

DOI:10.19651/j.cnki.emt.2209795

基于邻域降噪正交自编码器的工业过程故障检测^{*}

郭小萍 张志朋 李元

(沈阳化工大学信息工程学院 沈阳 110142)

摘要: 针对采用自编码器提取过程特征进行故障检测时,没有考虑数据的局部结构信息,提出邻域降噪正交自编码器(NDQAE)的方法。邻域保持嵌入算法提取数据的邻域信息作为权重对过程数据进行加权,强化数据局部结构信息。正交自编码器进一步提取带有局部信息加权的过数据非线性特征。通过加入噪声增强自编码器的鲁棒性,并采用反向传播算法训练网络参数,获得能够捕捉数据局部特性和全局特性的鲁棒自编码器模型。在该模型的隐特征和重构残差空间分别构建 T^2 和 SPE 统计量,并计算统计量控制限用于故障检测。在田纳西-伊斯曼(TE)化工过程和三相流过程进行仿真实验,结果表明了所提算法的有效性。

关键词: 自编码器;邻域保持嵌入;故障检测;多相流过程

中图分类号: TP277 **文献标识码:** A **国家标准学科分类代码:** 510.8060

Neighborhood denoising quadrature autoencoder based fault detection for industrial process

Guo Xiaoping Zhang Zhipeng Li Yuan

(School of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China)

Abstract: Aiming at using autoencoder to extract process features for fault detection, the local structure information of data is not considered, a method of neighborhood denoising quadrature autoencoder is proposed. The neighborhood preservation embedding algorithm extracts the neighborhood information of the data as a weight to weight the process data and strengthen the local structure information of the data. The orthogonal autoencoder further extracts the nonlinear features of the process data with local information weighting. The robustness of the autoencoder is enhanced by adding noise, and the network parameters are trained by the back-propagation algorithm to obtain a robust autoencoder model that can capture the local and global characteristics of the data. T^2 and SPE statistics are constructed in the latent feature and reconstructed residual space of the model, respectively, and the statistical control limits are calculated for fault detection. Simulation experiments are carried out on the Tennessee-Eastman process and the three-phase flow process, and the results show the effectiveness of the proposed algorithm.

Keywords: autoencoder; neighborhood preserving embedding; fault detection; multiphase flow process

0 引言

近年来,随着电子信息技术的飞速发展和综合自动化系统的不断普及,工业过程日趋于规模化和复杂化。运行过程中的任何异常都可能引发故障,造成无法估计的损失,这也使得人们对系统的安全性和可靠性提出了更高的要求^[1]。

数据驱动的多元统计过程检测方法已广泛用于各种工业过程^[2]。主成分分析(PCA)是应用最广泛的方法之一,但其是一种线性算法,难以处理不同过程变量之间的非线

性关系^[3]。文献[4]提出基于核主元分析(kernel principal component analysis, KPCA)残差方向梯度的故障检测方法,其具有较好的检测能力。但忽略了流形的非线性几何结构,且在应用之前必须定义核函数的参数,这限制了KPCA的性能。流形学习作为一种非线性方法,能有效地揭示高维数据中的非线性结构^[5]。目前流形学习主要有局部线性嵌入(locally linear embedding, LLE)、等距特征映射(ISO-MAP)等^[6-7]。经典的流形学习方法虽能够学习非线性流形结构,但对新样本的泛化能力较差。邻域保持嵌入(neighborhood preserving embedding, NPE)算法作为

