
	机构无制动行程开关失效	18
	开关分合闸不到位	19

2.2 预处理及特征提取

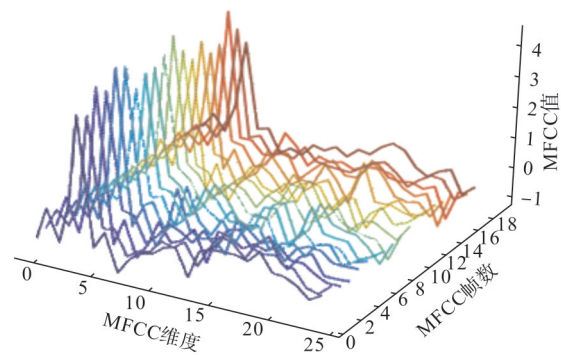
本节首先对采集到的音频样本进行分帧、加窗预处理操作,然后应用基于F比的加权MFCC提取音频信号的特征。本文设置的参数如下,输入音频信号长度为0.4 s,帧长为40 ms,帧移为20ms,重叠率为40%,由此计算可得加权MFCC向量组的帧数MFCC帧数为19,每帧包括26个MFCC维度特征,其数值用MFCC值表示。限于篇幅,本文仅以GIS设备中的弹簧机构的额定电压和机构拒分拒合两种情况为例分析加权MFCC特征提取结果,如图4所示。

2.3 OCSVM异常检测

为了进一步提升MFCC特征的可分性,提高模型性能和稳定性,利用OCSVM进行特征后处理,检测异常序列并去除异常点,以此达到净化数据集,提高后续SMOTE合成新样本质量的目的。



(a) 弹簧机构-拒分拒合



(b) 弹簧机构-储能电机额定电压

图4 利用加权MFCC提取的GIS声纹信号特征向量结果

Fig. 4 Results of feature vectors of GIS acoustic signals extracted using weighted MFCC

OCSVM的具体参数设置为:最多允许10%的样本被视为异常,使用系数为0.1的径向基函数核。为了可视化OCSVM的降噪结果,利用PCA降维技术,将高维MFCC特征投影到2维空间中,如图5所示,其中坐标轴PC1和PC2分别代表降维后的第一主成分和第二主成分,而右侧颜色轴的变化范围为(-1, 1),数据点越接近1表示数据点正常,越接近-1则表示可能是异常点或离群点。

2.4 SMOTE样本均衡

在消除异常噪声数据之后声纹数据仍然面临着样本不平衡的挑战。因此,通过SMOTE生成的合成样本可以增加少数类别的样本量,提高模型对少数类别的识别性能,从而提高声纹识别的准确性和鲁棒性。本文选择SMOTE的合成采样策略为自动,随机种子设为42,最近邻数量 $k=5$,样本均衡前后每类声纹的数量如表2所示。

