

刘蒙,赵晨晓,朱乔波,等.基于 t -SNE和ECOC-ISSA-SVM的变压器故障诊断[J].辽宁工程技术大学学报(自然科学版),2025,44(5):606-613.DOI:10.11956/j.issn.1008-0562.20250166
LIU Meng,ZHAO Chenxiao,ZHU Qiaobo,et al.Transformer fault diagnosis based on t -SNE and ECOC-ISSA-SVM[J].Journal of Liaoning Technical University(Natural Science),2025,44(5):606-613.DOI:10.11956/j.issn.1008-0562.20250166

基于 t -SNE和ECOC-ISSA-SVM的变压器故障诊断

刘蒙¹,赵晨晓¹,朱乔波¹,李梁¹,姚旭²,李鑫^{3*},赵明³

(1. 国网冀北电力有限公司,北京 100089;2. 北京科东电力控制系统有限责任公司,北京 100089;
3. 辽宁工程技术大学电气与控制工程学院,辽宁 葫芦岛 125105)

摘要:为解决电力变压器故障诊断中支持向量机(support vector machine, SVM)超参数优化和多分类性能不足的问题,采用 t -分布的随机邻居嵌入(t -distributed stochastic neighbor embedding, t -SNE)对26维溶解气体分析(DGA)数据进行非线性降维,引入纠错输出码(error correction output codes, ECOC),将改进麻雀搜索算法(improved sparrow search algorithm, ISSA)与切比雪夫混沌映射、柯西-高斯变分策略相结合,优化SVM超参数,处理多分类问题。研究表明:ECOC-ISSA-SVM(t -SNE)模型的诊断精度、召回率、特异性和F1值分别为95.6%、97.8%、99.6%和97.8%,各项指标较传统模型提升效果显著,诊断时间缩短至11 ms,诊断效率显著提高。研究结论为电力设备智能运维提供技术支持。

关键词:故障诊断;变压器;油中溶解气体;支持向量机;麻雀搜索算法; t -SNE降维;纠错输出码

中图分类号:TM407

文献标志码:A

文章编号:1008-0562(2025)05-0606-08

Transformer fault diagnosis based on t -SNE and ECOC-ISSA-SVM

LIU Meng¹, ZHAO Chenxiao¹, ZHU Qiaobo¹, LI Liang¹, YAO Xu², LI Xin^{3*}, ZHAO Ming³

(1. State Grid Jibei Electric Power Company Limited, Beijing 100089, China;2. Beijing Kedong Power Control Systems Company Limited, Beijing 100089, China;3. Faculty of Electrical and Control Engineering, Liaoning Technical University, Huludao 125105, China)

Abstract: To solve the problems of insufficient hyperparameter optimization and multi-classification performance of support vector machine (SVM) in the fault diagnosis of power transformers, the nonlinear dimension reduction of 26-dimensional dissolved gas analysis (DGA) data is carried out by using the t -distributed stochastic neighbor embedding (t -SNE). Error correction output codes (ECOC) are introduced, and the improved sparrow search algorithm (ISSA) was combined with Chebyshev chaotic mapping and Cauchy-Gaussian variational strategy to optimize the hyperparameters of SVM and handle multi-classification problems. The research results show that the diagnostic accuracy, recall rate, specificity and F1 value of the ECOC-ISSA-SVM(t -SNE) model are 95.6%, 97.8%, 99.6% and 97.8% respectively. Compared with the traditional model, the improvement effect of each index is significant. The diagnostic time is shortened to 11 ms and the diagnostic efficiency is significantly improved. The research conclusion provides technical support for the intelligent operation and maintenance of power equipment.

Keywords: fault diagnosis; transformer; dissolved gas in oil; support vector machine; sparrow search algorithm; t -SNE dimensionality reduction; error correction output codes

0 引言

电力变压器是电网核心设备,具有电压调控

与能量传输功能,但长期运行易发生故障,对故障类型进行精准诊断是保障系统稳定运行的关键

收稿日期:2025-04-16 修回日期:2025-05-21 接受日期:2025-06-04 责任编辑:黄琦

基金项目:国家自然科学基金项目(52104160)

作者简介:刘蒙(1980-),男,山东临沂人,硕士,高级工程师,主要从事电力电子技术方面的研究。E-mail: 2676515950@qq.com

通信作者:李鑫(1981-),女,辽宁阜新人,博士,教授,主要从事电力电子技术方面的研究。E-mail: lixinyz@126.com

环节^[1]。油中溶解气体分析 (DGA)^[2]是常用的诊断技术之一。传统方法 (例如三比值法、Duval 三角法)^[3]较依赖经验与历史数据, 存在编码缺失和过拟合问题。引入人工智能算法 (例如随机森林^[4]、支持向量机^[5]、神经网络^[6]) 可提升诊断效能, 其中, 支持向量机 (support vector machine, SVM) 因擅长处理小样本非线性数据且与 DGA 兼容性较强, 成为目前的研究热点。

文献[7]采用 ADASYN 平衡数据集改进支持向量机, 但检测的准确率提升有限。文献[8]利用秃鹰搜索算法 BES 优化 SVM, 虽效果显著, 但未解决算法种群多样性不足的问题。文献[9]结合核主成分分析及改进的海鸥算法来优化 SVM, 仍未能突破多分类问题的瓶颈。文献[10]和文献[11]分别通过改进麻雀搜索算法 ISSA 和组合模型 VMDSMA-LSSVM 来优化卷积神经网络 CNN 及最小二乘支持向量机 LSSVM, 但均未深入分析 DGA 数据特征的相关性。文献[12]采用基于混合策略的改进灰狼算法来优化 LSSVM 性能, 虽然提高了故障诊断的准确率, 但忽略了对多分类策略的改进。上述研究在模型优化方面已取得一定进展, 但存在数据特征利用不充分等问题。文献[13]通过堆叠残差网络模块改进 BP 神经网络, 实现了气体特征的高效融合, 但未对超参数进行优化。文献[14]采用过采样技术与深度置信网络 DBN 相结合的方法, 提升了诊断精度。文献[15]利用多尺度近似熵及 ISSA 来优化 CNN, 在离线数据方面表现优异, 但未考虑模型复杂性对实时诊断效果的影响。上述方法通过深度学习增强了特征提取和诊断能力, 但存在超参数调优不足、实时性考量缺失等局限, 需进一步探索能平衡故障诊断精度与效率的模型。

