

DOI: 10.13382/j.jemi.B2307027

基于 KPCA-CGSSA-KELM 的变压器故障识别方法*

江 兵 李 响 巢一帆 余子煜 陶 锴
(南京邮电大学自动化学院 南京 210046)

摘 要:针对冗余特征对变压器故障识别影响和传统方法识别准确率低的问题,提出一种基于核主成分分析(kernel principal component analysis, KPCA)与混沌麻雀搜索算法(chaos gauss sparrow search algorithm, CGSSA)优化核极限学习机(kernelized extreme learning machine, KELM)的变压器故障识别方法。首先,通过 KPCA 对变压器故障数据进行预处理,降低特征间相关性。其次,通过引入改进 Tent 映射和高斯变异策略优化麻雀搜索算法提高其搜索精度和收敛速度,并将 CGSSA 与麻雀搜索算法(SSA)、灰狼优化算法(GWO)及鲸鱼优化算法(WOA)效果进行对比。最后,利用经 KPCA 处理后的特征数据作为模型输入,并通过 CGSSA 准确选择 KELM 的核函数参数和正则化系数,建立 KPCA-CGSSA-KELM 变压器故障识别模型。实验结果表明,在相同输入数据的情况下,CGSSA 在收敛速度和寻优精度方面均有提升,且所提方法识别准确率为 95.7%,较 WOA-KELM、GWO-KELM、SSA-KELM 分别提高 18.6%、10%、15.7%。结果表明所提方法能有效处理冗余特征,提高故障识别准确率,证明了使用所提方法在在冗余特征影响的情况下进行变压器故障识别的有效性 & 可行性。

关键词:变压器故障识别;核主成分分析;混沌麻雀搜索算法;核极限学习机

中图分类号: TM407;TN06 **文献标识码:** A **国家标准学科分类代码:** 470.40

Transformer fault recognition method based on KPCA-CGSSA-KELM

Jiang Bing Li Xiang Chao Yifan Yu Ziyu Tao Kai
(College of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210046, China)

Abstract: To address the problems posed by redundant features in transformer fault recognition and the low accuracy of traditional methods, a transformer fault recognition method leveraging kernel principal component analysis (KPCA) in conjunction with chaotic sparrow search algorithm (CGSSA) is introduced. Initially, KPCA is employed to preprocess the transformer fault data, aiming to mitigate the correlations among features. Subsequently, CGSSA is improved by incorporating the improved Tent map and Gaussian mutation to increase the search accuracy and convergence speed of the algorithm. Comparing the results involving CGSSA, SSA, GWO and WOA. Utilizing the data extracted through the KPCA as the model input, CGSSA is then used to select the kernel function parameters and regularization coefficient of KELM, thereby establishing the KPCA-CGSSA-KELM transformer fault recognition model. The experimental results demonstrate that, with the identical input data, CGSSA has the best results in terms of convergence speed and optimization accuracy. In addition, the proposed method shows the fault recognition accuracy of 95.7%, which is 18.6%, 10%, and 15.7% higher than WOA-KELM, GWO-KELM, and SSA-KELM, respectively. These findings suggest that the proposed method effectively manages the impact of redundant features and enhances the precision of transformer fault recognition, thus verifying the validity and feasibility of the proposed method for transformer fault recognition under the feature redundancy.

Keywords: transformer fault recognition; KPCA; CGSSA; KELM

0 引言

变压器在电力系统输、变、配电工程中起至关重要的作用^[1]。因此,高效准确地识别变压器故障类型对于保障电力系统安全稳定运行具有重要意义^[2]。油中溶解气体分析(dissolved gas analysis, DGA)技术^[3]被广泛应用到变压器故障识别领域。油中溶解气体体积分数比值会随变压器状态而改变。通过分析气体类型和含量^[4],可以及时发现故障,但传统 DGA 方法操作复杂,故障识别率较低,所以常采用无编码比值法解决问题。

随着近年来人工智能技术的发展,DGA 与人工智能结合的变压器故障识别方法得以应用,例如支持向量机(support vector machines, SVM)^[5]、神经网络(neural network, NN)^[6]、极限学习机(extreme learning machine, ELM)^[7]等。欧阳鑫等^[8]针对传统方法识别率低的问题,通过特征优选的方法对灰狼算法(grey wolf optimizer, GWO)进行优化,并通过该方法对惩罚因子进行选择以提高分类准确率;李黄曼等^[9]用反向学习策略对麻雀搜索算法(sparrow search algorithm, SSA)进行优化,并通过优化核函数参数提高识别准确率。通过智能优化算法对支持向量机的参数进行优化并建立故障识别模型,但 SVM 本质上为二分类器,并且该模型对于参数选择要求较高,当面对变压器故障这一多分类任务时效果不佳。陈里里等^[10]提出一种基于堆栈稀疏自编码器与 Softmax 函数结合的深度神经网络模型,结果相较于 SVM 等方法在精度上有较大提升,但模型存在训练所需数据量大、学习成本高且易出现过拟合现象等问题,而变压器故障识别数据量较少,使用神经网络对变压器故障进行识别的准确率可能不高。Zhang 等^[11]提出一种基于改进完备集合经验模态分解(improved complete ensemble empirical mode decomposition, ICEEMD)与 ELM 结合的故障识别方法。ELM 解决了训练速度慢、参数选择难等问题,具有学习速度快、泛化能力强等优点,但 ELM 中参数的随机性会增加变压器故障识别模型的输出不确定性,从而影响变压器故障的最终识别结果。核极限学习机(kernel extreme learning machine, KELM)将核函数应用在 ELM 模型中,在克服 ELM 缺点的同时保留学习速度快、稳定性强等优点,有效降低了分类器参数产生过程中的随机性,从而提高了变压器故障识别模型的分类准确率与泛化能力。

