

DOI:10.19651/j.cnki.emt.2415978

基于对抗二次自编码器和集成学习的 工业过程早期故障检测*

刘喜平 高丙朋 蔡鑫

(新疆大学电气工程学院 乌鲁木齐 830017)

摘要: 由于工业过程早期微小故障存在数据振幅小,特征强耦合的特点,导致传统自编码器模型对此类故障的检测效果不佳,对此,提出一种基于对抗二次自编码器和集成学习的工业过程早期故障检测方法。首先引入一种二次型神经元嵌入普通自编码器模型的隐藏层中,增加模型的表达能力,其次提出一种对抗性的二次自编码器,在训练过程中引入 GAN 网络,使自编码器的特征学习遵循特定的概率分布。然后利用集成学习思想对正常工况数据进行采样,给每个采样的子集训练一个对抗二次自编码器,利用每个子模型的 SPE 和 T^2 统计量分别生成两个矩阵,接着在生成的矩阵上使用单步滑动窗口内奇异值分解的融合策略,将每个窗口内的最大奇异值作为检测统计量。使用一个数值例子和 TE 过程对所提方法进行验证。实验结果表明所提方法具有良好的早期微小故障检测性能。

关键词: 早期故障检测;二次自编码器;对抗自编码器;集成学习

中图分类号: TP277;TN06 **文献标识码:** A **国家标准学科分类代码:** 510.80

Incipient fault detection of industrial processes based on adversarial quadratic autoencoder and ensemble learning

Liu Xiping Gao Bingpeng Cai Xin

(College of Electrical Engineering, Xinjiang University, Urumqi 830017, China)

Abstract: Due to the characteristics of early minor faults in industrial processes, such as small data amplitudes and strong feature coupling, the detection performance of traditional autoencoder models for these faults is poor. In response, a method for early fault detection in industrial processes based on adversarial quadratic autoencoders and ensemble learning is proposed. Initially, a quadratic neuron is introduced into the hidden layer of a conventional autoencoder model to enhance its expressive power. Subsequently, an adversarial quadratic autoencoder is introduced, incorporating a GAN network during training to enforce feature learning to adhere to specific probability distributions. Then, employing the concept of ensemble learning, normal operational data is sampled, and each subset is used to train an adversarial quadratic autoencoder. Subsequently, two matrices, SPE and T^2 statistical quantities, are generated for each subset model. A fusion strategy utilizing singular value decomposition within a single-step sliding window is employed to utilize the maximum singular value within each window as a detection statistic. The proposed method is validated using a numerical example and the TE process, demonstrating its effectiveness in early detection of minor faults in industrial processes.

Keywords: incipient fault detection; quadratic autoencoder; adversarial autoencoder; ensemble learning

0 引言

随着工业技术的发展,工业过程越来越复杂,导致故障概率增加。故障检测与诊断技术是保证生产过程安全的重要

技术之一^[1]。故障检测是其中的重要组成部分,其目的是将故障与正常状态区分开。有效的故障检测方法可以避免事故的发生,保证工业过程的安全稳定,早期故障若不能及时的检测,将发展为重大故障,造成人员伤亡与经济损

收稿日期:2024-05-06

* 基金项目:国家自然科学基金(62303394)、新疆维吾尔自治区自然科学基金(2022D01C694)、新疆维吾尔自治区高校基本科研业务费科研项目(XJEDU2023P025)资助

失。由于计算机技术和数字存储技术在现代工业中的应用,数据驱动的故障检测方法受到了人们的关注^[1-3]。最初,基于数据的过程监控方法以统计学习方法为主,例如主成分分析(principal component analysis, PCA)^[4]和偏最小二乘(partial least squares, PLS)^[5]。这些方法都是在数据服从高斯分布的前提下,对数据进行降维处理,提取原始数据的主要特征,降低高维数据的计算复杂度。然而,它们也有局限性。如果数据不服从高斯分布,那么提取的特征就不是最优的。近来,随着深度学习的兴起,以神经网络为主要模型,将原始数据特征转化为更抽象的表征^[6],比如深度信念网络(deep belief network, DBN)等^[7]。自编码器(autoencoder, AE)由于其具有良好的降维能力,被广泛用于过程监控^[8]。对于大规模数据,浅层 AE 不足以获得良好的数据表示。Wu 等^[9]提出了一种堆叠注意力自编码器(stacked attention autoencoder, SAE),其可以更好的捕获数据的语义信息,利用少量的故障样本,实现了更好的故障检测效果。郭小萍等^[10]提出了一种邻域降噪正交自编码器的方法,在正交自编码器中加入邻域保持嵌入算法来提取数据的邻域信息,达到了较好的检测效果。Jang 等^[11]将对抗自编码器(adversarial autoencoder, AAE)引入工业过程故障检测领域,由于给特征学习添加了先验分布,其可以获得数据的确定流形特征,提高了故障检出率,降低了故障检测延迟。尽管 AAE 在一般故障的检测中展现出了优异的性能,但其对于早期微小故障的检测基本是无效的。

为了解决早期故障检测的问题,在上述已有的基础检测器的基础上,Liu 等^[12]提出了一种集成的 PCA 检测器,使用奇异值分解的方法,显著提高了早期故障的检出率,但是对于复杂的非线性工业过程,该方法的使用会受到很大的制约。Gao 等^[13]提出了一种分布式自适应加权堆叠自编码器(decentralized adaptively weighted stacked autoencoder, DAWSAE)的早期故障检测方法,将系统分为若干个子块分别进行检测,并在特征向量上使用自适应奇异值加权的策略,增加了早期微小故障的检测率,但是分布式的检测方式,忽略了各子块之间的联系,并且使模型变得相当复杂。李元等^[14]提出一种基于分数阶傅里叶变换与卷积神经网络的工业过程故障检测方法,使用分数阶傅里叶变换(fractional fourier transform, FRFT)放大微小故障的特征,然后使用卷积神经网络(convolutional neural network, CNN)进行检测,提高了对部分故障检测的灵敏度,但是对于更加微小的故障,该方法产生了很高的误报率。

显然,早期微小故障振幅小,特征强耦合的问题,极大的增加了故障检测的难度。为了改善自编码器模型对于早期微小故障检测的低效性,在 AAE 框架下,引入嵌入了二次神经元的二次自编码器作为基础检测器,提出一种对抗性的二次自编码器,在提取数据的流形特征的同时,利用二次神经元优异的表达能,改善特征强耦合情况下,故障特