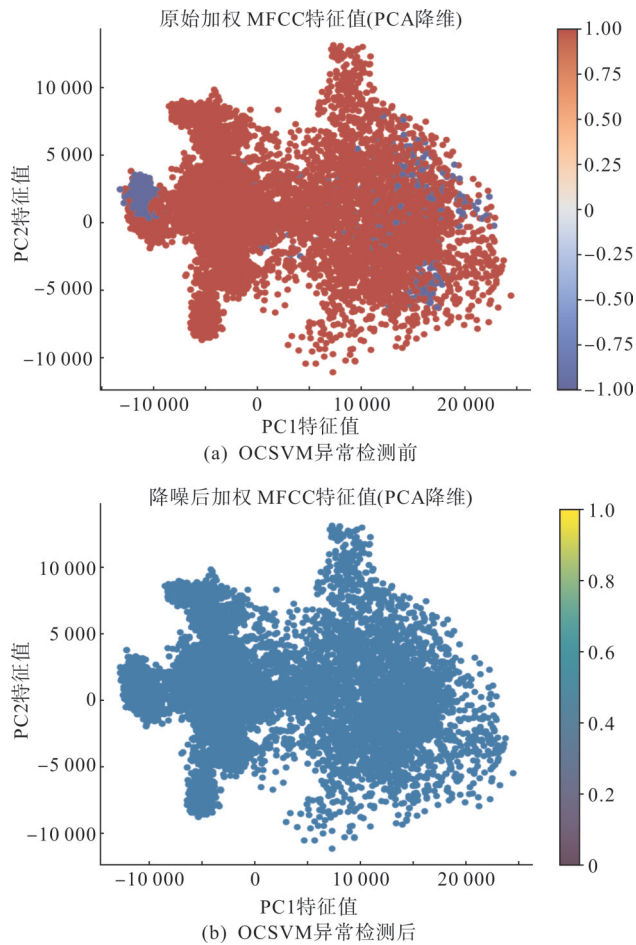


图5 通过PCA比较使用OCSVM进行异常点检测前后的加权MFCC特征图

Fig. 5 Comparison of weighted MFCC feature maps before and after anomaly detection using OCSVM by PCA

表2 样本均衡前后每类声纹的数量

Tab. 2 Numbers of voiceprint of each type before and after sample equalization

编号	均衡前数量	均衡后数量	编号	均衡前数量	均衡后数量
0	2 000	1 500	10	1 234	1 500
1	1 312	1 500	11	1 466	1 500
2	1 738	1 500	12	1 521	1 500
3	1 457	1 500	13	1 374	1 500
4	1 635	1 500	14	2 000	1 500
5	996	1 500	15	1 267	1 500
6	1 551	1 500	16	804	1 500
7	1 234	1 500	17	1 456	1 500
8	1 343	1 500	18	734	1 500
9	1 058	1 500	19	1 012	1 500

2.5 贝叶斯优化的BiGRU识别模型

本文采用贝叶斯优化对BiGRU模型进行超参

数选择,旨在获得更高的识别准确率。

首先定义了BiGRU模型的超参数搜索空间,为每个超参数指定范围或可能的值。这包括BiGRU层的数量(1~4)、每个层的隐藏单元数(1~100)、dropout率(0~0.5)、批处理大小(0~24)等。准确率作为优化的目标函数。

经过优化后的模型最佳超参数如表3所示,模型使用稀疏分类交叉熵作为损失函数,并采用自适应矩(adaptive moment, Adam)估计优化器优化模型参数。

表3 BiGRU模型超参数

Tab. 3 Hyperparameters of BiGRU

超参数	具体数值
层数	2
第1隐藏层神经元数	26
第2隐藏层神经元数	30
批处理大小	6
Dropout	0.1
训练次数	50
学习率	0.001
学习衰减率	0.1
衰减步数	10

3 结果分析

3.1 对比算法与评价指标

为了验证本文模型的有效性,在保持相同声纹数据预处理和特征提取的基础上,引入了十一种经典的分类方法和模型作为对比算法。这些方法包括最近邻算法(KNN)、朴素贝叶斯算法(naïve bayes, NB)、逻辑回归(logistic regression, LR)、径向基函数支持向量机(RBF SVM)、线性判别分析(Linear discriminant analysis, LDA)、二次判别分析(quadratic discriminant analysis, QDA)、决策树(decision tree, DT)、随机森林(random forest, RF)、自适应增强(adaptive boosting, AB)、卷积神经网络(CNN)以及长短时记忆网络(LSTM)。其中神经网络类的实验在Python 3的TensorFlow和Keras框架下完成,并在一台搭载NVIDIA A100的机器上进行训练。训练集、验证集与测试集的比例为7:2:1,其中,测试集在模型的训练和验证阶段未被使用,因此对模型来说是未知的,完全符合盲设工况的要求。利用这个独立的测试集能够客观

评估模型在处理未见过工况时的表现,验证所提方法的准确性和有效性。

在实验中采用了准确率^[28]与混淆矩阵^[29]两种评价指标对预测结果进行分析。准确率代表正确分类的样本数量占测试集总样本数量的比例。混淆矩阵 M_{co} 是用于衡量分类模型性能的一种工具。

$$M_{co} = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1N} \\ m_{21} & m_{22} & \cdots & m_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N1} & m_{N2} & \cdots & m_{NN} \end{bmatrix} \quad (17)$$