本文提出一种改进的变压器故障诊断模型 ECOC-ISSA-SVM (t -SNE)。通过 t -分布的随机邻居嵌入 (t -SNE) 对采用相关比值法构建的 26 维 DGA 数据进行降维处理, 利用 ISSA 优化 SVM 超参数, 借助纠错输出码 ECOC 构建多分类 SVM 模型, 以期提升算法诊断精度和效率。

1 DGA 数据集

1.1 DGA 数据的前期预处理

采用国家电网某公司提供的变压器故障数据集, 包含 5 种特征气体, 分别为乙炔 (C_2H_2)、乙烯 (C_2H_4)、乙烷 (C_2H_6)、甲烷 (CH_4) 和氢气 (H_2)。数据集涵盖 6 种故障模式, 即 T1 (中低温过热)、T2 (高温过热)、PD (局部放电)、D1 (低能放电)、D2 (高能放电) 和 NC (正常)。故障类型编号见表 1。每种故障包含 45 组样本且数据集按 2 : 1 进行划分, 即训练集 (30 组/类) 和测试

集 (15 组/类) 分别包含 180 组和 90 组样本。

表 1 故障类型编号
Tab.1 fault type number

编号	故障类型	编号	故障类型
1	T1	4	D1
2	T2	5	D2
3	PD	6	NC

1.2 DGA 数据的重构与降维

油浸式变压器中的气体含量可直接反映设备健康状况。采用相关比值法将 5 种特征气体数据扩展成 26 种故障特征类型, 具体见表 2。

表 2 故障特征类型
Tab.2 types of fault characteristics

编号	故障特征	编号	故障特征
1	乙炔(C_2H_2)/氢气(H_2)	14	甲烷(CH_4)/G1
2	乙炔(C_2H_2)/甲烷(CH_4)	15	氢气(H_2)/G1
3	乙炔(C_2H_2)/乙烷(C_2H_6)	16	乙炔(C_2H_2)/G2
4	乙炔(C_2H_2)/乙烯(C_2H_4)	17	乙烯(C_2H_4)/G2
5	乙烯(C_2H_4)/氢气(H_2)	18	乙烷(C_2H_6)/G2
6	乙烯(C_2H_4)/甲烷(CH_4)	19	甲烷(CH_4)/G2
7	乙烯(C_2H_4)/乙烷(C_2H_6)	20	氢气(H_2)/G2
8	乙烷(C_2H_6)/氢气(H_2)	21	乙炔(C_2H_2)/G3
9	乙烷(C_2H_6)/甲烷(CH_4)	22	乙烯(C_2H_4)/G3
10	甲烷(CH_4)/氢气(H_2)	23	乙烷(C_2H_6)/G3
11	乙炔(C_2H_2)/G1	24	甲烷(CH_4)/G3
12	乙烯(C_2H_4)/G1	25	氢气(H_2)/G3
13	乙烷(C_2H_6)/G1	26	(G2+G3)/G1

注: /表示比值; G1=氢气+甲烷+液化乙烷+乙炔+乙烷; G2=甲烷+液化乙烷+乙烯+乙炔; G3=液化乙烷+乙烯+乙炔。

虽然相关比值法扩展的 DGA 数据提升了故障区分度, 但故障特征维度过高会影响 SVM 计算效率。鉴于此, 采用 t -SNE 算法^[16]对 26 维特征进行降维处理, 得到基于非线性降维技术的维度特征值, 见表 3。

表 3 基于非线性降维技术的维度特征值
Tab.3 dimension eigenvalue based on nonlinear dimensionality reduction technology

特征值	取值	特征值	取值	特征值	取值
X_1	-13.011	X_2	-16.823	X_3	-2.630
	-8.599		-11.202		5.649
	-0.652		-19.256		-5.339
	4.955		-11.712		-4.410
	4.936		-14.375		-7.159
	3.085		-14.795		-6.002
	4.219		-1.856		4.368
	1.336		-1.654		-7.885
	5.811		-2.891		4.763
	5.300		0.535		4.989
	5.086		-0.399		2.672
	5.383		-8.919		7.875
	6.442		-6.489		10.295
	5.125		1.137		7.289
	5.485		0.762		6.112

对维度特征值进行归一化处理来消除量纲影响,即

$$X'_i = \frac{X_i - X_{i,\min}}{X_{i,\max} - X_{i,\min}}, \quad (1)$$

式中: X_i 为原始数据集中的第 i 个样本值, $i=1,2,3$; $X_{i,\max}$ 为原始数据集中的最大值; $X_{i,\min}$ 为原始数据集中的最小值; X'_i 为归一化后的第 i 个样本值。

2 改进麻雀搜索算法

2.1 麻雀搜索算法

麻雀搜索算法 (sparrow search algorithm, SSA) [17] 是一种新型的元启发式算法,可模拟麻雀觅食和躲避捕食者行为。该算法假设存在一个由 n 只麻雀组成的种群,满足

