

# 基于多特征选择算法的功率变换器故障分类方法

池程芝<sup>1</sup>, 潘震<sup>1</sup>, 徐钊<sup>2</sup>, 张一童<sup>2</sup>

(1.中国航空无线电电子研究所 航空电子系统综合技术重点实验室, 上海 200233;  
2.西北工业大学 电子信息学院, 陕西 西安 710072)

**摘要:**对综合模块化航电电源转换模块的核心部件进行故障诊断的过程中,选择合适的特征能够有效提高模型的效率和分类准确率,极大地降低学习算法的计算复杂度。设计了典型的 Sepic 结构 DC-DC 变换器模型,对 DC-DC 变换器的典型故障类型进行故障模拟;通过仿真获取相应的原始数据,采用数据进行预处理、特征提取与多特征选择融合;利用 BP 神经网络方法对 DC-DC 变换器进行故障诊断分析,仿真验证了该方法的有效性。

**关键词:**特征选择;BP 神经网络;故障诊断;功率变换器;Sepic 结构

中图分类号:V240.2

文献标志码:A

文章编号:1000-2758(2022)03-0645-06

在某些特定的作战环境中,极端的作战环境会使航电系统中的器件参数更容易产生变化,进而影响飞机系统的安全性和可靠性<sup>[1-2]</sup>,如何保障作战效率成为目前亟待解决的问题。在航电系统中,电源转换模块(power conversion module, PCM)的功能是将飞机电源系统的供电电压转换为飞机内部机架的直流电压,通过控制电源阵列开关给各模块供电;DC-DC 功率变换器是 PCM 模块中的核心部分,其健康状态会影响到其他所有模块,是其他模块正常工作的基础。在 DC-DC 功率变换器中有各种不同的电压电流信号,如果不能对电压电流信号进行有效处理,会使得宝贵的监测信号资源变为“数据爆炸”,且占用航电系统中较大的计算资源,整个航电系统的时效性和稳定性会受到影响。因此对功率变换器进行特征选择至关重要<sup>[2-3]</sup>。

为此国内外学者开展了大量的研究工作,文献[3-5]分析了特征选择和特征降维在故障诊断方向的作用和意义。文献[6]提出了基于主成分分析和极限学习机相结合的故障诊断算法,利用主成分分析和方差贡献率对原始数据进行降维。文献[7]对

特征信号进行了微分,是一种快速的检测和识别方法。文献[8]提出堆叠降噪自动编码器的特征降维方法,对原始特征信号进行数据压缩,从而达到降维效果。文献[9]基于一维卷积神经网络进行自适应特征提取。随着大规模集成系统的使用,传感器数量指数型增长,特征降维所需时间大大增加,传统的方法在处理集成度较高的系统时,会存在精度较低,泛化能力较弱等问题,特征选择能够大大缩短对大规模数据进行数据挖掘的时间。本文利用统计分析、智能优化算法等不同技术之间的差异性和互补性,提出更为有效的多模型融合的特征选择技术,使得所选特征对于不同分类器能有更好的鲁棒性。

## 1 DC-DC 变换器失效器件分析

在 DC-DC 变换器中常见且易损器件有铝电解电容、开关晶体管和功率二极管,上述元器件是功率变换器中较易出现故障的部分。铝电解电容吸收输出电压的波动从而使得输出电压更加稳定,且去除噪声导致的输出电压中的交流部分。开关晶体管是

收稿日期:2021-09-03

基金项目:航空自然科学基金(20185553034)、国防基础科研计划(JCKY2019205C615007)、中国博士后科学基金(2018M633574)及国家自然科学基金(61803309,61603303)资助

作者简介:池程芝(1984—),中国航空无线电电子研究所高级工程师,主要从事预测与健康管理和航电综合研究。

通信作者:徐钊(1982—),女,西北工业大学副教授,主要从事寿命预测与健康管理和多无人机网络控制研究。

e-mail:zhaoxu@nwpu.edu.cn

整个变换器的核心器件,该器件作用是功率转换。功率二极管用于整波整流。上述器件对于 DC-DC 变换器的正常工作有着致命性的影响,因此利用特征选择和故障诊断算法对上述器件进行分类研究非常有必要。

本文以图 1 所示的 DC-DC 变换器电路为研究对象,在 Cadence 仿真环境中搭建电路。通过仿真软件中“探针”采集原始数据信息,用于后续数据挖掘与特征选择工作。

