

DOI: 10.12158/j.2096-3203.2024.02.018

## 基于 ADASYN 数据平衡化的 PSO-BPNN 变压器套管故障诊断

杨昊<sup>1</sup>, 胡文秀<sup>1</sup>, 张璐<sup>2</sup>, 陈晋鹏<sup>1</sup>, 周思佳<sup>1</sup>, 赵思瑞<sup>1</sup>

(1. 西安工程大学电子信息学院, 陕西 西安 710600;

2. 国网陕西省电力有限公司电力科学研究院, 陕西 西安 710100)

**摘要:** 变压器套管作为设备重要的绝缘部件, 其绝缘性能直接影响着设备的安全运行。为诊断变压器套管绝缘状态, 改善变压器套管油中溶解气体的小样本不平衡数据对变压器套管故障诊断结果的影响, 使用粒子群优化结合反向传播神经网络 (particle swarm optimization combined with back propagation neural network, PSO-BPNN) 和自适应综合过采样 (adaptive synthetic sampling, ADASYN) 算法对变压器套管进行故障诊断。首先收集变压器套管的历史故障数据, 建立具有明确故障类别的变压器套管油中溶解气体样本集, 并通过 ADASYN 算法对原始数据中的少数类样本进行合成, 得到平衡后的故障数据, 然后将平衡后的油中溶解气体作为模型输入, 故障状态作为标签输出, 通过 PSO-BPNN 模型对变压器套管进行诊断, 最后在原始样本集下使用反向传播神经网络 (back propagation neural network, BPNN)、遗传结合反向传播神经网络 (genetic combined with back propagation neural network, G-BPNN) 算法、布谷鸟搜索结合反向传播神经网络 (cuckoo search combined with back propagation neural network, CS-BPNN) 算法以及 PSO-BPNN 模型对套管进行诊断。结果表明, 针对变压器油纸套管绝缘状态进行故障诊断的多个模型中, 基于 ADASYN 平衡数据后的 PSO-BPNN 模型和其他模型相比准确度最高, 能有效减小小样本不平衡数据对诊断结果的影响, 为判断变压器油纸套管绝缘性能提供了有效方法。

**关键词:** 变压器套管; 故障诊断; 油中溶解气体; 反向传播神经网络 (BPNN); 不平衡数据; 自适应综合过采样 (ADASYN)

中图分类号: TM85

文献标志码: A

文章编号: 2096-3203(2024)02-0170-09

## 0 引言

变压器套管复杂的运行环境会影响其绝缘性能, 进一步发展可能会引起变压器着火、爆炸等严重事故。据 CIGRE 工作组统计的世界范围内变压器套管运行结果显示, 变压器套管的使用比例中, 油浸纸套管占 69%, 而故障比例中, 油浸纸套管高达 62%。据全国电网 13 类输变电设施统计, 因充油式套管导致 220 kV 及以上变压器非计划停运时间为 109.85 h, 造成非计划停运次数最多<sup>[1-2]</sup>。变压器运行状态会直接影响到电力系统安全稳定运行<sup>[3]</sup>。因此, 研究变压器套管故障诊断方法, 不仅为我国智能电网开发建设提供理论支持, 也对电力系统的安全运行有着实际的工程指导意义。

目前, 变压器套管的运维手段有很多, 如红外检测、电气试验、频域介电谱、特高频检测等<sup>[4-6]</sup>。这些运维手段在一定程度上监测套管运行状态的同时也存在一些不足之处, 如易受外界环境干扰、易对设备绝缘造成一定损伤、不能实现设备在线检测、难以检测在运设备内部故障等。

变压器油浸套管运行过程中内部的油中溶解

气体数据更能直接反映其运行状态<sup>[7-8]</sup>。收集大量套管故障历史数据, 其中套管温度、微水含量、介损等数据稀少且杂乱、分布不均恐难以利用, 而套管油中溶解气体在故障诊断方面颇具利用价值。有研究表明, 油中溶解气体分析 (dissolved gases analysis, DGA) 技术可以对变压器内部绝缘状态进行诊断<sup>[9-12]</sup>。然而, 传统 DGA 技术无法表征故障特征量与故障类型间的映射关系, 因此, 将机器学习与 DGA 方法相结合是一种有效手段<sup>[13-15]</sup>。已有研究表明, DGA 技术在有油设备故障识别中被广泛应用, 其在变压器故障识别中效果突出, 而变压器套管属于少油设备, 其运行状态、充油量、电场分布、芯子结构以及油纸比率和变压器本体存在较大差异。因此, 基于 DGA 技术的变压器本体故障诊断方法不能直接用于变压器套管。

同时, 变压器套管故障数据少于正常运行数据, 因此, 现有的变压器套管数据集存在小样本不平衡问题, 会在一定程度上对诊断结果产生影响。国内外学者针对这种不平衡数据问题研究数据预处理方法<sup>[16-18]</sup>, 从数据预处理入手, 改善变压器套管故障数据不平衡问题, 典型的处理方法有过采样/欠采样。根据样本特点, 通过对少数类样本进行过采样处理来达到平衡化目的。自适应综合过

收稿日期: 2023-10-10; 修回日期: 2023-11-21

基金项目: 国家自然科学基金资助项目 (52007138)

采样(adaptive synthetic sampling, ADASYN)可以自适应地偏移决策边界,聚焦在较难学习的样本上,十分适用于套管故障数据,改善数据不平衡导致的诊断结果不佳问题,提高诊断模型性能。针对小样本现象,反向传播神经网络(back propagation neural network, BPNN)在解决此类问题中独具优势<sup>[19]</sup>。但同时也存在易陷入局部最优的缺点,采用新型群智能算法对其进行改进是一种较为可行的方法<sup>[20-22]</sup>。粒子群优化(particle swarm optimization, PSO)算法采用实数编码,能很好地解决高维问题。

综上所述,变压器套管作为设备关键部件之一,需要研究其故障诊断方法,及时掌握其绝缘状态。文中针对变压器套管故障数据不平衡的问题,采用 ADASYN 方法对更难学习的少数类故障数据进行合成,增加少数类样本数,改善不平衡问题。针对变压器套管故障数据少的问题,使用 BPNN 作为诊断模型,并通过新型群智能算法对其进行优化,从而达到提高诊断模型准确度的目的。为判断变压器油纸套管绝缘性能提供有效方法,为电力系统的安全稳定运行提供技术保障。