式中:对角线元素 m_{ii} 为各类工况声纹信号被模型正确识别的样本数量;非对角线元素 m_{ij} 为第 j 类声纹信号被模型错误识别为第 i 类声纹信号的样本数量; N 为总声纹信号种类数量。

3.2 声纹识别结果

各种分类算法对GIS设备不同工况声纹信号的总体识别准确率由表4所示。可以看到,准确率最高的是本文提出的BiGRU模型,识别率达到了92.8%,相比于CNN、LSTM、GRU神经网络分别提升了3.6%、2.5%和1.4%。这可以解释为:相较于GRU,BiGRU结合了前向和后向的信息流,能同时捕捉这两种方向的信息,有助于更好地理解声纹信号的动态特征;相较于CNN、LSTM,GRU类模型更适合处理声纹信号等时间序列数据,参数数量更少,计算速度更快且不容易过拟合。

表4 各分类算法在识别GIS设备不同工况声纹信号上的总体识别准确率

Tab. 4 Overall accuracies of various classification algorithms on recognizing voiceprint signals from different operating conditions of GIS

算法	KNN	NB	LR	LDA	QDA	DT
准确率/%	89.2	80.9	85.7	83.2	89.4	87.3
算法	RF	AB	CNN	LSTM	GRU	BiGRU
准确率/%	91.1	71.3	89.6	90.5	91.5	92.8

表4的结果还显示,在识别GIS设备不同工况的声纹信号这个任务上经典分类机器学习算法(如KNN、NB等)的表现整体上不如神经网络算法。自适应增强法的总体准确率只有71.3%,这可能是因为声纹类别过多,难以拟合复杂的声纹特征,导致其弱分类器之间的性能差异较大;相反,随机森林法却有高达91.1%的准确率,甚至超越了部分神经网络的结果,这可能是随机森林法本身就适用

于处理高维稀疏的数据,如声纹的加权MFCC特征,将其和随机子空间方法相结合可以有效抑制过拟合。

通常,参数越多的模型越复杂,学习能力越强,但也更容易过拟合,尤其是当训练数据有限时。因此,在评估模型性能时需要权衡参数数量与模型的泛化能力。本文将神经网络类模型的具体超参数和总参数数量作了对比,如表5所示,其中最佳超参数是通过本文所提的贝叶斯优化得出的。

表5 神经网络类模型的具体超参数和模型总参数结果

Tab. 5 Specific hyperparameters and total model parameter results for neural network models

模型	层数	第1层神经元数	第2层神经元数	第3层神经元数	Dropout	模型总参数
CNN	2	26	35		0.2	26 504
LSTM	2	32	48		0.2	25 428
GRU	3	20	25	25	0.1	16 965
BiGRU	2	26	30		0.1	24 764

从表5中可以看到,4种神经网络模型的总参数数均在同一数量级上,GRU的待定参数最少,CNN的待定参数最多。值得关注的是,在相同超参数的情况下,BiGRU的总参数大致为GRU的两倍,然而,双向模型由于能够获取更多上下文信息,可以在更少的超参数情况下就能达到相似甚至更优的性能。本文所提出的BiGRU模型通过贝叶斯优化调整超参数,相比GRU,在适度增加参数数量的情况下进一步提升了准确率,更好地平衡了模型性能和泛化能力。

为了进一步探究本文模型的声纹辨识结果,准确率最高的BiGRU模型辨识结果的混淆矩阵如图6所示,对角线元素代表该类工况被正确识别的数量,非对角线元素则代表该类工况被误识为另一类工况的数量。

通过计算可以确定BiGRU模型在识别GIS设备不同工况下的准确率,如表6所示。

由图6及表2可见,在采集到的20种不同工况的GIS设备声纹数据中,本文所提BiGRU模型表现出极高的分类准确率(平均为92.8%)。特别值得注意的是,BiGRU模型在检测液压机构的储能电机空转、储能电机额定电压运行以及弹簧机构的储能电机电压过低、储能电机异常等9种工况时表现出