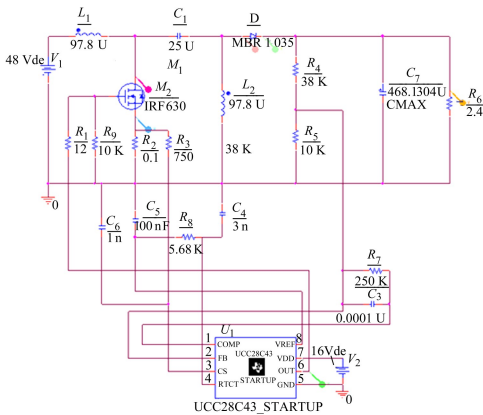


图 1 基于 Sepic 拓扑结构的电路模型

## 2 多模型融合框架

本文的故障诊断框架如图 2 所示。主要包含了数据预处理、特征提取、多特征选择算法融合、故障诊断。数据预处理对异常数据清洗和缺失数据添补整理;特征提取利用不同算法对原始信号进行分析及数据挖掘。特征选择模块是本文的重点部分,利用不同的选择算法对提取出的大量特征进行选择;最后通过 BP 神经网络算法建立所选特征变量与故障类型之间的对应关系。

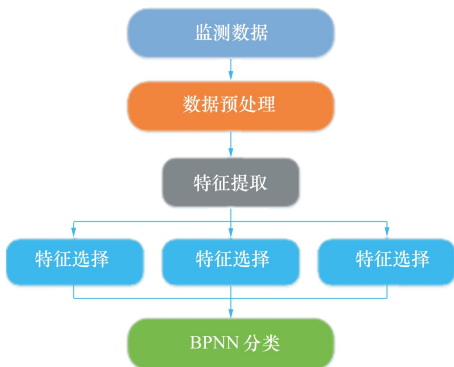


图 2 故障诊断框架

## 3 仿真实验验证

### 3.1 基于 ReliefF 算法的特征选择

ReliefF 算法是多分类特征选择算法。ReliefF 给特征以不同的权重,选择权重较大的特征,进行分类。ReliefF 在处理多类问题时,先从训练集中随机选取一个样本  $R$ ,从  $R$  相同类别的样本中找  $K$  个近邻样本,从不同类别的样本中找  $K$  个近邻样本,更新不同特征权重,其权重计算准则如(1)式所示

$$W_I(A) = W_I(A) - \sum_{j=1}^k \Delta(A, R, H_j) / (mk) + \sum_{C \in C_{class}(R)} \left[ \frac{P(C)}{1 - P(C_{class}(R))} \sum_{j=1}^k \Delta(A, R, M_j(C)) \right] / (mk) \quad (1)$$

式中:  $\Delta(A, R, H_j)$  是样本  $R$  与  $H_j$  在特征  $A$  上的差;  $C$  为与样本  $R$  不同的类别;  $P(C)$  为第  $C$  类的概率;  $C_{class}(R)$  为样本  $R$  所在的类;  $M_j$  为类别  $C$  中与样本  $R$  的第  $j$  个最近邻样本;  $\Delta(A, R, M_j(C))$  为样本  $R$  与样本  $M_j$  在特征  $A$  上的差。

### 3.2 基于 SVM-RFE 算法的特征选择

递归特征消除法 (recursive feature elimination, RFE) 反复多次创建模型,并在每次迭代时保存最好的特征,在下一次迭代时,采用上一次建模中没有用过的特征来构建下一个模型,最终找到一个性能最优的特征子集。

SVM-RFE 特征选择算法利用支持向量机 (support vector machine, SVM) 的分类性能作为特征选择评价标准。SVM-RFE 起始于全部特征,然后每次移去一个特征直到特征集合为空,移去的特征是所有特征中  $\| \omega \|^2$  最小的一个,  $\omega$  为支持向量机中的超平面法向量。这样对某一变量  $i$ ,排序评价准则如(2)式所示。

$$R_c(i) = \| w^{(i)} \|^2 = \sum_{k,j} a_k^{*(i)} a_j^{*(i)} y_k y_j K^{(i)}(x_k, x_j) \quad (2)$$

式中:  $K^{(i)}$  为变量  $i$  被移除之后训练数据的 Gram 矩阵;  $a_k^{*(i)}$  为支持向量的相关系数。算法的流程图如图 3 所示。