征难提取的问题,然后使用集成学习与滑动窗口奇异值分解的方法,放大检测统计量中包含的异常信息,改善早期微小故障振幅小的问题。最后通过数值例子和 TE 过程仿真实验,验证所提方法的有效性。

1 相关理论

1.1 二次自编码器

自编码器^[15]是一种无监督的神经网络模型。它可以学习到输入数据的隐含特征,同时用学习到的新特征可以重构出原始输入数据。自编码器的结构一般由两部分组成:编码器和解码器。编码器的作用是将输入特征空间中的向量压缩到隐藏特征空间中得到隐藏变量,而解码器则是将隐藏特征空间的表示恢复到原有的输入空间中。自编码器的目标是使得重构的输出尽可能接近原始的输入,从而学习到数据的有用特征。其一般采用均方误差作为损失函数,如式(1)所示。

$$MSE_{loss} = \frac{1}{N} \sum_{i=1}^N (x_i - x'_i)^2 \quad (1)$$

其中, x_i 为原始数据, x'_i 为重构数据。

重构误差即平方预测误差,如式(2)所示。

$$SPE = (\mathbf{X} - \hat{\mathbf{X}})^T (\mathbf{X} - \hat{\mathbf{X}}) \quad (2)$$

其中, \mathbf{X} 是输入向量, $\hat{\mathbf{X}}$ 为重构向量。

自编码器的训练只需要正常工况数据,然后根据测试样本的重构误差或其他统计量来进行异常检测,正是这一无监督训练的特性,使得其在历史故障样本较少的工业过程故障检测中成为了重要的数据模型建立方法。

二次自编码器(quadratic autoencoder, QAE)^[16-17]即在自编码器模型的基础上,嵌入了由二次神经元组成的隐藏层,这样的隐藏层称作二次输出层。相较于传统神经元(图1),二次神经元将内积输出替换为二次输出,增强了单个神经元的表达能力,结构图如图2所示。

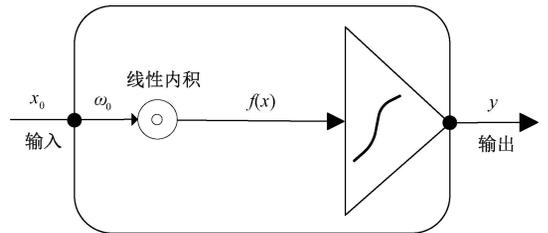


图 1 传统神经元

Fig. 1 Traditional neuron

传统神经元与二次神经元的输出如式(3)和(4)所示。

$$f(x) = \sum_{i=0}^n \omega_i x_i \quad (3)$$

$$f(x) = \left(\sum_{i=0}^n \omega_{i1} x_i \right) \left(\sum_{i=0}^n \omega_{i2} x_i \right) + \sum_{i=0}^n \omega_{i3} x_i^2 + c \quad (4)$$

二次自编码器的构型可以分为两种形式,一种完全由二次输出层组成,另一种则由传统层与二次输出层交叠而

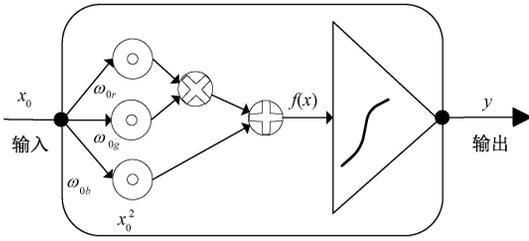


图 2 二次神经元
Fig. 2 Quadratic neuron

成,在下面的内容中,使用交叠形式的二次自编码器^[18],其结构图如图 3 中自编码器部分所示,隐藏层里实线矩形框中的神经元表示二次神经元,二次网络的训练过程与传统网络一样,其优化过程重复逐层链式规则引导的反向传播直到收敛。

1.2 集成学习

集成学习(ensemble learning)旨在通过结合多个基本模型的预测结果来提高整体性能。它通过构建多个模型,并将它们的预测结果进行组合,以获得比任何单个模型更好的预测能力。常见的集成学习方法主要包括:

Voting(投票):多个基本模型进行预测,然后通过多数投票或加权投票来确定最终预测结果。

Bagging(自举聚合):使用随机抽样的方式从训练集中构建多个基本模型,每个模型都是独立训练的,最终预测结果通过投票或平均值进行决定。

Boosting(提升):通过顺序训练多个基本模型,每个模型都试图修正前一个模型的错误,最终将它们组合成一个强大的模型。

2 基于对抗二次自编码器和集成学习的早期故障检测方法

2.1 对抗二次自编码器

受对抗自编码器^[19]的启发,提出一种对抗性的二次自编码器,称为对抗二次自编码器(adversarial quadratic autoencoder, AQAE),其结构图如图 3 所示。训练阶段在 QAE 的潜在空间引入 GAN 网络。AQAE 的结构包括 3 个部分:编码器、解码器和鉴别器。编码器将输入数据编码成潜在空间中的表示,解码器将潜在表示还原成原始数据,而鉴别器则试图区分编码后的潜在表示是否来自真实数据或是编码器生成的假数据。关键在于将编码器当作生成器,通过给的任意分布的真实数据(通常将这个先验分布设置为标准正态分布),在对抗训练过程中,使得编码器可以生成特定概率分布的特征。对抗过程的优化问题可以表示为:

$$\min_G \max_D E_{x \sim P_{data}} [\log D(x)] + E_{z \sim p(z)} [\log(1 - D(G(z)))] \quad (5)$$

