

DOI:10.19651/j.cnki.emt.2108251

基于 EEMD 融合 BAS-CNN 的高压电缆故障诊断*

肖 旰 周 莉 李敬兆 刘泽朝 张 珂
(安徽理工大学电气与信息工程学院 淮南 232000)

摘要: 针对高压电缆故障复杂程度高,实时监测成本过高的问题,提出一种基于集合经验模态分解(EEMD)和天牛须搜索算法优化卷积神经网络(CNN)的组合诊断方法。将电缆护层电流数据经 EEMD 分解为若干个本征模态分量(IMF),结合相关系数选取与原信号相关度最大的分量,作为 CNN 网络的输入。为了提高网络模型的分类精度,使用天牛须算法(BAS)优化 CNN 诊断模型的超参数。以淮南某煤矿高压电缆电流数据为例,实验结果表明,EEMD 有效的将电流信号进行分解,所设计的 BAS-CNN 网络与 2 组人为确定 CNN 超参数的网络对比,BAS-CNN 具有最高的分类精度,监测准确率达到 96.95%。

关键词: 护层电流;EEMD;BAS-CNN;故障检测

中图分类号: TM73 **文献标识码:** A **国家标准学科分类代码:** 470.4017

High-voltage cable fault diagnosis based on EEMD and BAS-CNN

Xiao Gan Zhou Li Li Jingzhao Liu Zechao Zhang Ke

(School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232000, China)

Abstract: For the problem of high complexity of high-voltage cable faults and high cost of real-time monitoring, a combined diagnosis method based on ensemble empirical modal decomposition (EEMD) and tennies whisker search algorithm optimized convolutional neural network (CNN) is proposed. The cable sheath current data are decomposed into several eigenmodal components (IMF) by EEMD, and the component with the highest correlation with the original signal is selected by combining the correlation coefficients as the input of the CNN network. In order to improve the classification accuracy of the network model, the hyperparameters of the CNN diagnostic model are optimized using the aspen whisker algorithm (BAS). Taking the high-voltage cable current data of a coal mine in Huainan as an example, the experimental results show that EEMD effectively decomposes the current signal, and the designed BAS-CNN network has the highest classification accuracy with 96.95% monitoring accuracy compared with 2 groups of networks with artificially determined CNN hyperparameters.

Keywords: sheath current; EEMD; BAS-CNN; troubleshooting

0 引 言

由于国家电网特高压投资规模的不断扩大,对设备能否稳定运行的要求也越来越高,电缆作为电力网络中不可或缺或电气设备的,其正常运行直接影响着电网的安全性^[1]。但随着运行时间增加,受外力损坏、绝缘层腐蚀、异常运行等因素影响,电力电缆发生使用故障的频次增长迅速升高,可靠性和稳定性逐渐被下降,成为破坏输电网络健康安全快速运行的重大隐患^[2],如何快速实现故障在线诊断成为一大难题。因此,高压电缆的在线监测系统势在必行。

当前国内外高压交流电缆故障诊断方法主要有直流叠

加法、直流分量法、介质损耗法、局部放电检测法和红外测温法等^[3]。直流叠加法无法准确评估电缆绝缘的剩余寿命;直流分量法测量精度不高;介质损耗法适用范围有限;局部放电检测法采集的信号量微弱;红外测温法中仪器电池使用寿命短,维护成本高。并且上述方法在故障排查效率和准确性上有待提高。此外电缆线路发生故障具有不确定性,而且故障类型多,线路分布广。目前已有许多专家学者对接地线电流及接地线电流法进行了研究。杨静等^[4]在接地线电流的计算模型中考虑了加入了绝缘层泄漏电流和电容电流,建立了基于接地线电流数值变化分析的故障点和故障类型判据。Wang 等^[5]提出一种基于无监督学习和

收稿日期:2021-11-03

* 基金项目:国家自然科学基金(51874010,61170060)、北京理工大学高精尖机器人开放性研究项目(2018IRS16)、物联网关键技术研究创新团队(201950ZX003)资助

时域分析的数据挖掘方法应用于电缆护层环流在线监测,为环流判据的设定提供参考依据。Pang 等^[6]提出了一种新的绝缘在线监测方法——二端同步 $\tan\delta$ 测量,但谐波和频率波动对方法依旧有着影响。杨隆凯等^[7]通过电缆护层电流构建利萨如图形结合线芯温度变化的 Pettit 检测结果,提高了故障类型识别的准确性,但对复杂线路结构检测准确率不高。上述研究成果虽然都对护层电流的时域信息进行了分析,但并未研究如何对故障情况下的护层电流包含的频域特征进行深度分析。

由于电缆运行环境的复杂性,且电缆电流基频干扰较大,采用普通的时、频分析无法很好地提取电流信号特征,导致电缆故障检测准确率不高。本文以护层电流作为故障诊断的特征数据,提出了基于集合经验模态分解改进的卷积神经网络的高压电缆实时故障诊断方法。该方法通过采集高压电缆首末端接地线护层电流,并对电流信号降噪处理,采用集合经验模态分解(ensemble empirical mode decomposition, EEMD)提取本征模态分量,选取相关系数大的本征模态分量(intrinsic mode function, IMF)作为卷积神经网络的输入,自动提取特征;最后,为了提高精度,使用天牛须优化卷积神经网络的超参数,实现对故障检测并快速分类,从而达到高压电缆故障诊断智能化的目的。

