

基于 LE-ISTOA-SVM 的聚合釜 化工过程故障诊断

高淑芝^{1†}, 范策²

(1. 沈阳化工大学 装备可靠性研究所, 辽宁 沈阳 110142; 2. 沈阳化工大学 信息工程学院, 辽宁 沈阳 110142)

摘要: 聚合釜是制备高分子化合物的最主要设备, 其能否稳定的运行关系到产品的质量以及人员、设备的安全。但是, 现代化工生产流程非常复杂, 化工过程诊断数据具有高维非线性的特点。提出了基于 LE-ISTOA-SVM 的聚合釜化工过程故障诊断方法。首先, 使用拉普拉斯特征映射算法 (Laplace Feature Mapping Algorithm, LE) 对故障数据进行降维。然后, 使用改进乌燕鸥优化算法 (Improved Sooty Tern Optimization Algorithm, ISTOA) 优化 SVM (Support Vector Machine, SVM) 的参数来提高其性能。最后, 利用聚合釜的实验数据做仿真测试。测试结果表明, 该方法有较好的诊断效果。

关键词: 故障诊断; 乌燕鸥优化算法; 聚合釜

中图分类号: TQ32

文献标识码: A

Fault Diagnosis of Polymerization Kettle Chemical Process Based on LE-ISTOA-SVM

GAO Shuzhi^{1†}, FAN Ce²

(1. Equipment Reliability Research Institute of Shenyang University of Chemical Technology, Shenyang, Liaoning 110142, China;
2. School of Information Engineering, Shenyang University of Chemical Technology, Shenyang, Liaoning 110142, China)

Abstract: Polymerization kettle is the most important equipment for preparing polymer compounds. Its stable operation is related to the quality of products and the safety of personnel and equipment. However, the modern chemical production process is very complex, and the chemical process diagnosis data has the characteristics of high-dimensional nonlinearity. This paper proposes a fault diagnosis method based on LE-ISTOA-SVM. Firstly, the Laplace feature mapping algorithm is used to reduce the dimension of fault data. Then, the improved sooty tern optimization algorithm is used to optimize the parameters of SVM to improve its performance. Finally, the experimental data of the polymerizer is used for simulation test. The test results show that the diagnosis effect of this method is good.

Key words: fault diagnosis; black tern optimization algorithm; polymerization kettle

聚合釜是制备高分子化合物的主要设备。聚合釜在运行过程中一旦发生故障, 就会危及设备和人身安全^[1-2]。准确且及时发现聚合釜出现的故障是很重要的问题。现代化工生产过程非常复杂, 所产生的数据具有高维非线性的特点, 这大大

增加了故障诊断的难度^[3-4]。针对化工过程数据的高维非线性的问题, 文献[5]提出了一种多尺度卷积神经网络与矩阵图相结合的故障诊断方法。文献[6]提出了一种基于动态扩展信念网络的故障诊断方法。深度神经网络虽然有诊断结果准确, 可

收稿日期: 2023-02-18

基金项目: 国家自然科学基金资助项目(52275156); 辽宁省教育厅重点(一般)资助项目(LJKZ0435); 国家重点研发计划资助项目(2019YFB2004400)