其中, G 为生成器/编码器, D 为鉴别器。

AQAE 的训练包括 2 个阶段,重构阶段和正则化阶

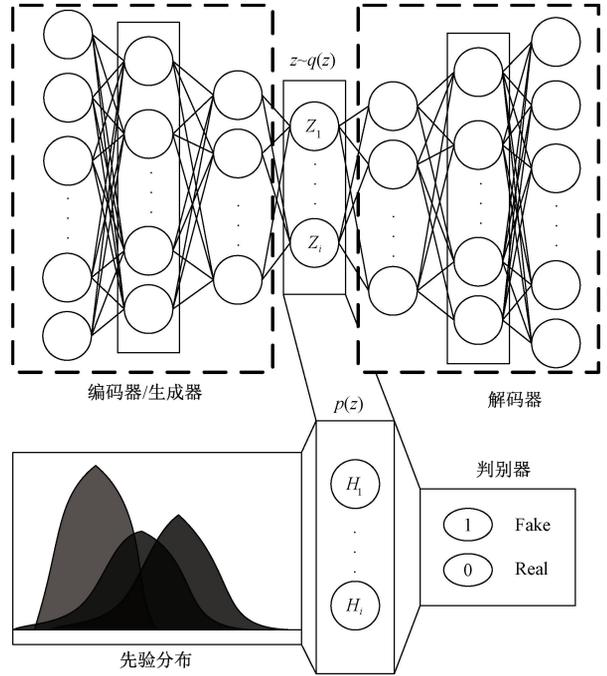


图 3 对抗二次自编码器
Fig. 3 Adversarial quadratic autoencoder

段,在重构阶段,编码器和解码器被用来最小化重构误差。正则化阶段,鉴别器被训练来最大化区分真实数据与编码器生成的假数据之间的差异,如式(6)所示,然后训练编码器生成相似数据来欺骗鉴别器,计算损失如式(7)所示。

$$-\nu = -\log(d_\lambda(p(z))) - \log(1 - d_\lambda(q(z))) \quad (6)$$

$$-\eta = \log(d_\lambda(q(z))) \quad (7)$$

其中, $p(z)$ 为先验分布, $q(z)$ 为编码器学习的特征分布。

由于引入了二次输出层,反向传播参数优化时采用分组优化策略,设置不同的学习率进行参数的更新,将二次层的二次输出权重参数分为 $Group_r, Group_g$ 两组,定义学习率分别为 lr_r, lr_g , 幂项的权重参数设置为 $Group_b$, 学习率为 lr_b , 定义普通层的学习率为 lr , 学习率的大小根据具体实验设置。经过训练, AQAE 能够学习到数据的高级表示,并生成具有多样性的新样本,具有很好的数据生成能力。

过程监控往往根据测量变量与正常状态之间的距离来判断系统的异常情况,同 AAE 一样,由于 AQAE 模型提取的特征向量非常接近高斯分布,因此构造 PCA 中使用的类似检测指标将不会影响故障检测的性能。所以,将训练好的 AQAE 模型应用于故障检测时,分别在潜在空间和重构空间计算两个统计量:

$$T^2 = z^T \Delta^{-1} z \quad (8)$$

$$SPE = (x - \hat{x})^T (x - \hat{x}) \quad (9)$$

其中, z 为特征向量, Δ 为训练数据在潜在空间的协方差矩阵, \hat{x} 为重构向量。

2.2 集成学习与奇异值分解融合策略

采用集成学习中的 Bagging 方法,在正常工况数据中进行随机采样,训练多个对抗二次自编码器。最后使用单步滑动窗口内奇异值分解的融合策略^[12],假设子检测器的检测结果为 $\mathbf{h} = (h_1, h_2, \dots, h_n)$, 其中 h 为异常信息, n 为样本数。将所有子检测器的检测结果合入一个矩阵:

$$\mathbf{P} = \begin{bmatrix} h_{1,1} & \dots & h_{n,1} \\ \vdots & \ddots & \vdots \\ h_{1,j} & \dots & h_{n,j} \end{bmatrix} \quad (10)$$

其中, j 表示子检测器个数。接着沿样本方向执行单步滑动窗口内的奇异值分解,得到每个窗口内数据的奇异值矩阵如下:

$$\theta_c = f(\mathbf{P}) = f \begin{pmatrix} h_{c,1} & h_{c+1,1} & \dots & h_{c+\omega-1,1} \\ \vdots & \vdots & \vdots & \vdots \\ h_{c,j} & h_{c+1,j} & \dots & h_{c+\omega-1,j} \end{pmatrix} =$$

$$\begin{bmatrix} \lambda_{c,1} & & & \\ & \ddots & & \\ & & \lambda_{c,j} & \end{bmatrix} \quad (11)$$

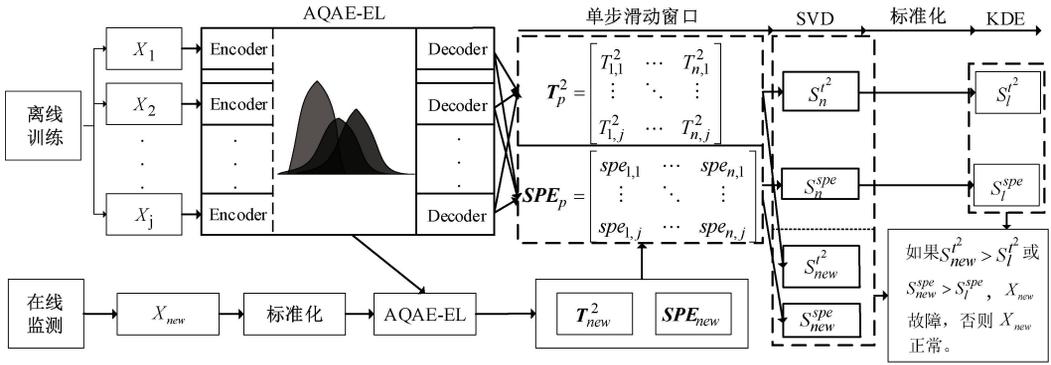


图 4 基于对抗二次自编码器和集成学习的检测方法流程图

Fig. 4 Flow chart of detection method based on adversarial quadratic autoencoder and ensemble learning

(3)为每个子集训练一个对抗二次自编码器,完成数据模型的建立。

(4)分别计算训练数据在每个子检测器潜在空间的 T^2 和重构空间的 SPE , 然后生成 2 个矩阵:

$$\mathbf{T}_p^2 = \begin{bmatrix} T_{1,1}^2 & \dots & T_{n,1}^2 \\ \vdots & \ddots & \vdots \\ T_{1,j}^2 & \dots & T_{n,j}^2 \end{bmatrix} \quad (14)$$

$$\mathbf{SPE}_p = \begin{bmatrix} spe_{1,1} & \dots & spe_{n,1} \\ \vdots & \ddots & \vdots \\ spe_{1,j} & \dots & spe_{n,j} \end{bmatrix} \quad (15)$$