针对上述问题,为了消除冗余特征的同时提高故障识别准确率,本文提出一种基于 KPCA-CGSSA-KELM 的变压器故障识别方法。通过核主成分分析提取主要特征,同时利用混沌高斯麻雀搜索算法优化 KELM 的核函数参数和正则化系数,最后利用该模型对变压器故障进

行识别。结果表明本文所提方法的故障识别准确率为 95.7%,能够有效识别变压器故障。

1 算法原理

1.1 核主成分分析

直接对变压器故障高维原始数据进行线性运算会影响模型对特征识别的准确性。因此,为避免变压器故障数据失真,本文采用核主成分分析(kernel principal component analysis, KPCA)进行特征降维。KPCA 是通过非线性核函数将输入空间数据 x_i 变换到高维空间 ϕ , 即 $x_i \rightarrow \phi(x_i)$; 然后在高维空间 ϕ 进行主成分分析来获取特征信息与主要特征,避免特征相似导致识别准确率下降^[12]。步骤如下:

1) 将变压器原始故障数据映射到高维特征空间 ϕ 形成高维故障数据 $\phi(x_i)$, $i = 1, 2, \dots, n$ 。

$$2) \sum_{i=1}^n \phi(x_i) = 0, \text{高维故障数据协方差矩阵 } \mathbf{P} \text{ 为:}$$

$$\mathbf{P} = \frac{1}{n} \sum_{i=1}^n \phi(x_i) \mathbf{T} \phi(x_i) \quad (1)$$

3) 设核函数 $M^* = \phi^T \phi$, 故障特征信息为求解 M^* 特征向量与特征值。对 $\mathbf{P}$ 进行 PCA 求解,有:

$$M^* \boldsymbol{\eta} = \lambda \boldsymbol{\eta} \quad (2)$$

式中: λ 为故障特征值; $\boldsymbol{\eta}$ 为特征值对应故障特征向量。

4) 设变压器各故障特征的累计贡献率为 90%。取前 s 个满足贡献率标准的故障特征值 λ_j ($j = 1, 2, \dots, s$) 及对应特征向量 $\boldsymbol{\eta}_j$ ($j = 1, 2, \dots, s$), 并降序排列:

$$\frac{\sum_{j=1}^s \lambda_j}{\sum_{i=1}^s \lambda_i} \geq 90\% \quad (3)$$

5) 变压器故障特征累计贡献率不小于设定值时,经降维映射后非线性样本 G 为所求特征主成分集合:

$$G = \left[\sum_{i=1}^n \boldsymbol{\eta}_i \phi(x_i) \right]^T = \boldsymbol{\eta}^T \left[\phi(x_i, x) \right]^T \quad (4)$$

1.2 混沌麻雀搜索算法(CGSSA)

1) 麻雀搜索算法

通过 SSA 实现变压器故障识别模型参数优化。SSA 是基于自然界麻雀觅食行为的智能优化算法^[13]。该算法将种群分为探索者、跟随者和预警者。麻雀个体位置是初始解,食物分布是函数在三维空间的值,而个体觅食行为是函数在三维空间内搜索全局最优解的过程。

假设在一个 j 维的搜索空间中有 N 只麻雀个体,则第 i 只麻雀在 j 维搜索空间中的位置为 $X_i = [X_{i,1}, \dots, X_{i,j}]$, $i = 1, 2, \dots, N$ 。

探索者引导种群进行觅食。位置更新如下:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(\frac{-i}{\alpha \cdot iter_{\max}}\right), & R_2 < ST \\ X_{i,j}^t + Q \cdot L, & \text{其他} \end{cases} \quad (5)$$

式中: t 为当前迭代次数; $iter_{\max}$ 为最大迭代次数; $X_{i,j}^t$ 表示第 i 只麻雀在第 j 维位置信息; α 是 $(0,1)$ 随机数; L 是 $1 \times j$, 元素均为 1 的矩阵; Q 是服从正态分布的随机数; $ST \in [0.5, 1]$, $R_2 \in [0, 1]$ 分别为安全值、预警值。

跟随者监督探索者并争夺食物。位置更新如下:

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{t^2}\right), & i > \frac{N}{2} \\ X_{\text{best}}^t + |X_{i,j}^t - X_{\text{best}}^t| \cdot A^+ \cdot L, & \text{其他} \end{cases} \quad (6)$$

式中: N 为种群规模; X_{best}^t , X_{worst}^t 分别表示当前探索者搜索的最优和最劣全局位置; A 是元素只包括 1 或 -1 的 $1 \times d$ 矩阵, 且 $A^+ = A^T(AA^T)$ 。

警戒者警戒并提醒种群反捕食。位置更新如下:

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta \cdot |X_{i,j}^t - X_{\text{best}}^t|, & f_i \neq f_g \\ X_{i,j}^t + K \cdot \left(\frac{|X_{i,j}^t - X_{\text{worst}}^t|}{(f_i - f_w) + \varepsilon}\right), & \text{其他} \end{cases} \quad (7)$$

式中: β 为服从正态分布的步长控制参数; K 是 $[-1, 1]$ 的随机数; f_i , f_g 和 f_w 分别表示第 i 个个体适应度、当前种群最佳适应度值和最劣适应度值; ε 为无限小正数。

2) 改进策略

SSA 具有实现简单、优化性能好等优点, 但在变压器故障识别模型参数优化的具体问题时, 易陷入局部最优导致收敛精度不高。为提高精度与全局寻优能力, 引入改进 Tent 混沌映射和高斯变异策略改进 SSA^[14]。