收稿日期:2022-04-29

^{*} 基金项目:国家自然科学基金(61490701,61673279)、辽宁省教育厅项目(LJ2020021)资助

LLE 的线性近似,能够解决样本外的泛化问题^[8]。但其局部子空间中的保持向量并非正交基向量,这会导致局部空间的流形失真。

近年来,基于神经网络的方法在非线性过程检测中表现出优越的效果^[9]。源于神经网络的深度学习能够对复杂的数据进行建模,挖掘数据中隐藏特征^[10]。其中,自编码器(autoencoder, AE)模型应用较广^[11]。但是在训练的过程中,非常容易出现过拟合,并且优化过程容易陷入局部最优。文献[12]提出了一种鲁棒自编码器模型(robust autoencoder, RAE),提高了模型鲁棒性。文献[13]在 RAE 基础上引入正交性限制,提出降噪正交自编码器(denoising quadrature autoencoder, DQAE)用于故障检测,解决自编码器算法的过拟合问题,并提升对关键特征的提取和学习能力。但该算法面向数据全局提取特征,难以有效地保留数据的局部结构信息。

本文提出一种基于邻域降噪正交自编码器(NDQAE)的故障检测方法。首先,通过 NPE 算法提取过程数据的邻域信息,作为权重加权数据,突出数据的局部信息。然后,采用正交自编码器进一步提取加权数据的非线性特征。加入噪声增强自编码器的鲁棒性,通过反向传播算法,进行网络参数训练,获得能够捕捉数据局部结构的鲁棒自编码器模型。采用该模型提取的隐特征和重构残差构建 T^2 和 SPE 统计量,通过核密度估计法确定统计量控制限用于过程故障检测。TE 过程和三相流过程的仿真验证实验结果说明了本文算法的有效性,并与 PCA、AE 及 DQAE 算法进行对比。

1 邻域保持嵌入算法

NPE 的思想主要是将高维数据 $X = [x_1, x_2, \dots, x_n] \in R^{m \times n}$ 映射到低维空间 $Y = [y_1, y_2, \dots, y_n] \in R^{p \times n}$ ($p < m$) 中, p 为低维空间维数,并保持数据的局部结构不变,这里的局部结构是通过重构权重矩阵来表征的。算法具体步骤如下^[14]:

1) 利用 K 近邻算法来构建邻接图。

2) 计算权重矩阵。令权重矩阵为 C , C 通过最小化目标函数得到:

$$\Phi(C) = \min \sum_{i=1}^n \|x_i - \sum_{j \in Q(i)} C_{ij} x_j\|^2 \quad (1)$$

式中: C_{ij} 表示节点 i 和节点 j 之间的权重值,若 i 和 j 不相邻,则 $C_{ij} = 0$; Q_i 为样本 x_i 的 k 近邻的集合;并且 $\sum_j C_{ij} = 1, j = 1, 2, \dots, n$ 。

3) 计算特征映射矩阵。根据高维空间中的样本点与其近邻点之间的权重矩阵 C , 通过求解以下优化问题来计算最终嵌入的低维空间:

$$\Phi(Y) = \min \sum_{i=1}^n \|y_i - \sum_{j=1}^n C_{ij} y_j\|^2 \quad (2)$$

式中: y_i 为低维空间样本点, y_j 为样本 y_i 的第 j 个近邻。

通过拉格朗日函数法可将优化问题转化为广义特征值求解问题:

$$X^T M X a = \lambda X^T X a \quad (3)$$

式中: $M = (I - C)^T (I - C)$, $Y^T Y = I$, 求解式(3)中最小的 p 个特征值所对应的特征向量组成映射矩阵 $A = [a_1, a_2, \dots, a_p] \in R^{m \times p}$ 。

2 邻域降噪正交自编码器的故障检测

2.1 邻域降噪正交自编码器

自编码器是一种无监督学习算法^[15]。AE 与有监督神经网络不同,其输入和输出有着相同数量的节点,其目的在于最小化输入和输出之间的重建误差。它由编码器 ϕ 和解码器 ψ 组成,编码器用于将输入数据 X 映射到低维特征空间,而解码器将低维特征空间中数据 h 映射为重构后的数据 Y 。对 AE 进行改进,本文提出的 NDQAE 网络结构如图 1 所示。