(5)执行滑动窗口奇异值分解策略,根据式(11)和(12)计算训练样本数据的统计量 S_n^2 和 S_n^{spe} 。

(6)标准化两个统计量如下:

$$S_i^{2/spe} = \frac{S_i^{2/spe} - \bar{S}^{2/spe}}{\sigma^{2/spe}} \quad (16)$$

其中, $\bar{S}^{2/spe}$ 为训练数据每个滑动窗口内最大奇异

其中, c 表示当前窗口, f 表示奇异值分解, ω 表示窗口宽度且 $\omega < n$, λ 表示奇异值。定义一个新的检测统计量:

$$S_c = \max(\lambda_{c,1} \dots \lambda_{c,j}) \quad (12)$$

将其作为当前窗口的异常信息。

2.3 检测流程

将整个检测过程分为 2 个阶段,分别为离线建模与在线监测阶段。整体流程如图 4 所示。

1)离线建模:

(1)对于正常工况数据 $X = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^{n \times m}$, 其中 n 为样本个数, m 为特征维度。在正常工况数据中进行随机采样,子集 $X_i \in \mathbb{R}^{v \times m}$, 其中 v 为子集的样本个数,由于随机性,每个子集是不同的,这样可以有效的避免过拟合,并且在实际工业过程中可以并行执行,在本文中,采样 3 个子集。

(2)将每个子集数据进行标准化,如下:

$$\bar{X}_i = \frac{X_i - \bar{X}}{\sigma} \quad (13)$$

其中, \bar{X} 为子集均值, σ 为方差。

值的均值, $\sigma^{2/spe}$ 为方差。

(7)使用核密度估计确定控制极限 S_i^2 和 S_i^{spe} 。

2)在线监测:

(1)新采样的样本 X_{new} 根据式(13)标准化以后输入集成模型中,得到对应于每个子模型的 \mathbf{T}_{new}^2 和 \mathbf{SPE}_{new} , 拼接到 \mathbf{T}_p^2 和 \mathbf{SPE}_p 之后:

$$\mathbf{T}_{p_{new}}^2 = \begin{bmatrix} T_{1,1}^2 & \dots & T_{n,1}^2 & T_{new,1}^2 \\ \vdots & \vdots & \vdots & \vdots \\ T_{1,j}^2 & \dots & T_{n,j}^2 & T_{new,j}^2 \end{bmatrix} \quad (17)$$

$$\mathbf{SPE}_{p_{new}} = \begin{bmatrix} spe_{1,1} & \dots & spe_{n,1} & spe_{new,1} \\ \vdots & \vdots & \vdots & \vdots \\ spe_{1,j} & \dots & spe_{n,j} & spe_{new,j} \end{bmatrix} \quad (18)$$

(2)执行下一步的滑动窗口奇异值分解,使用式(11)和(12)计算 X_{new} 的 S_{new}^2 和 S_{new}^{spe} 并标准化。

(3)若 $S_{new}^2 > S_i^2 \vee S_{new}^{spe} > S_i^{spe}$, 则系统故障,否则系统正常。

3 实验验证

3.1 数值例子

首先通过一个数值例子进行仿真验证^[20],具体模型表示如下:

$$\begin{cases} x_1 = t + e_1 \\ x_2 = x_1 + 0.1t + e_2 \\ x_3 = 0.01t + e_3 \end{cases} \quad (19)$$

其中, $e_i (i = 1, 2, 3)$ 是分布在 $[-1, 1]$ 上的独立白噪声变量; t 服从 $[0, 1]$ 分布。首先采样 1 000 个正常数据组成训练集,接着采样 800 个数据组成带故障样本的测试集,故障在第 300 个样本之后引入,故障 1 为加在变量 x_1 上的大小为 1.5 的阶跃信号,故障 2 为加在变量 x_1 上的大小为 0.3 的阶跃信号。使用 AAE, AQAE, AQAE-EL 对该数值例子进行检测。自编码器输入输出结构均为 $[4 \ 3 \ 2 \ 3 \ 4]$,设置 AAE 的学习率为 $lr = 0.0005$,AQAE 不同组的学习率分别为 $lr = lr_r = 0.0005, lr_g = lr_b = 0.0075$ 。先验分布均设置为标准正态分布且均使用核密度估计设置控制阈值,置信水平为 0.99。故障 1 的检测结果如图 5,6 所示。