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{bmatrix}, \quad (2)$$

式中: $\mathbf{X}$ 为麻雀种群的位置集合; n 为种群中麻雀总数量; m 为优化变量的维数; $x_{n,m}$ 为第 n 只麻雀在第 m 个维度上的位置。

$$\mathbf{F}_x = \begin{bmatrix} f(x_{1,1} & x_{1,2} & \cdots & x_{1,m}) \\ f(x_{2,1} & x_{2,2} & \cdots & x_{2,m}) \\ \vdots & \vdots & & \vdots \\ f(x_{n,1} & x_{n,2} & \cdots & x_{n,m}) \end{bmatrix}, \quad (3)$$

式中: $\mathbf{F}_x$ 为麻雀种群中所有个体的适应度值集合; 矩阵中每行数据为每个个体的适应度。

发现者位置更新满足关系式

$$x_{i,j}^{t+1} = \begin{cases} x_{i,j}^t \cdot \exp\left(-\frac{i}{at_{\max}}\right) & R_2 < S_T \\ x_{i,j}^t + QL & R_2 \geq S_T \end{cases}, \quad (4)$$

式中: $x_{i,j}^{t+1}$ 为第 $t+1$ 次迭代时, 第 i 只麻雀 (发现者) 在第 j 个维度上的位置; $x_{i,j}^t$ 为第 t 次迭代时, 第 i 只麻雀在第 j 个维度上的位置; a 为衰减系数; Q 为服从正态分布的随机数; $\mathbf{L}$ 为 $1 \times d$ 的矩阵, 矩阵内每个元素均为 1, d 为待优化的变量的维度; R_2 为预警值, 用于判断麻雀所处环境的安全状态; S_T 为安全阈值, 与 R_2 共同控制侦察行为; $t_{\max}$ 为最大迭代次数。

追随者的位置更新关系式为

$$x_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{\mathbf{X}_{\text{worst}}^t - x_{i,j}^t}{i^2}\right) & i > \frac{n}{2} \\ \mathbf{X}_p^{t+1} + |x_{i,j}^t - \mathbf{X}_p^{t+1}| \cdot \mathbf{A}^+ \cdot \mathbf{L} & i \leq \frac{n}{2} \end{cases}, \quad (5)$$

式中: $\mathbf{X}_{\text{worst}}^t$ 为第 t 次迭代时, 种群中适应度最差的

麻雀位置 (全局最劣解); $\mathbf{X}_p^{t+1}$ 为第 $t+1$ 次迭代时, 更新后的发现者最优位置; $\mathbf{A}$ 为方向向量, 其元素是取值 ± 1 的随机数; $\mathbf{A}^+$ 为 $\mathbf{A}$ 的 Moore-Penrose 伪逆, $\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1}$, 用于调整方向。

侦察者的位置更新关系式为

$$x_{i,j}^{t+1} = \begin{cases} \mathbf{X}_{\text{best}}^t + \beta \cdot |x_{i,j}^t - \mathbf{X}_{\text{best}}^t| & f_i > f_g \\ x_{i,j}^t + k \cdot \frac{|x_{i,j}^t - \mathbf{X}_{\text{worst}}^t|}{f_i - f_w + \varepsilon} & f_i = f_g \end{cases}, \quad (6)$$

式中: $\mathbf{X}_{\text{best}}^t$ 为第 t 次迭代时, 种群中适应度最优的麻雀位置; β 为步长控制参数, 取 $0 \sim 1$ 内服从正态分布的随机数; f_g 和 f_w 分别为当前全局最优适应度和当前全局最差适应度; f_i 为第 i 只麻雀的适应度; k 为缩放因子; ε 为极小的正实数。

2.2 多策略改进的麻雀搜索算法

对 SSA 进行改进。进行 Chebyshev 混沌映射初始化, 通过混沌序列增强种群多样性, 扩大全局探索范围并加速收敛。采用 Salp 群算法更新发现者位置。融合群体智能思想, 动态平衡全局搜索与局部开发能力。采用 Cauchy-Gaussian 变分策略, 通过混合变异机制提升种群多样性, 有效抑制早熟收敛。

均匀分布切比雪夫映射系统 [18] 满足

$$\begin{cases} \mathbf{X}_{i+1} = \cos[t \cdot \arccos(\mathbf{X}_i)] & -1 \leq \mathbf{X}_i \leq 1 \\ \mathbf{Y}_t = \mathbf{X}_{t+1} & t = 0, 1, 2, \dots, n-1 \\ \mathbf{Y}_{t+1} = (2/\pi) \arcsin(\cos \mathbf{Y}_t) + M \end{cases}, \quad (7)$$

式中: $\mathbf{X}_{t+1}$ 和 $\mathbf{Y}_{t+1}$ 分别为第 $t+1$ 次和第 $t+2$ 次映射迭代后的麻雀位置; $\mathbf{Y}_t$ 为第 $t+1$ 次映射迭代后的麻雀位置; t 为当前迭代次数; $\mathbf{X}_i$ 为第 i 次映射迭代后的麻雀位置; $\mathbf{X}_{t+1}$ 为第 $t+1$ 次映射迭代后的麻雀位置; M 为偏移量, 用于对 $\mathbf{Y}_{t+1}$ 的计算结果进行调整。

均匀分布切比雪夫映射初始化麻雀种群的具体步骤如下。

步骤 1 在 $[-1, 1]$ 内随机构建 1 个 m 维向量, 表示初始个体。

步骤 2 通过迭代替换式 (7) 的第 1 个方程, 生成 $n-1$ 个新个体。

步骤 3 引入比例因子和偏移量, 将标准切比雪夫映射 [19] 的输出转换至 SSA 搜索空间, 保证初始个体在搜索空间内均匀分布。

麻雀算法在早期阶段搜索范围较狭窄, 全局搜索能力较弱, 容易错过非零点位置的最优解。为解决上述问题, 参考尊海鞘算法 [20] 中 leaderd 的更新策略, 即

$$\mathbf{X}_d^1 = \begin{cases} \mathbf{F}_d + c_1[(u_b - l_b)c_2 + l_b] & c_3 \geq 0.5 \\ \mathbf{F}_d - c_1[(u_b - l_b)c_2 + l_b] & c_3 < 0.5 \end{cases}, \quad (8)$$