(1) Tent 混沌序列

种群初始化目的是生成一组符合要求的随机解, 而 Tent 混沌映射因为混沌变量的随机性、遍历性和规律性, 常被用于种群参数初始化。

Tent 映射方程为:

$$X_{n+1} = \begin{cases} 2X_n, & 0 \leq X_n < 0.5 \\ 2(1 - X_n), & 0.5 \leq X_n < 1 \end{cases} \quad (8)$$

式中: n 为映射次数, X_n 为第 n 次映射函数值。

张娜等^[15]发现 Tent 混沌序列易落入小周期与不稳定周期点。为了在这种情况下不破坏混沌变量 3 大特性, 故对 Tent 混沌映射策略提出改进。在常规 Tent 映射中引入随机变量 $rand(0, 1) \times \frac{1}{N_T}$ 。

改进 Tent 混沌映射表达式为:

$$X_{n+1} = \begin{cases} 2X_n + rand(0, 1) \times \frac{1}{N_T}, & 0 \leq X_n < 0.5 \\ 2(1 - X_n) + rand(0, 1) \times \frac{1}{N_T}, & 0.5 \leq X_n < 1 \end{cases} \quad (9)$$

式中: N_T 为序列粒子个数; $rand(0, 1)$ 为 $[0, 1]$ 内随机数。

将改进 Tent 混沌映射产生的解应用在麻雀个体位置以实现种群初始化。具体公式为:

$$X = X_{lb} + (X_{ub} - X_{lb})X_{n+1} \quad (10)$$

式中: X_{ub} 和 X_{lb} 为搜索空间上下限, X 为经改进 Tent 混沌映射后的麻雀个体位置。

(2) 混沌扰动

对式(9)进行贝努利变换得到混沌变量 Z_d :

$$Z_d = X_{n+1} = (2X_n) \bmod 1 + rand(0, 1) \cdot \frac{1}{N_T} \quad (11)$$

将混沌变量 Z_d 载波到解空间求解混沌扰动量 X_{new}^d :

$$X_{\text{new}}^d = d_{\min} + (d_{\max} - d_{\min})Z_d \quad (12)$$

式中: $d_{\min}$, $d_{\max}$ 为第 d 维变量 X_{new}^d 的最小值与最大值。

通过式(13)对麻雀个体进行混沌扰动, 即:

$$X'_{\text{new}} = (X' + X_{\text{new}})/2 \quad (13)$$

式中: X' 为需混沌扰动的个体位置; X_{new} 为混沌扰动量; X'_{new} 为经过混沌扰动后个体的新的位置信息。

(3) 高斯变异

指在变异操作时用一个符合正态分布的随机数代替原参数。变异过程具体公式为:

$$G(x) = x(1 + N(0, 1)) \quad (14)$$

式中: x 为原参数值; $N(0, 1)$ 表示满足标准正态分布的随机数; $G(x)$ 为变异后的函数值。

由正态分布特性可知, 高斯变异重点搜索变异前原个体附近局部区域。通过变异策略高效、精确搜寻全局极值点, 解决局部极小值问题, 同时能够有效防止 CGSSA 陷入局部最优导致收敛精度低^[16]。

3) 算法性能测试

为验证 CGSSA 寻优能力, 与 GWO^[16]、鲸鱼优化算法 (whale optimization algorithm, WOA)^[17]、SSA^[9] 通过单峰测试函数 F1 和多峰测试函数 F2 进行结果比较。测试函数如表 1 所示。

表 1 测试函数

Table 1 Test function

序号	函数名称	最优值
F_1	$\sum_{i=1}^n x_i^2$	0
F_2	$\sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	0

使用 MATLAB2020a 进行实验。最大迭代次数设置为 100, 种群规模设置为 30。通过绘制不同算法对于测试函数的寻优曲线, 可以直观看出 CGSSA 的收敛速度和寻优能力。

将 CGSSA 和 WOA、GWO、SSA 的寻优结果进行比

较,如图 1、2 所示。

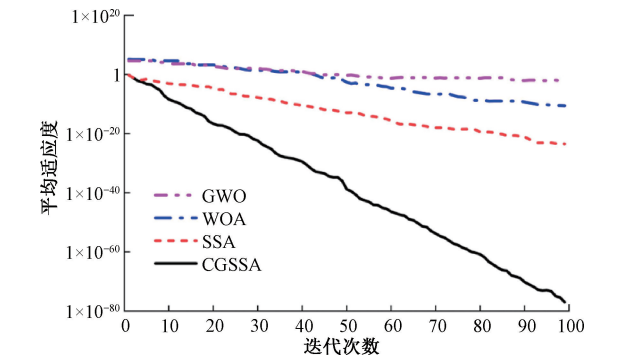


图 1 F_1 寻优曲线

Fig. 1 The optimization curve of F_1

图 1 实验结果表明,在单峰高维测试函数 F_1 寻优过程中,GWO、WOA、SSA 表现各有优劣,但 CGSSA 以最少迭代次数接近最优解,且迭代搜索性能和寻优性能也优于其他 3 种算法,表明了 CGSSA 在保证迭代速度的同时也充分保证了算法的搜索寻优能力。

图 2 实验结果表明,在多峰高维测试函数 F_2 寻优过

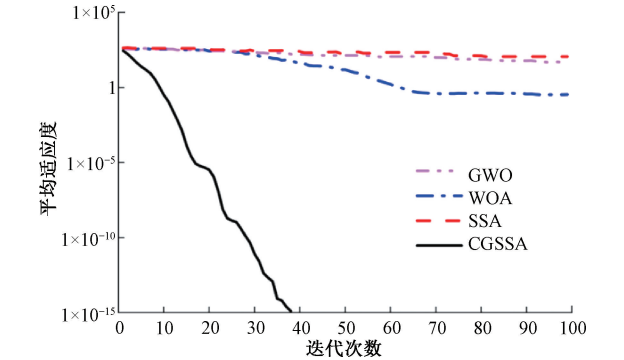


图 2 F_2 寻优曲线

Fig. 2 The optimization curve of F_2

程中,CGSSA 收敛速度与寻优精度明显优于其他 3 种算法,能够以最快收敛速度达到对应测试函数的最优适应度值,而其他 3 种算法因陷入局部最优趋于稳定。

为避免单次寻优结果具有偶然性,故针对测试函数 F_1 和 F_2 使用各算法分别独立运行 30 次。取平均值 $\bar{x}$ 与标准差 σ^2 作为评价指标。加粗结果为最优结果。具体寻优结果如表 2 所示。