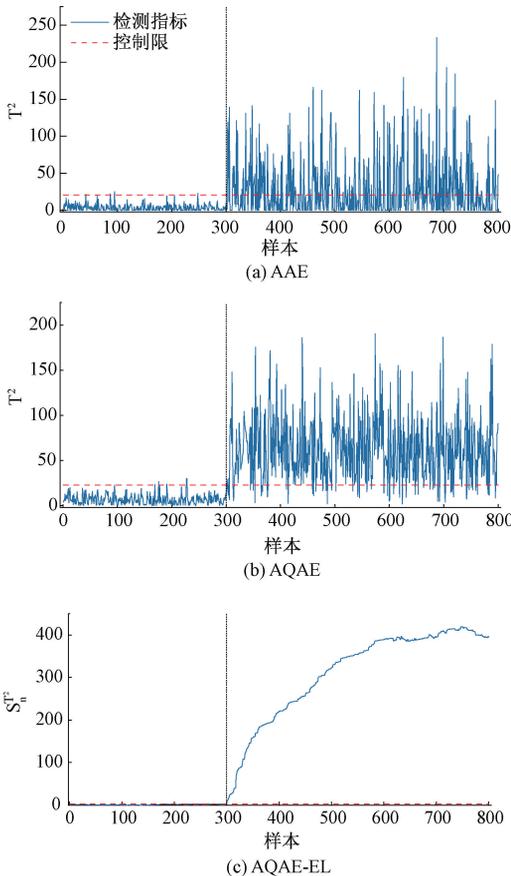


图 5 3 种方法 T^2 检测故障 1

Fig. 5 Three methods used T^2 to detect fault 1

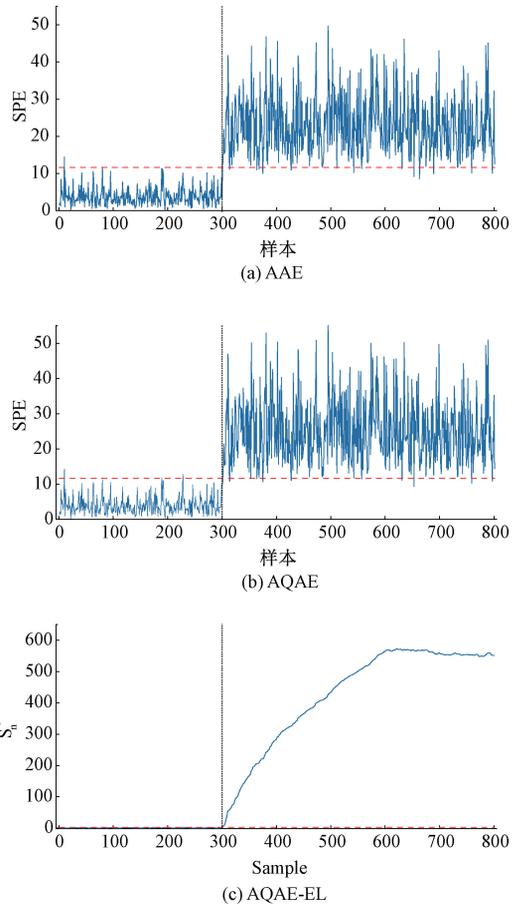


图 6 3 种方法 SPE 检测故障 1

Fig. 6 Three methods used SPE to detect fault 1

由故障 1 的检测结果可以看出,二次神经元的嵌入,对自编码器模型的特征提取能力有所提高,可以更好的表示故障样本的异常信息,对于 T^2 统计量的提升尤为明显,如图 5(a)、(b)所示,而 SPE 统计量相较于普通模型,变化比较小,如图 6(a)、(b)所示。AQAE-EL 由于综合了滑动窗口内样本的异常信息,显著的降低了误报率,并且将最大奇异值作为检测指标,有效的提高了故障样本的检出率,在两个统计量上的检测结果如图 5(c)与图 6(c)所示。

故障 2 的检测结果如图 7、8 所示,对于振幅变化较小的故障,AAE 和 AQAE 两种单模型的方法,都表现出了低效性,但从 T^2 统计量的检测图中可以看出,AQAE 要比 AAE 的检出率更高,两种方法的 SPE 统计量在较小故障的检测上仍然保持一致,提升不明显。而 AQAE-EL 在两种统计量上均可以及时检测出故障,并且有着非常低的误报率。

总体而言,二次层的嵌入提升了 AAE 模型 T^2 统计量的检测性能,对 SPE 统计量的提升不明显,AQAE-EL 由于使用了集成学习与奇异值分解的融合策略,放大了检测统计量中的异常信息,相较于单模型的方法,表现出了最优的检测性能,3 种方法对两种故障具体的检出率如表 1 所示。

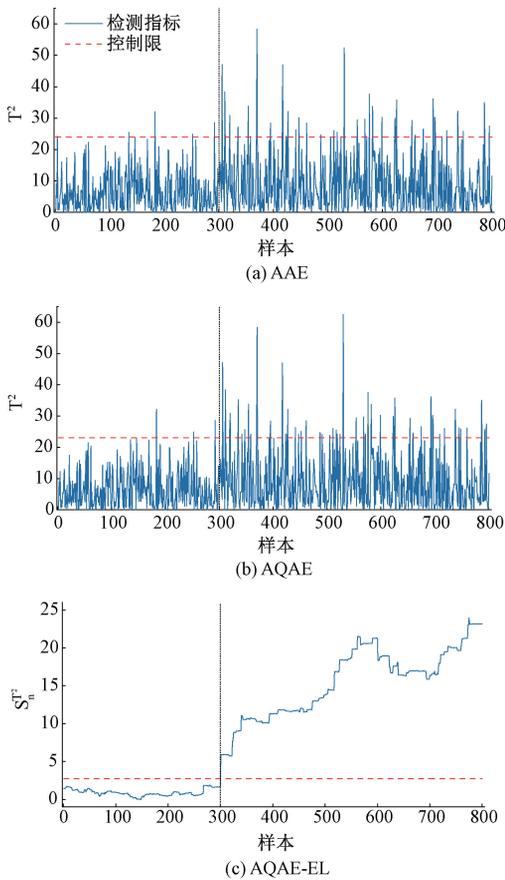


图 7 3 种方法 T^2 检测故障 2

Fig. 7 Three methods used T^2 to detect fault 2

3.2 TE 过程

TE 过程是 Downs 和 Vogel 提出的基准过程,如图 9 所示,它通常用于研究过程控制技术^[21]。大量关于过程监控和故障诊断的文献在验证模型性能时使用了 TE 过程生成的数据集。TE 过程的采样间隔为 3 min,有 52 个观测变量,共模拟了 21 种故障,故障说明见表 2。数据集中包含正常数据集,以及 21 个不同的故障数据集,每个故障数据集包含 960 个样本,故障在第 160 个样本之后引入^[22]。除了典型的较大故障,其中存在一些难以检测的早期微小故障,如故障 3、9、15 等,传统方法很难对这些故障进行有效的检测。

这里使用这几种故障来验证所提方法的有效性。自编码器模型输入输出结构为[52 26 13 5 13 26 52],其他参数与上一节一致。故障 3 和 9 均为物料 D 的温度变化引起的故障,不同的是故障 3 为阶跃型故障,故障 9 为随机变量型故障,图 10 和 11 显示了 AAE, AQAE, AQAE-EL 对故障 3 的检测结果。相较于 AAE 来说, AQAE 仍然保持着较好的 T^2 统计量检测性能,但是两种方法仍然不能较为有效的检测故障 3。而 AQAE-EL 则表现出稳定的早期微小故障检测性能,两个统计量的检测结果均达到了 90% 以上。

