

DOI: 10.13382/j.jemi.B2407634

# t-SNE 降维融合 SAPSO-BP 的飞机电弧故障识别<sup>\*</sup>

屈慧妍<sup>1</sup> 李娟<sup>1</sup> 戴洪德<sup>2</sup> 王希彬<sup>2</sup> 张依<sup>1</sup>

(1. 鲁东大学数学与统计科学学院 烟台 264025; 2. 海军航空大学航空基础学院 烟台 264001)

**摘要:**针对单个特征识别故障电弧时特征的阈值难以确定、且难以设置适用于不同负载的统一阈值等问题,为更准确高效地检测不同负载下的串联电弧故障,提出基于 t-分布随机邻域嵌入(t-SNE)和模拟退火粒子群算法优化 BP 神经网络(SAPSO-BP)结合的多特征故障电弧识别方法。首先,针对故障电弧电流高频分量丰富的特点,通过提取电流频率的变异系数改进传统的变异系数特征,构造时频域特征检测故障电弧,结果表明改进后的变异系数(CV)对不同负载的平均识别准确率达到 96%。其次,继续提取小波包细节分量以及能量熵等时频域特征与 CV 进行多特征融合,共同识别故障电弧。在融合过程中使用多种非线性降维算法对多维特征进行降维,并进行聚类可视化对比,发现使用 t-SNE 降维将多维特征降至三维空间对故障电弧的区分度最高。最后,将降维后的特征输入 SAPSO-BP 进行训练,并设计消融实验验证了提出方法的识别性能与鲁棒性。结果表明,融合算法 tSNE-SAPSO-BP 在不同负载上的识别性能较单个特征的识别准确率分别提升了 3.2%、16.8%、27.66%、33.5%。t-SNE 降维与聚类很好地处理了各特征间的非线性相关性,为融合机器学习方法识别故障电弧提供了关键特征信息。

**关键词:** 电弧故障识别;特征提取;改进的变异系数特征;t-SNE 算法;模拟退火粒子群算法

**中图分类号:** TN06; TM501.2 **文献标识码:** A **国家标准学科分类代码:** 510.40

## Aircraft arc fault identification based on t-SNE dimensionality reduction fusion with SAPSO-BP

Qu Huiyan<sup>1</sup> Li Juan<sup>1</sup> Dai Hongde<sup>2</sup> Wang Xibin<sup>2</sup> Zhang Yi<sup>1</sup>

(1. School of Mathematics and Statistical Science, Ludong University, Yantai 264025, China;

2. School of Aeronautical Foundation, Naval Aeronautical University, Yantai 264001, China)

**Abstract:** To more accurately and efficiently detect series arc faults under different loads and address the difficulty of setting a unified threshold for different loads when using a single feature, a multi-feature fault arc recognition method based on t-distributed stochastic neighbor embedding (t-SNE) and simulated annealing particle swarm optimization algorithm-optimized BP neural network (SAPSO-BP) is proposed. Firstly, considering the rich high-frequency components in arc fault currents, the traditional coefficient of variation (CV) feature is improved by extracting the CV of current frequency. The improved CV achieves an average recognition accuracy of 96% across different loads. Secondly, wavelet packet detail components and energy entropy, which are time-frequency domain features, are further extracted and fused with CV for the identification of arc faults. During the fusion process, various nonlinear dimensionality reduction algorithms are used, and clustering visualization comparisons are carried out. It is found that reducing the multidimensional features to a three-dimensional space using t-SNE dimensionality reduction provides the highest distinction for fault arcs. Finally, the reduced features are input into SAPSO-BP for training, and ablation experiments are designed to verify the recognition performance and robustness of the proposed method. The results show that the recognition performance of the fusion algorithm tSNE-SAPSO-BP is improved by 3.2%, 16.8%, 27.66%, and 33.5% respectively compared with the recognition accuracy of single features on different loads. t-SNE dimensionality reduction and clustering effectively deal with the nonlinear correlations between features, providing key feature information for the identification of fault arcs by the fusion machine learning method.

**Keywords:** arc fault identification; feature extraction; the improved coefficient of variation feature; t-SNE algorithm; simulates annealing particle swarm algorithm

收稿日期:2024-06-29 Received Date: 2024-06-29

<sup>\*</sup> 基金项目:鲁东大学研究生创新项目(IPGS2024-045)、工信部民机专项(MJ-2018-J-75)资助

## 0 引言

随着航空电气自动化技术的不断发展,飞机配电系统结构变得更加复杂,航空线缆数量也大大增加。由于飞行过程中气流的冲击,飞行状态变化引起的连续振动会导致电缆线束发生摩擦、弯曲甚至折断的现象,使得航空电缆损坏,继而发生航空交流电弧故障,甚至可能导致空难<sup>[1]</sup>。

电弧引发的空中灾难有 1998 年瑞士航空 111 号班机空难,1985 年日本航空 123 号班机空难,2009 年法国航空 447 号班机空难等,飞机一旦坠毁,机上乘客几乎全部遇难。电影《中国机长》的原型为 2018 年川航事件,据中国民用航空安全信息系统发布的调查报告称:“最大可能原因是 B-6419 号机右风挡封严可能破损,风挡内部存在空腔,外部水汽渗入并存留于风挡底部边缘。电源导线被长期浸泡后绝缘性降低,在风挡左下部拐角处出现潮湿环境下的持续电弧放电,电弧产生的局部高温导致双层结构玻璃破裂。风挡不能承受驾驶舱内外压差从机身爆裂脱落。”因此寻找一种能够准确识别电弧故障发生的方法对航空电弧故障检测技术研究和飞行安全具有重要意义。

按照电弧故障发生时电弧与电路连接关系,可将电弧故障分为串联电弧故障、并联电弧故障、接地电弧故障和复合电弧故障。其中,串联电弧故障作为低压配电系统电气火灾的重要诱因之一,使得其危害更大、隐蔽性更强、随机性大,且发生时电流幅值较小,容易被负载电流掩盖,使得其检测和预防更加困难。目前针对串联型电弧故障的检测方法主要分为两大类:一类是通过弧声、弧光、电磁辐射等物理特性检测故障电弧;另一类是通过分析线路中电流、电压的波形特征,提取电流的时频域特征来检测故障电弧,由于回路电流方便获取,因此这类方法是目前研究电弧故障检测的热点方法。

目前的串联电弧故障检测方法主要以分析和提取电流信号的时域、频域以及时频域特征为主,通过设置特征的阈值或将这些特征应用于机器学习算法来实现正常与故障电弧的区分。Park 等<sup>[2]</sup>和 Ning 等<sup>[3]</sup>提出了基于电流信号相邻波形的绝对差分 and 随机性、肩长的检测方法。刘艳丽等<sup>[4]</sup>和 Wang 等<sup>[5]</sup>通过对比不同矩阵参数,得到了最佳的滑动记忆矩阵尺寸,提出了一种正交方向改进局部三值化模式(OD-LTP)的特征提取与故障电弧检测方法,该方法在不同工况中表现出色,具有较高的鲁棒性。同时,Gong 等<sup>[6]</sup>和刘艳丽等<sup>[7]</sup>以电源端电压为对象,对信号进行小波分解得到的低频系数进行自定义差分阈值滤波,去除噪声,并通过阈值法检测串联型电弧故障。Jiang 等<sup>[8]</sup>提出了一种基于频谱特征的时间序列重