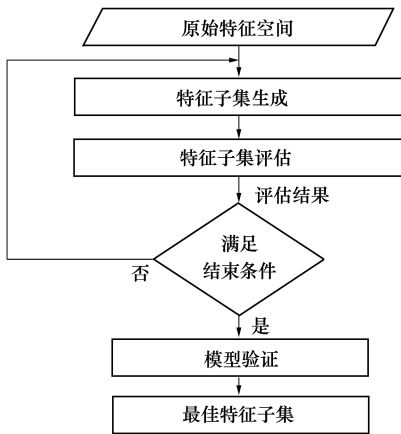


图3 SVM-RFE 算法流程图

### 3.3 基于随机森林算法的特征选择

随机森林算法(random forest, RF)是基于决策树(decision tree, DT)算法,其利用多棵决策树对样本进行训练并预测。算法的流程图如图4所示。

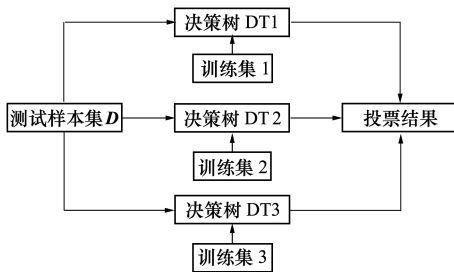


图4 随机森林算法基本思路

CART 决策树(classification and regression tree, CART)算法使用基尼系数来选择划分属性。基尼系数可表示为

$$G_{ini}(D) = \sum_{k=1}^{l_y} \sum_{k \neq k'} P_k P_{k'} = 1 - \sum_{k=1}^{l_y} P_k^2 \quad (3)$$

式中:  $D$  表示数据集;  $k$  表示第  $k$  类样本;  $P_k$  和  $P_{k'}$  分别表示不同类别标签的概率。可根据(3)式,计算出属性  $a$  的基尼系数,获得基尼指数定义

$$G_{ini\_index}(D, a) = \sum_{v=1}^V \frac{|D^v|}{|D|} G_{ini}(D^v) \quad (4)$$

式中:属性  $a$  有  $V$  个可能的取值  $\{a^1, a^2, \dots, a^V\}$ , 使用  $a$  来对样本集  $D$  进行划分,会产生  $V$  个分支节点,其中第  $v$  个分支结点包含了  $D$  中所有在属性  $a$  上取值为  $a^v$  的样本,记为  $D^v$ 。在候选属性集合  $A$  中,选择基尼指数最小的属性作为最优划分属性,既  $a_* = \operatorname{argmin} G_{ini\_index}(D, a)$ 。

### 3.4 特征系数加权融合

多特征选择算法融合模型可用(5)式表示

$$C = \sum_{i=1}^N w_i C_i(y, x) \quad (5)$$

式中:  $w_i$  表示第  $i$  个特征选择算法的权重系数;  $x$  表示训练数据;  $y$  表示测试数据;  $N$  表示模型总数,  $C_i(y, x)$  表示第  $i$  个特征选择算法选择结果。采取加权平均法对多类模型结果进行融合,既  $w_i = \frac{1}{N}$ 。

## 4 仿真验证

本部分的故障诊断算法研究主要针对于 DC-DC 变换器中硬故障类型,如易损器件的短路和开路故障,故对前述的 3 个易损器件分别进行开路 and 短路的故障仿真,由于开关晶体管的开路故障无法在仿真软件中实现,所以本部分对其余 5 种故障进行特征选择和分类研究。DC-DC 变换器的主要参数为输出电压,在正常状态和故障状态下的输出电压值会有明显的差异。采集不同故障类型下的输出电压值,可以根据输出电压的特征信息进行有效的分类。在实际电路中,电压信号可以采取直接测量的方式获取,而对于电流信号,电路中往往不存在采样输出接口,需要利用电磁感应的原理来获取,因此本文选择输出电压作为信号特征。

对输出电压提取时域特征如表 1 所示。

表 1 时域信号

特征名	公式
均值	$f_1 = \frac{1}{k} \sum_{i=1}^k x(i)$
标准差	$f_2 = \left( \frac{\sum_{i=1}^k (x(i) - f_1)^2}{k - 1} \right)^{\frac{1}{2}}$
偏度	$f_3 = \frac{\sum_{i=1}^N (X_i - \bar{X})^3}{N\sigma^3}$
峰度	$f_4 = \frac{\sum_{i=1}^N (X_i - \bar{X})^4}{N\sigma^4}$
峰峰值	$f_5 = \max(x_i) - \min(x_i)$
最大值	$f_6 = \max(x_i)$
最小值	$f_7 = \min(x_i)$