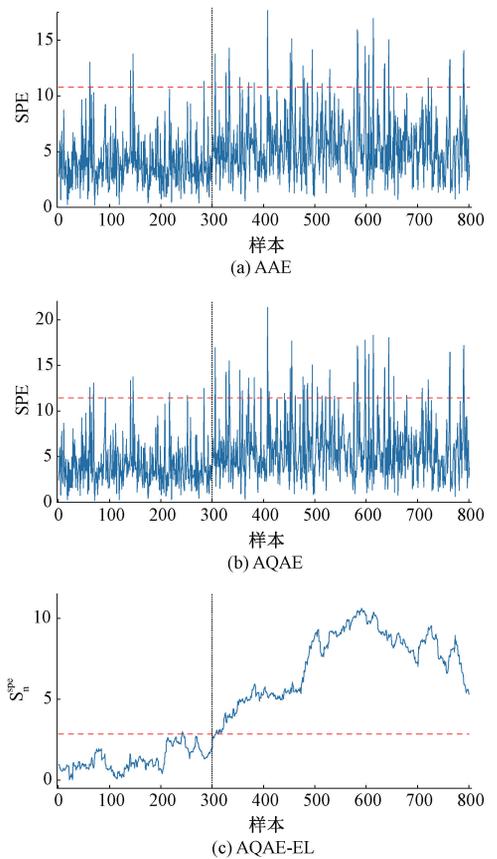


图 8 3 种方法 SPE 检测故障 2

Fig. 8 Three methods used SPE to detect fault 2

表 1 3 种方法对数值例子的故障检出率

Table 1 The FDR of three methods for numerical examples

方法	故障 1		故障 2	
	T^2	SPE	T^2	SPE
AAE	0.416	0.966	0.094	0.058
AQAE	0.874	0.980	0.110	0.070
AQAE-EL	1.000	1.000	1.000	0.990

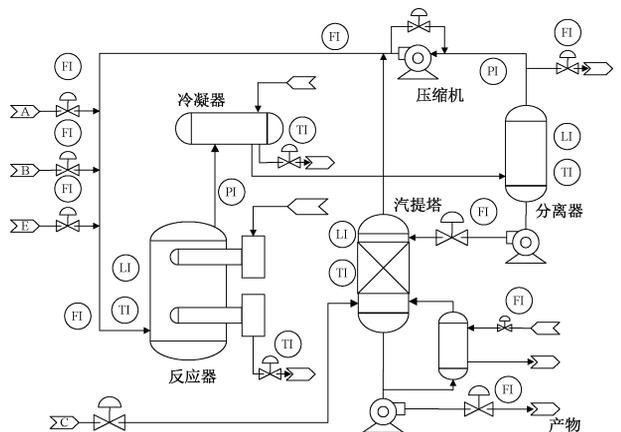


图 9 TE 过程

Fig. 9 TE process

表2 TE过程21种故障说明

Table 2 Description of 21 faults in TE process

故障编号	故障说明	故障类型
1	成分 A/C 进料比,成分 B 固定	阶跃
2	成分 B,成分 A/C 比固定	阶跃
3	成分 D 进料温度	阶跃
4	反应塔冷却水入口温度	阶跃
5	冷凝塔冷却水入口温度	阶跃
6	成分 A 进料丢失	阶跃
7	成分 C 进料压力损失	阶跃
8	成分 A,B,C 进料比	随机
9	成分 D 进料温度	随机
10	成分 C 进料温度	随机
11	反应塔冷却水入口温度	随机
12	冷凝塔冷却水入口温度	随机
13	反应塔动态	缓慢漂移
14	反应塔冷却水阀	阀门粘滞
15	冷凝塔冷却水阀	阀门粘滞
16~20	未知	未知
21	阀门固定在稳态位置	位置恒定

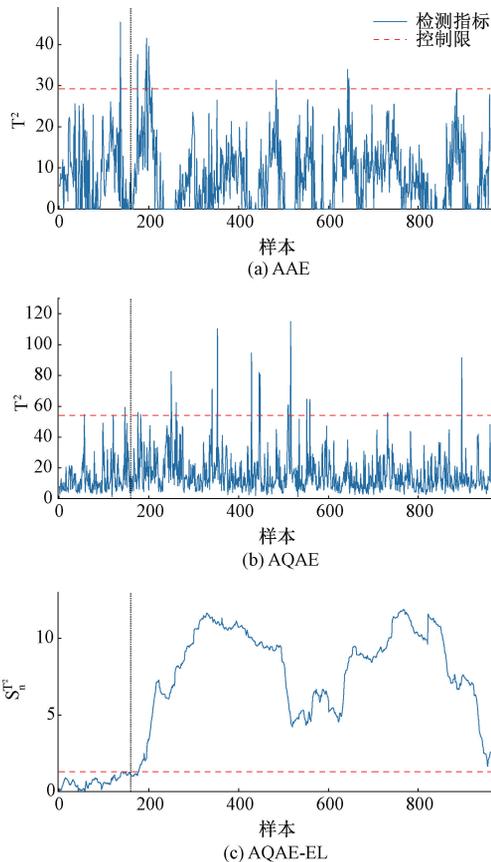
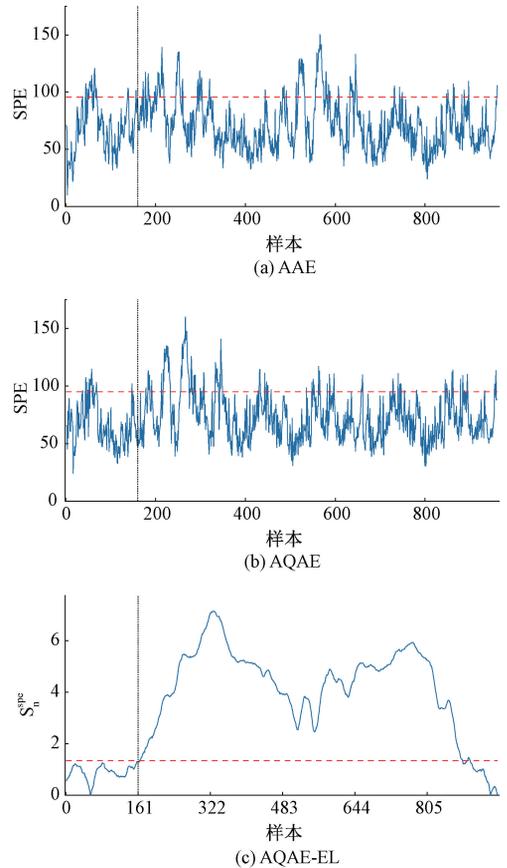
图10 3种方法 T^2 检测 TE 过程故障 3Fig. 10 Three methods used T^2 to detect TE process fault 3

图11 3种方法 SPE 检测 TE 过程故障 3

Fig. 11 Three methods used SPE to detect TE process fault 3

图12和13显示了3种方法对故障15的检测结果,故障15是由冷凝水阀门粘滞引起的故障,故障的机理表明了此类故障的发展较为缓慢,并且给整个系统带来的影响会越来越严重,3种方法的检测结果也证明了这一点,由于故障发生早期,对系统的影响非常小,3种方法均没有在第一时间检测到,但是相较于单模型的方法,AQAE-EL检测到故障的时间是最早的,并且整体的检出率也是最优的。