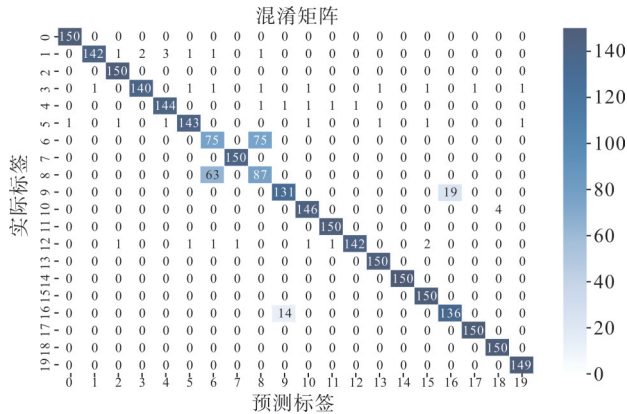


图6 GIS不同工况的识别结果的混淆矩阵

Fig. 6 Confusion matrix for the recognition results of different operating conditions in GIS

表6 BiGRU在识别GIS设备不同工况时的准确率

Tab. 6 Accuracies of BiGRU in recognizing different operating conditions of GIS equipment

工况编号	0	1	2	3	4
准确率/%	100	94.7	100	93.3	96
工况编号	5	6	7	8	9
准确率/%	95.3	50	100	58	87.3
工况编号	10	11	12	13	14
准确率/%	97.3	100	94.7	100	100
工况编号	15	16	17	18	19
准确率/%	100	90.7	100	100	99.3

100 %识别准确率。

其中, 仅对弹簧机构中的储能电机空转(编号6)和储能电机额定电压运行(编号8)这两类工况的识别结果相对较差, 准确率分别只有50 %和58 %。通过混淆矩阵可知, 它们倾向于互相识别成对方, 而没有被误识别为其他工况。这说明这两种工况下的声纹特征相似性较高, 模型难以区分。

进一步的解释是, 弹簧机构中储能电机在空载状态下的电压通常是接近额定电压的, 这一特性是通过电路设计实现的, 主要是为了确保断路器能够快速而可靠地回位, 切断电路并避免电气事故。因此, 用加权MFCC提取储能电机空载时的特征会与其额定运行时相似。而在液压机构中却没有这个问题, 这是因为液压机构中存在更多复杂的控制和反馈机制, 可以更灵活地处理不同的工况, 不必一直保持接近额定电压的水平。

3.3 消融实验

为了更详细地评估和验证所提出的本文算法对

GIS异常工况声纹识别的实际效果和性能影响, 针对算法的不同模块进行了一系列消融实验, 总体识别结果如表7所示。

表7中√表示在工况声纹识别中经历了相应的步骤。若加权MFCC栏中没有√, 则表示只使用MFCC进行特征提取; BiGRU栏中没有√, 则表示只使用GRU模型进行识别; 贝叶斯优化栏中没有√, 则表示手动进行参数调整。其中, 最后一列为基准模型, 其没有采用本文改进的任何方法, 仅利用MFCC进行特征提取, 采用手动参数调整的GRU模型进行识别。

表7 消融实验结果

Tab. 7 Results of ablation tests

加权MFCC	OCSVM异常检测	SMOTE样本均衡	贝叶斯优化	BiGRU	准确率/%
√	√	√	√	√	92.8
	√	√	√	√	88.5
√		√	√	√	89.1
√	√		√	√	86.6
√	√	√		√	92.0
√	√	√	√		91.5
基准模型(MFCC+GRU)					78.3

由表7可知本文提出的所有改进方法对GIS设备不同工况的声纹识别均有帮助, 准确率最大能提升18.5 %。值得注意的是, SMOTE样本均衡对模型准确率的提升起到了至关重要的作用。如果没有采用加权MFCC和OCSVM异常检测这两项改进方法, 模型性能将分别下降4.9 %和4.0 %。此外, 利用BiGRU代替GRU和应用贝叶斯优化调整网络参数能够进一步提高声纹识别的精度。