构方法,使用脉冲识别算法提取重构信号的电弧特征。

然而,一些研究表明,大部分时频域特征仅适用于少数特定类型的负载,针对于不同负载类型的统一阈值难以确定<sup>[9]</sup>,尤其是在非线性负载中,故障检测虚警率较高。针对这一问题,一些研究开始使用聚类算法和机器学习算法识别故障电弧,此时无需设置统一阈值。唐圣学等<sup>[10]</sup>、刘艳丽等<sup>[11]</sup>和 Gong 等<sup>[12]</sup>提出了将特征量进行多特征融合,通过极限学习机、支持向量机、轻量化卷积神经网络等算法诊断故障电弧,避免了单一特征造成的误判。董志文等<sup>[13]</sup>和 Yin 等<sup>[14]</sup>提出使用变分模态分解(VMD)对信号进行分解,提取能量熵和时域特征,并利用随机森林分类器进行故障电弧识别。江永鑫等<sup>[15]</sup>提出了一种改进的不完全 S 变换(IIST)时频分析方法并以此形成特征向量样本集,最后通过支持向量机(SVM)对样本集进行分类识别。郭洲等<sup>[16]</sup>提出通过完整集合经验模态分解(CEEMD)提取时频域特征,再利用 K 均值聚类(K-means++)进行分类识别。

上述方法均能够有效区分出故障电弧,为串联电弧故障检测技术的发展提供了重要的参考,但仍存在以下问题。1)单特征进行故障识别时,特征的阈值难以设置且无法对不同负载类型设置统一阈值。2)对于信号微弱且具有较强随机性的故障电弧,单纯依靠聚类算法难以有效识别。尤其是对于非线性负载,传统聚类算法对正常和故障电弧的区分度不高。

根据上述分析,本文提出一种融合 t-分布随机邻嵌入(t-SNE)降维和模拟退火粒子群算法优化 BP 神经网络(SAPSO-BP)的串联故障电弧检测方法。基于故障信号的频率特性,改进了传统的变异系数(CV)特征,提取电流频率的变异系数,从而构造了时频域特征,该特征便于对不同负载设置统一阈值。本文针对单纯的聚类算法无法准确区分正常与故障电弧问题,提出融合降维、聚类算法和机器学习算法进一步提高故障电弧检测的准确性和可靠性。首先,对原电流信号提取了 9 个时频域特征,将这些特征通过 t-SNE 算法映射到低维空间并进行聚类可视化,即降至 3 维特征并聚类,消除特征冗余信息的同时提高了故障电弧区分度。最后将 t-SNE 降维后的特征输入机器学习算法 SAPSO-BP,进一步地提高了故障电弧识别准确率。

## 1 航空串联电弧故障试验

### 1.1 电弧故障试验与数据采集

参照美国航空标准 SAE AS5692 搭建振动试验平台,试验原理如图 1 所示。振动试验电弧发生装置模拟了飞机在振动环境下,出现线缆接线柱松动而接触不良引发的串联电弧故障。试验选用 115 V/400 Hz 航空静变电

源作为试验电源,电弧故障发生器分别使用圆柱形碳棒和直径为 5 mm 的锥形铜棒作为电极。为了模拟电机故障引起的不同振动条件,采用音圈电机控制运动电极以不同的振动幅度和频率移动,从而模拟了电机故障引起的电接触点振动产生的电弧故障。音圈电机由 ADP-090-09 驱动器控制,驱动电源为 LRS-100-24 开关电源。最后由自制原理样机进行控制,进行数据采集,采样频率 2 525 kHz。

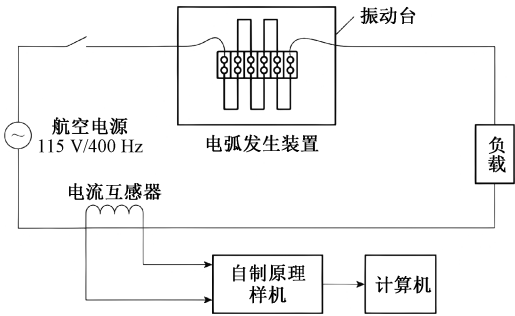


图 1 振动试验电路  
Fig. 1 Vibration test circuit

1.2 实验分析

经过多次试验获取了不同负载正常到故障的电流波形,实验中每个周期采集 6 250 个样本点,部分负载下的振动试验电流信号波形如图 2 所示,其中,负载 1~3 为线性负载,负载 4 为非线性负载。每个负载提取了 10 个周期的电流信号。通过图 2 可知,电弧故障发生前电流呈周期性的正弦波,而发生电弧故障时的电流过零点时会出现或长或短的平台部,也会夹杂着正常周期。尤其是负载 2、4,波形畸变更为严重,出现半波缺失情况甚至是“零休”现象持续多个周期。

2 故障电弧时频域特征的提取

2.1 改进的变异系数特征

为提取一个周期内的波形畸变现象,一些学者采用了统计学中常用的无量纲指标—变异系数来定量表征正常与故障电弧间的电流波形差异,以图 2 中的 4 个负载为例,以一个周期作为一组样本,计算各周期的变异系数如图 3 所示。可以发现变异系数特征对故障电弧有一定的识别能力,但仍存在无法对不同负载设置统一阈值以及对微弱的电弧故障辨识能力差等问题。且单纯的统计指标仅提取了电流信号的时域特征,其弊端在于它需要样本容量需要足够大。若将一个周期继续细分为 5 个时间样本,即一个样本仅包含 1 250 个采样点点,此时提取各样本变异系数数值如图 4 所示,发现正常样本与故障样本的变异系数值存在交叉,因此在样本量小的情况下提