1 数据预处理

本文主要研究高压电缆故障时的接地端护层电流信号,高压电缆在正常运行时产生的电流信号平稳,若高压电缆发生串联、并联和复合等故障,对应的电流信号也会发生变化,根据这一特性,利用电流信号可以判断高压电缆的运行状态,并结合网络模型对高压电缆的故障诊断,具体流程可以分为两个过程:数据处理和模型训练及分类,如图 1 所示。数据处理是对原始电流信号滤波去噪、特征提取,得到

电流信号的特征向量,模型训练及分类是以特征向量作为输入数据,通过天牛须搜索算法(beetle antennae search algorithm, BAS)优化的卷积神经网络(convolutional neural networks, CNN)的特征向量进行运算构成 BAS-CNN 网络,最终输出高压电缆运行故障诊断的结果。

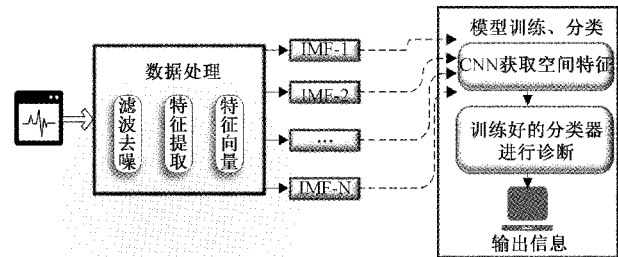


图 1 高压电缆运行故障诊断流程

1.1 滤波去噪

高压电缆运行环境复杂,运行中也会产生机械噪声,采集到的电流信号含有大量非平稳噪声干扰,需要对电流信号进行滤波去噪。通常噪声主要包含启停冲击、电磁干扰和环境干扰,环境干扰频率较低,启停冲击和电磁干扰的频率虽然分布较广,但信号能量低。根据信号中的噪声特性,文中选择高通滤波器和自适应滤波算法(least mean square, LMS)相结合的方法对电流信号进行滤波去噪。该方法先采用高通滤波器滤除信号中的环境噪声,再使用 LMS 算法进行整体噪声滤除。LMS 算法计算量小、性能稳定,通过设定理想的期望信号,利用要满足均方误差最小的判据对原始信号不断优化迭代在一定程度上减少信号中的噪声干扰^[8]。图 2、3 是原始电流信号与经过降噪处理后的信号。对比可知,采用此方法对电流信号噪声滤除明显,信号有用特征保留较好。

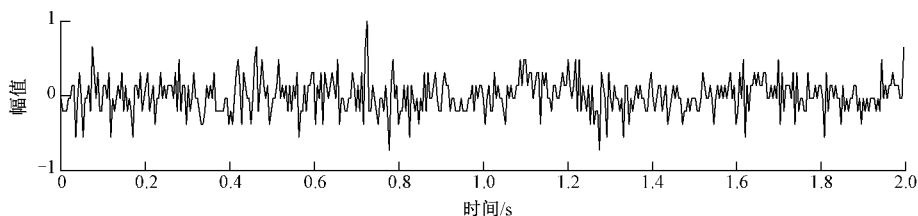


图 2 原始电流信号波形

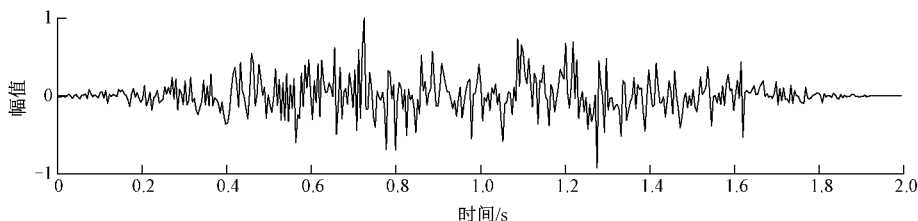


图 3 滤波降噪后的波形

1.2 电流信号非线性处理

采用时域分析法提取信号特征,只能得出信号幅值与时间之间的关系,无法获取电流信号的特征信息。EEMD 作为信号分析中常用的频域处理方法,通过加入白噪声,改变极值点结构,获得满足要求的信号特征,已广泛应用于电气设备电流信号特征分析,并取得了很好的结果^[9]。因此,文中使用 EEMD 提取电流信号的特征信息,得到高压电缆运行的故障数据。EEMD 特征提取操作步骤如图 4 所示。