## 1 PSO-BPNN 算法

BPNN 根据网络路径进行运算,包括正向传播和误差反向传播 2 个部分。由于神经网络越复杂其误差函数也越不规则,对误差函数进行逼近时,易出现梯度弥散现象,因此引入 PSO 算法进行改善。

### 1.1 算法总体框架

PSO 算法是一种经典的群智能算法,通过模拟鸟类的觅食行为以及社会活动,利用群体中个体对信息的共享,使整个求解问题有序化,从而获得最优解。利用 PSO 算法寻找 BPNN 算法中权重、阈值的最优值,减小 BPNN 算法陷入局部最优的现象,提高诊断模型的性能。

利用诊断模型对数据进行训练,需要先确定模型网络拓扑,利用特征参量的维数确定模型输入层数,根据故障类型确定模型输出层数,经过不断调整得到最佳隐藏层。图 1 为粒子群优化结合反向传播神经网络(particle swarm optimization combined with back propagation neural network, PSO-BPNN)算法总体框架。将 BPNN 中的权重、阈值赋随机数,粒子群算法中粒子速度、位置以及网络参数初始化,各个粒子在其搜索空间内搜索各层间的连接权值和节点阈值,用神经网络算法得出对应的输出。粒子的参数更新公式简单,计算速度较快,粒子群能够自适应地调整惯性权重从而跳出局部极小值。

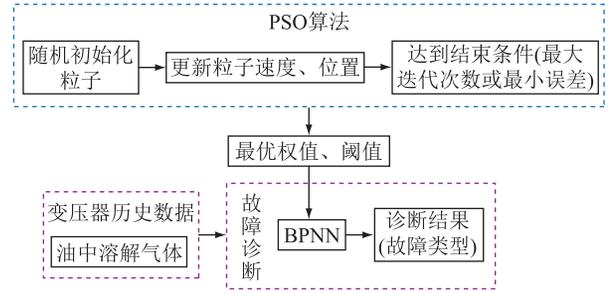


图 1 PSO-BPNN 算法总体框架

Fig.1 General framework of the PSO-BPNN algorithm

### 1.2 PSO 算法

PSO 算法中,每个粒子本身代表 1 个解,具有速度和位置 2 个属性,速度代表移动快慢,位置代表移动方向。解的优劣程度由适应度函数决定。PSO 算法流程如图 2 所示。

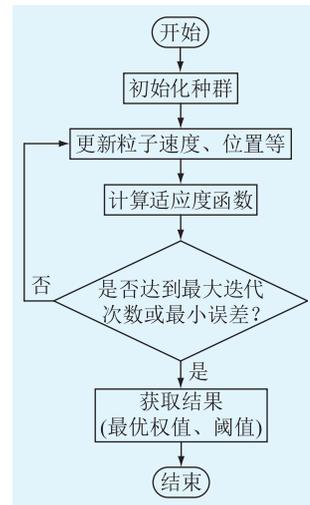


图 2 PSO 算法流程

Fig.2 PSO algorithm flow

随机初始化一群粒子,通过迭代找到最优解,在每一次的迭代中,粒子通过搜索跟踪来更新,在找到最优值后,粒子更新速度、位置为:

$$x'_o = x_o + v_o \quad (1)$$

$$v'_o = \omega v_o + c_1 \lambda (p_{o,best} - x_o) + c_2 \lambda (g_{o,best} - x_o) \quad (2)$$

式中:  $x_o$  为种群中的第  $o$  个粒子的位置;  $v_o$  为种群中的第  $o$  个粒子的速度;  $p_{o,best}$  为第  $o$  个粒子的个体最优位置;  $g_{o,best}$  为第  $o$  个粒子的全局最优位置;  $\lambda$  为  $(0,1]$  间的随机数;  $c_1$ 、 $c_2$  为学习因子,分别用来调整向  $p_{o,best}$  和  $g_{o,best}$  飞行的步长;  $\omega$  为惯性权重,描述粒子上一代速度对当代速度的影响,通过调节  $\omega$  即可调节 PSO 算法的全局与局部寻优能力。

## 2 ADASYN

为了提高变压器套管诊断模型的性能,改善数

据不平衡分布,通过 ADASYN 方法,将变压器套管故障原始数据中,较难学习的样本,合成更多的数据,可以减少原始不平衡数据分布导致的学习偏见。将 ADASYN 技术处理后的数据样本作为训练集输入变压器套管故障诊断模型,通过准确度、均方误差、所用时间等指标,对诊断模型性能进行分析评定。

ADASYN 方法是将不同少数样本学习的困难程度按权重分布,为少数较难学习的样本合成更多的训练数据。ADASYN 支持二分类和多分类,不仅可以减少原始不平衡数据分布导致的学习偏见,而且可以自适应偏移决策边界,使其聚焦于较难学习的样本。

训练样本  $D_{tr}$  包括  $m$  组样本  $\{x_i, y_i\}$ , 其中  $i = 1, 2, \dots, m$ ;  $x_i$  为  $n$  维特征空间  $X$  中的 1 组特征量;  $y_i \in Y = \{1, 2, 3, 4, 5, 6\}$  为与  $x_i$  相关的类别标签;  $m_s$ 、 $m_1$  分别为少数类样本的数量和多数类样本的数量,  $m_s \leq m_1$ ,  $m_s + m_1 = m$ ;  $x_{is}$  为少数类;  $x_{i1}$  为多数类;类不平衡度  $d = m_s/m_1, d \in (0, 1]$ 。