式中: $\mathbf{X}_d^1$ 为第1个个体在 d 维中的位置; $\mathbf{F}_d$ 为食物源在 d 维中的位置; u_b 和 l_b 分别为解空间的上限和下限; c_1 为动态调整参数; c_2 和 c_3 均为 $[0,1]$ 内的随机数。

在该算法的部分搜索和整体搜索中, c_1 满足关系式

$$c_1 = 2e^{-(4t/t_{\max})^2}. \quad (9)$$

由于 $R_2 \geq S_T$, 消去式 (4)、式 (5) 中的 L , 则发现者位置更新公式为

$$\mathbf{x}_{i,j}^{t+1} = \begin{cases} \mathbf{x}_{i,j}^t \cdot \frac{c_1[(u_b - l_b)c_2 + l_b]}{(1 + c_3)u_b} & R_2 < S_T \\ \mathbf{x}_{i,j}^t + Q & R_2 \geq S_T \end{cases}. \quad (10)$$

针对 SSA 后期易陷入局部最优的问题, 采用 Cauchy-Gaussian 突变策略^[21]进行改进, 其计算式为

$$\mathbf{U}_{\text{best}}' = \mathbf{X}_{\text{best}}' [1 + \lambda_1 C(0, \sigma^2) + \lambda_2 G(0, \sigma^2)], \quad (11)$$

$$\sigma = \begin{cases} 1 & f(\mathbf{X}_{\text{best}}) < f(\mathbf{X}_i) \\ \exp\left[\frac{f(\mathbf{X}_{\text{best}}) - f(\mathbf{X}_i)}{|f(\mathbf{X}_{\text{best}})|}\right] & \text{其他} \end{cases}, \quad (12)$$

式 (11)、式 (12) 中: $\mathbf{U}_{\text{best}}'$ 和 $\mathbf{X}_{\text{best}}'$ 分别为变异后和变异前最优个体的位置; σ^2 为柯西-高斯变分策略的标准差; $C(0, \sigma^2)$ 为满足柯西分布的随机变量; $G(0, \sigma^2)$ 为满足高斯分布的随机变量; λ_1 为随着算法迭代逐渐减小的系数; λ_2 为随着算法迭代逐渐增大的系数。通过动态调整 λ_1 、 λ_2 , 算法在搜索过程中更加灵活高效。

2.3 ISSA 算法

ISSA 算法具体实施步骤如下。

步骤1 参数初始化。设置算法初始参数。

步骤2 种群初始化。通过切比雪夫混沌映射生成初始种群。

步骤3 适应度排序。按适应度降序排列个体, 优先保留最优解。

步骤4 位置更新。发现者和追随者分别根据式 (6)、式 (7) 更新位置, 优化搜索方向; 侦察者按式 (8) 更新位置, 避免局部最优; 个体位置基于式 (12) 进行变异, 提升搜索能力。

步骤5 终止判断。若满足迭代条件, 输出最

优解; 否则, 返回步骤3继续优化。

2.4 ISSA 算法性能测试

为验证 ISSA 算法对基准测试函数最优问题的处理能力, 选取6个基准函数进行测试, 其中, $F_1 \sim F_3$ 为单峰测试函数, $F_4 \sim F_6$ 为多峰测试函数。具体见表4。实验操作系统为 Windows 10 (64位), 处理器为 AMD R7 5800H with Radeon Graphics. 3.20GHz, 计算机内存为 16 GB, 仿真平台为 Matlab 2020b。

表4 测试函数相关参数
Tab.4 test function related parameters

函数	维度	范围	最优值
F_1	30	$[-100.00, 100.00]$	0
F_2	30	$[-10.00, 10.00]$	0
F_3	30	$[-100.00, 100.00]$	0
F_4	30	$[-1.28, 1.28]$	0
F_5	30	$[-32.00, 32.00]$	0
F_6	30	$[-5.12, 5.12]$	0

$$F_1(x) = \sum_{i=1}^n x_i^2,$$

$$F_2(x) = \sum_{i=1}^n |x_i| + \prod_{i=1}^n |x_i|,$$

$$F_3(x) = \sum_{i=1}^n \left(\sum_{j=1}^n x_j \right)^2,$$

$$F_4(x) = \sum_{i=1}^n i x_i^4 + r_{\text{and}} [0, 1),$$

$$F_5(x) = -20 \exp \left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \right) - \exp \left[\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i) \right] + 20 + e,$$

$$F_6(x) = \sum_{i=1}^n x_i^2 - 10 \cos(2\pi x_i) + 10.$$

选取鲸鱼优化算法 (WOA)^[22]、海洋捕食者算法 (MPA)^[23]、正弦余弦算法 (SCA)、金豺优化算法 (GJO)^[24]、SSA 及 ISSA 进行对比测试。为减少随机误差, 提升测试结果的可靠性, 每种算法在6个基准函数上独立运行30次, 算法参数设定见表5。表中, A 为系数向量, 用于控制包围猎物、螺旋更新等搜索行为, 迭代过程中从2线性递减到0; b_w 为形状参数; F 为飞蛾的自适应步长因子; O 为螺旋形状控制参数; a_m 为控制参数, 平衡算法的探索 and 开发能力; C_1 为学习因子; P 为发现者 (探索者) 麻雀数量占麻雀总数量 (Q) 的比例; S 为加入者 (追随者) 麻雀数量占麻雀总数量的比例。