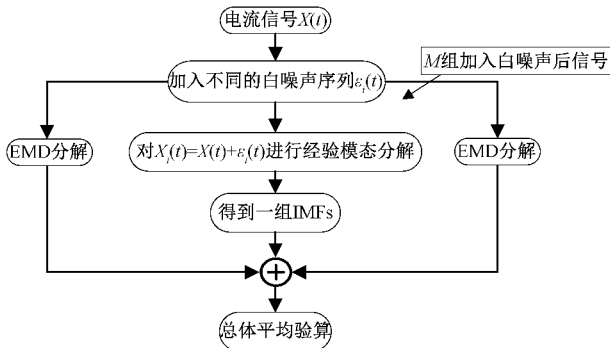


图 4 EEMD 分解流程

1) 设定原始信号的处理次数 M 。

2) 对一个原始信号 $X(t)$ 分别加上随机白噪声,通常加入的白噪声的标准差是原序列的标准差的 0.2 倍,构成一个新的信号如式(1)所示。

$$X_i(t) = X(t) + \epsilon_i(t) \tag{1}$$

式中: $X_i(t)$ 表示第 i 次实验中加入噪声后的新信号, $\epsilon_i(t)$ 表示第 i 次添加的随机白噪声序列, $i = 1, 2, \dots, M$ 。

3) 对所得含噪声的信号 $X_i(t)$ 分别进行 EMD 分解,得到一系列的 IMF 分量和的形式:

$$X_i(t) = \sum_{j=1}^J C_{i,j}(t) + r_{i,j}(t) \tag{2}$$

式中: $C_{i,j}(t)$ 为第 i 次加入白噪声后分解得到的第 j 个 IMF, $r_{i,j}(t)$ 是残余函数,代表信号的平均趋势, J 是 IMF 的数量, $j = 1, 2, \dots, J$ 。

4) 对步骤 2) 和 3) 循环 M 次,每次分解加入幅值不同的白噪声信号得到 IMF 的集合为:

$$C_{1,j}(t) C_{2,j}(t) \dots C_{m,j}(t) \tag{3}$$

5) 利用不相关序列的统计平均值为零的原理,将上述对应的 IMF 进行集合平均运算,得到 EEMD 分解后最终的 IMF,即:

$$C_j(t) = \frac{1}{m} \sum_{i=1}^m C_{i,j}(t) \tag{4}$$

式中: $C_j(t)$ 是 EEMD 分解的第 j 个 IMF。

相关系数是用以反映相关变量之间密切程度的统计指标,文中用本征模态分量与原始信号的相关系数作为判据,选取合适的 IMF 分量作为 CNN 输入。本征模态分量与原始信号的相关系数为:

$$r_j = \frac{\sum_{t=1}^N (y(t) - \bar{y})(c_j(t) - \bar{c}_j)}{\sqrt{\sum_{t=1}^N (y(t) - \bar{y})^2 \sum_{t=1}^N (c_j(t) - \bar{c}_j)^2}} \tag{5}$$

式中: r_j 为第 j 个 IMF 与原信号 $y(t)$ 的相关系数, $t = 1, 2, \dots, N$ 为信号的采样点。

2 网络模型设计

IMF 分量不仅在空间上存在数值表达特征,而且不同时间段的数值也存在紧密联系。通过设计 BAS-CNN 网络模型,利用 CNN 网络提取 IMF 分量的空间特征信息与时间特征信息,并且通过天牛须算法 BAS 自适应优化 CNN 模型结构使得模型诊断的精度、速度得以提升。不同卷积层使用不同尺寸的卷积核,对数据降维,提高模型分类准确率,网络模型结构如图 5 所示。