表 2 测试函数寻优结果比较

Table 2 Comparison of the optimization results of test function

序号	GWO		WOA		SSA		CGSSA	
	$\bar{x}$	σ^2	$\bar{x}$	σ^2	$\bar{x}$	σ^2	$\bar{x}$	σ^2
F_1	1.58×10^{-2}	9.78×10^{-3}	3.43×10^{-11}	8.56×10^{-11}	3.85×10^{-24}	2.258×10^{-23}	6.28×10^{-78}	3.95×10^{-77}
F_2	4.48×10^1	1.91×10^1	3.98×10^{-1}	1.176	1.09×10^2	3.85×10	0	0

4) 算法消融实验

在相同输入、输出下,对 CGSSA 进行消融实验。在 CGSSA 模型中分别除去改进混沌映射策略、混沌扰动策略及高斯变异策略进行实验。具体消融实验结果如表 3 所示。通过消融实验验证变压器故障识别模型中优化算法 CGSSA 的有效性 with 快速性。

表 3 消融实验结果

Table 3 Results of ablation experiment

模型	迭代次数	最优迭代次数	最优耗时/s
本文 (CGSSA)	100	37	3.741
Tent+高斯变异	100	55	5.522
改进 Tent	100	74	4.968
Tent	100	63	6.773

1.3 核极限学习机

ELM 隐含层权重由随机赋值确定或人工经验给出增加了模型的随机性。本文提出的变压器故障识别模型是通过采用核函数思想对 ELM 的隐含层特征映射进行

优化形成的 KELM^[17]建立,有效提高了变压器故障识别模型的分类准确率与处理效率。

对于 N 组变压器故障样本 $x_i, i = 1, \dots, N$, 有:

$$\begin{cases} f(x) = h(x)\boldsymbol{\rho} = H\boldsymbol{\rho} \\ \boldsymbol{\rho} = H^T(\frac{I}{C} + HH^T)^{-1}T \end{cases} \quad (15)$$

式中: x 为输入训练的变压器故障样本; $f(x)$ 为实际输出; $h(x)$ 为样本矩阵; H 为隐含层输出矩阵; $\boldsymbol{\rho}$ 为隐含层与输出层的权重向量; I 为对角阵; C 为变压器故障识别模型中分类器的正则化系数。

定义 KELM 的核矩阵 $\boldsymbol{\Omega}_{ELM}$ 代替 HH^T , 有:

$$\begin{cases} \boldsymbol{\Omega}_{ELM} = HH^T \\ \boldsymbol{\Omega}_{ELMij} = h(x_i)h(x_j) = K(x_i, x_j) \end{cases} \quad (16)$$

式中: x_i, x_j 为输入变压器故障数据训练样本。

$K(x_i, x_j)$ 为高斯核函数, 即:

$$K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{g^2}) \quad (17)$$

式中: g 为变压器故障识别模型核函数参数。

KELM 的输出为:

$$y(x) = h(x)b^* = h(x)H^T\left(HH^T + \frac{I}{c}\right)^{-1}s =$$
$$\begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_n) \end{bmatrix} \left(\Omega_{ELM} + \frac{I}{c}\right)^{-1}s \quad (18)$$

2 基于 KPCA-CGSSA-KELM 故障识别流程

基于 KPCA-CGSSA-KELM 的变压器故障识别流程如图 3 所示,具体步骤如下。

步骤 1) 将变压器故障数据集通过 KPCA 映射到高维空间消除冗余数据后在提取变压器故障数据特征主元,并按 8 : 2 划分训练集与测试集;

步骤 2) 初始化变压器故障识别模型中优化算法 (CGSSA) 的参数,包括种群规模 n 、迭代次数 t ;

步骤 3) 设置变压器故障识别模型中分类器 (KELM) 的参数优化范围,包括正则化系数和核函数参数,并根据改进策略的结果初始化种群位置;

步骤 4) 针对变压器故障识别模型中的优化算法 CGSSA,计算当前参数下适应度值并排序,记录当前局部最优适应度值及对应位置;

步骤 5) 根据式 (5) ~ (7) 更新 CGSSA 的探索者、跟随者和警戒者位置信息;

步骤 6) 比较麻雀个体 f_i 与种群适应度值 f_{avg} 。通过多种改进策略使个体摆脱局部最优,然后更新最优适应度值及其对应位置,迭代次数 $t = t + 1$;

步骤 7) 重复步骤 4) ~ 6),直到迭代次数达到设定的最大值或个体适应度达到最优适应度值;

步骤 8) 输出此时 KELM 的参数组合即为最优参数,构建 KPCA-CGSSA-KELM 变压器故障识别模型,并输出变压器故障识别的最终分类结果。

3 实验分析

为保证所提模型具有普适性,本文采用国际电工委员会液体和气体绝缘介质技术委员会公开 IEC TC 10 数据集,其中共 350 组人工标记好的变压器故障样本数据^[17]。选取变压器正常状态(N-C)数据 37 组、低能放电 (LE-D) 状态数据 58 组、低温过热(L-T) 状态数据 10 组、高能放电 (HE-D) 状态数据 95 组、高温过热 (H-T) 状态数据 106 组、中低温过热 (LM-T) 状态数据 35 组、中温过热(M-T) 状态数据 9 组,共 350 组数据。具体油中溶解气体浓度如表 4 所示。