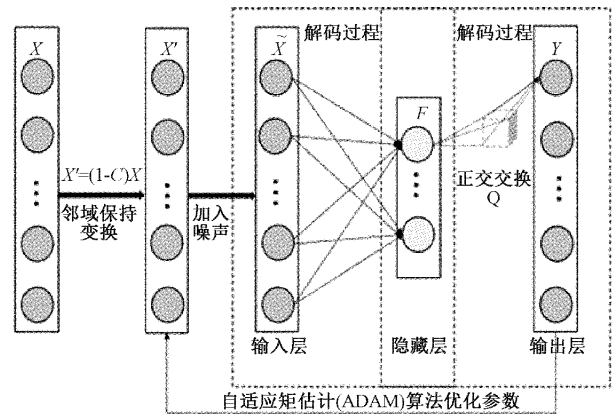


图 1 NDQAE 网络结构示意图

通过 NPE 算法构造一个表征局部结构的权重矩阵,对自编码器模型的输入数据进行变换,在自编码器模型中以回归的形式表示 NPE 的目标函数。令:

$$X' = (I - C) X \quad (4)$$

式中: C 为 NPE 算法中求得的权重矩阵, $I = \text{diag}(d_{11}, \dots, d_{NN})$ 。

设 $G = (I - C)$, 则邻域自编码器的优化问题可写成如下形式:

$$[W, b; W', b'] = \arg \min_{W, b, W', b'} \sum_{i=1}^N \|x'_i - \sigma'(W [\sigma(W x'_i + b)] + b')\|^2 \quad (5)$$

式中: N 为样本数, σ 是编码器的激活函数, W, b 是编码器权重矩阵和偏差向量; σ' 是解码器的激活函数, W', b' 是解码器权重矩阵和偏差向量。 $x'_i = G_i X$, G_i 是矩阵 G 的第 i 行数据。

通过优化参数来学习非线性映射,邻域自编码器的思想是重构数据近似包含局部流形结构的变换数据 x'_i 。

对邻域自编码器模型加上正交性限制。使用正交变换

Q 来代替非线性变换 ϕ , 即利用非线性变换 ϕ 作为编码器来提取特征, 然后采用正交变换 Q 将特征解码到原始数据空间, 优化问题转换成以下形式:

$$[W, b; Q] = \arg \min_{W, b, Q} \sum_{i=1}^N \|x'_i - f(x'_i; W, b) Q^T\|^2 \quad (6)$$

式中: x'_i 为包含局部流形结构的变换数据, $f(x'_i; W, b)$ 是 p ($p < m$) 维的潜在特征, 正交矩阵 $Q \in R^{m \times p}$ 满足约束条件 $Q^T Q = I_{p \times p}$ 。

加入正交性限制后, 再对该模型加权后的建模数据加入噪声来训练。降噪自编码器使用加入噪声的数据训练 DAE 网络以恢复真实输入, 具有较好的鲁棒性^[16]。因此本文所提 NDQAE 方法的优化问题如式(7):

$$[W, b, Q] = \arg \min_{W, b, Q} \sum_{i=1}^N \|x'_i - f(\tilde{x}_i; W, b) Q^T\|^2$$

$$\text{s. t. } Q^T Q = I_{p \times p} \quad (7)$$

其中, \tilde{x}_i 为包含噪声的数据, $\tilde{x}_i = x'_i + \epsilon_i$, ϵ_i 表示加入的高斯噪声。可以将式(7)写成以下形式:

$$[W, b, Q] = \arg \min_{W, b, Q} \|X' - F(\tilde{X}; W, b) Q^T\|^2$$

$$\text{s. t. } Q^T Q = I_{p \times p} \quad (8)$$

其中, $F \in R^{N \times p}$ 是建模数据的潜在特征。

2.2 故障检测统计量

Hotelling T^2 和 SPE 统计量对于生产过程的故障检测具有非常重要的作用。在特征子空间, 基于潜在特征 F , 给出 T^2 统计量的定义如下:

$$T_i^2 = f_i \sum_f f_i^T \quad (9)$$

其中, $f_i = f(\tilde{x}_i; W, b)$, ($i = 1, 2, \dots, N$), \sum_f 是 F 相关的协方差矩阵。

在残差子空间, 基于输入的重构误差, SPE 统计量的定义如下:

$$SPE_i = \|x'_i - f_i Q^T\|^2 \quad (10)$$

由于构造的统计量很难得到准确的概率分布, 本文利用核密度估计(kernel density estimation, KDE)法近似计算两个统计量的控制限^[17]。