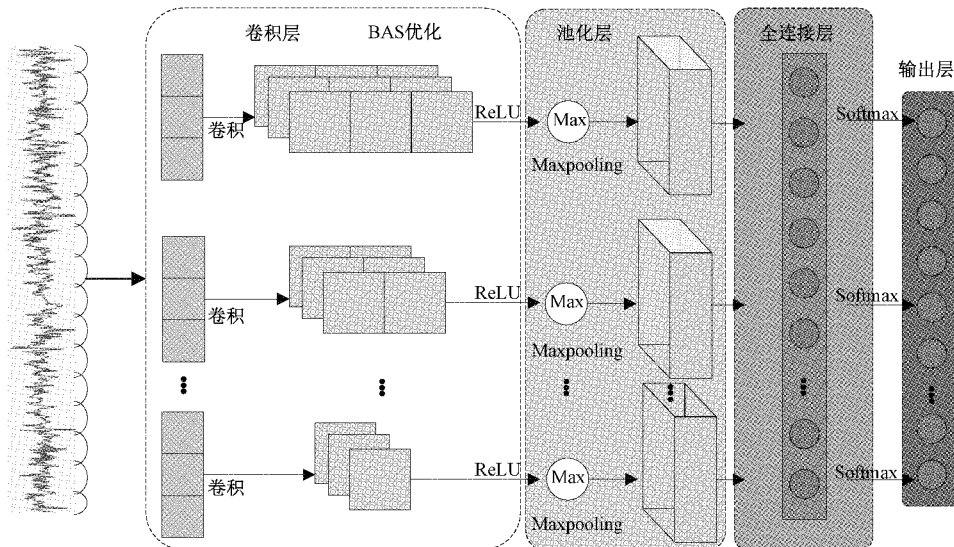


图 5 BAS-CNN 网络模型结构

2.1 BAS 优化的 CNN 网络

CNN 是近年来发展起来的一种高效特征识别方法,包括输入层、卷积层、池化层、全连接层和输出层,前一层的输出作为后一层的输入,逐级提取更深层次的有效特征^[10]。电流信号可以视为在时间轴上周期性采样形成的一维数据网格,不同卷积层中包含多种不同尺寸的卷积核,多个卷积核对应电流信号的 IMF 分量的局部具体特征进行观察,促使了局部特征向全局特征的不断抽象。为了增强算法模型特征表达能力,加速 CNN 模型的梯度下降,选用 ReLU 激活函数(rectified linear unit, ReLU),其表达式如下:

$$a_{l+1,m}(n) = f(y_{l+1,m}(n)) = \max\{0, y_{l+1,m}(n)\} \quad (6)$$

式中: $y_{l+1,m}(n)$ 为经过卷积运算得到的值, $a_{l+1,m}(n)$ 为 $y_{l+1,m}(n)$ 激活后的值。

最常用的池化层有平均池化和最大池化。文中采用最大池化层,对上一层数据图的缩放映射,以减少参数并且降低数据维度。最大池化操作如下:

$$P^{l+1}(n) = \max_{(n-1)H+1 \leq t \leq mH} \{q^l(t)\} \quad (7)$$

式中: $q^l(t)$ 池化表示第 l 层中第 t 个神经元的值, $t \in [(n-1)H+1, mH]$, H 是池化区域的宽度, $P^{l+1}(n)$ 表示池化操作后第 $l+1$ 层中神经元对应的值。

CNN 的分类阶段一般由全连接层和分类器组成,主要对上一层的信息进行整合、分类。

日前在故障诊断领域,提出了粒子群优化算法(particle swarm optimization, PSO)、灰狼优化算法(grey wolf optimizer, GWO)、鲸鱼优化算法(whale optimization algorithm, WOA)优化卷积神经网络超参数等提高诊断准确率的方法,这些算法取得了不错的成效^[11]。Jiang 等^[12]提出一种新的仿生智能全局优化算法——天牛须算法,该算法参数少、计算简单、鲁棒性强,适用于求解低维函数优化问题。文中采用 BAS 搜索算法来对 CNN 网络进行优化。

BAS 算法主要是根据天牛须左右触角对食物气味的感知程度的差别来判断下一步前进方向。在每一次迭代中,天牛任意方向飞行去搜索食物,利用式(8)和(9)定义天牛左右两只触角的位置坐标,便于检测下一个移动方向:

$$X'_l = X^l - l' * \vec{d} \quad (8)$$

$$X''_l = X^l + l' * \vec{d} \quad (9)$$

式中: X'_l, X''_l 分别表示第 l 次迭代左右两须位置坐标, l' 表示两须间距离, \vec{d} 是随机向量。根据左右两须触角感知的气味浓度差进行对比,利用变步长法更新天牛的位置:

$$X^{t+1} = X^t + \delta^{t+1} \vec{d} \text{sign}(f(X'_l) - f(X''_l)) \quad (10)$$

式中: δ^{t+1} 为天牛迭代 $t+1$ 次的探索步长, $\text{sign}(\cdot)$ 为符号函数, $f(X'_l), f(X''_l)$ 是左右两须的气味浓度。

2.2 网络模型设计流程

为了提高网络模型分类精度,使用反向传播算法(back-propagation through time, BPTT)实现网络模型的反向传播,逐层更新网络模型的权重矩阵,自适应提取相关系数大的

IMF 分量特征信息^[13]。具体流程如图 6 所示。

1) 使用滑动窗口数据^[14]处理本征模态分量分段制作数据集,将数据集按照一定的比例分为训练集、验证集和测试集。

2) 设定 BAS 参数,初始化天牛位置,探索步长,最大迭代次数。开始训练,将一只天牛作为一个待优化的卷积神经网络的结构个体,通过训练得到修正模型,以训练集的正确率作为该天牛的适应度函数^[15];训练过程中采用 BPTT 算法更新网络模型的权重参数,获得更全面的 IMF 特征信息;不断调试得到网络模型表现最佳的超参数。