若  $d$  不满足小于预设的最大类别不平衡度  $d_{th}$ , 则进行下列计算。

(1) 少数类别样本需要合成的总样本数量  $G_s$ :

$$G_s = (m_1 - m_s) \beta \quad (3)$$

式中:  $\beta \in [0, 1]$ , 用于制定所需平衡水平的参数。 $\beta = 1$  时,表示创建 1 个完全平衡的数据集。

(2) 对于少数类  $x_{is}$  中每个样本  $x_i$ , 根据  $n$  维空间中的欧几里得距离找到  $K$  个邻近样本,并计算比率  $r_i$  为:

$$r_i = \frac{\Delta i}{K} \quad (4)$$

式中:  $\Delta i$  为少数类样本  $x_i$  的  $K$  个邻近样本中属于多数类  $x_{i1}$  的样本数量;  $K$  取任意正整数即可。

(3) 对比率  $r_i$  进行标准化处理得:

$$\hat{r}_i = r_i / \sum_{i=1}^{m_s} r_i \quad (5)$$

(4) 计算每个少数类样本  $x_i$  最终需要合成的样本数量  $g_i$  为:

$$g_i = \hat{r}_i G_s \quad (6)$$

(5) 从  $x_i$  的  $K$  个邻近样本中随机选择 1 个少数类别  $x_{zi}$ , 合成数据样本如下:

$$s_i = x_i + (x_{zi} - x_i) \lambda \quad (7)$$

式中:  $x_{zi} - x_i$  为  $n$  维空间中的差分分量。

(6) 重复合成步骤直到满足每个少数类样本需要合成的样本数量  $g_i$ 。

### 3 基于 PSO-BPNN 的变压器套管不平衡数据故障诊断

套管在运行中可能会出现多种故障,如密封失效/破坏、绝缘老化、末屏故障、放电性故障、过热性故障等。套管中电容、电阻、介质损耗、温度、局放、油中溶解气体以及微水含量等特征量是判断设备绝缘状态的重要指标。

#### 3.1 特征量选取

收集套管油中溶解气体发现,新的套管经长时间运行后,油纸绝缘会发生老化,油中溶解气体会出现一些  $CO$ 、 $CO_2$ 、 $H_2$  等,但若发生放电、过热等故障时对应气体含量会迅速增加。如发生局部放电会产生  $H_2$ 、 $CH_4$ ;发生低能放电会产生  $H_2$ 、 $C_2H_2$ ;发生高能放电会产生  $C_2H_2$ 、 $C_2H_4$ ;发生油中热故障会产生  $C_2H_4$ 、 $C_2H_6$ ;发生纸中热故障会产生  $CO$ 、 $CO_2$ 。若套管内部受潮,水在电场作用下发生反应,也会产生大量  $H_2$ ,仅通过油中溶解气体难以区分局部放电和受潮 2 种情况,需要综合微水含量进行判断。因此,文中不对受潮缺陷进行诊断。

变压器套管经长时间运行后会发生老化或产生故障,不同故障类型下,套管绝缘物质裂解产生的气体种类以及气体含量各不相同。可以根据这一特点,利用 GB/T 24624—2009 中的准则,对油中溶解气体数据的故障类型进行划分,利用油中溶解气体对变压器套管绝缘状态进行诊断<sup>[22]</sup>。文中将 7 种特征气体 ( $H_2$ 、 $C_2H_2$ 、 $CH_4$ 、 $C_2H_6$ 、 $C_2H_4$ 、 $CO$ 、 $CO_2$ ) 及总烃含量作为特征量来判断变压器套管故障类型。

#### 3.2 数据来源及分类

文中所采用训练集和测试集是从公开发表的文献和 IEC TC 10 数据集中收集到的故障样本<sup>[24-28]</sup>。根据 GB/T 24624—2009,将收集到的油色谱数据划分故障类型,进行网络训练。整个数据集由 155 组已明确故障类型的样本组成,随机选取 70% 样本作为训练集,30% 作为测试集。表 1 为不同故障类型样本数量的分布情况。

表 1 不同故障类型下样本分布

Table 1 Sample distribution under different fault types

故障类型	故障标签	样本数量
正常状态	D1	58
局部放电	D2	29
低能放电	D3	42
高能放电	D4	17
中低温过热	D5	4
高温过热	D6	5

根据 GB/T 24624—2009 中的准则,将收集到的变压器套管故障样本除正常状态外划分为两大类,即放电性故障和过热性故障。前者包括局部放电、低能放电和高能放电,后者为中低温过热、高温过热。不同故障类型所对应的气体含量数据如表 2 所示。

表 2 不同故障类型对应气体含量数据

故障 标签	特征气体含量/ $10^{-6}(\mu\text{L}\cdot\text{L}^{-1})$							
	H <sub>2</sub>	C <sub>2</sub> H <sub>2</sub>	CH <sub>4</sub>	C <sub>2</sub> H <sub>6</sub>	C <sub>2</sub> H <sub>4</sub>	CO	CO <sub>2</sub>	总烃
D1	5.3	0	0.6	0.6	0.2	22.1	126.3	1.44
D2	53.3	24.3	0.8	8.5	2.1	321.2	423.6	33.00
D3	66.7	247.0	22.0	4.2	40.6	239.4	628.4	313.00
D4	174.0	64.4	24.0	93.0	293.0	1 016.0	11 262.0	474.00
D5	341.0	0	416.0	69.0	5.5	647.6	1 716.4	490.00
D6	124.0	5.2	296.0	84.0	481.0	446.7	1 603.2	867.00

### 3.3 数据预处理

神经网络模型输入为 7 种特征气体(H<sub>2</sub>、C<sub>2</sub>H<sub>2</sub>、CH<sub>4</sub>、C<sub>2</sub>H<sub>6</sub>、C<sub>2</sub>H<sub>4</sub>、CO、CO<sub>2</sub>)及总烃,故障类型(即标签)包含 6 个不同的类别(正常、局部放电、低能放电、高能放电、中低温过热、高温过热)。由于不同故障所对应的特征气体含量差异较大,为减小误差,提高模型性能,对数据进行归一化处理,使其处于同一数量级,如:

$$x'_{in} = \frac{x_{in} - \min x_{in}}{\max x_{in} - \min x_{in}} \quad (8)$$

式中:  $x_{in}$  为某组样本中 1 种特征气体的含量,预处理后的数据被限定在  $[0, 1]$  范围内,可以消除奇异样本导致的不良影响,加快梯度下降求解速度。