2.3 基于 NDQAE 的故障检测方法

该方法主要包括离线建模和在线检测两部分, 流程如图 2 所示。

离线建模主要步骤:

- 1) 收集正常过程数据作为训练数据, 对训练数据进行 z-score 标准化处理;
- 2) 通过式(4)将训练数据加权变换为 X' , 再对其加入噪声生成数据 \tilde{X} ;
- 3) 初始化 NDQAE 模型参数和正交矩阵 Q ;
- 4) 根据目标函数来训练模型, 优化参数 W, b ;
- 5) 固定 W 和 b , 计算 Q ;
- 6) 用模型学习出的潜在特征 F 和重构残差空间, 计算 T^2 、SPE 统计量;

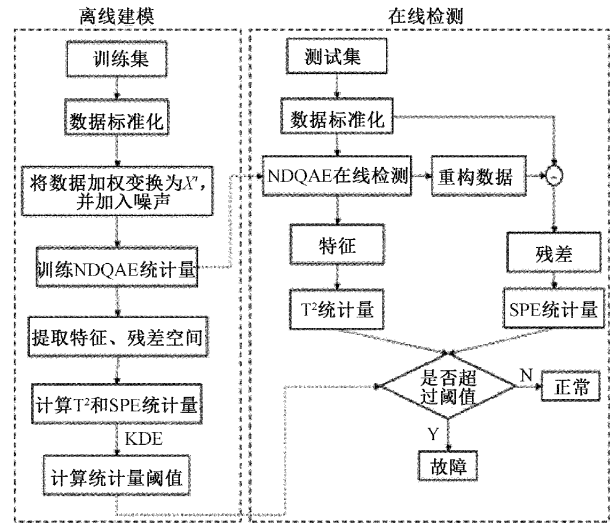


图 2 NDQAE 算法故障检测流程

7) 用 KDE 法计算统计量控制限。

在线故障检测步骤:

- 1) 获取测试数据, 用训练数据标准化的参数对其标准化;
- 2) 用离线建模的 NDQAE 模型提取测试数据潜在特征;
- 3) 根据提取的特征和残差分别计算测试数据的 T^2 、SPE 统计量;
- 4) 若统计量超过控制限, 则判定为发生故障; 若没有超过控制限, 则为正常。

3 实验验证

3.1 TE 过程

TE 过程仿真应用平台是一个基于实际工业过程的仿真例子, 用于测试过程系统工程中故障检测和诊断方法的有效性^[18]。其过程主要包括 5 个操作单元: 反应器、冷凝器、循环压缩机、气液分离器以及汽提塔。

TE 过程数据包括正常数据集与 21 个故障数据集, 每个数据集都包含 52 个变量。利用在正常操作条件下获得 500 个样本的数据作为训练数据建立模型, 21 个不同的过程故障数据作为测试数据, 每个故障数据包含 960 个样本, 这些故障在第 161 个样本处引入, 即前 160 个样本是正常的, 从第 161 开始到结束存在故障。

为了验证所提方法的有效性, 与 PCA、AE 和 DQAE 方法进行检测对比。根据主元贡献率 85% 确定主元个数为 27, 4 种方法均使用相同的主元个数。NDQAE 算法中近邻个数 k 设置为 33, 学习率设为 0.01, 检测统计量控制限置信度设置为 99%。NDQAE 模型输入层节点数为 52, 隐藏层节点数为 27, 输出层节点数为 52。