作者简介: 高淑芝(1968—), 女, 辽宁沈阳人, 博士研究生, 教授, 研究方向: 聚合釜化工过程智能控制。

† 通信联系人, E-mail: szg6868@126.com

有效处理海量、复杂数据的优点,但是,深度神经网络的训练需要大量的数据^[7]。文献[8]提出将 PCA 降维算法与 LVQ 神经网络相结合的故障诊断方法。针对化工过程过于复杂的问题,文献[9]提出先用小波分析对数据进行降噪,然后利用概率神经网络进行诊断的方法。文献[10]提出了一种基于径向基神经网络的故障诊断方法。但是,神经网络算法具有过度依赖初始值的选取,故障诊断结果不稳定的缺点^[11-13]。支持向量机是一种利用了数学方法的分类算法,在解决非线性数据方面有独特的优势^[14-15]。针对化工过程十分复杂,诊断难度大的问题,文献[16]提出了将网格巡优法、主元分析法和支持向量机相结合的故障诊断方法。文献[17]提出了一种 Zernike 矩阵与混沌算法优化的支持向量机相结合的诊断方法。针对化工过程中基于单变量的故障诊断问题,文献[18]提出了一种相空间重构技术与支持向量机相结合的故障诊断方法。参数选取的好坏对支持向量机的分类性能有较大的影响。而支持向量机参数的选取往往依靠经验,如果选取较差的参数会使支持向量机的分类性能较差。文献[19]提出了一种故障诊断方法,在使用多尺度模糊熵(MFE)提取风机轴承故障特征后,利用乌燕鸥优化算法(Sooty Tern Optimization Algorithm,STOA)对支持向量机(Support vector machine,SVM)的参数进行优化。然而,乌燕鸥优化算法的局部搜索能力比较差,这使得该算法容易出现局部极值点的问题且寻优速度较慢。本文提出了一种基于 LE-ISTOA-SVM 的聚合釜化工过程故障诊断方法,首先用拉普拉斯特征映射(Laplace Eigennaps,LE)对数据进行降维,然后利用改进乌燕鸥优化算法优化(Improved Sooty Tern Optimization Algorithm,ISTOA)的支持向量机作为诊断的模型。

1 拉普拉斯特征映射

拉普拉斯特征映射是一种流行学习方法,它主要作用是图拉普拉斯概念将高维数据进行降维^{[20][21]}。其主要思路是尽量减少有关系点在降维空间的距离,以保留原始数据信息。该算法的步骤如下:

步骤 1:将所有的点组成一个图,其中使用 KNN 算法连接这些点。

步骤 2:确定各个点之间的权重,计算公式如下:

$$w_{ij} = e^{-\frac{\|x_i - x_j\|_2^2}{2\sigma^2}} \quad (1)$$

步骤 3:计算拉普拉斯图矩阵 L 和它的特征值。

$$Ly = \lambda Dy \quad (2)$$

式中: D 表示对角矩阵($D_{ii} = \sum_{j=1}^n W_{ij}$), $L = D - W$ 。

LE 算法将降维问题转化为求矩阵的特征向量的问题。这使得计算更简单且运算速度更快。

2 改进乌燕鸥优化算法

2.1 乌燕鸥优化算法

乌燕鸥优化算法是一种新型的群智能优化算法,其模拟了乌燕鸥迁移和捕猎行为,这种算法全局搜索能力很强^{[22][23]}。

算法主要分为迁移行为和捕食行为两部分。

2.1.1 迁徙行为(全局搜索)

乌燕鸥群在迁移时会有不同起始的位置,以避免彼此之间的碰撞。同时,位置比较差的乌燕鸥也会朝着位置最好的方向移动。

避免碰撞:

$$\vec{C}_{st} = S_A \times \vec{P}_{st}(z) \quad (3)$$

式中: \vec{P}_{st} 表示乌燕鸥的旧位置; z 表示当前迭代次数; C_{st} 表示不与其他乌燕鸥发生碰撞的新位置; S_A 的更新公式为:

$$S_A = C_f - z \times (C_f / T_{max}) \quad (4)$$

式中: $C_f = 2$; T_{max} 表示最大迭代次数。

确定相对间距:

$$\vec{M}_{st} = C_B \times [\vec{P}_{bst}(z) - \vec{P}_{st}(z)] \quad (5)$$

其中: $\vec{P}_{bst}(z)$ 表示当前最好的位置; C_B 的引入可以提升全局搜索能力, $C_B = 0.5 \times s$, s 为 $[0, 1]$ 范围内的随机数; M_{st} 表示最佳位置的方位。向最优个体靠近:

$$\vec{D}_{st} = \vec{C}_{st} + \vec{M}_{st} \quad (6)$$

式中:表示乌燕鸥向最佳位置移动的轨迹。

2.1.2 捕食行为(局部搜索)