因此, 本文算法的有效性在于其强大的特征提取、异常处理技术、有效的样本均衡方法以及参数优化的BiGRU模型。

4 结论

本文针对GIS设备声纹信号存在的噪声干扰和样本不平衡问题提出了改进算法以增强GIS异常工况声纹的识别准确率, 得到结论如下。

1) 基于F统计量实现了对声纹样本的MFCC特征加权提取, 对重要特征加权突出, 同时降低噪声特征的权重, 增强了声纹信号的可区分性。

2) OCSVM异常检测算法实现了对声纹样本的

异常值检测与删除,有效净化了训练数据,提高了后续样本均衡和建模的稳定性。

3) SMOTE算法高效地平衡了声纹样本类别的分布,通过生成新样本消除了数据之间的不平衡。

4) 构建了基于贝叶斯优化的BiGRU声纹识别模型,取得了92.8%的平均识别准确率,相比多种经典算法均有显著提升。

实验结果表明本文构建的GIS异常工况声纹识别算法是有效和可靠的。各项改进的融合运用使识别性能得到大幅提升。本研究为GIS设备状态监测和故障预测提供了新的思路,具有重要的工程应用价值。后续将考虑进一步扩充样本量和类别,并探究不同模型集成的识别效果。

参考文献

- [1] 王先培,肖伟,胡明宇,等.基于SF₆分解产物融合判断的GIS绝缘裂化趋势划分[J].高电压技术,2016,42(6):1834-1840.
WANG Xianpei, XIAO Wei, HU Mingyu, et al. GIS insulation deterioration trend division based on SF₆ decomposition products fusion judgment[J]. High Voltage Engineering, 2016, 42(6): 1834-1840.
- [2] 骆立实,姚文军,王军,等.用于GIS局部放电诊断的SF₆分解气体研究[J].电网技术,2010,34(5):225-230.
LUO Lishi, YAO Wenjun, WANG Jun, et al. Research on partial discharge diagnosis of GIS by decomposed gas of SF₆ [J]. Power System Technology, 2010, 34(5): 225-230.
- [3] OYAMA M, HANAI E, AOYAGI H, et al. Development of detection and diagnostic techniques for partial discharges in GIS [J]. IEEE Transactions on Power Delivery, 1994, 9(2): 811-818.
- [4] 董子凡,任劭帅,殷建刚,等.GIS双断口隔离开关不停电试验电场分布及击穿特性[J].中国电力,2023,56(9):168-177.
DONG Zifan, REN Jieshuai, YIN Jiangang, et al. Research on electric field distribution and breakdown characteristics of GIS double-fracture disconnect switch under non-outage experiment [J]. Electric Power, 2023, 56(9): 168-177.
- [5] 蔡永翔,肖小兵,何肖蒙,等.影响半密闭腔后气体绝缘恢复的涡环现象观测与分析[J].南方电网技术,2024,18(8):1-8.
CAI Yongxiang, XIAO Xiaobing, HE Xiaomeng, et al. Observation and analysis of vortex ring phenomena affecting the insulation recovery of post-arc gas of semi-enclosed chamber [J]. Southern Power System Technology, 2024, 18(8): 1-8.
- [6] 王思涵,马宏忠,孙维,等.基于GSABO-ICEEMDAN-KELM在气体绝缘封闭开关设备局部放电故障诊断中的应用[J/OL].南方电网技术,1-11[2024-10-08].<http://kns.cnki.net/kcms/detail/44.1643.TK.20240925.1701.004.html>.
WANG Sihan, MA Hongzhong, SUN Wei, et al. Application of GSABO-ICEEMDAN-KELM in partial discharge fault diagnosis of gas-insulated enclosed switchgear [J]. Southern Power System Technology, 1-11[2024-10-08].<http://kns.cnki.net/kcms/detail/44.1643.TK.20240925.1701.004.html>.
- [7] 吕金壮,陈兵,张长虹,等.特高压直流高速开关研发关键技术研究[J].南方电网技术,2023,17(5):9-18.