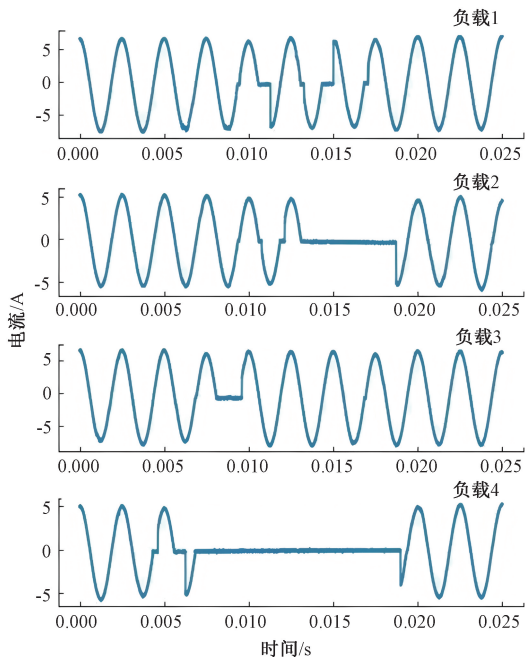


图 2 振动试验电流波形  
Fig. 2 Current waveform of vibration test

取变异系数无意义。

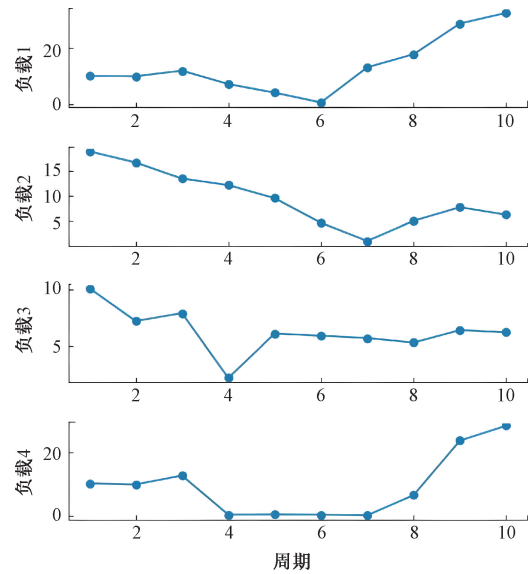


图 3 大样本情况下的变异系数值  
Fig. 3 Coefficient of variation values in large sample sizes

为克服变异系数的这一缺点,提出向变异系数这一时域特征中引入频率特点,将变异系数这一单纯的时域特征改进为时频域特征,进一步提高特征的识别准确率。同时通过直方图获取各窗口电流信号的频率大小  $c_i$ ,如图 5 所示。将各时间窗口分为 50 个区间,即各样本包含 1 250 个电流采样点,取各区间的概率密度值,使用这 50 个概率密度值代表该时间窗口的频率分布  $c_i, i = 1, 2, \dots$ ,

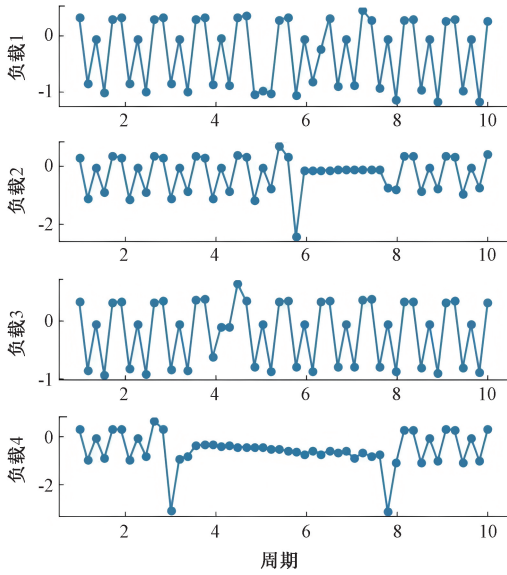


图4 小样本情况下的变异系数值

Fig. 4 Coefficient of variation in small sample sizes

50, 可以看到在过零点时故障电弧的电流频率明显高于正常运行时的电流频率。

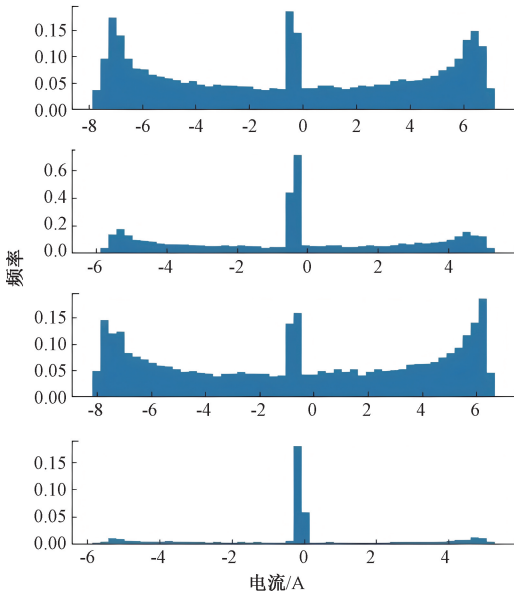


图5 故障电弧电流信号的频率分布直方图

Fig. 5 Frequency distribution histogram of fault arc current signals

最终,对各负载的 50 个时间窗口样本依次计算它们频率值的变异系数,计算过程如式(1)所示,其中  $\bar{x}$ 、 $\sigma$  分别代表各样本的均值和标准差,如式(2)、(3)所示。

$$CV = \frac{\sigma}{\bar{x}} \quad (1)$$

$$\bar{x} = \frac{\sum_{i=1}^N c_i}{N}, N = T \cdot fs \quad (2)$$

$$\sigma = \frac{\sum_{i=1}^N (c_i - \bar{x})^2}{N}, N = T \cdot fs \quad (3)$$

各负载改进后的 CV 特征值如图 6 所示,可以看出正常窗口的变异系数基本都稳定在 0~1.0 之间,而存在电弧故障的时间窗口的变异系数下降到了 1.0 以下,能明显地区分开正常与电弧故障。且改进后的变异系数在一定程度上弥补了变异系数对样本量敏感的缺陷,有利于进一步准确定位故障电弧发生的时间。

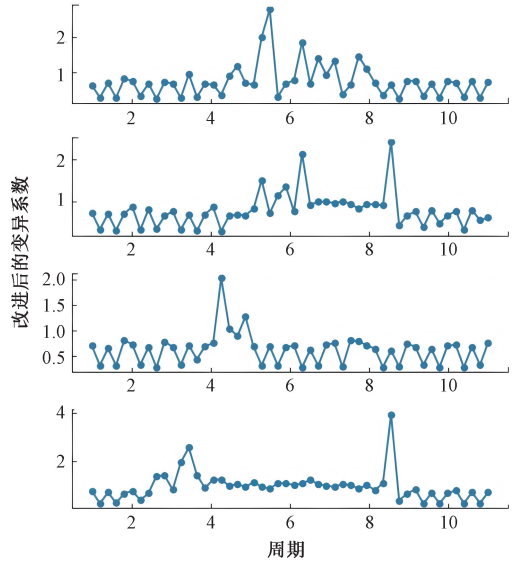


图6 改进后的变异系数值

Fig. 6 Improved coefficient of variation values

## 2.2 基于小波包分解(WPD)的频域特征提取

为进一步对比检验改进后变异系数特征的故障识别性能,提取故障电弧检测中一些常用的特征指标,对比各特征的识别性能,同时也为融合模型构建提供特征数据基础。WPD 作为一种信号处理工具,已被广泛应用于多个领域,其优势在于能够提供信号在时域和频域的多分辨率分析 WPD 在不同分解层的分解信号可按照式(4)经过逐层计算得到。

$$S_{i+1,2j}(n) = \sum_k g(k - 2n) S_{i,j}(k) \quad (4)$$

$$S_{i+1,2j+1}(n) = \sum_k g(k - 2n) S_{i,j}(k)$$

利用 db3 (Daubechies 3) 小波基函数对获取的各负载信号做 6 层小波分解。负载 1 电流采样数据各层细节信号能量如图 7 所示,其中 d1 表示第 1 层细节分量。