### 3.4 故障诊断流程

为了改善 BPNN 易陷入局部最优的问题,提高网络诊断精度,文中使用 PSO-BPNN 对变压器套管进行故障诊断。PSO-BPNN 流程如图 3 所示。

故障诊断的主要步骤如下:

(1) 设定神经网络各层层数、节点数以及参数(种群数量、惯性权重、学习因子及迭代次数),确定网络拓扑,初始化粒子速度、位置。

(2) 计算适应度值,选定激活函数,将每个粒子适应度值与历史最优的位置进行比较,选取最优位置进行保留;再将适应度值与全局历史最好位置进行比较,保留全局最优位置,并记录粒子序号(根据式(1)、式(2)更新每个粒子速度、位置)。

(3) 判断是否达到最大迭代次数或最小误差的结束条件,满足条件则结束循环,输出最优权值与阈值至 BPNN,否则循环继续。

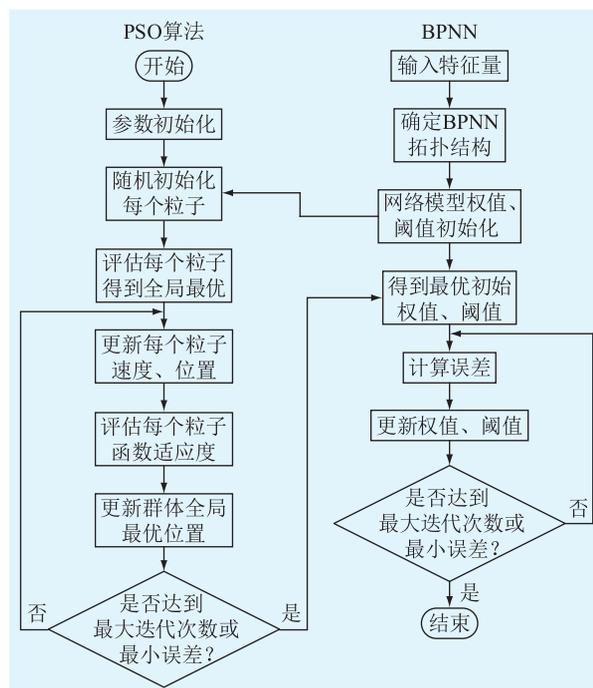


图 3 PSO-BPNN 流程

Fig.3 PSO-BPNN flow

## 4 模型算例仿真及对比分析

### 4.1 评价指标

为评估 PSO-BPNN 模型的性能,文中使用诊断准确率、误差进行衡量。准确率  $A_c$  是指诊断正确的样本占总样本个数的比例,表示为:

$$A_c = \frac{n_c}{N} \quad (9)$$

式中:  $n_c$  为正确诊断样本数;  $N$  为样本总数。

误差用均方误差  $M_{SE}$  来表示:

$$M_{SE} = (A - T)^2 / N \quad (10)$$

式中:  $A$  为实际输出;  $T$  为期望输出。

### 4.2 基于 ADASYN 的原始数据平衡化处理

由于变压器套管油中溶解气体故障样本分布极不平衡,正常状态样本约占 37.4%,而高温过热样本仅占 3.2%,这种不平衡的样本易使模型将“故障”误判为“正常”,可能会引发严重后果。为提高故障诊断的精确度,引入过采样方法对不平衡样本进行处理。

与原始样本对比,ADASYN 处理前后的局部放电故障后油中溶解气体数据如图 4 所示。其中 C<sub>2</sub>H<sub>2</sub>、C<sub>2</sub>H<sub>4</sub> 浓度较低难以辨识,因此,单独示意如图 4(a)、(b)所示,使用不同颜色的柱状区分合成前后的数据。图 4(c)为局部放电后油中溶解气体数据示意,将不同气体用不同颜色区分,同一颜色表示同类气体数据,分别使用棱锥和圆柱区分原始数据

与 ADASYN 数据平衡化后的合成数据。

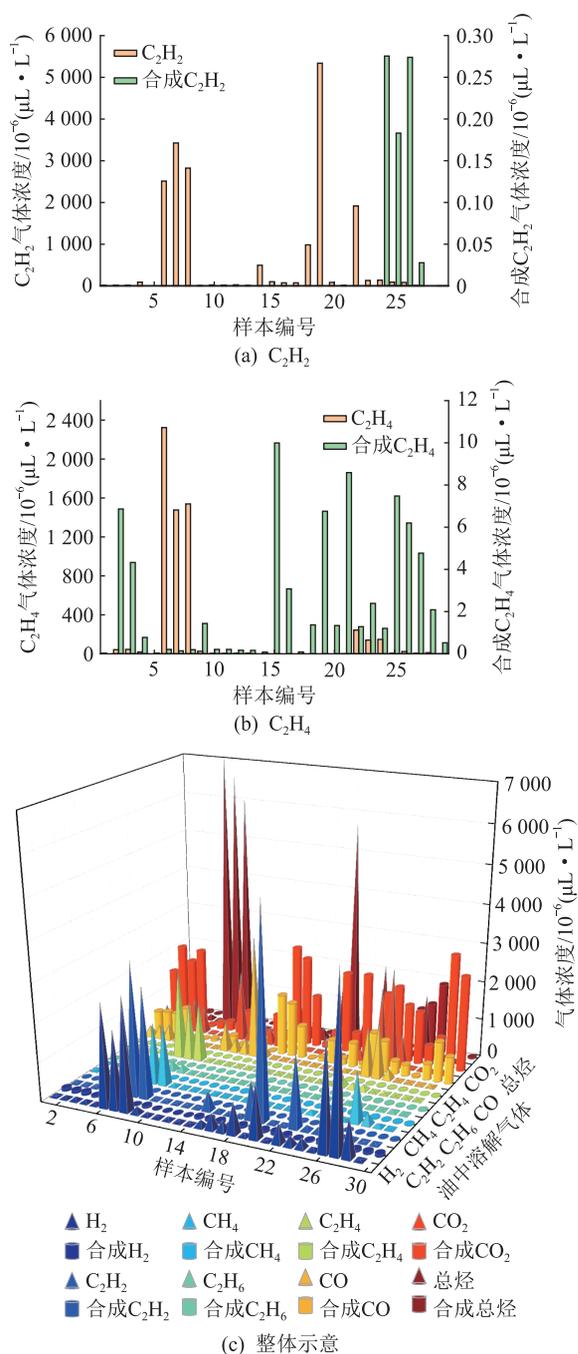


图4 ADASYN 前后局部放电数据

Fig.4 Partial discharge datas before and after ADASYN

通过 ADASYN 方法对少数类样本进行过采样处理, D2、D3、D4、D5、D6 分别增加 30、12、44、55、55 组故障样本, 样本总数由 155 组增加至 351 组, 改善原始数据分布不平衡问题。表 3 为数据处理前后少数类样本分布情况。