LÜ Jinzhuang, CHEN Bin, ZHANG Changhong, et al. Research on key technology of ultra high voltage DC high speed switch [J]. Southern Power System Technology, 2023, 17(5): 9-18.
- [8] KUSUMOTO S, ITOH S, TSUCHIYA Y, et al. Diagnostic technique of gas insulated substations by partial discharge detection [J]. IEEE Transactions on Power Apparatus and Systems, 1980, 99(4): 4196-4206.
- [9] 杨庆,董富宁,罗曼丹,等.宽频电压感知方法及其数据应用[J].高电压技术,2021,47(6):1969-1983.
YANG Qing, DONG Funing, LUO Mandan, et al. Wideband voltage sensing methods and sensing data applications [J]. High Voltage Engineering, 2021, 47(6): 1969-1983.
- [10] 刘创华,何金,张春晖,等.GIS局部放电时域波形图像的模式识别方法[J].电力系统及其自动化学报,2019,31(10):24-30.
LIU Chuanghua, HE Jin, ZHANG Chunhui, et al. Pattern recognition method for time-domain waveform images of GIS partial discharge [J]. Proceedings of the CSU-EPSA, 2019, 31(10): 24-30.
- [11] 张重远,岳浩天,王博闻,等.基于相似矩阵盲源分离与卷积神经网络的局部放电超声信号深度学习模式识别方法[J].电网技术,2019,43(6):1900-1906.
ZHANG Zhongyuan, YUE Haotian, WANG Bowen, et al. Pattern recognition of partial discharge ultrasonic signal based on similar matrix BSS and deep learning CNN [J]. Power System Technology, 2019, 43(6): 1900-1906.
- [12] 丛浩熹,李庆民,齐波,等.基于红外传感的GIS隔离开关触头温度在线监测技术研究[J].电力自动化设备,2014,34(3):144-148,161.
CONG Haoxi, LI Qingmin, QI Bo, et al. Online GIS switch contact temperature monitoring based on IR sensing [J]. Electric Power Automation Equipment, 2014, 34(3): 144-148, 161.
- [13] 赵红,李双田.改进的多级线性预测晚期混响抑制算法[J].信号处理,2014,30(6):674-682.
ZHAO Hong, LI Shuangtian. Improved late reverberation suppression algorithm using multiple step linear prediction [J]. Journal of Signal Processing, 2014, 30(6): 674-682.
- [14] WANG T, CAO J, PEI L. A novel underground pipeline surveillance system based on hybrid acoustic features [J]. IEEE Sensors Journal, 2021, 21(2): 1040-1050.
- [15] 王荣昊,李喆,孙正,等.基于FISVDD与GRU的变压器声纹识别技术[J].高电压技术,2022,48(11):4546-4556.
WANG Ronghao, LI Zhe, SUN Zheng, et al. Transformer voiceprint recognition technology based on FISVDD and GRU [J]. High Voltage Engineering, 2022, 48(11): 4546-4556.
- [16] 马宏忠,王健,杨启帆,等.基于声振特征区分的SMA优化SVM变压器机械松动识别方法[J].电机与控制学报,2023,27(10):42-53.
MA Hongzhong, WANG Jian, YANG Qifan, et al. SMA-optimized SVM transformer state identification method based on acoustic vibration feature differentiation [J]. Electric Machines and Control, 2023, 27(10): 42-53.
- [17] 万可力,马宏忠,崔佳嘉,等.基于Mel-GADF与ConvNeXt-T的变压器铁芯松动故障诊断方法[J].电力自动化设备,2024,44(3):217-224.