DC-DC 变换器是典型的数模混合电路,当电路发生故障时,信号各频段的能量较正常值必定发生

一定的变化,利用离散小波变换只保留了低频信息,丢失了高频信息,而利用小波包分解能同时保留低频和 高频信息,得到较为完备的故障信息。所以在时频域中,对输出电压进行 3 层小波包分解。

采用以上特征对各输出电压信号进行特征提取,对 5 种不同的故障类型各取 10 组进行示例,对特征提取部分获得的故障特征向量进行归一化处理,以消除不同特征分量之间的量纲之间的差异。

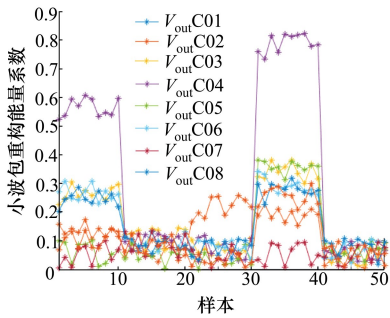


图 5 时域特征统计量曲线

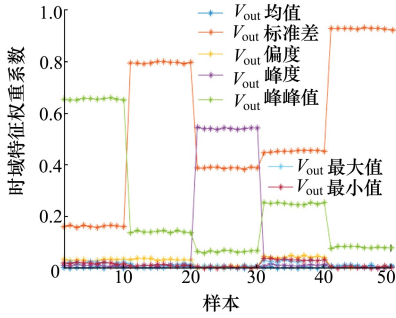


图 6 时频域特征统计量曲线

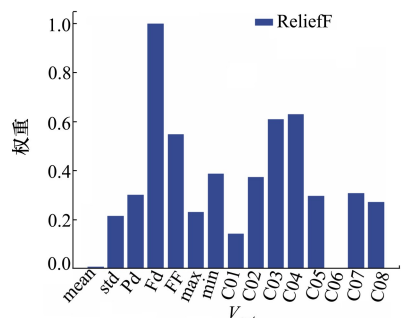


图 7 ReliefF 模型特征选择结果

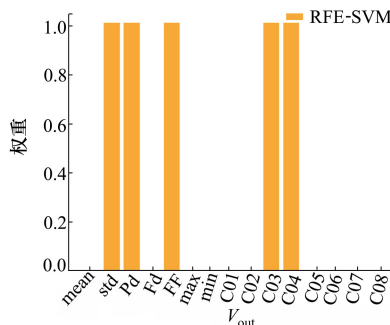


图 8 RFE-SVM 模型特征选择结果

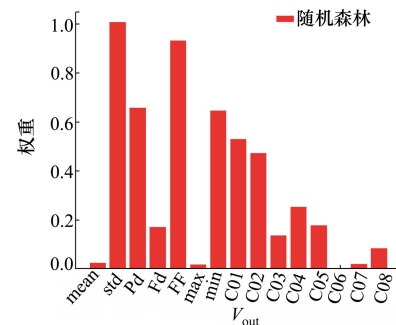


图 9 随机森林特征选择结果

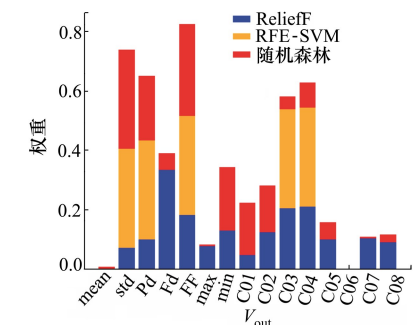


图 10 加权平均算法结果图

由图 7~9 可以看出,不同的选择算法对特征的选择有着相似性,也有差异性。在 ReliefF 算法中,时域信号中输出电压的偏度和峰峰值具有较大的贡献度,在时频域中,重构系数中第三和第四系数权重较大。RFE-SVM 算法的返回值为布尔值,可以看出除了 ReliefF 算法中权重较大部分外,输出电压的标准差也被选为主要特征之一。而随机森林算法和前者具有较大的差异,主要体现在时频域中,重构系数的第一和第二系数大于第三和第四系数;但在时域中,输出电压的标准差、偏度及峰峰值和其他算法结果相似。

仿真实验产生不同故障类型各 40 组,共 200 组为训练集,利用训练集对模型进行训练;留出不同故障类型各 10 组,共 50 组为测试集,进行故障诊断准确率的评估。如图 10 所示,根据加权平均算法的结