由图 7 可以看出,负载 1 在正常运行时其细节信号能量集中在零点附近,发生电弧故障时,细节信号突增,



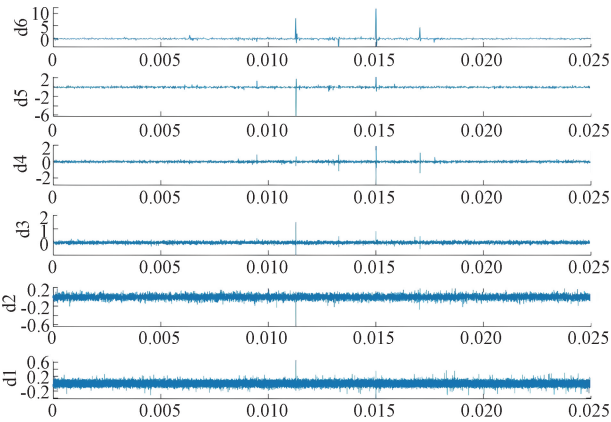


图 7 负载 1 的电流信号小波包分解细节分量

Fig. 7 Detail components of current signal WPD of load 1

可以选取这些突增点作为识别电弧故障的依据。同时为了更加准确的定位到每个故障点,对细节分量  $d1 \sim d6$  分别进行 3 次样条插值处理,使特征维度与样本容量一致。

小波包能量熵能够描述振动信号各频率成分的能量变化<sup>[17-18]</sup>,可以对故障信号进行定位和检测。通过 db3 小波基和式(5)对采集得到的电流信号进行 3 层小波分解,得到 8 个小波包信号,通过式(6)归一化能量熵公式对 8 个小波包进行重构,计算各自的重构信号能量熵。

$$E_{j,k} = \left| \int_{-\infty}^{\infty} S_{j,k} dt \right|^2 \sum_{j \in Z} \sum_{k \in Z} |d_{j,k}|^2 \quad (5)$$

$$\begin{cases} E\left(\sum_k^{2j-1} |E_{j,k}|^2\right)^{\frac{1}{2}} \\ \lambda_{j,k} = \frac{E_{j,k}}{E} \end{cases} \quad (6)$$

式中:  $S$  为原始信号,  $S_{j,k}$  为小波包重构后的信号,  $j, k$  分别表示小波包分解第  $j$  层的第  $k$  个节点;  $E_{j,k}$  为重构信号所对应的能量。

选取部分正常运行和存在电弧故障的电流信号,对它们最后一层能量熵进行对比,如表 1、2 所示,其中 E0 表示最后一层第 1 个小波的能量。分析发现,发生故障时,正常电弧与故障电弧的最后一层第 1 个小波的能量熵 E0 有明显的大小差别,而 E2~E8 未见这种差别,因此提取最后一层第 1 个小波的能量熵 E0 作为判断是否发生电弧故障的频域特征之一。

2.3 电流特征提取

分析各负载下的电流波形特点,发现发生电弧故障时,电流幅值会明显减小,但随后会有一个重燃的过程,即电流突增。针对这一现象,提取样本点间电流变化率反映的故障特征,计算每个周期内相邻两个离散点电流信号采样值的差值并取其绝对值,并除以采样间隔得到电流变化率,如式(7)所示。

$$g_i = | (c_{i+1} - c_i) \cdot fs | \quad (7)$$

表 1 正常电弧能量熵

Table 1 The energy entropy of a normal arc

	第 1 组	第 2 组	第 3 组	第 4 组
E0	4 876.39	4 928.36	130.05	299.31
E1	0.05	0.08	0.06	0.03
E2	0.04	0.22	0.03	0.02
E3	0.02	0.05	0.01	0.03
E4	0.03	0.16	0.02	0.01
E5	0.01	0.11	0.03	0.01
E6	0.04	0.05	0.07	0.11
E7	0.09	0.13	0.13	0.13

表 2 故障电弧能量熵

Table 2 Energy entropy of faulty arc

	第 1 组	第 2 组	第 3 组
E0	13.52	58.44	8.65
E1	0.04	0.07	0.05
E2	0.03	0.03	0.02
E3	0.03	0.03	0.02
E4	0.07	0.02	0.04
E5	0.05	0.02	0.01
E6	0.11	0.04	0.04
E7	0.03	0.1	0.15

式中:  $g_i$  为第  $i$  时刻的电流变化率;  $c_i$  为第  $i$  时刻的电流大小;  $fs$  为采样频率。

图 8 所示为各负载的电流变化率,可以看到发生电弧故障后电流变化率突增明显,可以作为时域特征之一。

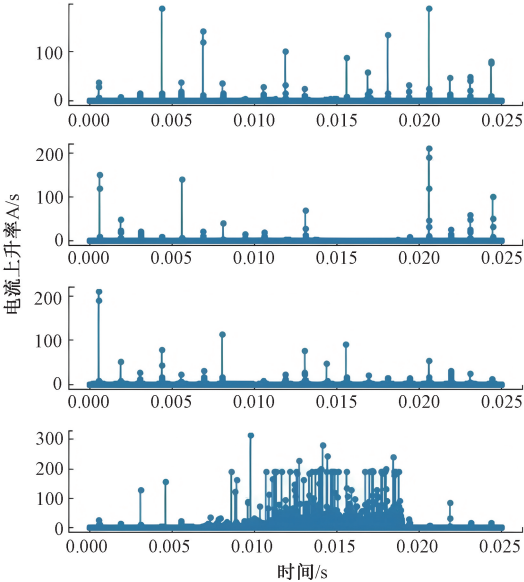


图 8 各负载电流变化率

Fig. 8 Change rate of each load current

3 基于 t-SNE 算法的特征降维

提取的特征包括改进的 CV、电流变化率、小波包变

换前6层细节分量 d1~d6、能量熵 E0 共9个特征。充分的特征变量可以提升预测精度,但过多的特征不仅容易导致特征冗余、信息重叠,还会增加故障识别模型的复杂性,影响识别精度。

### 3.1 特征降维可行性分析

首先使用 Pearson 相关系数和最大互信息系数(MIC)来衡量特征间的线性相关性和非线性相关性。Pearson 相关系数仅用来衡量两个随机变量之间的线性关系,而 MIC 系数衡量两个变量之间的关联程度,包括线性和非线性关系。