在锁定了猎物的位置后,乌燕鸥会不断地改变方向,以一种螺旋的方式发动进攻,其更新公式如下:

$$x' = R * \sin(i) \quad (7)$$

$$y' = R * \cos(i) \quad (8)$$

$$z' = R * i \quad (9)$$

$$P_{st}(z+1) = (D_{st} * (x' + y' + z')) * P_{bst}(z) \quad (10)$$

其中, R 表示每个螺旋的半径, $R = ue^{kv}$, u 和 v

为常数,其值都为1; i 表示 $[0, 2\pi]$ 范围内的随机数。

2.2 改进乌燕鸥优化算法

乌燕鸥优化算法虽然具有良好的全局搜索能力、结构简单等优点,但是局部搜索能力较差,在寻优过程中很容易出现局部最优解的情况。本文对其进行了如下改正。

2.2.1 改进局部搜索阶段

对乌燕鸥优化算法在局部搜索阶段的更新公式进行改进,引入轮盘赌的思想,随机选取一个数 r ($r \in [0, 1]$)。如果 $r > 0.5$,使用式(10)进行位置更新,否则采用式(11)作为位置更新。

$$P_{st}(z+1) = P_{st}(z) + \epsilon * \{\omega * [P_j(z) - P_k(z)] + (1 - \omega) * P_{bst}(z)\} \quad (11)$$

式中: ϵ, ω 表示 $[0, 1]$ 范围内的随机数; $P_j(z)$ 和 $P_k(z)$ 表示任意两只乌燕鸥的位置。

2.2.2 混沌扰动策略

乌燕鸥优化算法在迭代后期很容易出现局部极值的情况,而且一旦陷入局部最优值,算法将很难跳出。使用混沌扰动帮助算法摆脱局部极值点。Logistic映射、Tent映射等是比较常见的混沌映射。但因为Logistic映射很容易取到边缘值,而Tent映射又往往出现小周期或者不定周期的现象,会发生陷入不动点的情形;而Circle映射具有分布均匀而平稳的特性。故采用Circle映射对全局最优解进行扰动,计算公式如下:

$$x_{t+1} = \text{mod} \left[x_t + 0.2 + \frac{b}{2\pi} \sin(2\pi x_t), 1 \right] \quad (12)$$

$$P'_{st}(z) = P_{st}(z) * (1 + x_t) \quad (13)$$

式中, mod 函数为求余函数, $b = 0.5$ 。

用式(13)对适度最好的四分之一的种群进行扰动,如果新生成的种群的适应度好于原来的种群,则替换原来的种群。

2.2.3 边界策略

算法在寻优的过程中很容易出现解超出搜索范围的情况,为了避免这种情况的发生,本文提出了边界搜索策略。但是,乌燕鸥优化算法原有的边界搜索方式,同样会忽略对边界范围的探索。因此,改进了其边界搜索公式,以增强算法的边界寻优能力,计算公式如下:

$$P_{st} = \begin{cases} ub - (P_{st} - ub) * R / P_{st}, & \text{if } P_{st} > ub \\ lb + (lb - P_{st}) * R / P_{st}, & \text{if } P_{st} < lb \end{cases} \quad (14)$$

式中:搜索空间的范围为 $[lb, ub]$, R 是范围为

$[0, 1]$ 的随机数。

2.2.4 算法的流程

ISTOA算法的流程如下:

- (1)确定算法的参数;
- (2)设置各个乌燕鸥的初始位置;
- (3)计算所有乌燕鸥的适应度值,并进行排序,选出位置最好的乌燕鸥;
- (4)根据式(3)~式(6)进行全局搜索;
- (5)随机选取一个常数 r ,若 $r > 0.5$,则采用式(7)~式(10)进行局部搜索,否则,采用式(11)进行局部搜索;
- (6)根据式(12)、式(13),对位置最好的四分之一的乌燕鸥进行扰动。
- (7)判断是否满足迭代停止条件,若满足则停止迭代,否则返回步骤(3)。

算法流程如图1所示:

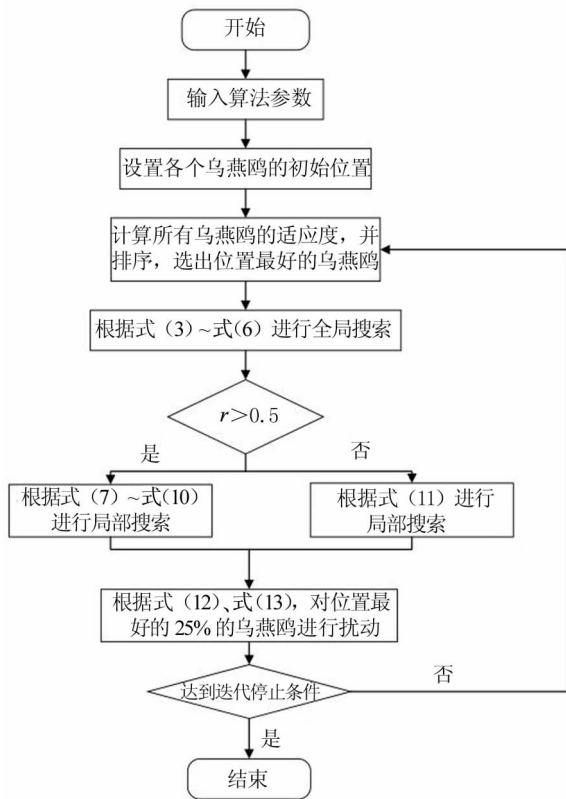


图1 ISTOA算法流程图

2.2.5 改进算法的性能测试

为了验证改进乌燕鸥优化算法的性能,我们让改进前后的两种算法对相同的测试函数进行寻优,以验证改进的效果。算法的种群数设定为50,最大迭代次数设定为500。如表1所示,对STOA算法的改进可以有效提升其寻优能力。测试的结果如图2所示,ISTOA算法的寻优能力和寻优效率都得到了提升。

表 1 算法的测试结果

函数	范围	指标	STOA	ISTOA
$f_1 = \sum_i^n x_i + \prod_{i=1}^n x_i $	$[-10, 10]$	最优值 平均值	0 1.56×10^{-193}	0 1.40×10^{-309}
$f_2 = \sum_{i=1}^n (x_i + 0.5)^2$	$[-100, 100]$	最优值 平均值	7.46 7.63	6.25 6.60
$f_3 = \sum_{i=1}^n i^2 x_i + rand[0, 1]$	$[-1.28, 1.28]$	最优值 平均值	4.07×10^{-5} 4.19×10^{-5}	3.91×10^{-6} 4.04×10^{-6}
$f_4 = \frac{1}{4000} \sum_{i=1}^n x_{2i} - \prod_{i=1}^n \cos(\frac{x_i}{\sqrt{i}}) + 1$	$[-600, 600]$	最优值 平均值	579.4 587.95	2.21×10^{-13} 2.28×10^{-13}
$f_5 = 4x_{21} - 2.1x_{21} + \frac{1}{3}x_{61} + x_{11}x_{22} - 4x_{22} + 4x_{42}$	$[-5, 5]$	最优值 平均值	1.08 1.14	0.34 0.38

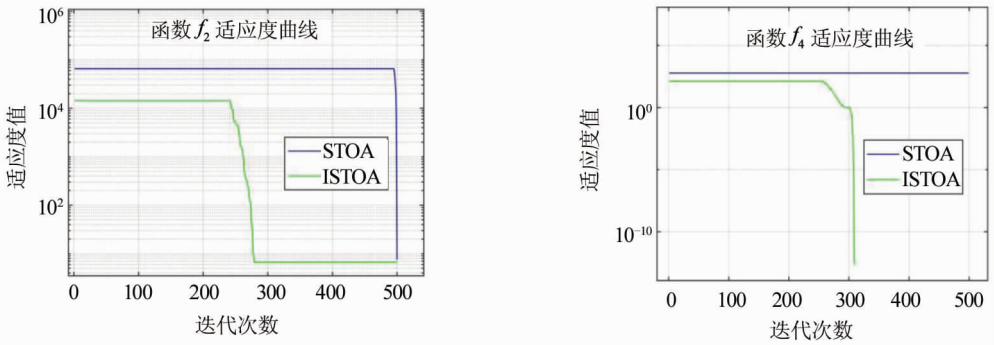


图 2 适应度曲线

3 LE-ISTOA-SVM 故障诊断模型

3.1 支持向量机

支持向量机是一种参考了数学中统计学思想的分类算法,具有优异的分类性能。主要分类过程是先将需要分类的数据映射到高维空间,然后找到一个将不同类型的数据分布在该平面两端的超平面。该算法属于线性分类器,但使用核函数替换其中的内核,则可以处理非线性问题。