表 5 算法参数设定
Tab.5 setting of algorithm parameters

算法类型	参数设定
WOA	A 从 2 到 0, $b_w=1$
MPA	$F=0.2, O=0.5$
SCA	$\alpha_m=2$ 或 $\alpha_m=0.5$
GJO	C_1 从 1 到 2
SSA	$P=0.2Q, S=0.2Q, R_2=0.8$
ISSA	$P=0.2Q, S=0.2Q, R_2=0.8$

设置实验参数如下：种群规模为 50，最大迭代次数为 500。不同算法的搜索结果对比见表 6。

在测试条件相同的情况下，ISSA 在 6 个测试函数中均优于其他算法。其中，在 $F_1 \sim F_3$ 函数中获

表 6 不同算法的搜索结果对比
Tab.6 comparison of search results of different algorithms

函数	WOA	MPA	SCA	GJO	SSA	ISSA
F_1	3.46×10^{-84}	4.23×10^{-25}	7.23×10^{-2}	1.58×10^{-57}	0	0
F_2	3.28×10^{-58}	1.58×10^{-14}	1.03×10^{-4}	7.65×10^{-34}	0	0
F_3	1.60×10^4	6.44×10^{-9}	8.33×10^2	5.09×10^{-23}	2.03×10^{-48}	0
F_4	0	0	1.51	0	0	0
F_5	8.88×10^{-16}	1.15×10^{-13}	7.05×10^{-2}	4.44×10^{-15}	8.88×10^{-16}	8.88×10^{-16}
F_6	0	0	1.08×10^{-2}	0	0	0

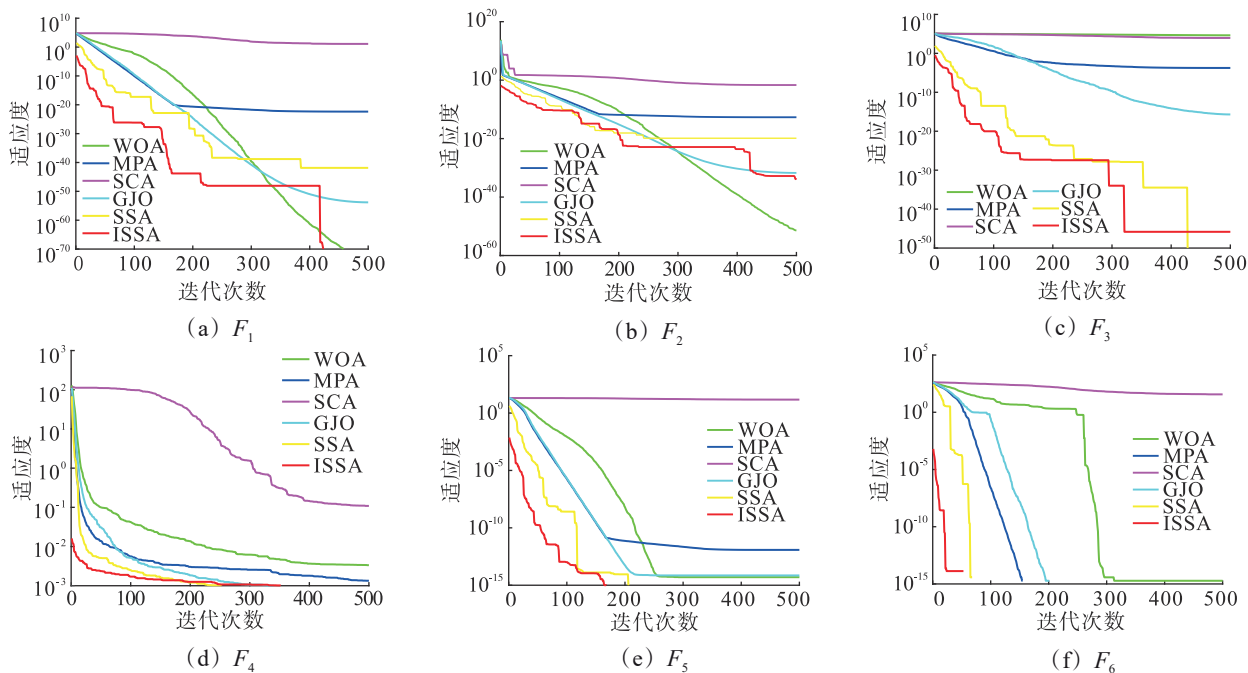


图 1 不同算法在 $F_1 \sim F_6$ 函数中的迭代曲线

Fig.1 iterative curves of different algorithms in $F_1 \sim F_6$ functions

3 变压器故障诊断模型

3.1 支持向量机

支持向量机是一种基于统计学习理论的二元分类模型，是机器学习方法的重要组成部分^[25]。它

得理论最优解，在 F_4 、 F_6 函数中 ISSA 与 SSA 均能稳定收敛至最优解，表明 ISSA 兼具优异的全局搜索和局部优化能力，能更有效地探索解空间，且稳定性更好，更接近理论最优值。不同算法在 $F_1 \sim F_6$ 函数中的迭代过程见图 1。由图 1 可以看出，ISSA 通过切比雪夫初始化和改进发现者机制，显著提升了优质个体比例，收敛最快且无早熟现象，寻优性能较好，对提升变压器故障诊断的准确性较有利，而其他算法均存在不同程度的早熟收敛问题。

通过以上分析可知，与 WOA、MPA、SCA、GJO、SSA 相比，ISSA 的局部极值逃避能力、整体择优协同能力和收敛性能更好，优势显著。

利用核函数将原始样本集映射到高维特征空间，从而构建最佳超平面。支持向量机在原始形式下的决策函数为

$$f(\mathbf{x}) = \text{sign}[\omega^T \varphi(\mathbf{x}_i) + b], \quad (13)$$

式中： $\mathbf{x}$ 为输入样本的特征向量； sign 为符号函数；

ω 为超平面法向量; $\varphi(\mathbf{x}_i)$ 为可将原始样本 $\mathbf{x}_i$ 映射到高维特征空间的映射函数; $\mathbf{b}$ 为分类阈值,可决定超平面的位置。