果,本文仿真实验中采用时域中输出电压标准差、偏度及峰峰值和频域中第三和第四系数作为特征矢量,首先将上述所选出的 5 个特征作为特征矢量,利用 KNN 算法进行分类研究。结果如图 11 所示。

由上述的时域特征矢量和小波包分解、重构各频段小波能量频域特征矢量可知,初选特征矢量维度较高,特征分量间存在较大的信息冗余,如果直接用于 BP 神经网络训练,得出的模型精度不高且计算时间较长。故采用多特征选择算法对其进行数据挖掘工作,ReliefF 算法、RFE-SVM 及随机森林 3 种特征选择算法的仿真结果如图 7~9 所示。

果,本文仿真实验中采用时域中输出电压标准差、偏度及峰峰值和频域中第三和第四系数作为特征矢量,首先将上述所选出的 5 个特征作为特征矢量,利用 KNN 算法进行分类研究。结果如图 11 所示。

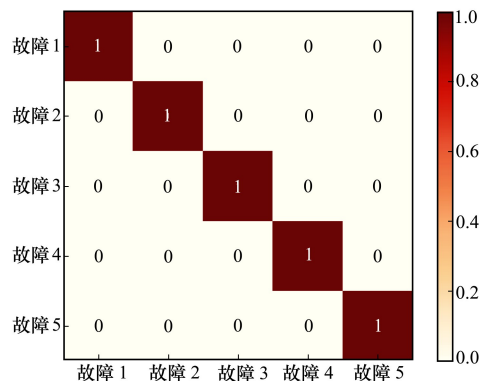


图 11 基于 KNN 模型硬故障分类结果示意图

选出的 5 个特征组成的特征矢量能够在 KNN 分类算法下实现 100% 准确率。为了更进一步验证算法效果,按照权重系数的大小,利用 KNN 算法分别对 5 组被选取特征进行分类研究。诊断结果如表 2 所示。

表 2 不同特征分类结果图

序号	选取特征	KNN 分类算法	
		准确率/%	特征权重
1	峰峰值	100	0.824 0
2	标准差	100	0.738 0
3	偏度	68	0.650 8
4	重构系数 04	40	0.626 5
5	重构系数 03	68	0.581 4

从表 2 可以看出,输出电压峰峰值和标准差都能在 KNN 算法中实现 100% 的准确率,为了验证多

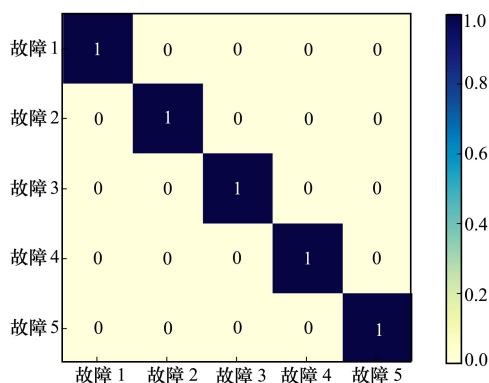


图 12 基于 BPNN 模型硬故障分类结果示意图

特征选择算法融合的鲁棒性,将输出电压峰峰值特征放入 BPNN 网络中进行验证,从图 12 可以看出,真实测试集和预测的测试集结果完全一致,在仿真实验的条件下,诊断精度也能达到 100%。结合图 11 和图 12 的分类结果可以得到,本文提出的多特征选择算法模型仅用最高权重系数的特征即输出电压的峰峰值,就能在不同的分类器中得到较高的准确率。

## 5 结 论

本文对 DC-DC 变换器不同故障进行失效模型分析,考虑到实际电路信号采集问题,选择了对 DC-DC 变换器影响最小的输出电压信号,通过仿真获取原始数据并进行特征提取、多特征选择算法降维,得到特征矢量,建立了 BP 神经网络分类模型。仿真结果表明本文提出的多特征选择模型融合方法能够进行多故障分类的任务。

相比于传统数据,高维数据中的数据量更多、特征之间的相关性也更为复杂,尽管部分机器学习的算法模型对高维数据也能实现分类预测等功能,然而在实际工程中,泛化性较弱的问题是非常常见的。本文在特征选择方面提出了多模型融合的算法,可以有效避免“过拟合”和“维数爆炸”问题,且增强了模型的泛化能力,相较于特征降维,对于研究人员而言,该方法更易理解。下一步工作是利用更多的数据集去验证算法模型的鲁棒性和准确性。