3.2 故障诊断流程

- (1) 利用化工厂 70 m³ 聚合釜作为实验对象,并获取原始数据。
- (2) 利用拉普拉斯特征映射对数据降维,并将其分为训练数据集和测试数据集;
- (3) 输入 ISTOA 算法和 LE 算法的参数;
- (4) 利用 ISTOA 算法优化 SVM 分类器的参数;
- (5) 输出 SVM 的最优参数,并训练 SVM 分类器;
- (6) 将测试数据输入训练好的故障诊断模型。
- (7) 进行故障诊断,并输出诊断结果。

4 仿真实验

4.1 数据获取

以某化工厂 70 m³ 聚合釜为研究对象,选取不同状态下的运行的数据作为实验数据。选取 500 组数据作为训练数据,100 组数据作为测试数据,不同运行状态下的数据如表 2 所示。

聚合釜在运行过程中,如果发生故障,则多个因素发生相应变化。经过反复实验测试,将减速机的中轴振动值、搅拌电流值、密封压力值、操作压力值、搅拌转速、减速机温度、聚合釜操作温度、机械密封温度作为故障特征。聚合釜的故障类型如表 3 所示。

表 2 故障数据集

故障类型	样本数	标签
正常运行	160	1
电机故障	80	2
减速机故障	80	3
机封故障	80	4

4.2 实验的参数设置

实验设备的配置为:Window 10 64 位操作系统, i5-9300H @2.40 GHz, 软件平台为 MATLAB 2018b。改进乌燕鸥优化算法的种群规模设定为 50, 最大迭代次数设定为 70, 学习率设定为 0.01。SVM 的惩罚系数和核参数的范围设置为(0,30)。

4.3 测试结果分析

为了验证模型的有效性, 分别用鲸鱼优化算法(Whale Optimization Algorithm, WOA)、灰狼优化算法(Grey Wolf optimization algorithm, GWO)与乌燕鸥优化算法与支持向量机相结合的故障诊断模型, 与本文所提出的诊断模型进行对比测试。将 GWO-SVM、WOA-SVM、STOA-SVM、LE-ISTOA-SVM 四种故障诊断模型用相同的数据集进行训练和测试。从图 3 可以看到, 正常运行的状态下, LE-ISTOA-SVM 故障诊断的准确率最高,

GWO-SVM、WOA-SVM、STOA-SVM 三种模型的准确率都为 100%, 而 ISTOA-SVM 模型有一组数据诊断错误, 准确率为 95%。在诊断电机故障时, WOA-SVM 与 GWO-SVM 模型的准确率都为 85%, 而另外两种模型, 20 组测试数据, 仅有一组数据被误诊为正常状态, 准确率为 95%。在诊断减速机故障时, 诊断率最低。WOA-SVM 模型有 4 组数据诊断错误, 准确率为 80%。GWO 有 6 组数据诊断错误, 准确率为 70%。STOA-SVM 模型有 3 组数据诊断错误, 准确率为 85%, 而 LE-ISTOA-SVM 模型准确率为 100%。在诊断机封故障时, WOA-SVM 与 GWO-SVM 模型有 2 组数据诊断错误, 诊断准确率为 90%, 而另外两组只有 1 组数据诊断错误, 准确率为 95%。经过对比分析, LE-ISTOA-SVM 模型比其他几种诊断模型的诊断效果更好。