3) 一旦已达到最佳正确率或最大迭代次数,则停止计算;如果未达到最佳正确率或最大迭代次数,则更新天牛位置,回到步骤 2) 继续循环。

4) 通过测试集和相关判断指标评估网络模型的性能。

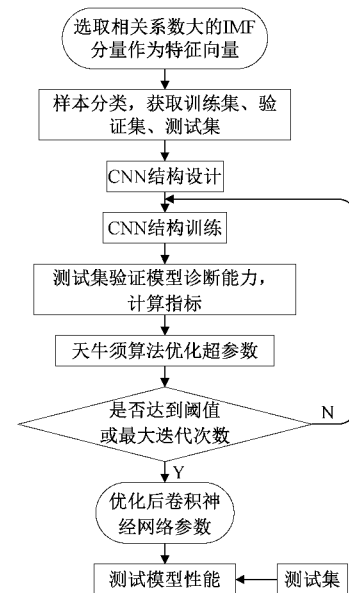


图 6 网络模型设计流程

2.3 网络模型分类效果评价指标

以平均绝对百分比误差(mean absolute percentage error, MAPE)作为网络性能的评判指标。MAPE 可以反映分类结果与原始数据之间的相对误差^[16]。假设网络模型分类结果为 $y = \{y_1, y_2, y_3, \dots, y_n\}$, 实际电流信号为 $x = \{x_1, x_2, x_3, \dots, x_n\}$, 那么 MAPE 的表达式为:

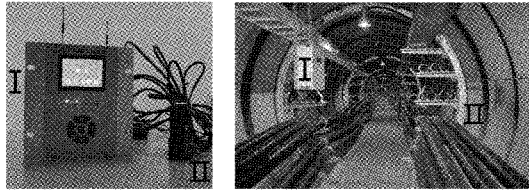
$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - x_i}{y_i} \right| \times 100\% \quad (11)$$

3 实验与结果分析

3.1 实验说明

进行故障植入试验,采集电流数据并进行处理来验证本文所提出的故障诊断方法。电流信号处理软件采用 MATLAB2019a;模型框架为 pytorch。实验平台和装置如图 7 所示,其中图 7(a)中的 I 为分析数据处理装置,II 为

微电流高精度电流互感器;图 7(b)中的 I 为通信节点, II 为高压电缆。



(a)高压电缆在线监测装置 (b)现场实验场景

图 7 实验平台和装置

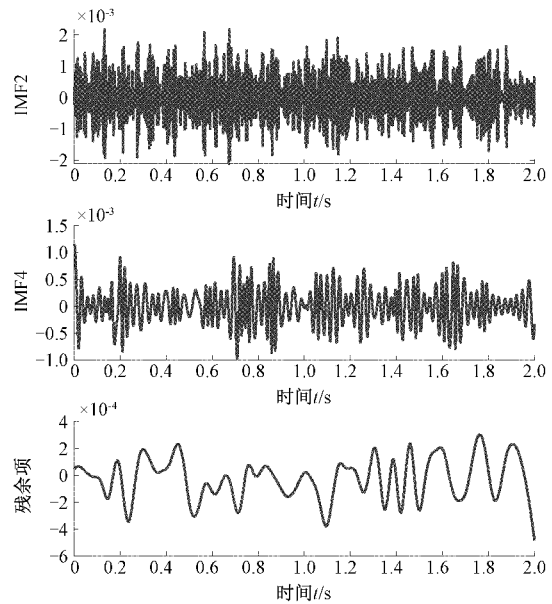
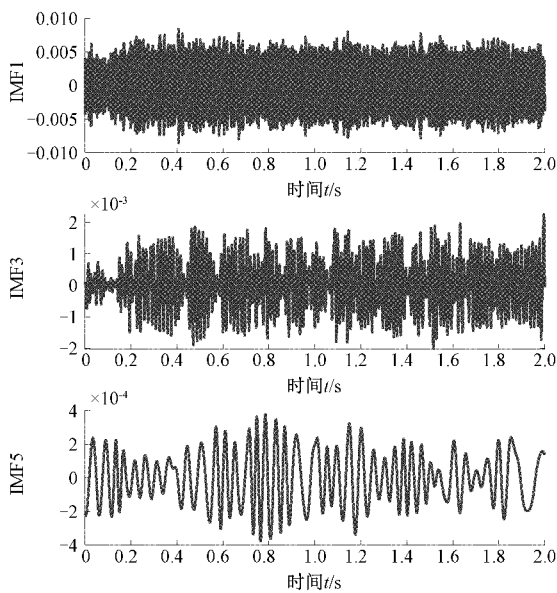
3.2 数据集制作