4.3 运行环境及仿真结果

在 1 台配置 Intel(R) Core(TM) i5-9300H 中央处理器、8 GB 内存的 Windows 操作系统的电脑上进行仿真实验, 软件环境为 MATLAB 2018a。

表 3 ADASYN 数据平衡化后少数类样本分布

Table 3 Minority class sample distribution after ADASYN data balancing

数据类型	不同故障类型下的样本分布					
	D1	D2	D3	D4	D5	D6
原始数据	58	29	42	17	4	5
ADASYN 处理后的数据	58	59	54	61	59	60

实验过程中使用 BPNN 对变压器套管进行故障诊断, 如图 5 所示, 其中 D0 为模型无法判断出故障类型的情况。训练函数采用 trainbfg 准牛顿 BPNN 算法, 此函数收敛时所需的迭代次数相对较少, 比较适合小型网络。输入、输出以及隐含层节点数分别为 9、6、8, 样本总数为 155 组, 训练集为 109 组, 测试集为 46 组。

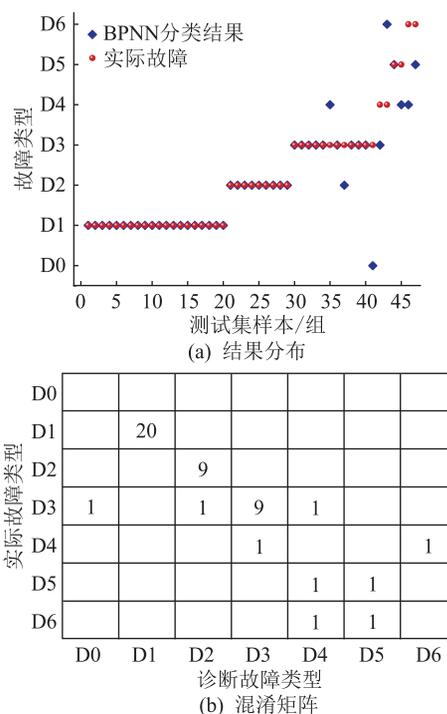


图5 BPNN 分类结果

Fig.5 BPNN classification results

BPNN 对变压器套管中油中溶解气体诊断结果如图 5(a) 所示。6 种设备状态中有 4 种存在预测错误, 其中低能放电误判次数最多, 将低能放电错误判断为局部放电和高能放电甚至无法判断。这可能是由于局部放电能量持续增大导致低能放电甚至产生高能放电, 产生的气体种类、浓度有相似之处, 即数据特征存在相似, 因此产生误判问题较多; 其次是高能放电误判为低能放电和高温过热; 高温过热误判为中低温过热及高能放电; 中低温过热误判为高能放电。图 5(b) 为测试集的混淆矩阵。

基于 BPNN 模型的变压器套管故障诊断中, 输

入原始不平衡小样本数据,模型能学到的特征较少,易出现欠拟合等问题,从而导致诊断结果并不理想。因此,对 BPNN 进行优化是非常有必要的。

为表明 ADASYN 数据平衡化的有效性,基于 PSO-BPNN 模型,使用原始数据对变压器套管进行诊断,测试集诊断结果如图 6 所示。从图 6 测试集诊断结果中可以看出,经过 PSO 优化的 BPNN 诊断准确度有所提升,但仍有误判,相较 BPNN 诊断结果,错误率较低,参数经过调整后的模型准确度显著提高,参数设置为:种群大小为 10;迭代次数为 50;学习因子  $c_1$ 、 $c_2$  为 1.5;惯性权重  $\omega$  为 0.8 时,模型性能最佳,测试集准确率为 97.7%,用时 54.657 s,表明该优化算法有较高可靠性。

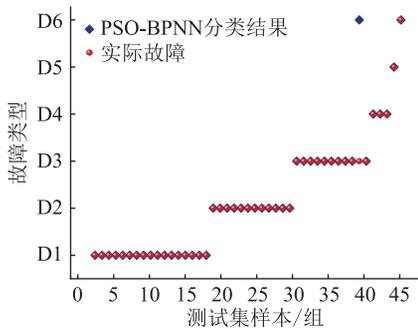


图 6 PSO-BPNN 分类结果

Fig.6 PSO-BPNN classification results

使用 PSO-BPNN 对变压器套管故障进行诊断,能快速达到收敛。适应度曲线见图 7,适应度值是以均方差作为适应度曲线的目标函数进行计算。

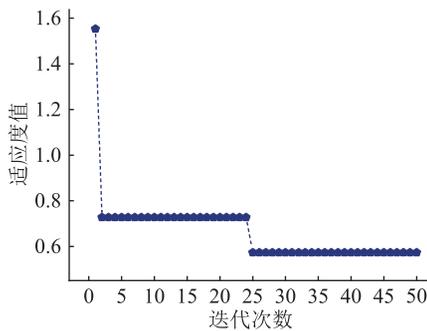


图 7 PSO-BPNN 适应度曲线

Fig.7 PSO-BPNN adaptation curve

文中使用 ADASYN 方法平衡后的故障数据作为 PSO-BPNN 诊断模型的训练集,测试集仍使用原始数据集。对变压器套管进行故障诊断。测试集诊断结果如图 8 所示。

基于 ADASYN 平衡数据后的 PSO-BPNN 模型中粒子群初始参数设定不变,测试集中变压器套管诊断的准确率为 99.07%。从图 8 分类结果中可以看出,仅存在 1 处错误位于高能放电,对比 BPNN 诊