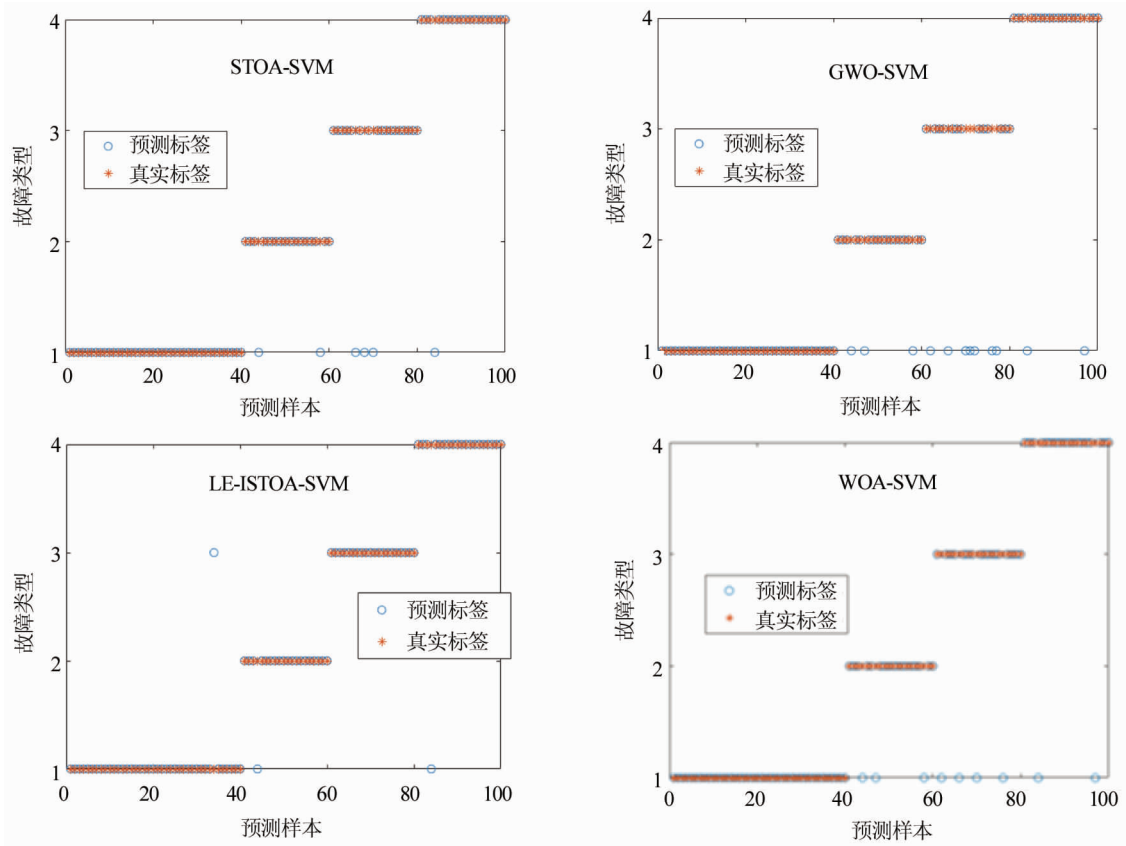


图3 几种故障诊断方法的比较

从表 3 可以看出, WOA-SVM 模型的诊断准确率为 90%, 运行时间为 11.67s。GWO-SVM 模型的诊断准确率为 88%, 运行时间为 15.21s。STOA-SVM 模型的诊断准确率为 93%, 运行时间为 9.64 s。而 LE-ISTOA-SVM 模型的诊断效果

最好, 诊断准确率为 97%, 运行时间为 5.12s。由此可以看出, 拉普拉斯映射结合改进乌燕鸥优化算法优化的支持向量机, 不但诊断准确率最高, 而且诊断效率也优于其他几种方法。

表 3 故障诊断结果

模型	惩罚因子	核函数参数	运行时间/s	准确率
WOA-SVM	30.373	14.553	11.67	90%
GWO-SVM	19.958	21.626	15.21	88%
STOA-SVM	21.708	17.792	9.64	93%
LE-ISTOA-SVM	10.017	5.134	5.12	97%

5 结 论

提出了一种 LE-ISTOA-SVM 的故障诊断方法,并用于判断聚合釜生产聚氯乙烯时是否处于正常运行状态。首先,使用拉普拉斯特征映射算法对数据降维。然后用改进乌燕鸥优化算法优化支持向量机的参数。对乌燕鸥算法的改进:首先,改进算法的局部搜索策略,然后利用 Circle 映射对种群中适应最好的四分之一进行混沌扰动,最后改进边界搜索机制。测试结果表明,本文所提的方法具有较好的故障诊断效果。