软边缘分类器是对原始决策函数的拓展,可适应更复杂的分类场景,其表达式为

$$g(\mathbf{x}) = \text{sign} \left[\sum_{i=1}^N \lambda_i y_i K(\mathbf{x}_i, \mathbf{x}_j) + \mathbf{b} \right], \quad (14)$$

式中: λ_i 为拉格朗日乘子,可反映支持向量对分类超平面的重要程度; y_i 为样本 $\mathbf{x}_i$ 的类别标签; $K(\mathbf{x}_i, \mathbf{x}_j)$ 为核函数,其表达式为

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp \left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2} \right), \quad (15)$$

式中, σ 为核函数参数。

3.2 ECOC编码策略

ECOC采用二值编码策略,结合机器学习方法与纠错输出编码,将 n 类问题分解为 $n(n-1)/2$ 个二分类子任务,可在传输过程中实现模型的多类识别和纠错功能^[26]。采用ECOC策略下的OVO(one versus one)编码来构建多类模型,将6类故障诊断转化为15个二元分类问题,见表7。其中,1表示正样本问题;-1表示负样本问题;0表示忽略该问题。在进行故障诊断时,分类器 C_i ($i=1,2,\dots,15$)首先生成测试样本的预测编码,再通过汉明距离匹配预设故障编码库,确定距离最小的编码组合,最终输出该编码组合所对应的故障类别。

表7 OVO编码
Tab.7 OVO encoding

分类器	T1	T2	PD	D1	D2	NC
C_1	1	-1	0	0	0	0
C_2	1	0	-1	0	0	0
C_3	1	0	0	-1	0	0
C_4	1	0	0	0	-1	0
C_5	1	0	0	0	0	-1
C_6	0	1	-1	0	0	0
C_7	0	1	0	-1	0	0
C_8	0	1	0	0	-1	0
C_9	0	1	0	0	0	-1
C_{10}	0	0	1	-1	0	0
C_{11}	0	0	1	0	-1	0
C_{12}	0	0	1	0	0	-1
C_{13}	0	0	0	1	-1	0
C_{14}	0	0	0	1	0	-1
C_{15}	0	0	0	0	1	-1

3.3 ECOC-ISSA-SVM(t-SNE)故障诊断模型

ECOC-ISSA-SVM(t-SNE)故障诊断模型通过

改进后的ISSA对SVM中的核参数 σ 、惩罚系数 C 进行优化,提升诊断精度。故障诊断流程图2。

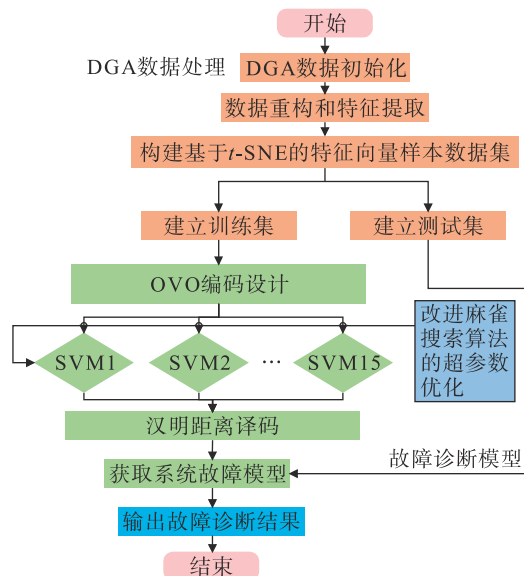


图2 故障诊断流程

Fig.2 fault diagnosis process

3.4 案例分析

(1) 与其他常见故障诊断模型比较

采用26维DGA特征数据集,对比ECOC-SVM与SVM、KNN(K邻近)、RF(随机森林)、ELM(极限学习机)的故障诊断效果。各模型参数设定如下:①SVM采用高斯核, $C=1.0$, $\gamma=1/d$, γ 为高斯核函数的参数,控制核函数的宽度, d 为特征数;②KNN采用欧氏距离, $k=5$, k 为进行分类或回归预测时,与待预测样本距离最近的邻居样本个数;③RF采用Gini分裂准则,包含100棵树;④ELM隐含层节点数为50,激活函数为Sigmoid。选取召回率、特异性、F1值以及诊断精度作为诊断效果评估指标。各模型故障诊断结果对比见表8。由表8可以看出,ECOC-SVM模型综合性能优势显著,召回率、特异性、F1值以及诊断精度分别高达93.5%、98.9%、91.9%和88.7%,在识别变压器放电故障等关键场景中表现较好。

表8 基础模型故障诊断结果对比

模型	召回率/%	特异性/%	F1值/%	诊断精度/%
SVM	85.7	96.9	84.2	83.3
KNN	87.1	97.3	86.6	82.2
RF	88.1	97.5	87.8	84.4
ELM	89.6	97.8	88.5	85.5
ECOC-SVM	93.5	98.9	91.9	88.7

(2) 与文献中的故障诊断模型对比

选取 BES-SVM、KPCA-TISOA-SVM、MISSA-BiLSTM 和 ECOC-ISSA-SVM 模型, 采用 26 维 DGA 特征数据集, 开展故障诊断效果对比实验, 验证 ECOC-ISSA-SVM (t -SNE) 模型的可靠性。模型故障诊断结果对比见表 9, 故障诊断时间见表 10。

表 9 模型故障诊断结果对比

Tab.9 comparison of model fault diagnosis results

模型	召回率/%	特异性/%	F1 值/%	诊断精度/%
BES-SVM	91.6	98.5	90.9	88.8
KPCA-TISOA-SVM	92.2	98.9	92.8	91.9
MISSA-BiLSTM	91.3	98.2	91.7	91.5
ECOC-ISSA-SVM	96.8	99.3	96.6	96.7
ECOC-ISSA-SVM(t -SNE)	97.8	99.6	97.8	95.6