Pearson 相关系数的热力图如图9所示,各特征间的线性相关程度均低于0.17,说明特征间线性相关程度很低。

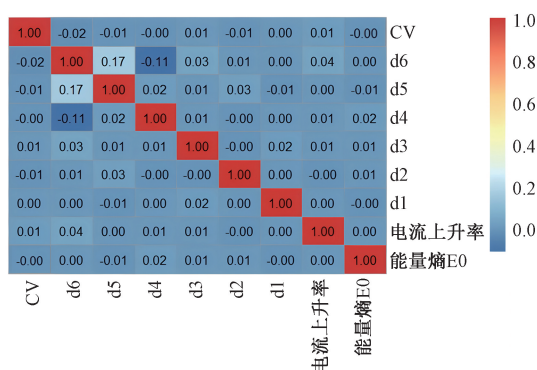


图9 Pearson 相关系数热力图

Fig. 9 Pearson correlation coefficient heat map

MIC-R<sup>2</sup> 系数热力图如图10所示,它给出了非线性关系的最高值。可以看出改进后的 CV 与 d6、d5、能量熵 E0 间的系数均大于0.7,呈高度相关,说明这些特征之间存在高度的非线性相关性。

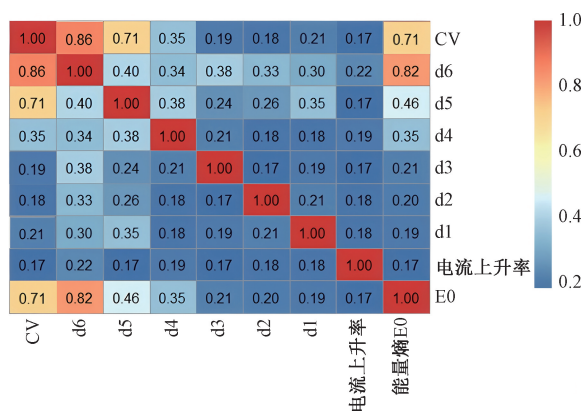


图10 MIC-R<sup>2</sup> 系数热力图

Fig. 10 Coefficient heat map

### 3.2 特征降维

t-SNE 是一种集降维与可视化于一体的非线性降维算法,与其他降维算法(如主成分线性降维)相比,t-SNE

适用于具有非线性相关数据的降维与聚类可视化。对于大多数应用场景,通常将数据通过 t-SNE 降维到2维或3维。这样既能更好地保留数据的局部结构,又能提高计算效率。

t-SNE 由 SNE 发展而来,SNE 的优化目标是让低维空间上的高斯分布尽量逼近高维空间上的高斯分布,SNE 原理如下:

1) 点之间的相似度是通过欧氏距离计算得出的条件概率来表达的,此时每个高维特征空间下每个电流样本点  $x_i$  都是一个聚类的类中心,高维空间中使用高斯概率密度函数衡量其他样本点  $x_j$  属于这一类的概率,如式(8)所示。

$$p_{j|i} = \frac{\exp(-\|x_i - x_j\|^2 / (2\sigma_i^2))}{\sum_{k \neq i} \exp(-\|x_i - x_k\|^2 / (2\sigma_i^2))} \quad (8)$$

2) 离聚类中心越远,则属于这一类的概率越小。同样可以在低维空间上计算聚类中心  $y_i$  与其他点  $y_j$  的距离,此时的相似度如式(9)所示。

$$q_{j|i} = \frac{\exp(-\|y_i - y_j\|^2)}{\sum_{k \neq i} \exp(-\|y_i - y_k\|^2)} \quad (9)$$

3) 最终希望样本点在高维空间和低维空间上的分布尽量相似,即  $p_{j|i} = q_{j|i}$ 。此时将 SNE 算法的优化目标转化为优化两个概率分布间的距离,使式(10)中信息散度(information divergence)为0。

$$KL(P_j \| Q_i) = \sum_i \sum_j p_{j|i} \log \frac{p_{j|i}}{q_{j|i}} \quad (10)$$

通过 SNE 算法将电流信号投影至二维空间中,效果如图11所示。其中 SNE\_1、SNE\_2 代表降维后的二维特征1代表正常电弧,2代表故障电弧。可以看到通过 SNE 降至二维特征后,在聚类时出现了拥堵问题,即正常电弧和故障电弧两个簇中部聚集在一起,两簇交集部分区分度不高。这可能是由于降到二维后无法得到可信映射,数据点在低维空间上仅仅保留了局部特征,但丢失了全局特征。

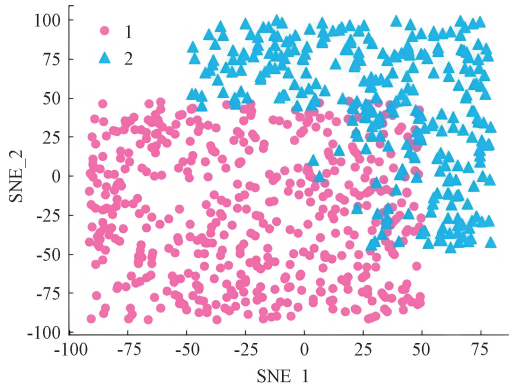


图11 SNE 算法聚类可视化

Fig. 11 Visualization of SNE algorithm clustering

针对拥堵问题,进一步提出了 t-SNE 算法:在高维空间下仍使用高斯分布来度量点与点间的距离,在低维空间下则选择自由度为 1 的  $t$  分布将距离转换为概率分布,而不是高斯分布。即式(8)中的  $q_{ji}$  更新为:

$$q_{ji} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_i - y_k\|^2)^{-1}}$$

(11)

图 12 所示为  $t$  分布的尾端比高斯分布在同等情况下概率更高,可以使得高维空间下  $x_i$  与  $x_j$  间的中低等距离在映射后距离更大,使低维空间中  $y_j$  被推的更远,从而实现全局结构的保留。

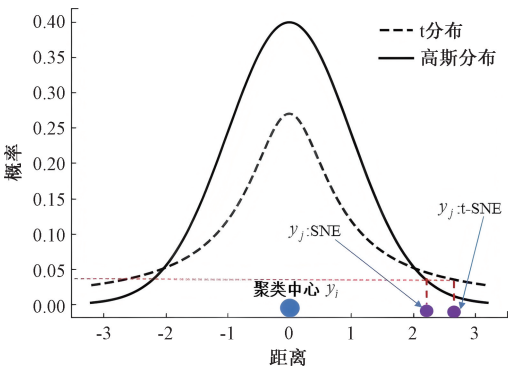


图 12 t-SNE 与 SNE 低维空间距离对比  
Fig. 12 Comparison of low-dimensional space distance between t-SNE and SNE