3.1 故障数据处理

变压器发生故障时,绝缘油发生热裂解产生多种气

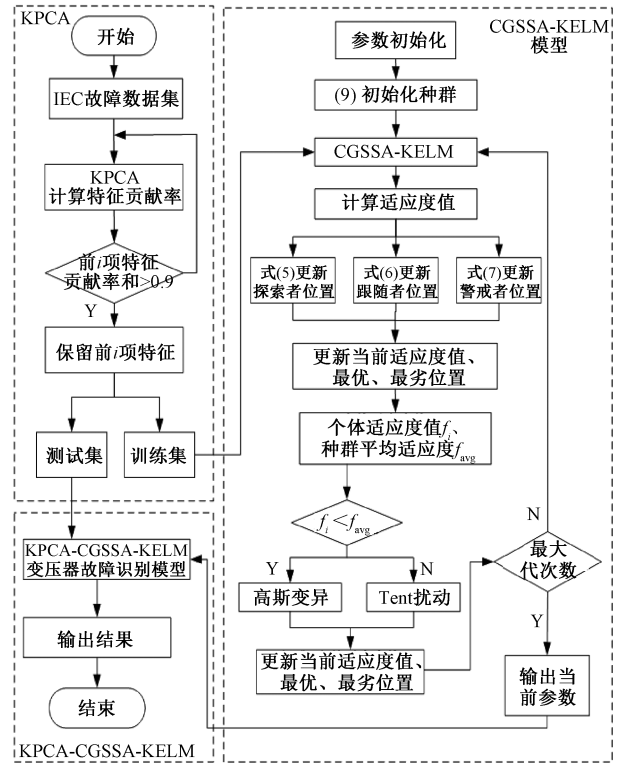


图 3 故障识别流程

Fig. 3 Fault identification process

表 4 变压器原始故障数据

Table 4 Original transformer fault data

类型	油中溶解气体浓度/($\mu\text{L}\cdot\text{L}^{-1}$)				
	H ₂	CH ₄	C ₂ H ₂	C ₂ H ₄	C ₂ H ₆
N-C/1	14. 67	3. 68	0. 2	2. 71	10. 54
LE-D/2	345. 3	112. 25	58. 75	51. 5	27. 5
L-T/3	247. 7	148. 7	0. 9	192. 7	78. 7
HE-D/4	1678	652. 9	419. 1	1005. 9	80. 7
H-T/5	172. 9	334. 1	37. 7	812. 5	172. 9
LM-T/6	170. 1	320. 6	3. 2	520. 3	53. 7
M-T/7	181. 2	262. 7	37. 7	528. 9	210. 4

体。传统方法使用 H₂、CH₄、C₂H₂、C₂H₄、C₂H₆ 的浓度作为特征进行识别。仅使用上述气体进行故障识别可能影响分类准确度,故本文采用无编码比值法。

构造 14 维变压器故障特征参量集作为输入。具体故障特征参量如表 5 所示。其中,C1、C2、C3、C4、C5 为特征气体 H₂、CH₄、C₂H₂、C₂H₄、C₂H₆ 浓度,TH 为烃类气体(即 CH₄、C₂H₂、C₂H₄、C₂H₆) 浓度总和,TE 为特征气体(即 H₂、CH₄、C₂H₂、C₂H₄、C₂H₆) 浓度总和。

高维故障数据会造成模型复杂、训练时间增长。同时冗余数据会影响模型收敛速度与结果精度^[18]。因此,本文利用 KPCA 将原始 14 维故障特征参量映射到高维空间后进行降维,并选择累计贡献率超 90% 的主成分作

为特征输入以减少模型训练时间。

表 5 故障特征量

Table 5 Fault characteristics

编号	特征	编号	特征
1	C1/(TE)	8	C4/C5
2	C2/(TE)	9	C2/(TH)
3	C3/(TE)	10	C3/(TH)
4	C4/(TE)	11	C4/(TH)
5	C5/(TE)	12	C5/(TH)
6	C2/C1	13	(C2+C4)/(TH)
7	C3/C4	14	TH/TE

对 3.1 节构建的 14 维故障特征参量进行处理。主成分特征贡献率如图 4 所示,具体结果如表 6 所示。

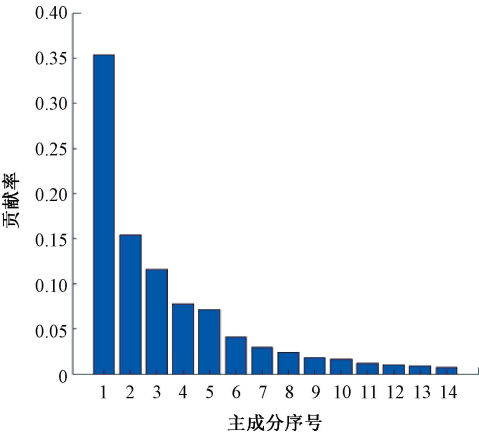


图 4 主成分特征贡献率

Fig. 4 Feature contribution rate

表 6 KPCA 特征提取结果

Table 6 Feature extraction results of KPCA

主成分	特征值	贡献率/%	累计贡献率/%
1	6.31	36.82	36.82
2	5.63	19.96	56.78
3	4.75	12.40	69.18
4	3.94	10.62	79.80
5	3.39	6.58	86.38
6	2.68	5.45	91.83