本文选取淮南某煤矿高压电缆工作时电流信号作为

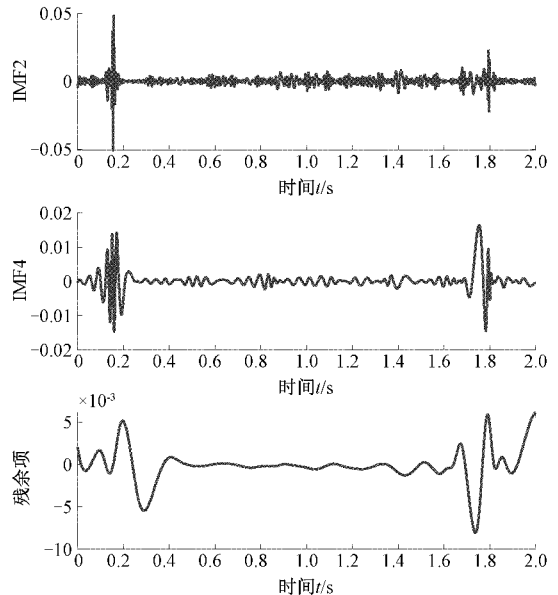
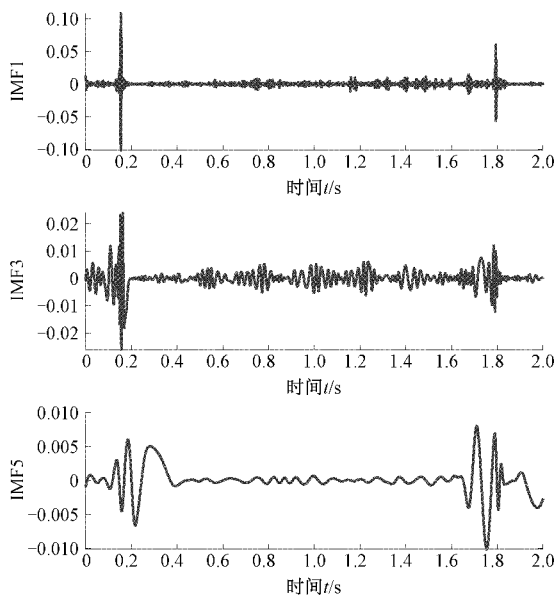
实验数据,包括电缆在 4 种不同状态下采集到的电流信号,分别为正常状态、串联故障状态、并联故障状态以及复合故障状态。对每段数据首先进行归一化,减少数据之间的相互干扰,再采用第 1 节介绍的数据处理方法,提取各数据样本的 IMF 特征分量,并将处理后的数据样本分为 2 组,70%作为训练集,30%为测试集。

3.3 EEMD 分解

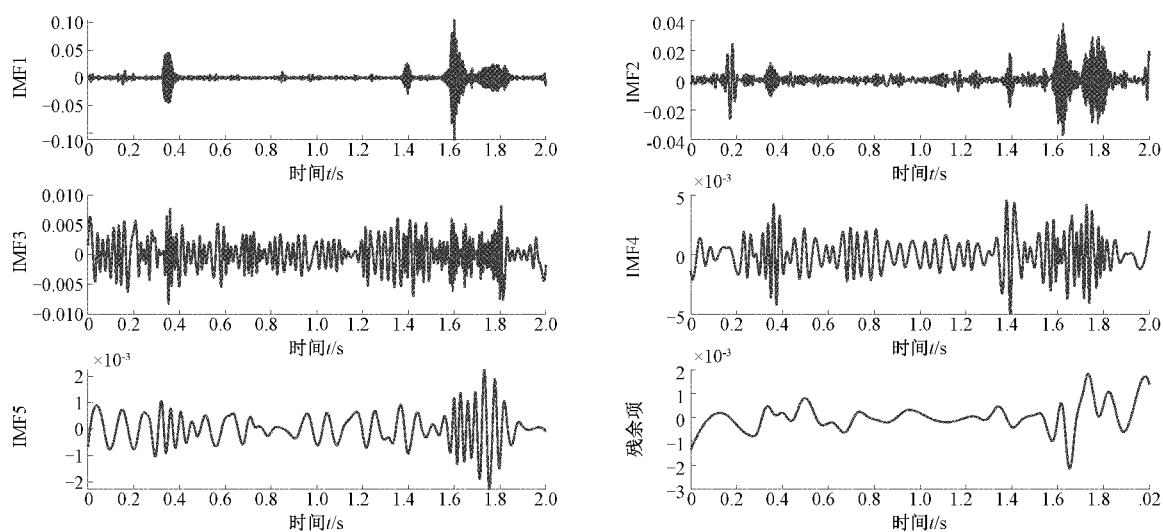
为了确定 IMF 的个数,本文研究中对 4 类数据分别进行了 EEMD 分解,每一种故障数据的序列长度为 2 000,并选出相关系数大的 IMF 分量。所得到的实验结果如图 8 所示。4 种状态下电流信号 IMF 分量相关系数如表 1 所示。