故障 4 的数据和 NDQAE 方法所提取的特征在三维空间中的可视化结果如图 3(a) 和 (b) 所示。在这些图中, 三

角形点表示正常数据,圆点代表故障数据,从图中可以看出,故障4数据的三维投影难以分辨正常数据和故障数据;而经过 NDQAE 模型提取特征后,正常数据和故障数据可以较好地分开。

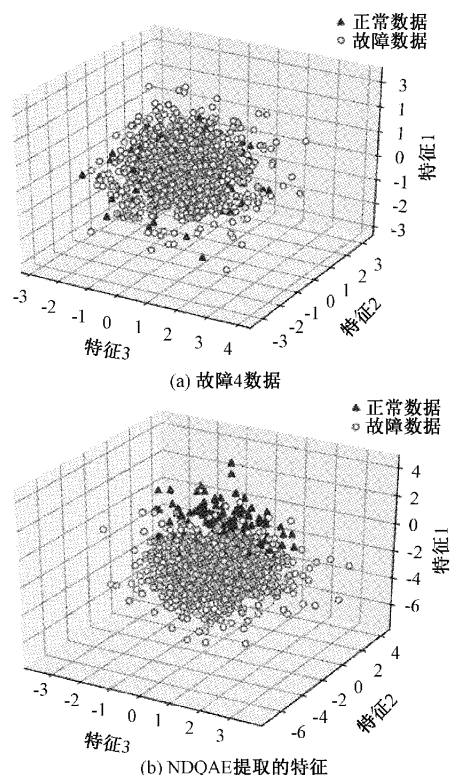


图3 故障4的特征在三维空间的可视化

表1为4种方法对TE过程21种故障的 T^2 和SPE统计量的故障检测结果。对于每一个故障,检测性能最佳的在表中用粗体表示。故障3、9、15和19在众多文献中难以检测出来,在本文的实验中同样得到证实。对于 T^2 和SPE统计量的检测结果,NDQAE在16个故障中检测性能最佳,特别是故障4、10和20,基于NDQAE方法的故障检测率比DQAE方法高出很多。

各种方法对故障4的详细检测结果如图4所示。故障4是TE过程的反应器冷却入口温度改变而产生的阶跃故障。从图4可以看出,所提NDQAE方法的SPE统计量的变化比其他方法统计量的变化要显著得多。在SPE统计量中,NDQAE的检测率远高于其他方法(如表1),可以发现局部几何信息对故障4的故障检测非常关键。

由表1可知,所提方法的SPE统计量在TE过程的大多数故障中提供了最佳的故障检测性能。基于局部几何结构的设计,在某些故障中的检测效果得到了非常明显的提升,说明了由该算法投影得到的特征空间包含了更多有效的信息,具有更好的特征提取能力,使得其具有更好的故障检测性能。