从表 6 计算得主成分 1~6 累计贡献率为 91.83%,满足标准。所以本文取前 6 个成分对应向量作为输入。

3.2 参数优化的改进模型

为验证不同的参数优化算法对变压器故障识别模型分类性能的影响,本文以 3.1 节经 KPCA 处理后的结果作为输入特征向量,分别利用 GWO^[17]、WOA^[19]、SSA^[9]和 CGSSA 这 4 种算法对 KELM 进行全局参数寻优。对各算法的参数寻优能力进行对比。

各算法进行迭代 100 次后的结果如图 5 所示。

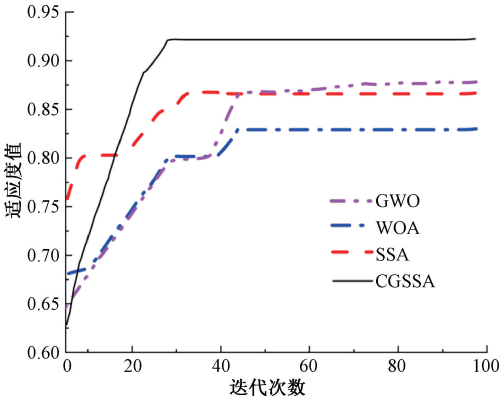


图 5 函数寻优曲线

Fig. 5 Function optimization curve

如图 5 所示,WOA 算法不适合处理离散化问题,迭代 48 次达到收敛状态,且适应度值最低;GWO 算法较 WOA 算法适应度值有所提升,最终适应度值为 0.877,但迭代 75 次才达到收敛状态;SSA 算法较 GWO 算法效果有所提升,迭代 37 次达到收敛状态,最终适应度值为 0.863;而 CGSSA 收敛速度最快,仅迭代 23 次达到收敛状态,且适应度值达到 0.935。说明本文所提 CGSSA 在收敛速度和全局寻优方面较其他 3 种算法更具优势。

3.3 不同模型分类结果对比

将经过 CGSSA 优化后参数输入到故障识别模型进行训练,并将预测的故障状态与数据集中实际出现的故障类型进行对比,验证相应模型的分类准确性。

对比 GWO-KELM、WOA-KELM、SSA-KELM、CGSSA-KELM 4 种模型^[20]对于 IEC 公开变压器故障数据集的识别准确率。将测试集共计 70 组 DGA 故障数据经过 KPCA 处理后,输入上述 4 类训练好的模型中进行故障分类。模型识别结果如图 6~9 所示,具体的故障识别准确率如表 7 所示。

表 7 不同方法故障识别结果对比

Table 7 Comparison of fault identification results of different methods (%)

故障类型	故障识别准确率			
	WOA-KELM	GWO-KELM	SSA-KELM	CGSSA-KELM
N-C	85.7	85.7	100	100
LE-D	83.3	91.7	75	91.7
L-T	100	100	100	100
HE-D	60	66.7	80	93.3
H-T	66.7	83.3	66.7	94.4
LM-T	71.4	100	71.4	100
M-T	100	100	100	100
识别正确率	75.7	85.7	80.0	95.7

由图 6~9 可得,在 70 组经过 KPCA 后的测试数据中,CGSSA-KELM 正确预测的变压器故障个数为 67 个。

预测分类	1	6 8.6%	1 1.4%	0 0.0%	1 1.4%	1 1.4%	1 1.4%	0 0.0%	60.0% 40.0%
	2	1 1.4%	10 14.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	90.9% 9.1%
	3	0 0.0%	1 1.4%	6 8.6%	2 2.9%	1 1.4%	1 1.4%	0 0.0%	54.5% 45.5%
	4	0 0.0%	0 0.0%	0 0.0%	9 12.9%	3 4.3%	0 0.0%	0 0.0%	75.0% 25.0%
	5	0 0.0%	0 0.0%	0 0.0%	1 1.4%	12 17.1%	0 0.0%	0 0.0%	92.3% 7.7%
	6	0 0.0%	0 0.0%	0 0.0%	1 1.4%	1 1.4%	5 7.1%	0 0.0%	71.4% 28.6%
	7	0 0.0%	0 0.0%	0 0.0%	1 1.4%	0 0.0%	0 0.0%	5 7.1%	83.3% 16.7%
		85.7% 14.3%	83.3% 16.7%	100% 0.0%	60.0% 40.0%	66.7% 33.3%	71.4% 28.6%	100% 0.0%	75.7% 24.3%
		实际分类							

图 6 WOA-KELM 分类结果

Fig. 6 WOA-KELM classification results

预测分类	1	6 8.6%	1 1.4%	0 0.0%	1 1.4%	0 0.0%	0 0.0%	0 0.0%	75.0% 25.0%
	2	1 1.4%	11 15.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	91.7% 8.3%
	3	0 0.0%	0 0.0%	6 8.6%	1 1.4%	0 0.0%	0 0.0%	0 0.0%	85.7% 14.3%
	4	0 0.0%	0 0.0%	0 0.0%	10 14.3%	2 2.9%	0 0.0%	0 0.0%	83.3% 16.7%
	5	0 0.0%	0 0.0%	0 0.0%	1 1.4%	15 21.4%	0 0.0%	0 0.0%	93.8% 6.3%
	6	0 0.0%	0 0.0%	0 0.0%	1 1.4%	1 1.4%	7 10.0%	0 0.0%	77.8% 22.2%
	7	0 0.0%	0 0.0%	0 0.0%	1 1.4%	0 0.0%	0 0.0%	5 7.1%	83.3% 16.7%
		85.7% 14.3%	91.7% 8.3%	100% 0.0%	66.7% 33.3%	83.3% 16.7%	100% 0.0%	100% 0.0%	85.7% 14.3%
		实际分类							