## 参考文献:

[1] 文成林,吕菲亚,包哲静,等. 基于数据驱动的微小故障诊断方法综述[J]. 自动化学报, 2016, 42(9): 1285-1299  
 WEN Chenglin, LYU Feiya, BAO Zhejing, et al. A review of data-driven micro-fault diagnosis methods[J]. Acta Automatica Sinica, 2016, 42(9): 1285-1299 (in Chinese)

[2] 毛鹏. 飞机航电系统故障分析方法与故障诊断技术探讨[J]. 电子测试, 2020(3): 106-107  
 MAO Peng. Discussion on fault analysis method and fault diagnosis technology of aircraft avionics system[J]. Electronic Testing, 2020(3): 106-107 (in Chinese)

[3] 赵燕辉. 基于大数据的航电综合分析系统的设计与实现[D]. 成都: 电子科技大学, 2020  
 ZHAO Yanhui. Design and implementation of avionics comprehensive analysis system based on big data[D]. Chengdu: University of Electronic Science and Technology of China, 2020 (in Chinese)

[4] 唐贵基, 王晓龙. 参数优化变分模态分解方法在滚动轴承早期故障诊断中的应用[J]. 西安交通大学学报, 2015, 49(5): 73-81  
 TANG Guiji, WANG Xiaolong. Application of variational mode decomposition method for parameter optimization in early fault diagnosis of rolling bearings[J]. Journal of Xi'an Jiaotong University, 2015, 49(5): 73-81 (in Chinese)

- [5] 赵光权, 葛强强, 刘小勇, 等. 基于 DBN 的故障特征提取及诊断方法研究[J]. 仪器仪表学报, 2016, 37(9): 1946-1953  
ZHAO Guangquan, GE Qiangqiang, LIU Xiaoyong, et al. Research on fault feature extraction and diagnosis method based on DBN[J]. Journal of Instrumentation, 2016, 37(9): 1946-1953 (in Chinese)
- [6] 秦琪. 航电电路模块 PHM 技术研究[D]. 西安: 西北工业大学, 2018  
QIN Qi. Research on PHM technology of avionics circuit module[D]. Xi'an: Northwestern Polytechnical University, 2018 (in Chinese)
- [7] 曾易寒. 基于输出电压特征的 DC-DC 变换器故障识别与寿命预测方法研究[D]. 成都: 电子科技大学, 2020  
ZENG Yihan. Research on fault identification and life prediction method of DC-DC converter based on output voltage characteristics[D]. Chengdu: University of Electronic Science and Technology of China, 2020 (in Chinese)
- [8] 李运富. 飞机二次电源故障模式分析与故障特征提取[D]. 天津: 中国民航大学, 2020  
LI Yunfu. Failure Mode analysis and fault feature extraction of aircraft secondary power supply[D]. Tianjin: Civil Aviation University of China, 2020 (in Chinese)
- [9] ZOU Fengqian, ZHANG Haifeng, SANG Shengtian, et al. An anti-noise one-dimension convolutional neural network learning model applying on bearing fault diagnosis[J]. Measurement, 2021, 186: 110236

## Power converter fault classification method based on multi-feature selection algorithm

CHI Chengzhi<sup>1</sup>, PAN Zhen<sup>1</sup>, XU Zhao<sup>2</sup>, ZHANG Yitong<sup>2</sup>

(1. Science and Technology on Avionics Integration Laboratory, China Institute of Aeronautical Radio Electronics, Shanghai 200233, China;  
2. School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China)

**Abstract:** In the process of fault diagnosis for the core components of the integrated modular avionics power conversion module, selecting appropriate features can effectively improve the efficiency and classification accuracy of the model, and greatly reduce the computational complexity of the learning algorithm. This paper first designs a typical Sepic structure DC-DC converter model to simulate the typical fault types of the DC-DC converter; secondly, the corresponding original data is obtained through simulation; after data preprocessing, feature extraction and using multiple feature selection fusion algorithm, BP neural network method is used finally for fault diagnosis analysis of DC-DC converter. The simulation verifies the effectiveness of the above method.

**Keywords:** feature selection; BP neural network; fault diagnosis; power converter; Sepic structure

**引用格式:** 池程芝, 潘震, 徐钊, 等. 基于多特征选择算法的功率变换器故障分类方法[J]. 西北工业大学学报, 2022, 40(3): 645-650

CHI Chengzhi, PAN Zhen, XU Zhao, et al. Power converter fault classification method based on multi-feature selection algorithm[J]. Journal of Northwestern Polytechnical University, 2022, 40(3): 645-650 (in Chinese)