3.2 多相流过程

克兰菲尔德大学的多相流设备是设计用于监控水、油

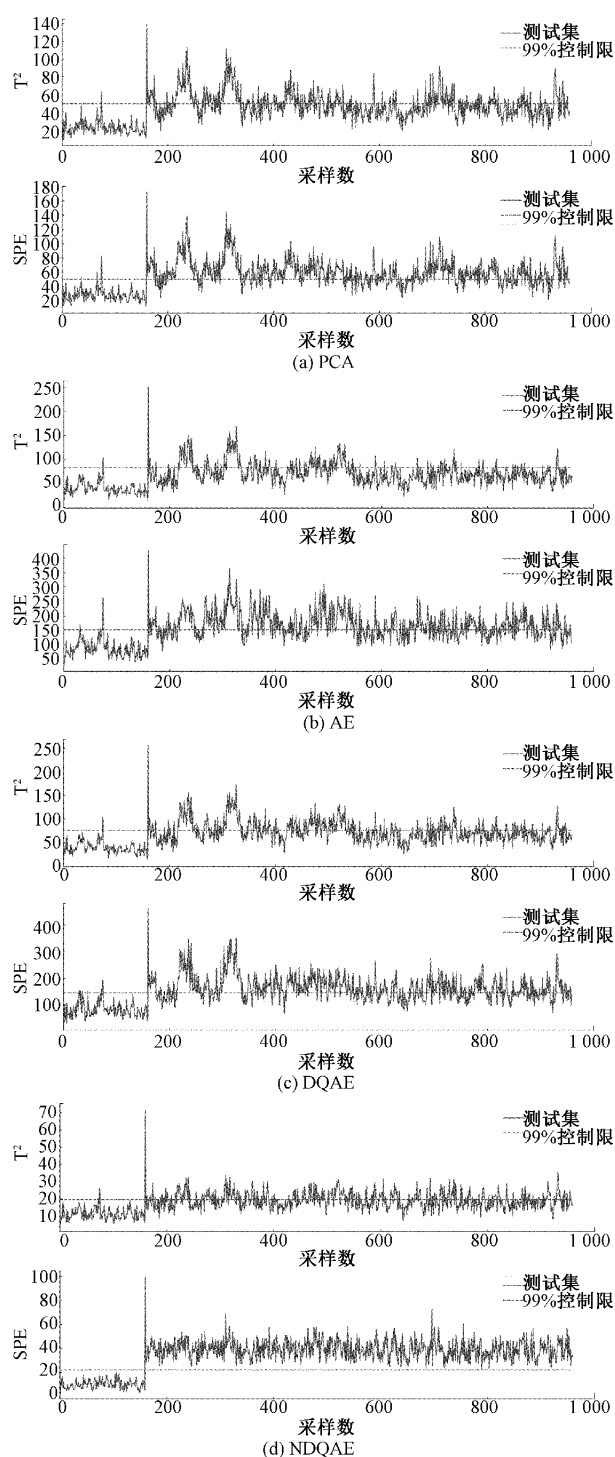


图4 故障4的检测结果图

和空气进入加压系统时的流量^[19]。其试验区由管道和气液两相分离器组成,它能够以要求的速率供应单相空气、水和油,或这些液体的混合物,最后在地面的水平的三相分离器中流体混合物被分离。空气排放到大气中,油和水的乳状液在各自的聚结器中分离,然后传送至各自的储罐。

多相流过程数据包括正常数据(T_1 、 T_2 、 T_3)与6种故

表 1 TE 过程 21 种故障 T² 和 SPE 统计量的故障检测率

%

Fault No.	PCA		AE		DQAE		NDQAE	
	T ²	SPE	T ²	SPE	T ²	SPE	T ²	SPE
1	99.25	99.00	99.50	99.25	99.50	99.25	99.25	99.50
2	98.37	98.62	98.50	98.00	98.25	98.37	98.25	98.65
3	5.87	21.87	3.60	0.10	1.60	0.25	13.25	15.00
4	43.15	72.62	30.87	59.87	43.75	64.87	43.75	100.00
5	27.87	40.75	26.70	22.10	23.00	23.75	32.62	31.87
6	99.12	99.65	99.00	98.87	99.25	99.50	99.62	100.00
7	100.00	100.00	100.00	69.90	100.00	100.00	100.00	100.00
8	97.75	98.00	97.50	95.62	96.37	97.25	98.50	98.25
9	4.75	17.87	4.00	0.62	0.37	0.37	12.88	15.25
10	45.50	57.37	49.50	43.25	49.25	50.87	62.87	63.12
11	54.50	72.50	50.40	28.10	39.62	35.12	56.50	55.12
12	98.25	98.75	100.00	97.37	97.87	97.00	99.25	99.12
13	94.37	95.12	94.25	94.12	94.00	94.25	94.62	95.00
14	99.87	100.00	99.62	95.12	97.87	100.00	99.75	100.00
15	9.62	11.37	5.70	1.90	2.62	2.00	20.50	10.87
16	31.75	50.12	17.90	9.00	16.00	14.5	44.25	44.62
17	84.62	92.00	79.90	72.60	74.12	83.70	82.62	86.37
18	89.25	91.25	89.00	88.37	90.12	89.37	89.50	89.12
19	2.50	19.37	8.20	0.20	0.00	0.62	3.62	11.87
20	44.25	58.25	28.75	36.62	42.37	45.37	62.25	60.25
21	34.75	42.12	34.40	20.40	35.20	25.25	34.37	44.00