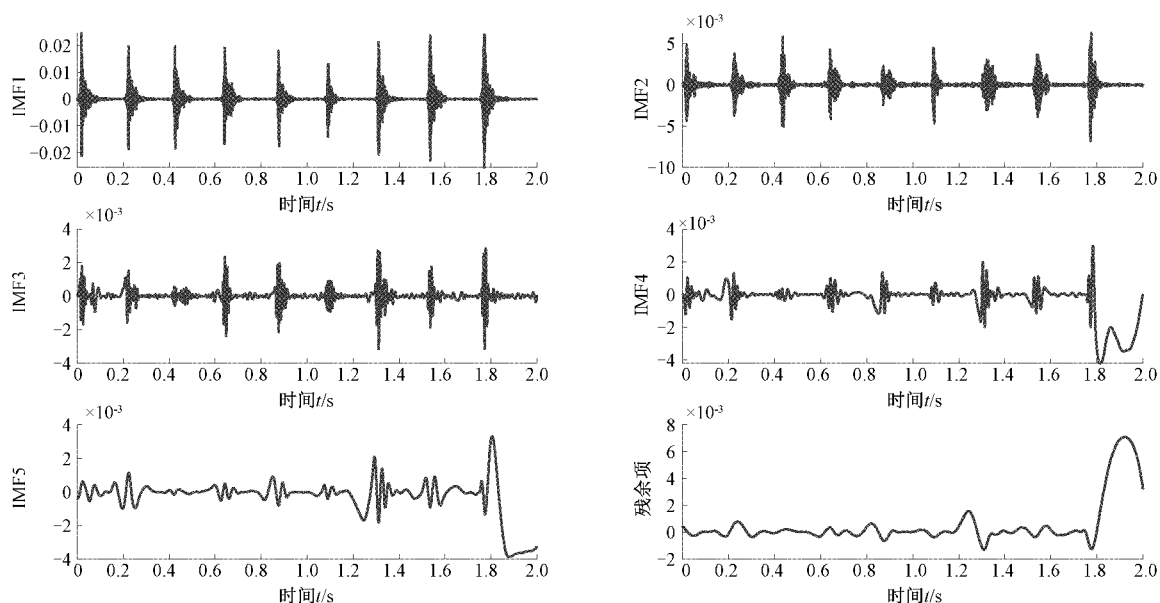
(a) 正常电流信号IMF分量图



(b) 串联故障电流信号IMF分量图



(c) 并联故障电流信号IMF分量图



(d) 复合故障电流信号IMF分量图

图 8 4 种信号 EEMD 分解图

表 1 4 种状态下电流信号 IMF 分量相关系数

IMF 分量	正常	串联故障	并联故障	复合故障
IMF1	0.002 104 7	0.410 01	0.197 69	0.010 916
IMF2	0.016 246	0.745 17	0.458 34	0.224 24
IMF3	0.004 465 6	0.240 63	0.850 63	0.945 49
IMF4	0.968 31	0.064 025	0.082 571	0.002 363 3
IMF5	0.001 555	0.018 218	0.024 53	0.000 359 06

由图 8 可知, 电流故障信号经 EEMD 分解后, 得到 5 个频率由高到低的 IMF 分量和 1 个余项。其中, IMF1 到 IMF4 所包含的有效信息较之 IMF5 要多。为了提高故障诊断准确率, 对不同的 IMF 分量进行相关系数计算, 根

据计算结果选取更加能够达到准确表征原始信号的目标。

由表 1 数据可知, 正常工作状态下选取 IMF4, 串联故障工作状态下选取 IMF2, 并联故障及复合故障工作状态下选取 IMF3 作为 CNN 网络输入。

3.4 网络模型参数设计

1) 天牛须算法参数设定

对于上文中提到的 CNN 网络结构,采用天牛须算法优化。其中天牛须算法中的各项参数为两须间距离 $l^0=4$, 初始步长 $\delta^0=2$,最大迭代次数 $T=40$ 、初始位置 $[a_0, k_0]$; 确定准确率为 BAS 适应度函数,对 CNN 网络进行迭代,计算 CNN 网络分类结果的准确率。

2) 优化后 CNN 结构

用 BAS 优化 CNN 超参数,其准确率随天牛更新迭代次数收敛趋势如图 9 所示,当迭代次数达到 16 时最大准确率值为 0.977,对应最优食物位置即最优参数组合 $[5, 50]$, 以同样的方式对 CNN 不同卷积层的卷积核尺寸进行确定,结果分别为 3 和 1。

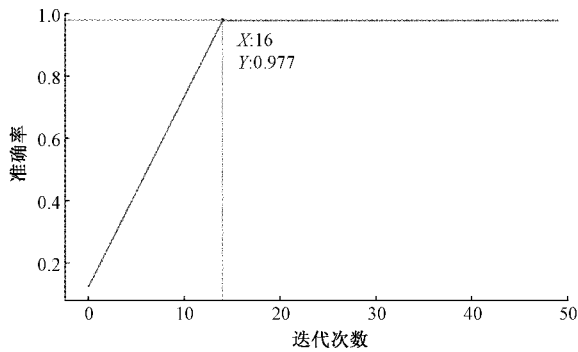


图 9 优化后准确率变化图

3.5 模型性能对比及评估

为了验证 BAS-CNN 网络的优越性,在相同测试集下使用人工随机选择 CNN 参数的 CNN1、CNN2 进行对比实验。表 2 所示为 3 种方法的实验数据(C1、C2、C3 为每层卷积核的大小,E1 为模型训练的迭代次数),损失函数如图 10 所示。

表 2 BAS 算法及人工选择的 CNN 超参数

超参数	CNN1	CNN2	BAS-CNN
C1	7	9	5
C2	5	7	3
C3	3	5	1
E1	30	100	50

从图 10 可以看出,BAS-CNN 模型损失函数值最低,且下降速度快,波动小。说明模型对故障检测的性能较好。

计算 10 次取均值作为计算结果。优化结果如表 3 所示。

可以看出,对比其他几组模型,文中提出的 BAS-CNN 网络模型计算的 MAPE 值低于其他模型,这说明 BAS-CNN 网络能够很好地检测出高压电缆故障运行,且故障监测的实时性很高。