参考文献

- [1] 高淑芝,赵娜. 基于 DKPCA 的聚合釜故障诊断研究[J]. 沈阳化工大学学报, 2015, 29(2): 178-182.
- [2] 夏永彬. 聚合釜粗糙集及神经网络故障诊断研究[D]. 秦皇岛:燕山大学, 2018.
- [3] KANG Jialin. Visualization analysis for fault diagnosis in chemical processes using recurrent neural networks[J]. Journal of the Taiwan Institute of Chemical Engineers, 2020, 112: 137-151.
- [4] PAN Zhuofu, WANG Yalin, YUAN Xiaofeng, et al. A classification-driven neuron-grouped SAE for feature representation and its application to fault classification in chemical processes[J]. Knowledge-Based Systems, 2021.
- [5] SONG Qiusheng, PENG Jiang. A multi-scale convolutional neural network based fault diagnosis model for complex chemical processes[J]. Transactions of The Institution of Chemical Engineers. Process Safety and Environmental Protection, Part B, 2022, 159: 575-584.
- [6] WANG Yalin, PAN Zhuofu, YUAN Xiaofeng, et al. A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network[J]. ISA Transactions, 2020, 96: 457-467.
- [7] 衷路生,夏相明. 基于深度残差网络的化工过程故障诊断[J]. 过程工程学报, 2020, 20(12): 8.
- [8] 谭莉,于春梅. 基于 PCA-LVQ 神经网络的化工过程故障诊断[J]. 工业控制计算机, 2016, 29(11): 86-87.
- [9] 谷雷,杨青,王大志. 基于小波分析与概率神经网络的化工过程故障诊断[J]. 工业仪表与自动化装置, 2008(3): 4.
- [10] 谷雷. 基于人工神经网络的化工过程的故障诊断[D]. 沈阳:沈阳理工大学, 2008.
- [11] 陈剑雪. ACO-BP 算法在化工过程故障诊断中的应用[J]. 化工自动化及仪表, 2012(7): 4.
- [12] 李国友,张新魁,才士文,等. 基于改进 KFDA 与 DE 优化 SOM 的故障诊断模型及其化工过程诊断[J]. 化工进展, 2022(4): 041.
- [13] 张卫华,吴重光,王春利,等. 基于神经网络的化工过程故障诊断[J]. 计算机与应用化学, 2010(7): 5.
- [14] 李国友,杨梦琪,杭丙鹏,等. 基于模糊粗糙集和鲸鱼优化支持向量机的化工过程故障诊断[J]. 振动与冲击, 2022, 41(2): 8.
- [15] 冀丰恩,余云松,张早校. LDA_SVM 方法在化工过程故障诊断中的应用[J]. 高校化学工程学报, 2020, 34(2): 8.
- [16] 高欣. 基于支持向量机的田纳西化工过程故障诊断研究[D]. 锦州:渤海大学, 2016.
- [17] ZHENG Nan, ZHANG Guoying, ZHANG Yang, et al. Brain tumor diagnosis based on Zernike moments and support vector machine optimized by chaotic arithmetic optimization algorithm[J]. Biomedical Signal Processing and Control, 2023, 82: 104543.
- [18] 郑尧,王奎升,王耘. 基于 PSR 和 SVM 方法的化工过程单变量故障诊断[J]. 北京化工大学学报(自然科学版), 2008(1): 93-96.
- [19] 汤占军,孙润发. 基于多尺度模糊熵和 STOA-SVM 的风机轴承故障诊断[J]. 电机与控制应用, 2021, 48(12): 5.
- [20] 韩保金,曾岳. 基于拉普拉斯特征映射的免疫算法[J]. 哈尔滨商业大学学报(自然科学版), 2019, 35(3): 316-321.
- [21] 孙伟伟,刘春,李巍岳. 联合改进拉普拉斯特征映射和 k-近邻分类器的高光谱影像分类[J]. 武汉大学学报(信息科学版), 2015, 40(9): 1151-1156.
- [22] 孙珂琪,陈永峰. Lévy 飞行的正余弦乌燕鸥混合算法及应用[J]. 机械设计与制造, 2023(1): 212-217.
- [23] 乔夏君,薛薇,王浩,等. 混沌乌燕鸥算法优化发动机参数自整定 PID 控制[J]. 计算机测量与控制, 2022, 30(6): 132-137.