相较于AAE,嵌入了二次神经元的AQAE具有更好的早期故障特征提取能力的原因如图14所示。将两个模型使用正常工况数据训练,并提取特征向量,使用核密度估计计算每个特征的概率分布并绘图,AQAE提取的特征概率分布更接近于标准正态分布,原因在于嵌入了二次神经元,相比AAE有了更好的表达能力,从而在对抗训练过程中增强了模型的流形嵌入能力,最终增加了 T^2 统计量的稳定性。

为了进一步说明所提方法的有效性,采用同样使用自编码器模型为基础检测器的DAWSAE^[13]在TE过程中的早期故障检测效果作为对比,由于该方法为分布式检测方式,这里选择各个子块中最优的检测结果作为对比,表3显示了几种方法使用 T^2 和SPE两个统计量对故障3、9、15

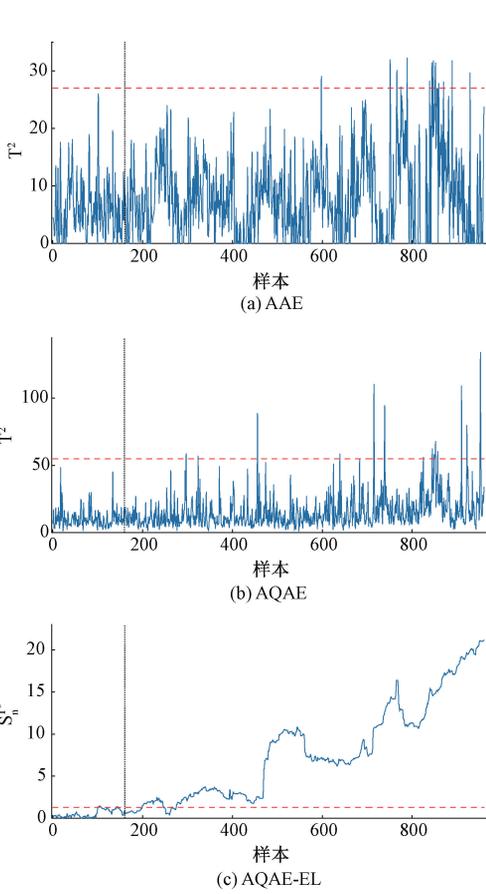


图 12 3 种方法 T^2 检测 TE 过程故障 15

Fig. 12 Three methods used T^2 to detect TE process fault 15

的检出率对比。整体来说在自编码器模型中嵌入二次输出层对故障检测是有效的,并且结合了集成学习的AQAE-EL在保持低误报的同时,相较于分布式检测的DAWSAE方法,实现了更高的检测正确率,有效的检测到了早期微小故障。

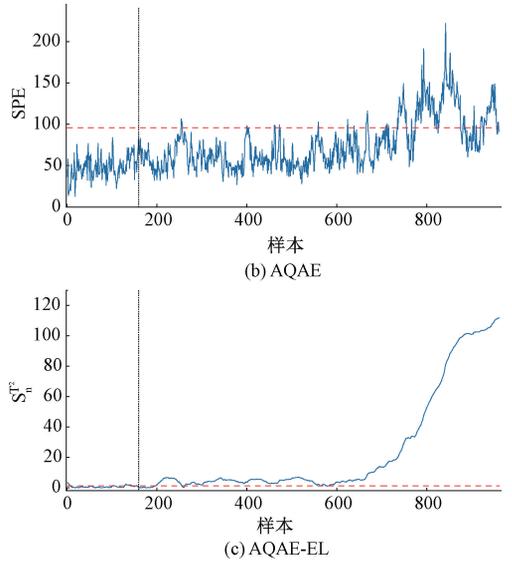
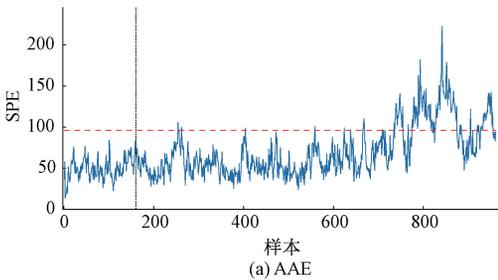


图 13 3 种方法 SPE 检测 TE 过程故障 15

Fig. 13 Three methods used SPE to detect TE process fault 15

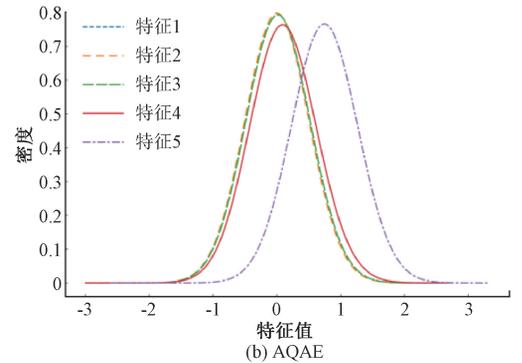
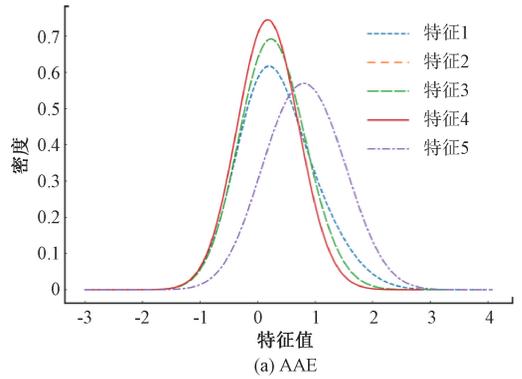


图 14 正常样本特征向量的概率密度图

Fig. 14 Probability density plot of normal sample eigenvector

表3 4种方法对TE过程3类早期故障的检出率

Table 3 The FDR of three kinds of incipient faults in TE process by four methods

方法	故障 3		故障 9		故障 15	
	T^2	SPE	T^2	SPE	T^2	SPE
AAE	0.013	0.140	0.125	0.100	0.018	0.204
AQAE	0.020	0.183	0.118	0.113	0.021	0.217
DAWSAE	0.783	0.766	0.486	0.930	0.275	0.280
AQAE-EL	0.980	0.906	0.900	0.925	0.925	0.945