除 SNE 与 t-SNE 外,常用的非线性降维技术还包括核主成分分析(Kernel PCA),与局部线性嵌入(Locally linear embedding, LLE)等,采用这些方法算法将多维特征投影至二维与三维空间中,它们的聚类效果如表 3 所示,包含使用各种方法降至二维以及三维空间的轮廓系数(Silhouette Coefficient),该指标展示了簇内相似度(内聚度)和簇间相似度(分离度)。其取值范围为  $[-1, 1]$ , 值越接近 1 表示聚类效果越好,值越接近 -1 表示聚类效果越差。

表 3 不同特征降维方法效果对比

Table 3 Comparative analysis of the effects of different feature dimensionality reduction methods

降维方法	二维空间	三维空间
SNE	0.32	0.49
Kernel PCA	0.19	0.34
LLE	-0.82	-0.23
t-SNE	0.43	0.73

经试验可知,4 种降维方法在二维空间中均存在不同程度的拥堵问题,而 t-SNE 方法三维空间更好地保留和揭示了数据的局部结构,因此选用 t-SNE 方法将多维特征降维至三维空间,聚类效果如图 13 所示。

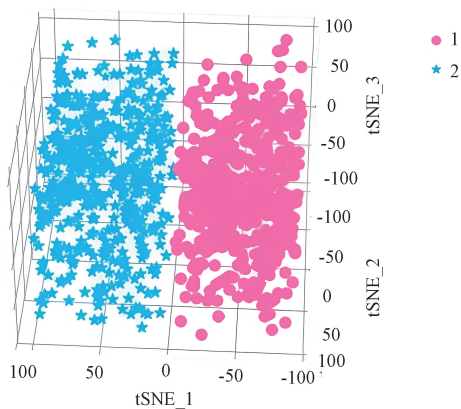


图 13 t-SNE 算法聚类可视化  
Fig. 13 Visualization of t-SNE algorithm clustering

其中,tSNE\_1、tSNE\_2、tSNE\_3 代表降维后的三维特征,1 代表正常电弧,2 代表故障电弧。可以看出 t-SNE 算法在三维空间的聚类效果明显优于 SNE 算法在二维空间中的聚类,t-SNE 算法有效减轻了拥堵问题,实现了全局特征的保留,使正常点与故障点的区分度提高,为使用神经网络准确识别电弧故障提供了数据条件。

4 基于 TSNE-SAPSO-BP 的电弧故障识别

4.1 TSNE-SAPSO-BP 模型构建

将特征进行非线性降维后,提出将 t-SNE 与 SAPSO 优化 BP 神经网络融合的模式构建方法。

BP 神经网络是一种多层前馈神经网络,由输入层、隐藏层和输出层组成。每个神经元都与前一层的所有神经元相连,通过加权和的方式传递信号,并经过激活函数进行非线性变换。

在设置 BP 神经网络的初始权值和阈值时,可以随机初始化或采用其他方法寻找最优值,将最优值代入神经网络训练。粒子群算法(PSO)<sup>[19]</sup>和 SAPSO 算法便是目前寻优效果最好的智能优化算法之一。针对 PSO 能快速收敛但易陷入局部最优的缺点,SAPSO 通过模拟退火算法(SA)进一步优化 PSO,提高了模拟退火的全局搜索能力。它来源于固体退火原理,在问题所在的解空间内生成初始解,模拟粒子在一定温度下的运动状态,记录各粒子的位置和全局最优位置,不断产生新解,并通过式(12)对比新旧解,选择最优解。

$$P = \exp\left(-\frac{\Delta f}{T}\right)$$

(12)

式中:  $\Delta f$  是适应度的变化值,当  $\Delta f < 0$  时,物体从高内能状态向低内能状态转变,则一定接受新解;当  $\Delta f > 0$  时,系统以概率  $P$  接受这个差的解,可以跳出局部最优解,避免模型早熟<sup>[20]</sup>。



将 t-SNE 降维后的三维特征输入故障识别算法 SAPSO-BP,建立 tSNE-SAPSO-BP 融合算法,融合算法框架如图 14 所示。

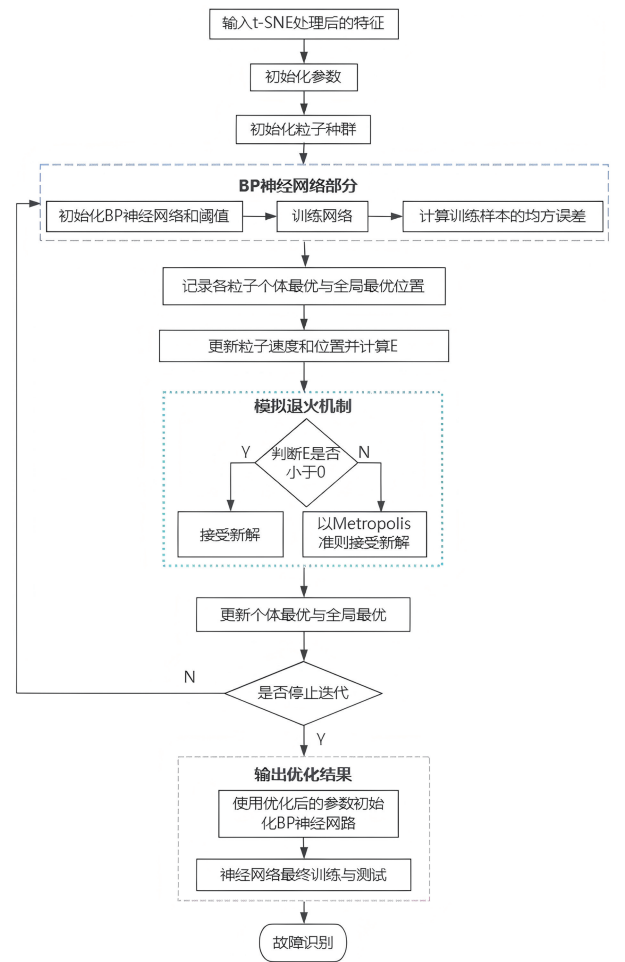


图 14 TSNE-SAPSO-BP 模型框架

Fig. 14 TSNE-SAPSO-BP model framework

4.2 故障识别

在识别过程中,各负载选取 5 000 个样本点作为训练集,其余样本点作为测试集。输入降维后的 3 个特征后,使用模拟退火粒子群算法初始化参数,得到一组神经网络的权值及阈值的最优解并代入神经网络训练。最终模型在第 13 次退火后,模型收敛,训练结束。以负载 1 为例,模型的部分输入量和输出量如表 4 所示。

训练集中负载 1 的辨识结果混淆矩阵如图 15 所示,该模型的输出辨识率达到 99.5%。

为更好地验证降维后的 tSNE-SAPSO-BP 融合模型对故障识别性能提升的有效性,这里将融合模型与提取的单个特征的识别效果进行对比,分别比较它们的识别准确率、虚警率和漏警率,各负载的识别性能如表 5 所示。由表 5 可见,tSNE-SAPSO-BP 融合模型在不同负载

表 4 神经网络输入和输出量(部分)