图 7 GWO-KELM 分类结果

Fig. 7 GWO-KELM classification results

预测分类	1	7 10.0%	1 1.4%	0 0.0%	0 0.0%	0 0.0%	1 1.4%	0 0.0%	77.8% 22.2%
	2	0 0.0%	9 12.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	3	0 0.0%	2 2.9%	6 8.6%	0 0.0%	1 1.4%	1 1.4%	0 0.0%	60.0% 40.0%
	4	0 0.0%	0 0.0%	0 0.0%	12 17.1%	3 4.3%	0 0.0%	0 0.0%	80.0% 20.0%
	5	0 0.0%	0 0.0%	0 0.0%	1 1.4%	12 17.1%	0 0.0%	0 0.0%	92.3% 7.7%
	6	0 0.0%	0 0.0%	0 0.0%	1 1.4%	2 2.9%	5 7.1%	0 0.0%	62.5% 37.5%
	7	0 0.0%	0 0.0%	0 0.0%	1 1.4%	0 0.0%	0 0.0%	5 7.1%	83.3% 16.7%
		100% 0.0%	75.0% 25.0%	100% 0.0%	80.0% 20.0%	66.7% 33.3%	71.4% 28.6%	100% 0.0%	80.0% 20.0%
		实际分类							

图 8 SSA-KELM 分类结果

Fig. 8 SSA-KELM classification results

预测分类	1	7 10.0%	1 1.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	87.5% 12.5%
	2	0 0.0%	11 15.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	3	0 0.0%	0 0.0%	6 8.6%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	4	0 0.0%	0 0.0%	0 0.0%	14 20.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	5	0 0.0%	0 0.0%	0 0.0%	1 1.4%	17 24.3%	0 0.0%	0 0.0%	94.4% 5.6%
	6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 1.4%	7 10.0%	0 0.0%	87.5% 12.5%
	7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5 7.1%	100% 0.0%
		100% 0.0%	91.7% 8.3%	100% 0.0%	93.3% 6.7%	94.4% 5.6%	100% 0.0%	100% 0.0%	95.7% 4.3%
		实际分类							

图 9 CGSSA-KELM 分类结果

Fig. 9 CGSSA-KELM classification results

故障诊断正确率为 95.7%,比相同条件下的 WOA-KELM、GWO-KELM、SSA-KELM 变压器故障识别准确率分别提高 18.6%、10%、15.7%。结果表明, KPCA-CGSSA-KELM 可以更加精确地识别变压器故障状态。

4 结 论

针对变压器故障识别准确率低以及冗余特征对于识别过程的干扰,本文提出一种基于核主成分分析与混沌高斯麻雀搜索算法优化核极限学习机的变压器故障识别方法。该方法通过构建故障特征集并进行核主成分分析,在有效消除冗余特征的同时找到提高故障识别准确率的关键特征;通过引入改进 Tent 混沌映射和高斯变异策略的 CGSSA 算法对 KELM 的正则化系数 α 和核函数

参数 g 进行优化,建立 KPCA-CGSSA-KELM 故障识别模型。实验结果证明 KPCA-CGSSA-KELM 与 WOA-KELM、GWO-KELM、SSA-KELM 相比,具有更高的故障识别准确率与泛化能力,可以对变压器异常状态进行判定来及时发现运行过程中存在的故障。但在变压器发生低能放电、高能放电以及高温过热故障时的识别准确率有待提高。下一步会深入研究变压器多重故障识别方法,进一步提高变故障识别准确率。

参考文献

[1] 李典阳,张育杰,冯健,等. 变压器故障样本多维诊断及结果可信度分析[J]. 电工技术学报,2022,37(3): 667-675.

LI D Y, ZHANG Y J, FENG J, et al. Multidimensional diagnosis of transformer fault samples and reliability