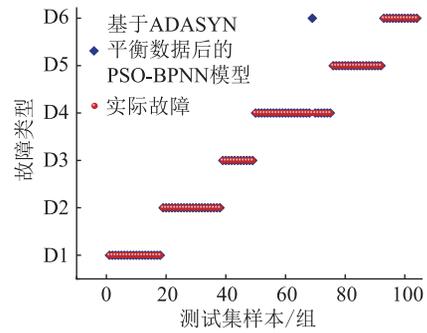


图 8 基于 ADASYN 平衡数据后的 PSO-BPNN 模型分类结果

Fig.8 PSO-BPNN model based on ADASYN balanced data classification results

断结果来看,错误显著减少,表明文中所提方法能有效提高故障诊断模型性能。

#### 4.4 算法对比

为了进一步验证 PSO-BPNN 提高变压器套管故障诊断结果的有效性,从原始数据(155 组)中随机挑选 109 组样本作为训练集,46 组作为测试集,分别将 BPNN、PSO-BPNN、遗传结合反向传播神经网络(genetic combined with back propagation neural network, G-BPNN)算法及布谷鸟搜索结合反向传播神经网络(cuckoo search combined with back propagation neural network, CS-BPNN)算法模型的诊断结果进行对比。

图 9、图 10 分别为 G-BPNN、CS-BPNN 模型测试集中实际值与期望值的对比示意。其中, G-BPNN 模型的参数设定为:迭代次数为 50;种群规模为 10;选择函数为 0.09;交叉函数为 2。

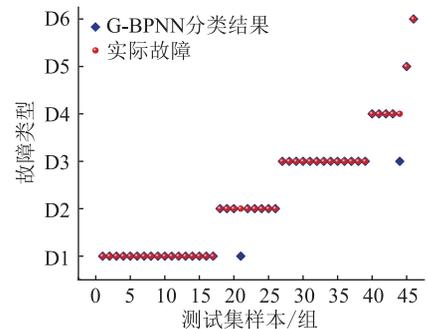


图 9 G-BPNN 分类结果

Fig.9 G-BPNN classification results

G-BPNN 模型对变压器套管故障状态,诊断精度有所提升,但仍存在误判,主要分布在局部放电及高能放电,准确度略低于 PSO-BPNN 模型,所用时长相差不大。

CS-BPNN 模型的参数设定为:迭代次数为 50;概率为 0.25;鸟巢数为 10。CS-BPNN 模型的误判主

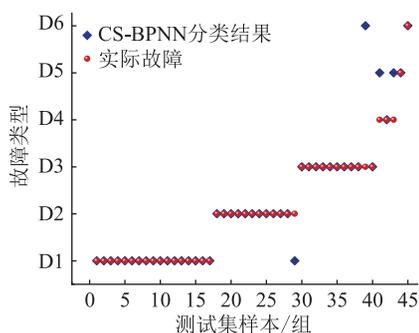


图 10 CS-BPNN 分类结果

Fig.10 CS-BPNN classification results

要位于高能放电,误判次数略高于 G-BPNN 及 PSO-BPNN 模型。

不同模型下准确率及所用时间对比如表 4 所示,从结果可以看出,CS-BPNN 算法诊断准确率为 91.2%;BPNN 诊断的准确率最低,仅为 82.9%;而 PSO-BPNN 以及 G-BPNN 算法诊断准确率分别达到 97.7%和 95.6%。

表 4 不同模型评价指标结果对比  
Table 4 Evaluation indicator results comparison of the different models

故障诊断模型	准确率/%	用时/s
BPNN	82.9	2.670
PSO-BPNN	97.7	54.657
G-BPNN	95.6	55.240
CS-BPNN	91.2	29.930

将 BPNN、G-BPNN、PSO-BPNN、CS-BPNN 4 种模型进行对比可以看出,BPNN 准确度最低,但所用时间最短,仅需 2.670 s;PSO-BPNN 模型准确度最高,表明该模型实际值与期望值最为接近,最可靠,但用时大约需要 55 s;CS-BPNN 模型准确度比单一 BPNN 高,与 PSO-BPNN 模型相比准确度较低,用时较短;G-BPNN 所用时间与 PSO-BPNN 模型相近,但准确度相较 PSO-BPNN 低约 2.1%。

## 5 结论

(1) BPNN 对变压器套管故障进行诊断的性能较差,使用 PSO 对 BPNN 中的权重及阈值进行优化,能有效提高变压器套管故障诊断的准确度。

(2) 将基于 ADASYN 平衡数据后的 PSO-BPNN 模型与 BPNN、G-BPNN、CS-BPNN、PSO-BPNN 诊断结果进行对比,结果表明:BPNN 诊断模型准确度最低,为 82.9%;其次是 CS-BPNN 准确度为 91.2%;G-BPNN 准确度为 95.6%;PSO-BPNN 准确度为 97.7%;文中模型准确度最高约为 99%。

(3) 使用 ADASYN 方法对训练样本进行预处理,

可以有效提高模型性能,综合考虑评价指标,变压器套管油中溶解气体小样本集适用的最佳模型为基于 ADASYN 平衡数据后的 PSO-BPNN 模型,与其他模型相比更为可靠,准确度高,性能最佳。

### 参考文献:

[1] 于群英,赵天成,何秋月,等. 油浸电容式套管典型故障分析及仿真研究[J]. 变压器,2020,57(9):80-84,54.  
YU Qunying,ZHAO Tiancheng,HE Qiuyue, et al. Analysis and simulation of typical faults of oil-immersed capacitor bushing [J]. Transformer,2020,57(9):80-84,54.

[2] 国江,倪学锋,林浩,等. 油纸电容式套管故障统计分析及模拟测试研究[J]. 电测与仪表,2014,51(24):113-119.  
GUO Jiang,NI Xuefeng,LIN Hao, et al. Research on the statistical analysis and simulation tests of the oil-impregnated paper capacitance graded bushing faults [J]. Electrical Measurement & Instrumentation,2014,51(24):113-119.