Table 4 Neural network inputs and outputs (Partial)

特征 1	特征 2	特征 3	预测类别
6.564 0	-52.431 7	22.793 3	1
11.521 9	-51.145 4	31.430 9	1
-4.428 0	-53.644 7	37.522 9	1
32.021 5	-50.934 0	32.939 1	1
-2.742 8	-52.516 1	25.001 7	1
73.921 0	26.247 7	51.755 6	2
-38.878 6	-7.636 5	65.869 2	2
-37.205 4	-7.540 3	67.296 2	2
-25.595 0	-13.034 5	75.737 4	2
91.721 0	24.247 7	49.755 6	2

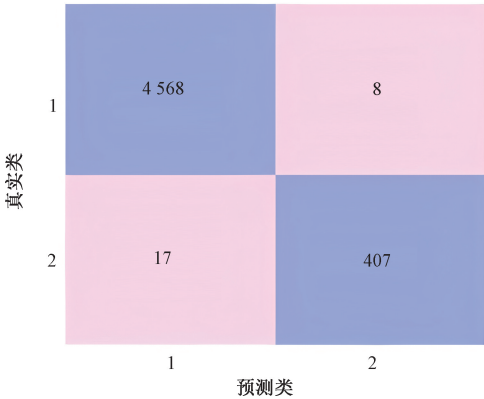


图 15 负载 1 辨识结果混淆矩阵

Fig. 15 Confusion matrix of load 1 identification results

上的平均识别准确率较改进后的变异系数特征、小波包变换细节分量、能量熵 E0 和电流变化率分别提升了 3.2%、16.8%、27.66%、33.5%。这说明 t-SNE 融合 SAPSO-BP 方法通过结合多种特征和数据模态,能更全面地捕捉故障特征,从而在电弧故障识别中表现出更高的精度和适应性,进一步提高了故障电弧识别准确率。同时,改进后的变异系数在单个特征中识别效果最好,说明将原始的变异系数特征构建为时频域特征的改进方法有助于进一步提高电弧故障识别的性能。

表 5 降维后的融合模型与单个特征的识别准确率对比

Table 5 Comparison of recognition accuracy between the dimensionality-reduced fusion model and

	single-feature models			(%)
	识别准确率	虚警率	漏警率	
tSNE-SAPSO-BP	99.2	0.17	4.01	
改进后的变异系数	96	2.51	6.74	
小波包变换细节分量	82.4	17.02	16.13	
能量熵 E0	71.54	28.35	25.15	
电流变化率	65.7	22.93	36.72	



将负载 1 诊断结果进行可视化,如图 16 所示,红色点为通过该模型诊断出的故障点,可以看到存在“平肩部”现象的区域均被标出,验证了该方法的有效性。

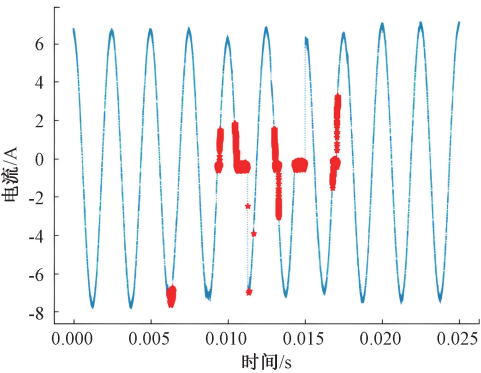


图 16 负载 1 故障点可视化  
Fig. 16 Load 1 fault point visualization

4.3 消融实验

为验证本文提出的改进的 CV 特征与故障识别算法的有效性与鲁棒性,在测试集上设置一定滑动步长的滑动窗口,以每个窗口为一个测试样本,在各窗口内提取各特征并进行故障电弧辨识,分别验证 CV 和降维方法在融合模型中的贡献和必要性。

在计算标准准确率(ACC)时,分别计算测试集上不同负载的识别准确率,以各负载的平均识别准确率作为标准 ACC。同时,为验证特征与模型的鲁棒性,设置窗口大小为 6 250 个电流采样点,滑动步长为 1 250 的滑动窗口,取各窗口的平均识别准确率为稳健 ACC 作为衡量鲁棒性的评价指标。稳健 ACC 的计算公式如式(13)所示。

$$\text{稳健 ACC} = \frac{\sum_{i=1}^l \text{ACC}_i}{l}$$

(13)

式中:  $l$  表示滑动窗口的数量;  $\text{ACC}_i$  表示第  $i$  个窗口的准确率。

在故障识别过程中,通过分别移除 CV 和 t-SNE 特征降维模块设计不同组合的消融实验,并在本文提出的故障识别算法 SAPSO-BP 基础上进行结果对比。消融实验的测试结果如表 6 所示。通过实验可以看出,将能量熵、小波包变换和电流上升率等特征进行特征降维后的识别准确率上升了 2.05%,减少了各特征间的信息冗余;引入改进后的变异系数 CV 后,识别精度提高了 6.62%;最后,在同时引入改进特征和特征降维模块后,稳健 ACC 提高了 15.56%,进一步验证了 CV 和 t-SNE 降维对融合模型的贡献,且该融合模型在不同负载和滑动窗口中均表现出良好的鲁棒性,能够适应不同负载下的电流数据特征。

表 6 消融实验结果  
Table 6 Results of ablation experiment

模块		实验结果/%	
CV	t-SNE 降维	标准 ACC	稳健 ACC
×	×	87.76	83.4
×	✓	89.81	86.14
✓	×	94.38	91.22
✓	✓	99.2	98.96

注:×代表移除。

5 结 论

为了解决单特征在识别飞机串联故障电弧时,特征的统一阈值难以设置的问题,基于故障电弧的频率分布特性,提出改进变异系数特征,构建了识别性能更好、适应性更强的时频域特征;并提出了一种 t-SNE 非线性特征降维结合 SAPSO-BP 进行多特征融合的串联电弧故障识别方法。单纯的聚类算法对正常和故障电弧的区分度不高,利用 t-SNE 对多维特征进行非线性降维和聚类可视化,不仅可以有效降低各种时频域特征之间的冗余度,且将 t-SNE 降维后的 3 维特征输入机器学习算法 SAPSO-BP 进行训练和更新,进一步提高了单特征故障识别的效率。结果表明 t-SNE 降维后的融合模型 tSNE-SAPSO-BP 在不同负载上的平均识别性能达到 99.2%,该模型具有更强的泛化性和鲁棒性,为更精准地识别故障点提供了条件,为进一步完善航空交流串联故障电弧检测系统和开展在线电弧故障定位提供了参考。未来,将围绕提出方法的实时性,降低模型训练、识别的复杂度和运算时间两方面开展优化研究,为研发串联故障电弧实时在线检测装置提供思路。

参考文献

[ 1 ] WU N, PENG M, WANG J, et al. Research on voltage-current type series arc fault detection method based on composite multi-scale entropy [ J ]. Electric Power Systems Research, 2025, 242: 111445.