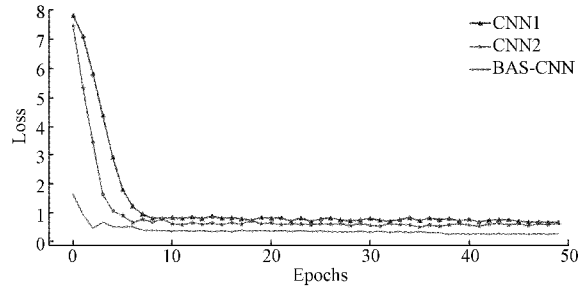


图 10 不同模型在同一验证集下的损失函数曲线

表 3 不同模型分类对比

模型	MAPE/%	时间 T/s
CNN1	29.38	109.769
CNN2	17.09	114.93
BAS-CNN	3.05	98.7

4 结 论

本文提出了 BAS-CNN 网络模型进行特征分类,首先通过 EEMD 提取高压电缆 4 类状态下电流信号特征作为网络模型的输入,再利用 BAS 算法适合全局寻优的特性来选择 CNN 最佳结构。与人工依据大量经验和实验来选择 CNN 参数相比,该算法能够自动选择 CNN 最佳结构。经实验验证,所设计的模型对高压电缆运行状态的诊断准确率更高,且诊断结果稳定性更好,实现了高压电缆故障诊断的智能化。

参考文献

- [1] 胡丽斌,张传升,谭笑,等. 退役电缆附件微观结构与电荷特性研究[J]. 中国电机工程学报, 2021, 41(2): 770-780.
- [2] BICEN Y. Trend adjusted lifetime monitoring of underground power cable [J]. Electric Power Systems Research, 2017, 143(2):189-196.
- [3] 郭卫,周松霖,王立,等. 电力电缆状态在线监测系统的设计及应用[J]. 高电压技术, 2019, 45(11): 3459-3466.
- [4] 杨静,朱晓岭,董翔,等. 基于护层电流的高压电缆故障在线监测和诊断[J]. 高电压技术, 2016, 42(11): 3616-3625.
- [5] WANG Y, YE H, ZHANG T, et al. A data mining method based on unsupervised learning and spatiotemporal analysis for sheath current monitoring[J]. Neurocomputing, 2019, 35(4):54-63.
- [6] PANG B, ZHU B, WEI X, et al. On-line monitoring method for long distance power cable insulation[J]. IEEE Transactions on Dielectrics & Electrical Insulation, 2016, 23(1):70-76.

- [7] 杨隆凯,夏向阳,周正雄. 基于利萨如图形与温度 Pettitt 检测的高压电缆故障诊断方法[J]. 电力科学与工程,2021,37(7):34-41.
- [8] 石欣,范智瑞,张杰毅,等. 基于 LMS-随机森林的肌电信号下肢动作快速分类[J]. 仪器仪表学报,2020,41(6):218-224.
- [9] 施杰,伍星,刘韬. 基于 MPDE-EEMD 及自适应共振解调的轴承故障特征提取方法[J]. 电子测量与仪器学报,2020,237(9):47-54.
- [10] 杨小东,纪国宜. 基于融合 CNN 的滚动轴承剩余寿命预测[J]. 国外电子测量技术,2021,316(3):62-67.
- [11] 程换新,黄震. 基于改进 PSO 优化 RNN 的短期电力负荷预测模型[J]. 电子测量技术,2019,42(20):94-98.
- [12] JIANG X Y, LI S. BAS: Beetle antennae search algorithm for optimization problems[J]. International Journal of Robotics and Control, 2018, 1(1):1-5.
- [13] 郭晨霞,刘佑祺,杨瑞峰. 基于萤火虫算法优化 BP 神经网络的光强度补偿方法[J]. 电子测量技术,2021,369(13):6-10.
- [14] ZHANG W, PENG G, LI C, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals[J]. Sensors, 2017, 17(2):425-435.
- [15] 闫重熙,陈皓. 基于改进天牛须搜索算法优化 LSSVM 短期电力负荷预测方法研究[J]. 电测与仪表,2020,57(6):6-11,18.
- [16] ZHENG J Q, DU J, LIANG Y T, et al. Deeppipe: Theory-guided lstm method for monitoring pressure after multi-product pipeline shutdown [J]. Process Safety and Environmental Protection, 2021, 155: 518-531.

作者简介

肖旸,硕士研究生,主要从事电气设备故障诊断。

E-mail:xgan0220@163.com

周莉,副教授,博士,主要从事智能控制、运动控制、智能检测与信息处理。

E-mail:359915143@qq.com

李敬兆,教授,博士,主要从事矿山物联网技术及应用。

E-mail:jzhli@aust.edu.cn

刘泽朝,博士研究生,主要从事电气设备故障诊断与健康预测。

E-mail:499224516@qq.com

张珂,硕士研究生,主要从事电力负荷预测。

E-mail:314850336@qq.com