[3] 李继攀,刘宏领,郭奇军,等. 基于功率因数角的接地变压器匝间短路故障辨识[J]. 供用电,2023,40(9):50-57.  
LI Jipan,LIU Hongling,GUO Qijun, et al. Identification of inter turn short circuit fault of grounding transformer based on power factor angle [J]. Distribution & Utilization,2023,40(9):50-57.

[4] 刘云鹏,董王英,许自强,等. 基于卷积神经网络的变压器套管故障红外图像识别方法[J]. 高压电器,2021,57(10):134-140.  
LIU Yunpeng,DONG Wangying,XU Ziqiang, et al. Infrared image recognition method on fault of transformer bushing based on convolutional neural networks [J]. High Voltage Apparatus,2021,57(10):134-140.

[5] 陈晓琳,符小桃,吴乾东,等. 基于频域介电谱的油纸绝缘套管局部受潮诊断研究[J]. 电力工程技术,2022,41(2):149-155.  
CHEN Xiaolin,FU Xiaotao,WU Qiantong, et al. Partial damp diagnostic method of oil-paper insulating bushing based on frequency domain spectroscopy [J]. Electric Power Engineering Technology,2022,41(2):149-155.

[6] 陈钜栋,江军,杨小平,等. 油浸式高压套管局部放电非接触式特高频检测[J]. 电力工程技术,2021,40(1):147-154.  
CHEN Judong,JIANG Jun,YANG Xiaoping, et al. Non-contact partial discharge detection of high voltage oil-impregnated-paper bushing based on UHF technology [J]. Electric Power Engineering Technology,2021,40(1):147-154.

[7] CHEN J Q. Fault prediction of a transformer bushing based on entropy weight TOPSIS and gray theory [J]. Computing in Science & Engineering,2019,21(6):55-62.

[8] 赵洪山,王奎,王震,等. 基于 VMD-MGRU 的箱式变压器高压套管温度预测[J]. 电机与控制学报,2021,25(8):18-28.  
ZHAO Hongshan,WANG Kui,WANG Zhen, et al. Temperature prediction of box-type transformer high-voltage bushing based on VMD-MGRU [J]. Electric Machines and Control,2021,25(8):18-28.

[9] 王晶,许素安,洪凯星,等. 基于 DGA 特征量优选与 GA-SVM

- 的变压器故障诊断模型[J]. 变压器, 2020, 57(12): 36-40, 46.
- WANG Jing, XU Suan, HONG Kaixing, et al. Transformer based on DGA feature quantity optimization and GA-SVM fault diagnosis model[J]. Transformer, 2020, 57(12): 36-40, 46.
- [10] 张育杰, 冯健, 李典阳, 等. 基于油色谱数据的变压器故障征兆新优选策略[J]. 电网技术, 2021, 45(8): 3324-3332.
- ZHANG Yujie, FENG Jian, LI Dianyong, et al. New feature selection method for transformer fault diagnosis based on DGA data[J]. Power System Technology, 2021, 45(8): 3324-3332.
- [11] 王旭, 生西奎, 慕锐, 等. 基于多频超声波技术和 GWO-BP 算法的变压器油电气性能检测方法研究[J]. 高压电器, 2020, 56(8): 114-120.
- WANG Xu, SHENG Xikui, MU Kun, et al. Electrical performance testing method of transformer oil based on multi-frequency ultrasonic technology and grey wolf-optimized neural network[J]. High Voltage Apparatus, 2020, 56(8): 114-120.
- [12] 王文森, 杨晓西, 刘阳, 等. 基于层次聚类分析的变压器油中溶解气体在线监测数据异常检测[J]. 高压电器, 2023, 59(1): 142-147.
- WANG Wensen, YANG Xiaoxi, LIU Yang, et al. Anomaly detection of online monitoring data of dissolved gases in transformer oil based on hierarchical cluster analysis[J]. High Voltage Apparatus, 2023, 59(1): 142-147.
- [13] 康佳宇, 张沈习, 张庆平, 等. 基于 ANOVA 和 BO-SVM 的变压器故障诊断方法[J]. 高电压技术, 2023, 49(5): 1882-1891.
- KANG Jiayu, ZHANG Shenxi, ZHANG Qingping, et al. Fault diagnosis method of transformer based on ANOVA and BO-SVM[J]. High Voltage Engineering, 2023, 49(5): 1882-1891.
- [14] 代杰杰, 宋辉, 盛戈啤, 等. 采用 LSTM 网络的电力变压器运行状态预测方法研究[J]. 高电压技术, 2018, 44(4): 1099-1106.
- DAI Jiejie, SONG Hui, SHENG Gehao, et al. Prediction method for power transformer running state based on LSTM network[J]. High Voltage Engineering, 2018, 44(4): 1099-1106.
- [15] DE ANDRADE LOPES S M, FLAUZINO R A, ALTAFIM R A C. Incipient fault diagnosis in power transformers by data-driven models with over-sampled dataset[J]. Electric Power Systems Research, 2021, 201: 107519.
- [16] YU S, TAN W M, ZHANG C M, et al. Research on hybrid feature selection method of power transformer based on fuzzy information entropy[J]. Advanced Engineering Informatics, 2021, 50: 101433.
- [17] XU C, LI X L, WANG Z H, et al. Improved BLS based transformer fault diagnosis considering imbalanced samples[J]. Energy Reports, 2022, 8: 1446-1453.
- [18] 王艳, 李伟, 赵洪山, 等. 基于融合少数类过采样均衡多分类数据的改进极限学习机的变压器故障诊断方法[J]. 电网技术, 2023, 47(9): 3799-3807.
- WANG Yan, LI Wei, ZHAO Hongshan, et al. Transformer fault diagnosis fused with synthetic minority over-sampling balanced multi-classification data based on improved extreme learning machine[J]. Power System Technology, 2023, 47(9): 3799-3807.
- [19] 余长厅, 黎大健, 汲胜昌, 等. 基于振动噪声及 BP 神经网络的变压器故障诊断方法研究[J]. 高压电器, 2020, 56(6): 256-261.
- YU Zhangting, LI Dajian, JI Shengchang, et al. Research on transformer fault diagnosis method based on vibration noise and BP neural network[J]. High Voltage Apparatus, 2020, 56(6): 256-261.
- [20] 周文潮, 柴淑颖. 基于 Python 技术 GA-BP 神经网络的变压器故障诊断研究[J]. 变压器, 2022, 59(6): 61-64.
- ZHOU Wenchao, CHAI Shuying. Research on transformer fault diagnosis based on Python technology GA-BP neural network[J]. Transformer, 2022, 59(6): 61-64.
- [21] 郭秀才, 刘冰冰, 王力立. 基于小波包和 CS-BP 神经网络的矿用电力电缆故障诊断[J]. 计算机应用与软件, 2021, 38(9): 105-110.
- GUO Xiucan, LIU Bingbing, WANG Lili. Fault diagnosis of mining power cable based on wavelet packet and CS-BP neural network[J]. Computer Applications and Software, 2021, 38(9): 105-110.
- [22] 王享, 黄新波, 朱永灿. PSO-IGWO 优化混合 KELM 的变压器故障诊断方法[J]. 西安工程大学学报, 2019, 32(2): 154-160.
- WANG Xiang, HUANG Xinbo, ZHU Yongcan. The transformer fault diagnosis method based on PSO-IGWO optimized hybrid KELM[J]. Journal of Xi'an Polytechnic University, 2019, 32(2): 154-160.
- [23] 全国绝缘子标准化技术委员会. 绝缘套管油为主绝缘(通常为纸)浸渍介质套管中溶解气体分析(DGA)的判断导则: GB/T 24624—2009[S]. 北京: 中国标准出版社, 2010: 1.
- National Insulator Standardization Technical Committee. Insulated bushings guide for the interpretation of dissolved gas analysis (DGA) in bushings where oil is the impregnating medium of the main insulation (generally paper): GB/T 24624-2009 [S]. Beijing: Standards press of China, 2010: 1.
- [24] 洪志湖, 颜冰, 钱国超, 等. 一起 110 kV 油-油-油纸绝缘变压器套管故障原因分析[J]. 变压器, 2021, 58(12): 72-77.
- HONG Zhihu, YAN Bing, QIAN Guochao, et al. Cause analysis of a 110 kV oil-oil oil paper insulation transformer bushing fault[J]. Transformer, 2021, 58(12): 72-77.
- [25] 张晨晨, 丁国成, 陈庆涛, 等. 基于频域介电谱的一起 110 kV 变压器套管故障分析[J]. 变压器, 2019, 56(5): 77-80.
- ZHANG Chenchen, DING Guocheng, CHEN Qingtao, et al. Fault analysis of 110 kV transformer bushing based on FDS[J]. Transformer, 2019, 56(5): 77-80.
- [26] 李丰, 蚁天军, 何建明, 等. 一起 220 kV 变压器套管渗油故障原因分析及处理[J]. 电瓷避雷器, 2021(2): 90-94.
- LI Feng, YI Tianjun, HE Jianming, et al. Analysis and treatment of an oil-leakage fault of 220 kV transformer bushing[J]. Insulators and Surge Arresters, 2021(2): 90-94.