障数据。单个数据集可能无法涵盖所有可能的正常情况，为了获得更丰富和可靠的训练集，将这 3 个数据集两两混合，生成 3 个不同的组合集，在监测期间使用剩余数据集时，观察哪种训练数据集组合会产生较低的误报率，结果证明 T2、T3 组合的数据集具有较好的效果^[20]。本文中也将数据集 T2、T3 进行组合作为训练集。该多相流模型中的故障 1、2 和 4 是微小故障，在故障引入后的一段时间内对整个过程行为的影响不大，故难以进行故障检测。采用典型故障 3、5 和 6 作为测试数据用于模型测试。故障 3 包

括 3 种条件的数据(命名 3_1,3_2,3_3);故障 5 包括 2 种条件的数据(命名 5_1,5_2);故障 6 包括 2 种条件的数据(命名 6_1,6_2)。每个数据集都包括 24 个过程变量。

将所提方法与 PCA 和 DQAE 方法对多相流过程的故障 3、5 和 6 进行检测对比。根据主元贡献率 85% 确定主元个数为 5,NDQAE 中学习率设为 0.01,检测统计量控制限置信度设置为 99%。NDQAE 模型输入层节点数为 24,隐藏层节点数为 5,输出层节点数为 24。3 种方法对多相流过程故障 3、5 和 6 的检测结果汇总在表 2 中。

表 2 多相流过程中 PCA、DQAE 和 NDQAE 的故障检测率(%)和误报率(%) (括号内)

故障类别	PCA		DQAE		NDQAE	
	T ²	SPE	T ²	SPE	T ²	SPE
3_1	97.37(15.95)	99.72(32.97)	97.74(9.40)	98.58(8.53)	97.82(12.25)	98.59(4.03)
3_2	81.41(51.77)	OT(OT)	OT(OT)	OT(OT)	OT(OT)	OT(OT)
3_3	97.22(22.29)	99.65(12.31)	97.85(15.67)	99.03(10.65)	98.28(13.25)	98.93(3.52)
5_1	70.63(70.08)	OT(OT)	OT(OT)	OT(OT)	OT(OT)	OT(OT)
5_2	83.34(11.47)	89.41(22.53)	79.89(4.65)	70.39(4.12)	79.38(6.44)	85.86(4.76)
6_1	99.72(1.27)	99.81(5.97)	99.72(0.00)	99.81(0.00)	99.72(0.00)	99.92(0.00)
6_2	99.86(0.00)	99.94(0.28)	99.86(0.00)	99.92(0.00)	100.00(0.00)	99.95(0.00)

在故障引入之前,就已经超过控制限的情况记为“OT”,检测性能最佳的在表中用粗体表示。由表2可以看出,PCA方法不适用于故障3和5, T^2 和SPE统计量的误报率都非常高。所提方法对于这几种故障SPE统计量的检测率高于DQAE方法,并且有着较低的误报率。本文所提方法的SPE统计量在多相流过程的故障3、5和6中表现出了最佳的故障检测性能,说明通过保留过程数据的局部结构信息,增强了模型的有效性。

4 结 论

本文提出了一种NDQAE的方法,并将其用于工业过程的故障检测中。该方法将过程数据局部结构信息的保留和非线性特征的提取相结合。解决了传统自编码器在特征提取中不能保留数据局部结构信息的问题,能够有效地捕捉数据中固有的局部结构信息,构建具有鲁棒特性的故障检测模型。将该算法应用于TE化工过程和多相流过程中,与其他几种算法相比,所提算法故障检测率最高,验证了其有效性。自编码器及其优化方法在工业过程故障检测领域具有巨大的潜力可供挖掘。