[ 2 ] PARK H P, KWON G Y, LEE C K, et al. AI-enhanced time-frequency domain reflectometry for robust series arc fault detection in DC grids [ J ]. Measurement, 2024, 238: 115188.

[ 3 ] NING X, SHENG D, ZHOU J, et al. Arc\_EffNet: A novel series arc fault detection method based on lightweight neural network[J]. Electronics, 2023, 12(22): 4617.

[ 4 ] 刘艳丽,张凌玮,吕正阳,等. 基于 OD-LTP 的振动条件下串联型故障电弧检测方法研究[J]. 电子测量与仪器学报,2024,38(9):203-211.

- LIU Y L, ZHANG L W, LYU ZH Y, et al. Research on series fault arc detection method based on OD-LTP under vibration condition [J]. Journal of Electronic Measurements and Instrumentation, 2024, 38 (9): 203-211.
- [5] WANG Y, HOU L, PAUL K C, et al. ArcNet: Series AC arc fault detection based on raw current and convolutional neural network[J]. IEEE Transactions on Industrial Informatics, 2021, 18(1): 77-86.
- [6] GONG Q, PENG K, GAO Q, et al. Series arc fault identification method based on wavelet transform and feature values decomposition fusion DNN[J]. Electric Power Systems Research, 2023, 221: 109391.
- [7] 刘艳丽, 杨贺允, 刘乙雁, 等. 基于电源端电压的电动汽车电弧故障检测[J]. 仪器仪表学报, 2024, 45(5): 262-270.
- LIU Y L, YANY H Y, LIU Y Y, et al. Electric vehicle arc fault detection based on power supply terminal voltage [J]. Chinese Journal of Scientific Instrument, 2024, 45(5): 262-270.
- [8] JIANG J, LI W, WEN Z, et al. Series arc fault detection based on random forest and deep neural network[J]. IEEE Sensors Journal, 2021, 21(15): 17171-17179.
- [9] JIANG R, WANG Y, GAO X, et al. AC series arc fault detection based on RLC arc model and convolutional neural network [J]. IEEE Sensors Journal, 2023, 23(13): 14618-14627.
- [10] 唐圣学, 马晨阳, 勾泽. 基于时频特征融合与 GWO-ELM 的棒控电源早期故障状态辨识方法[J]. 仪器仪表学报, 2023, 44(1): 121-130.
- TANG SH X, MA CH Y, GOU Z. Early Fault State identification method of rod-controlled power supply based on time-frequency feature fusion and GWO-ELM [J]. Chinese Journal of Scientific Instrument, 2023, 44(1): 121-130.
- [11] 刘艳丽, 王浩, 张帆. 电动汽车串联型电弧故障检测方法[J]. 电子测量与仪器学报, 2023, 37(6): 222-231.
- LIU Y L, WANG H, ZHANG F. Electric vehicle series arc fault detection method [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37 (6): 222-231.
- [12] GONG Q, PENG K, WANG W, et al. Series arc fault identification method based on multi-feature fusion[J]. Frontiers in Energy Research, 2022, 9: 824414.
- [13] 董志文, 苏晶晶. 基于变分模态分解能量熵混合时域特征和随机森林的故障电弧检测方法[J]. 电气技术, 2024, 25(1): 1-7.
- DONG ZH W, SU J J. Arc fault detection method based on variational mode decomposition energy entropy hybrid time domain feature and random forest [J]. Electrical Technology, 2024, 25(1): 1-7.
- [14] YIN Z, WANG L, ZHANG B, et al. An integrated DC series arc fault detection method for different operating conditions [J]. IEEE Transactions on Industrial Electronics, 2020, 68(12): 12720-12729.
- [15] 江永鑫, 陈丽安. 基于 IIST 与 SVM 的串联故障电弧诊断方法研究[J]. 电力科学与技术学报, 2023, 38(5): 159-168.
- JIANG Y X, CHEN L AN. Research on series fault arc diagnosis method based on IIST and SVM [J]. Journal of Electric Power Science and technology, 2023, 38(5): 159-168.
- [16] 郭洲, 张军, 李隆, 等. 基于 Kmeans++ 聚类的光伏系统直流电弧故障检测研究[J]. 太阳能学报, 2024, 45(11): 320-329.
- WU ZH, ZHANG J, LI L, et al. Research on DC arc fault detection of photovoltaic system based on Kmeans++ clustering [J]. Journal of Solar Energy, 2024, 45(11): 320-329.
- [17] 聂兴毅, 黄华, 李旭东, 等. 基于改进小波包能量熵和阈值自适应的切削颤振在线监测[J]. 仪器仪表学报, 2024, 45(5): 227-238.
- NIE X Y, HUANG H, LI X D, et al. On-line monitoring of cutting chatter based on improved wavelet packet energy entropy and adaptive threshold [J]. Chinese Journal of Scientific Instrument, 2024, 45(5): 227-238.
- [18] 崔芮华, 李泽, 佟德栓. 基于三维熵距和熵空间的航空电弧故障检测与分类技术[J]. 电工技术学报, 2021, 36(4): 869-880.
- CUI R H, LI Z, TONG D SH. Aviation arc fault detection and classification based on three-dimensional entropy distance and entropy space [J]. Transactions of China Electrotechnical Society, 2021, 36(4): 869-880.
- [19] 佟德栓. 航空交流电弧故障试验及诊断方法研究[D]. 天津: 河北工业大学, 2021.
- TONG D SH. Research on aero AC arc fault Test and diagnosis method [D]. Tianjin: Hebei University of Technology, 2021.
- [20] 刘锋. 基于 PCA-SAPSO-BP 神经网络的瓦斯涌出量预

测研究[J]. 煤矿安全, 2023, 54(4): 60-68.

LIU F. Research on gas emission prediction based on PCA-SAPSO-BP neural network [J]. Coal Mine Safety, 2023, 54(4): 60-68.

## 作者简介



屈慧妍, 2023 年于鲁东大学获得学士学位, 现为鲁东大学硕士研究生, 主要研究方向为电弧故障识别技术。

E-mail: 2220599185@qq.com

**Qu Huiyan** received her B. Sc. degree from Ludong University in 2023. Now she is a M. Sc. candidate at Ludong University. Her main research interests include arc fault identification technology.



E-mail: daidaiquanquan123@126.com

**Li Juan** (Corresponding author) received her B. Sc. degree from Shanxi University in 2003, M. Sc. degree from Northwestern Polytechnical University in 2006, and Ph. D. degree from Air Force Engineering University in 2018, respectively. Now she is an associate professor in Ludong University. Her main research interests include life expectancy prediction and health management.

李娟(通信作者), 2003 年于山西大学获得学士学位, 2006 于西北工业大学获得硕士学位, 2018 年于空军工程大学获得博士学位, 现为鲁东大学副教授, 主要研究方向为寿命预测与健康管理。