[27] DUVAL M, DEPABLA A. Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases [J]. IEEE Electrical Insulation Magazine, 2001, 17(2): 31-41.

[28] 廖伟涵, 郭创新, 金宇, 等. 基于四阶段预处理与 GBDT 的油浸式变压器故障诊断方法 [J]. 电网技术, 2019, 43(6): 2195-2203.

LIAO Weihao, GUO Chuangxin, JIN Yu, et al. Oil-immersed transformer fault diagnosis method based on four-stage preprocessing and GBDT [J]. Power System Technology, 2019, 43

(6): 2195-2203.

作者简介:



杨昊

杨昊(1988),男,博士,副教授,研究方向为输变电设备状态监测与故障诊断(E-mail: yanghao@xpu.edu.cn);

胡文秀(1997),女,硕士在读,研究方向为高电压与绝缘技术;

张璐(1987),男,博士,高级工程师,从事电力变压器运行状态监测工作。

### Fault diagnosis of transformer oil-paper bushings in PSO-BPNN algorithm based on ADASYN data balancing

YANG Hao<sup>1</sup>, HU Wenxiu<sup>1</sup>, ZHANG Lu<sup>2</sup>, CHEN Jinpeng<sup>1</sup>, ZHOU Sijia<sup>1</sup>, ZHAO Sirui<sup>1</sup>

(1. School of Electronic and Information, Xi'an Polytechnic University, Xi'an 710600, China;

2. State Grid Shaanxi Electric Power Co., Ltd. Electric Power Research Institute, Xi'an 710100, China)

**Abstract:** The insulation performance of transformer bushings is a crucial aspect that directly affects the safe operation of equipment. To diagnose the insulation status of transformer bushings and mitigate the impact of small-sample imbalanced data on diagnostic results, a particle swarm optimization combined with back propagation neural network (PSO-BPNN) and adaptive synthetic sampling (ADASYN) method are employed to fault diagnosis of transformer bushing. Initially, historical fault data of transformer bushings are gathered, and a sample set of dissolved gases in transformer oil with distinct fault categories is established. The ADASYN algorithm is used to synthesize the minority class samples in the original data, which allowed for obtaining balanced fault data. The balanced dissolved gases in oil served as the model input, and the fault status is used as the label output to diagnose the transformer bushings using the PSO-BPNN model. To diagnose the bushings under the original sample set, the back propagation neural network (BPNN), genetic combined with back propagation neural network (G-BPNN), cuckoo search combined with back propagation neural network (CS-BPNN), and PSO-BPNN models are used. The results reveal that the PSO-BPNN model based on ADASYN balanced data exhibited the highest accuracy among the various models for fault diagnosing the insulation status of transformer bushings. This approach effectively mitigate the impact of small sample imbalanced data on diagnostic results, and provide an effective method for assessing the insulation performance of transformer bushings.

**Keywords:** transformer bushing; fault diagnosis; dissolved gas in oil; back propagation neural network (BPNN); unbalanced data; adaptive synthetic sampling (ADASYN)

(编辑 吴昊)