参考文献

- [1] 曹伟嘉,杨留方,徐天奇,等. 基于精英反策略麻雀搜索优化随机森林的变压器故障诊断[J]. 国外电子测量技术, 2022, 41(2): 138-143.
- [2] 牛玉广,王世林,林忠伟,等. 基于多元统计过程监控的锅炉过程故障检测[J]. 动力工程学报, 2017, 37(10): 829-836.
- [3] 常玉清,王姝,王福利,等. 基于多PCA模型的过程监测方法[J]. 仪器仪表学报, 2014, 35(4): 901-908.
- [4] 周卫庆,司凤琪,徐洽皋,等. 基于KPCA残差方向梯度的故障检测方法及应用[J]. 仪器仪表学报, 2017, 38(10): 2518-2524.
- [5] 李智豪. 基于流形学习的流程工业近红外故障检测[D]. 无锡:江南大学, 2021.
- [6] 王广斌,罗军,贺旖琳,等. 基于局部线性嵌入的多流形学习故障诊断方法[J]. 湖南科技大学学报(自然科学版), 2015, 30(4): 34-39.
- [7] 孙劲光,丁胜锋. 鲁棒性监督等距特征映射方法[J]. 中国矿业大学学报, 2017, 46(4): 932-938.
- [8] 赵小强,姚红娟. 基于DDNPE算法的间歇过程故障诊断[J]. 兰州理工大学学报, 2020, 46(4): 96-102.
- [9] XIAO Y, HE Y. A novel approach for analog fault diagnosis based on neural networks and improved

kernel PCA [J]. Neurocomputing, 2011, 74(7): 1102-1115.

- [10] 唐颖川,黄姣茹,钱富才. 基于深度学习与注意力机制的化工故障分类[J]. 电子测量技术, 2022, 45(4): 168-174.
- [11] ZHAO H. Neural component analysis for fault detection [J]. Chemometrics and Intelligent Laboratory Systems, 2018, 176: 11-21.
- [12] 蒋立. 基于自编码器模型的非线性过程监测[D]. 杭州:浙江大学, 2018.
- [13] 冯倩玉. 基于降噪正交自编码器的TE过程故障诊断[D]. 上海:上海交通大学, 2018.
- [14] 李元,姚宗禹. 基于邻域保持嵌入的主多项式非线性过程故障检测[J]. 上海交通大学学报, 2021, 55(8): 1001-1008.
- [15] 薛南,吕柏权,倪陈龙. 基于自编码器和填充函数的深度学习优化算法[J]. 电子测量技术, 2019, 42(23): 79-84.
- [16] 景军锋,党永强,苏泽斌. 基于改进SAE网络的织物疵点检测算法[J]. 电子测量与仪器学报, 2017, 31(8): 1321-1329.
- [17] GRAMACKI A. Nonparametric kernel density estimation and its computational aspects[M]. Cham: Springer International Publishing, 2018.
- [18] 王培良,夏春江. 基于PCA-PDBNs的故障检测与自主学习辨识[J]. 仪器仪表学报, 2015, 36(5): 1147-1154.
- [19] STIEF A, TAN R, CAO Y, et al. A heterogeneous benchmark dataset for data analytics: Multiphase flow facility case study[J]. Journal of Process Control, 2019, 79: 41-55.
- [20] RUIZ-CAECEL C, CAO Y, MBA D, et al. Statistical process monitoring of a multiphase flow facility[J]. Control Engineering Practice, 2015, 42: 74-88.

作者简介

郭小萍,博士,教授,主要研究方向为基于数据驱动的复杂过程故障诊断。

E-mail: gxp2001@sina.com

张志朋,硕士研究生,主要研究方向为基于数据驱动的复杂过程故障诊断。

E-mail: 2277464871@qq.com

李元,博士,教授,主要研究方向为基于数据驱动的复杂过程故障诊断。

E-mail: li-yuan@mail.tsinghua.edu.cn