- analysis of results [J]. Transactions of China Electrotechnical Society, 2022, 37 (3): 667-675.
- [2] 栗磊,王廷涛,赫嘉楠,等. 考虑过采样器与分类器参数优化的变压器故障诊断策略[J]. 电力自动化设备, 2023, 43(1): 209-217.
- LI L, WANG T T, HE J N, et al. Transformer fault diagnosis strategy considering optimization of sampler and classifier parameters [J]. Electric Power Automation Equipment, 2023, 43 (1): 209-217.
- [3] 李雷军,吴超,付华,等. 基于油中溶解气体分析的 ISSA 优化 LSSVM 变压器故障诊断研究[J]. 电工电能新技术, 2023, 42(10): 84-94.
- LI L J, WU CH, FU H, et al. Research on ISSA optimized LSSVM transformer fault diagnosis based on dissolved gas analysis in oil [J]. Advanced Technology of Electrical Engineering, 2023, 42(10): 84-94.
- [4] 周光宇,马松龄. 基于机器学习与 DGA 的变压器故障诊断及定位研究[J]. 高压电器, 2020, 56 (6): 262-268.
- ZHOU G Y, MA S L. Research on transformer fault diagnosis and location based on machine learning and DGA [J]. High Voltage Apparatus, 2020, 56 (6): 262-268.
- [5] 石延辉,杨洋,廖毅,等. 基于改进粒子群算法优化 SVM 的变压器故障诊断[J]. 武汉大学学报(工学版), 2023, 56(10): 1238-1244.
- SHI Y H, YANG Y, LIAO Y, et al. Transformer fault diagnosis based on improved particle swarm optimization SVM [J]. Engineering Journal of Wuhan University, 2023, 56 (10): 1238-1244.
- [6] 王艳,王寅初,赵洪山,等. 基于 AdaBoost. M2-ISSA-ELM 算法的电力变压器故障诊断方法[J/OL]. 电力自动化设备, 1-8.
- WANG Y, WANG Y CH, ZHAO H SH, et al. Based on AdaBoost M2-ISSA-ELM algorithm for power transformer fault diagnosis [J/OL]. Electric Power Automation Equipment, 1-8.
- [7] 王利福,魏玉琪,刘屹江泽. 基于 KPCA 与 IPFA-KELM 的变压器故障识别方法[J]. 控制工程, 2023, 30(7): 1180-1189.
- WANG L F, WEI Y Q, LIU Y J Z. Transformer fault identification method based on KPCA and IPFA-KELM [J]. Control Engineering of China, 2023, 30 (7): 1180-1189.
- [8] 欧阳鑫,李志斌. 基于样本扩充和特征优选的 IGWO 优化 SVM 的变压器故障诊断技术[J]. 电力系统保护与控制, 2023, 51(18): 11-20.
- OUYANG X, LI ZH B. Transformer fault diagnosis technology based on IGWO optimized SVM with sample expansion and feature optimization [J]. Power System Protection and Control, 2023, 51 (18): 11-20.
- [9] 李黄曼,张勇,张瑶. 基于 ISSA 优化 SVM 的变压器故障诊断研究[J]. 电子测量与仪器学报, 2021, 35(3): 123-129.
- LI H M, ZHANG Y, ZHANG Y. Research on transformer fault diagnosis based on ISSA optimized SVM [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35 (3): 123-129.
- [10] 陈里里,何颖,董绍江. 基于深度神经网络的液压泵泄漏状态识别[J]. 仪器仪表学报, 2020, 41(4): 86-94.
- CHEN L L, HE Y, DONG SH J. Identification of hydraulic pump leakage status based on deep neural networks [J]. Chinese Journal of Scientific Instrument, 2020, 41 (4): 86-94.
- [11] 张淑清,苑世钰,姚玉永,等. 基于 ICEEMD 及 AWOA 优化 ELM 的机械故障诊断方法[J]. 仪器仪表学报, 2019, 40(11): 172-180.
- ZHANG SH Q, YUAN SH Y, YAO Y Y, et al. Mechanical fault diagnosis method based on ICEEMD and AWOA optimized ELM [J]. Chinese Journal of Scientific Instrument, 2019, 40 (11): 172-180.
- [12] 刘光达,董梦坤,张守伟,等. 基于 KPCA-SVM 的表面肌电信号疲劳分类研究[J]. 电子测量与仪器学报, 2021, 35(10): 1-8.
- LIU G D, DONG M K, ZHANG SH W, et al. Fatigue classification of surface electromyography signals based on KPCA-SVM [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35 (10): 1-8.
- [13] 王雨虹,王志中. 基于 RFRFE 与 ISSA-XGBoost 的变压器故障辨识方法[J]. 电子测量与仪器学报, 2021, 35(12): 142-150.
- WANG Y H, WANG ZH ZH. Transformer fault identification method based on RFRFE and ISSA-XGBoost [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35 (12): 142-150.
- [14] 吕鑫,慕晓冬,张钧,等. 混沌麻雀搜索优化算法[J]. 北京航空航天大学学报, 2021, 47(8): 1712-1720.
- LYU X, MU X D, ZHANG J, et al. Chaotic sparrow search optimization algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47 (8): 1712-1720.
- [15] 张娜,赵泽丹,包晓安,等. 基于改进的 Tent 混沌万有引力搜索算法[J]. 控制与决策, 2020, 35(4): 893-900.
- ZHANG N, ZHAO Z D, BAO X AN, et al. Improved tent chaotic universal gravitational search algorithm [J].

- Control and Decision, 2020, 35 (4): 893-900.
- [16] RUDOLPH G. Local convergence rates of simple evolutionary algorithms with Cauchy mutations[J]. IEEE Transactions on Evolutionary Computation, 1997, 1(4): 249-258.
- [17] 李云溟, 咸日常, 张海强, 等. 基于改进灰狼算法与最小二乘支持向量机耦合的电力变压器故障诊断方法[J]. 电网技术, 2023, 47(4): 1470-1478.
LI Y H, XIAN R CH, ZHANG H Q, et al. A power transformer fault diagnosis method based on the coupling of improved grey wolf algorithm and least squares support vector machine [J]. Power System Technology, 2023, 47 (4): 1470-1478.
- [18] 汪可, 李金忠, 张书琦, 等. 变压器故障诊断用油中溶解气体新特征参量[J]. 中国电机工程学报, 2016, 36(23): 6570-6578, 6625.
WANG K, LI J ZH, ZHANG SH Q, et al. New characteristic parameters of dissolved gases in oil for transformer fault diagnosis [J]. Proceedings of the CSEE, 2016, 36 (23): 6570-6578, 6625.
- [19] 安国庆, 史哲文, 马世峰, 等. 基于 RF 特征优选的 WOA-SVM 变压器故障诊断[J]. 高压电器, 2022, 58(2): 171-178.
- AN G Q, SHI ZH W, MA SH F, et al. WOA-SVM transformer fault diagnosis based on RF feature optimization [J]. High Voltage Apparatus, 2022, 58 (2): 171-178.
- [20] 谢国民, 黄睿灵, 丁会巧. 基于 VMD 样本熵和 KELM 的输电线路故障诊断[J]. 电子测量与仪器学报, 2019, 33(5): 73-79.
XIE G M, HUANG R L, DING H Q. Transmission line fault diagnosis based on VMD sample entropy and KELM [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33 (5): 73-79.

作者简介



江兵(通信作者), 2007 年于中国科学技术大学获博士学位, 现为南京邮电大学副教授, 主要研究方向为智能仪器与测控系统。

E-mail: jiangb@njupt.edu.cn

Jiang Bing (Corresponding author) received his Ph. D. degree from University of Science and Technology of China in 2007. Now he is an associate professor at Nanjing University of Posts and Telecommunications. His main research interests include intelligent instruments and measurement & control systems.