- WAN Keli, MA Hongzhong, CUI Jiajia, et al. Fault diagnosis method of transformer core loosening based on Mel-GADF and ConvNeXt-T[J]. Electric Power Automation Equipment, 2024, 44(3): 217 – 224.
- [18] 孙汉文, 李喆, 林睿, 等. 基于新奇检测的两级电气故障声纹识别算法[J]. 电网技术, 2021, 45(7): 2888 – 2895.
- SUN Hanwen, LI Zhe, LIN Rui, et al. Two-stage voiceprint recognition algorithm of electrical faults based on novelty detection[J]. Power System Technology, 2021, 45(7): 2888 – 2895.
- [19] 路智欣, 邹璟, 方勇, 等. 基于优化 MFCC 参数的 GIS 绝缘缺陷声纹诊断模型[J]. 电子技术与软件工程, 2020(18): 223 – 226.
- LU Zhixin, ZOU Jing, FANG Yong, et al. Acoustic diagnosis model of GIS insulation defects based on optimized MFCC parameters[J]. Electronic Technology and Software Engineering, 2020(18): 223 – 226.
- [20] 田莎莎, 唐菡, 余纬. 改进 MFCC 参数在非特定人语音识别中的研究[J]. 科技通报, 2013, 29(3): 139 – 146.
- TIAN Shasha, TANG Wan, SHE Wei. Research of improved MFCC parameters in signer-independent speech recognition[J]. Bulletin of Science and Technology, 2013, 29(3): 139 – 146.
- [21] SAHIDULLAH M, SAHA G. A novel windowing technique for efficient computation of MFCC for speaker recognition[J]. IEEE Signal Processing Letters, 2013, 20(2): 149 – 152.
- [22] 王丰华, 王邵菁, 陈颂, 等. 基于改进 MFCC 和 VQ 的变压器声纹识别模型[J]. 中国电机工程学报, 2017, 37(5): 1535 – 1543.
- WANG Fenghua, WANG Shaojing, Cheng Song, et al. Voiceprint recognition model of power transformers based on improved MFCC and VQ[J]. Proceedings of the CSEE, 2017, 37(5): 1535 – 1543.
- [23] CHANG C C, LIN C J. Training v-support vector classifiers: theory and algorithms[J]. Neural Computation, 2001, 13(9): 2119 – 2147.
- [24] 谢桦, 陈俊星, 赵宇明, 等. 基于 SMOTE 和决策树算法的电力变压器状态评估知识获取方法[J]. 电力自动化设备, 2020, 40(2): 137 – 142.
- XIE Hua, CHEN Junxing, ZHAO Yuming, et al. Knowledge acquisition method of power transformer condition assessment based on SMOTE and decision tree algorithm[J]. Electric Power Automation Equipment, 2020, 40(2): 137 – 142.
- [25] CHUNG J, GULCEHRE C, CHO H K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. arXiv preprint arXiv:1412.3555, 2014.
- [26] SHU X, ZHANG L, QI G J, et al. Spatiotemporal co-attention recurrent neural networks for human-skeleton motion prediction[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(6): 3300 – 3315.
- [27] 毛明轩, 冯心营, 陈思宇, 等. 基于贝叶斯优化卷积神经网络的路面光伏阵列最大功率点电压预测方法[J]. 中国电机工程学报, 2024, 44(2): 620 – 631.
- MAO Mingxuan, FENG Xinying, CHEN Siyu, et al. A novel maximum power point voltage forecasting method for pavement photovoltaic array based on Bayesian optimization convolutional neural network[J]. Proceedings of the CSEE, 2024, 44(2): 620 – 631.
- [28] ZHONG Y, HAO J, WANG X, et al. Classification method for mechanical defects in GIS equipment based on mode function analysis and improved relevance vector machines[J]. CSEE Journal of Power and Energy Systems, 2023, 9(2): 790 – 801.
- [29] LIU Z, PAN Q, DEZERT J, et al. Combination of classifiers with optimal weight based on evidential reasoning[J]. IEEE Transactions on Fuzzy Systems, 2018, 26(3): 1217 – 1230.

收稿日期: 2023-11-21; 网络首发日期: 2024-06-26

作者简介:

庄小亮(1988), 男, 高级工程师, 硕士, 从事高压直流输变电设备故障诊断及状态检修研究, xl.zhuang@163.com;

李乾坤(1985), 男, 高级工程师, 学士, 从事高压直流输变电设备检修维护及试验管理研究, 15815826866@163.com;

季天瑶(1981), 女, 通信作者, 教授, 博士生导师, 博士, 研究方向为信号与信息处理及其在电力系统和电力市场中的应用, tyji@scut.edu.cn。