表 10 故障诊断时间

Tab.10 fault diagnosis time

模型	最优值/ms	平均值/ms
ECOC-ISSA-SVM(t -SNE)	11	14
ECOC-ISSA-SVM	97	101

由表 9 可知, ECOC-ISSA-SVM (t -SNE) 整体诊断性能较好。在 6 类变压器故障诊断中, 召回率、特异性、F1 值和诊断精度分别高达 97.8%、99.6%、97.8% 和 95.6%, 但在局部放电和低能放电时存在误判。在处理 26 维 DGA 数据集时, 经 t -SNE 降维处理后, ECOC-ISSA-SVM (t -SNE) 的诊断精度略低于 ECOC-ISSA-SVM。由表 10 可知, ECOC-ISSA-SVM (t -SNE) 的平均诊断时间为 11 ms, 约为 ECOC-ISSA-SVM 诊断时间的 1/10, 表明 ECOC-ISSA-SVM (t -SNE) 能够在保证诊断精度的前提下, 显著提升处理速度, 满足实时诊断需求。

4 结论

(1) 通过切比雪夫混沌映射初始化种群, 融合樽海鞘算法的领导者更新策略及柯西-高斯变分策略, 构建 ISSA。ISSA 的全局搜索能力、收敛速度及抗早熟性能显著优于 WOA、MPA、SCA、GJO 和 SSA。在单峰函数 $F_1 \sim F_3$ 中可实现理论最优解的收敛, 在多峰函数 F_4 和 F_6 中均收敛至最优解, 在 F_5 中收敛精度提升明显, 可为 SVM 超参数的高效优化提供技术支持。

(2) 基于 t -SNE 算法的 ECOC-ISSA-SVM (t -SNE) 模型, 通过数据降维、超参数优化和多分类建模的技术路径, 在保留关键故障特征的前提下缩短诊断时间, 较好地平衡变压器故障诊断精度

与效率, 满足电力变压器实时监测需求。

参考文献(References):

- [1] 李斌, 白翔旭. 多策略改进 COA 算法优化 LSSVM 的变压器故障诊断研究[J]. 电工电能新技术, 2025, 44(4): 112-119.
LI Bin, BAI Xiangxu. Research on transformer fault diagnosis of LSSVM optimized by multi-strategy improved COA algorithm[J]. Advanced Technology of Electrical Engineering and Energy, 2025, 44(4): 112-119.
- [2] 谢国民, 齐晓亮. 多策略融合改进 AO 优化 SVM 的变压器故障诊断研究[J]. 控制工程, 2024, 31(11): 2000-2009.
XIE Guomin, QI Xiaoliang. Research on transformer fault diagnosis based on multi-strategy fusion and improved AO-optimized SVM [J]. Control Engineering of China, 2024, 31(11): 2000-2009.
- [3] 赵鑫, 王东丽, 彭泓, 等. 基于多策略改进蜣螂算法优化的变压器故障诊断[J]. 电力系统保护与控制, 2024, 52(6): 120-130.
ZHAO Xin, WANG Dongli, PENG Hong, et al. Transformer fault diagnosis based on a multi-strategy improved dung beetle optimizer [J]. Power System Protection and Control, 2024, 52(6): 120-130.
- [4] 赵超, 张迅, 王冕, 等. 基于改进随机森林的电力变压器试验和监测数据气体浓度预测[J]. 电测与仪表, 2024, 61(11): 205-210.
ZHAO Chao, ZHANG Xun, WANG Mian, et al. Gas concentration prediction of power transformer test and monitoring data based on improved random forest[J]. Electrical Measurement & Instrumentation, 2024, 61(11): 205-210.
- [5] 杨旭, 周文, 程林, 等. 基于 KPCA-SVM 的变压器多源信息融合故障诊断研究[J]. 高压电器, 2025, 61(2): 54-62.
YANG Xu, ZHOU Wen, CHENG Lin, et al. Research on multi-source information fusion fault diagnosis of transformer based on KPCA-SVM [J]. High Voltage Apparatus, 2025, 61(2): 54-62.
- [6] 王英洁, 曹铁男. 基于改进 Elman 神经网络的变压器故障诊断方法[J]. 沈阳工业大学学报, 2021, 43(3): 254-258.
WANG Yingjie, CAO Tienan. Transformer fault diagnosis method based on improved Elman neural network[J]. Journal of Shenyang University of Technology, 2021, 43(3): 254-258.
- [7] 刘迪迪, 王洋, 刘辉乾, 等. 基于 ADASYN 平衡化数据集的 POA-SVM 变压器故障诊断[J]. 电网与清洁能源, 2023, 39(8): 36-44.
LIU Didi, WANG Yang, LIU Huiqian, et al. POA-SVM transformer fault diagnosis based on ADASYN balanced data set[J]. Power System and Clean Energy, 2023, 39(8): 36-44.
- [8] 周晓华, 冯雨辰, 胡旭初, 等. 基于秃鹰搜索算法优化 SVM 的变压器故障诊断[J]. 南方电网技术, 2023, 17(6): 99-106, 116.
ZHOU Xiaohua, FENG Yuchen, HU Xuchu, et al. Transformer fault diagnosis based on SVM optimized by bald eagle search algorithm [J]. Southern Power System Technology, 2023, 17(6): 99-106, 116.
- [9] WU Y H, SUN X B, ZHANG Y, et al. A power transformer fault diagnosis method-based hybrid improved seagull optimization algorithm and support vector machine[J]. IEEE Access, 2021, 10: 17268-17286.
- [10] LI J F, HAI C, FENG Z, et al. A transformer fault diagnosis method based on parameters optimization of hybrid kernel extreme