4 结 论

针对传统自编码器模型对工业过程早期微小故障检测效果不佳的问题,提出了一种基于对抗二次自编码与集成学习的早期故障检测方法,通过在普通自编码器模型中嵌入二次神经元,增强了模型的表达能力,并且在训练过程中引入对抗网络,使编码器的特征学习遵循先验分布,提升了 T^2 统计量的检测性能,最后使用集成学习与奇异值分解的融合策略,使用对故障特征更为敏感的最大奇异值作为检测统计量,提高了早期故障的检出率。通过数值例子和TE过程验证了所提方法的有效性。

参考文献

- [1] MIAN Z B, DENG X F, DONG X H, et al. A literature review of fault diagnosis based on ensemble learning[J]. Engineering Applications of Artificial Intelligence, 2024, 127: 107357.
- [2] WANG J J, YE L K, GAO R X, et al. Digital twin for rotating machinery fault diagnosis in smart manufacturing[J]. International Journal of Production Research, 2019, 57(12): 3920-3934.
- [3] 姚羽曼,罗文嘉,戴一阳.数据驱动方法在化工过程故障诊断中的研究进展[J].化工进展,2021,40(4): 1755-1764.
YAO Y M, LUO W J, DAI Y Y. Research progress of data-driven methods in fault diagnosis of chemical process [J]. Chemical Industry and Engineering Progress, 2021, 40(4): 1755-1764.
- [4] 姚远,佟佳蓉,高军,等.基于分层分块DLPPCA-SVM的复杂工业过程监测与故障诊断方法[J].控制与决策,2022,37(5): 1402-1408.
YAO Y, TONG J R, GAO J, et al. Monitoring and fault diagnosis method of complex industrial process based on DLPPCA-SVM[J]. Control and Decision, 2022, 37(5): 1402-1408.
- [5] HARROU F, NOUNOU M N, NOUNOU H N, et al. PLS-based EWMA fault detection strategy for process monitoring[J]. Journal of Loss Prevention in the Process Industries, 2015, 36: 108-119.
- [6] LI W, HUANG R Y, LI J P, et al. A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges [J]. Mechanical Systems and Signal Processing, 2022, 167: 108487.
- [7] NIU G Q, YI X H, CHEN C, et al. A novel effluent quality predicting model based on genetic-deep belief network algorithm for cleaner production in a full-scale paper-making wastewater treatment [J]. Journal of Cleaner Production, 2020, 265: 121787.
- [8] SAKURADA M, YAIRI T. Anomaly detection using autoencoders with nonlinear dimensionality reduction[C]. Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, 2014: 4-11.
- [9] WU Q Q, YAN X F. Interval-valued-based stacked attention autoencoder model for process monitoring and fault diagnosis of nonlinear uncertain systems[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 72: 1-10.
- [10] 郭小萍,张志朋,李元.基于邻域降噪正交自编码器的工业过程故障检测[J].电子测量技术,2022,45(21): 142-147.
GUO X P, ZHANG ZH P, LI Y. Neighborhood denoising quadrature autoencoder based fault detection for industrial process [J]. Electronic Measurement Technology, 2022, 45(21): 142-147.
- [11] JANG K J, HONG S Y, KIM M S, et al. Adversarial autoencoder based feature learning for fault detection in industrial processes [J]. IEEE Transactions on Industrial Informatics, 2021, 18(2): 827-834.
- [12] LIU D C, SHANG J, CHEN M Y. Principal component analysis-based ensemble detector for incipient faults in dynamic processes [J]. IEEE Transactions on Industrial Informatics, 2020, 17(8): 5391-5401.
- [13] GAO H H, HUANG W J, GAO X J, et al. Decentralized adaptively weighted stacked autoencoder-based incipient fault detection for nonlinear industrial

- processes[J]. ISA Transactions, 2023, 139: 216-228.
- [14] 李元, 辛梦媛. 基于分数阶傅里叶变换与卷积神经网络的工业过程故障检测[J]. 电子测量技术, 2024, 47(2): 1-8.
LI Y, XIN M Y. Fault detection of industrial processes based on fractional order Fourier transform and convolutional neural network [J]. Electronic Measurement Technology, 2024, 47(2): 1-8.
- [15] BI G, ZHENG S H, ZHOU L. Online monitoring of diamond grinding wheel wear based on linear discriminant analysis[J]. The International Journal of Advanced Manufacturing Technology, 2021, 115(7): 2111-2124.
- [16] FAN F L, CONG W X, WANG G. A new type of neurons for machine learning[J]. International Journal for Numerical Methods in Biomedical Engineering, 2018, 34(2): e2920.
- [17] FAN F L, SHAN H M, KALRA M K, et al. Quadratic autoencoder (Q-AE) for low-dose CT denoising[J]. IEEE Transactions on Medical Imaging, 2019, 39(6): 2035-2050.
- [18] LIAO J X, HOU B J, DONG H C, et al. Heterogeneous autoencoder empowered by quadratic neurons[J]. ArXiv preprint arXiv:2204.01707, 2022.
- [19] MAKHZANI A, SHLENS J, JAITLEY N, et al. Adversarial autoencoders[J]. ArXiv preprint arXiv: 1511.05644, 2015.
- [20] 谢彦红, 杨滕, 贾冬妮, 等. 基于主元分析的滑动窗口累积和的微小故障检测[J]. 计算机应用与软件, 2023, 40(5): 60-66, 96.
XIE Y H, YANG T, JIA D N, et al. Minor fault detection based on PCA and moving window cumulative sum [J]. Computer Applications and Software, 2023, 40(5): 60-66, 96.
- [21] ZHU Q X, DING W J, HE Y L. Novel multimodule Bayesian network with cyclic structures for root cause analysis: Application to complex chemical processes[J]. Industrial & Engineering Chemistry Research, 2020, 59(28): 12812-12821.
- [22] ZHANG C, GUO Q X, LI Y. Fault detection in the Tennessee Eastman benchmark process using principal component difference based on k-nearest neighbors[J]. IEEE Access, 2020, 8: 49999-50009.

作者简介

刘喜平(通信作者), 硕士研究生, 主要研究方向为基于数据驱动的工业过程早期故障检测及根源诊断。

E-mail: 1422100233@qq.com

高丙朋, 副教授, 主要研究方向为流程工业过程预测控制、间歇故障诊断。

E-mail: xjugaobp@xju.edu.cn