- learning machine[J]. IEEE Access, 2021, 9: 126891-126902.
- [11] DING C, DING Q C, FENG L, et al. Prediction model of dissolved gas in transformer oil based on VMD-SMA-LSSVM[J]. IEEE Transactions on Electrical and Electronic Engineering, 2022, 17 (10): 1432-1440.
- [12] 李云渠, 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 Yunhao, XIAN Richang, ZHANG Haiqiang, et al. Fault diagnosis for power transformers based on improved grey wolf algorithm coupled with least squares support vector machine[J]. Power System Technology, 2023, 47(4): 1470-1478.
- [13] LIU Y, WU G N, GUO Y J, et al. A new noncontact detection method for assessing the aging state of composite insulators[J]. IEEE Transactions on Industrial Informatics, 2024, 20(4): 6802-6813.
- [14] 周萱, 吴伟丽. 基于改进 SMOTE 不均衡样本处理和 IHPO-DBN 的变压器故障诊断方法研究[J]. 电力系统保护与控制, 2024, 52 (11): 21-30.
- ZHOU Xuan, WU Weili. Transformer fault diagnosis method based on improved SMOTE unbalanced sample processing and IHPO-DBN[J]. Power System Protection and Control, 2024, 52(11): 21-30.
- [15] SHANG H K, LIU Z D, WEI Y L, et al. A novel fault diagnosis method for a power transformer based on multi-scale approximate entropy and optimized convolutional networks[J]. Entropy, 2024, 26 (3): 186.
- [16] 王宁, 郭梓昱, 田淑珂, 等. 基于融合特征 t -SNE 降维的控制图质量异常模式识别[J]. 系统工程理论与实践, 2024, 44(7): 2381-2393.
- WANG Ning, GUO Ziyu, TIAN Shuke, et al. Control chart quality abnormal pattern recognition based on hybrid feature t -SNE fusion [J]. Systems Engineering-Theory & Practice, 2024, 44(7): 2381-2393.
- [17] 王艳, 李伟, 赵洪山, 等. 基于油中溶解气体分析的 DBN-SSAELM 变压器故障诊断方法[J]. 电力系统保护与控制, 2023, 51(4): 32-42.
- WANG Yan, LI Wei, ZHAO Hongshan, et al. Transformer DGA fault diagnosis method based on DBN-SSAELM[J]. Power System Protection and Control, 2023, 51(4): 32-42.
- [18] 徐慧玲, 刘升, 李安东. 基于双向局部开发和黄金正弦的异构导向的鲸鱼优化算法[J]. 计算机工程与科学, 2024, 46(6): 1128-1140.
- XU Huiling, LIU Sheng, LI Andong. A heterogeneous guided whale optimization algorithm based on forward-reverse local exploitation and the golden sine algorithm[J]. Computer Engineering & Science, 2024, 46(6): 1128-1140.
- [19] 郁书好, 苏守宝. 混沌萤火虫优化算法的研究及应用[J]. 计算机科学与探索, 2014, 8(3): 352-358.
- YU Shuhao, SU Shoubao. Research and application of chaotic glowworm swarm optimization algorithm[J]. Journal of Frontiers of Computer Science and Technology, 2014, 8(3): 352-358.
- [20] 梁成龙, 陈志环. 基于混合策略改进的樽海鞘群算法[J]. 控制与决策, 2024, 39(8): 2541-2550.
- LIANG Chenglong, CHEN Zhihuan. Improved salp swarm algorithm based on hybrid strategy[J]. Control and Decision, 2024, 39(8): 2541-2550.
- [21] LI K W, LI S H, HUANG Z C, et al. Grey Wolf Optimization algorithm based on Cauchy-Gaussian mutation and improved search strategy[J]. Scientific Reports, 2022, 12: 18961.
- [22] 唐虎, 李霞, 王安阳, 等. 基于 WOA-VMD 算法的地铁钢轨波磨识别[J]. 噪声与振动控制, 2025, 45(3): 209-215.
- TANG Hu, LI Xia, WANG Anyang, et al. Rail corrugation identification method based on WOA-VMD algorithm[J]. Noise and Vibration Control, 2025, 45(3): 209-215.
- [23] 陈丽芳, 曹柯欣, 张思鹏, 等. 群智能优化算法最新进展[J]. 计算机工程与应用, 2024, 60(19): 46-67.
- CHEN Lifang, CAO Kexin, ZHANG Sipeng, et al. Recent progress of swarm intelligent optimization algorithms[J]. Computer Engineering and Applications, 2024, 60(19): 46-67.
- [24] 冯新, 杨雄, 曾豫豪. 多场景下基于 IGJO-DWA 算法的机器人路径规划[J]. 传感器与微系统, 2024, 43(11): 131-134, 138.
- FENG Xin, YANG Xiong, ZENG Yuhao. Robot path planning based on IGJO-DWA algorithm in multiple scenarios[J]. Transducer and Microsystem Technologies, 2024, 43(11): 131-134, 138.
- [25] 李耀华, 缙应祺, 潘成, 等. 基于 ISTOA-SVM 的发动机故障诊断方法研究[J]. 机械设计, 2025, 42(6): 154-164.
- LI Yaohua, GOU Yingqi, PAN Cheng, et al. Study on aeroengine fault-diagnosis method based on ISTOA-SVM[J]. Journal of Machine Design, 2025, 42(6): 154-164.
- [26] 李海, 田众, 钱君. 基于 ECOC 平衡随机森林的雷达降水粒子分类[J]. 系统工程与电子技术, 2024, 46(5): 1599-1606.
- LI Hai, TIAN Zhong, QIAN Jun. Hydrometeor classification for radar based on ECOC-balanced random forest[J]. Systems Engineering and Electronics, 2024, 46(5): 1599-1606.