



## 基于双向生成对抗网络的工业过程故障检测

牟建鹏, 刘文韬, 熊伟丽

引用本文:

牟建鹏, 刘文韬, 熊伟丽. 基于双向生成对抗网络的工业过程故障检测[J]. *智能系统学报*, 2024, 19(5): 1199–1208.  
MOU Jianpeng, LIU Wentao, XIONG Weili. Industrial process fault detection based on bidirectional generative adversarial networks[J]. *CAA Transactions on Intelligent Systems*, 2024, 19(5): 1199–1208.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202306011>

## 您可能感兴趣的其他文章

### 基于生成对抗网络的人脸口罩图像合成

Masked face image synthesis based on a generative adversarial network

智能系统学报. 2021, 16(6): 1073–1080 <https://dx.doi.org/10.11992/tis.202012010>

### 基于互信息的多块 $k$ 近邻故障监测及诊断

Multiblock  $k$ -nearest neighbor fault monitoring and diagnosis based on mutual information

智能系统学报. 2021, 16(4): 717–728 <https://dx.doi.org/10.11992/tis.202007035>

### 利用残差密集网络的运动模糊复原方法

Image restoration with residual dense network

智能系统学报. 2021, 16(3): 442–448 <https://dx.doi.org/10.11992/tis.201912002>

### 生成对抗网络辅助学习的舰船目标精细识别

Fine-grained inshore ship recognition assisted by deep-learning generative adversarial networks

智能系统学报. 2020, 15(2): 296–301 <https://dx.doi.org/10.11992/tis.201901004>

### 基于生成对抗网络的机载遥感图像超分辨率重建

Super-resolution reconstruction of airborne remote sensing images based on the generative adversarial networks

智能系统学报. 2020, 15(1): 74–83 <https://dx.doi.org/10.11992/tis.202002002>

### 基于小样本学习的LCD产品缺陷自动检测方法

An automatic small sample learning-based detection method for LCD product defects

智能系统学报. 2020, 15(3): 560–567 <https://dx.doi.org/10.11992/tis.201904020>

DOI: 10.11992/tis.202306011

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20240828.1655.034>

# 基于双向生成对抗网络的工业过程故障检测

牟建鹏<sup>1</sup>, 刘文韬<sup>1</sup>, 熊伟丽<sup>1,2</sup>

(1. 江南大学 物联网工程学院, 江苏 无锡 214122; 2. 江南大学 轻工过程先进控制教育部重点实验室, 江苏 无锡 214122)

**摘要:** 标准双向生成对抗网络的模型结构由全连接层构成, 在进行故障检测时仅使用单个样本的过程特征进行统计量构建。因此, 提出了一种改进双向生成对抗网络的工业过程故障检测方法。该方法用降噪自编码器对样本进行预处理, 构建重构误差作为双向生成对抗网络的输入, 以减少异常样本中正常信息对异常信息的淹没, 增强模型对微小故障的检测能力; 并将长短时编解码器引入双向生成对抗网络模型中, 使得生成器在生成虚假样本的同时还可以关注当前时刻样本的过程特征和历史时刻样本间的关联性, 增强了模型对时间序列数据的检测能力。将所提故障检测方法应用于田纳西伊斯曼过程和实际磨煤机工业过程, 其在保证低误报率的同时, 提升了报警率, 并且具有良好的泛化性能。

**关键词:** 故障检测; 生成对抗网络; 长短时记忆网络; 自编码器; 重构误差; 时间序列; 田纳西伊斯曼过程; 磨煤机过程

中图分类号: TP277 文献标志码: A 文章编号: 1673-4785(2024)05-1199-10

中文引用格式: 牟建鹏, 刘文韬, 熊伟丽. 基于双向生成对抗网络的工业过程故障检测 [J]. 智能系统学报, 2024, 19(5): 1199-1208.

英文引用格式: MOU Jianpeng, LIU Wentao, XIONG Weili. Industrial process fault detection based on bidirectional generative adversarial networks[J]. CAAI transactions on intelligent systems, 2024, 19(5): 1199-1208.

## Industrial process fault detection based on bidirectional generative adversarial networks

MOU Jianpeng<sup>1</sup>, LIU Wentao<sup>1</sup>, XIONG Weili<sup>1,2</sup>

(1. School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China; 2. Key Laboratory of Advanced Process Control for Industry (Ministry of Education), Jiangnan University, Wuxi 214122, China)

**Abstract:** The standard bidirectional generative adversarial network (BiGAN) model structure consists of a fully connected layer and only considers the process characteristics of a single sample for statistical construction during fault detection. Therefore, an improved bidirectional generative adversarial network for industrial process fault detection is proposed. The method preprocesses the samples using a denoising autoencoder and constructs the reconstruction error as the input for the BiGAN. This approach reduces the masking of abnormal information by normal information in abnormal samples and enhances the ability of the model to detect small faults. Additionally, a long short-term codec is introduced into the BiGAN model, which enables the generator to focus on both the current process characteristics and the correlation between historical samples while generating false samples. This improves the ability of the model to detect time-series data. The proposed fault detection method was applied to the Tennessee Eastman process and an actual coal mill industrial process. This method improved the alarm rate while maintaining a low false alarm rate, which demonstrates good generalization performance.

**Keywords:** fault detection; generative adversarial networks; long short-term memory network; autoencoders; reconstruction error; time series; TE process; coal pulverization process

工业生产过程中发生故障会导致产品质量下降, 影响系统的稳定运行, 甚至危害人身安全。

因此故障检测在现代工业中发挥着越来越重要的作用, 及时准确地判断系统是否运行在正常工况下对保证生产过程安全性及产品质量具有重要意义<sup>[1]</sup>。工业过程中常用的故障检测方法之一是统计分析方法<sup>[2]</sup>, 主要包括主成分分析 (principle com-

收稿日期: 2023-06-07. 网络出版日期: 2024-08-29.

基金项目: 国家自然科学基金项目 (61773182).

通信作者: 熊伟丽. E-mail: [greenpre@163.com](mailto:greenpre@163.com).

©《智能系统学报》编辑部版权所有

ponent analysis, PCA)<sup>[3]</sup>、独立成分分析<sup>[4]</sup>和偏最小二乘法<sup>[5]</sup>等。并在此基础上衍生出了多种适用于非线性系统的方法,如核 PCA<sup>[6]</sup>、动态 PCA<sup>[7]</sup>和加权 PCA<sup>[8]</sup>等。这些基于投影的方法其性能取决于特征提取结果,从复杂过程中提取有效特征的能力受限,因此故障检测的能力受到影响。

深度学习作为一种基于数据驱动的故障检测方法,凭借其强大的特征提取能力引起了学者们的极大关注。典型的深度学习方法有卷积神经网络<sup>[9]</sup>、深度自编码器<sup>[10]</sup>和生成对抗网络 (generative adversarial network, GAN)<sup>[11]</sup>等。GAN 模型通过博弈对抗思想进行网络训练,具有模拟真实数据复杂分布的生成器及无需先验知识的判别器。文献 [12] 利用 GAN 模型生成数据,在小数据量故障样本的基础上,得到大量可靠数据集,从而实现基于小数据量样本的故障检测与诊断。文献 [13] 提出一种基于 Wasserstein 距离条件梯度惩罚生成对抗网络,以实现不平衡数据集的故障诊断。针对 GAN 模型存在的训练困难以及模式崩溃的问题, Yang 等<sup>[14]</sup>将双向生成对抗网络 (bidirectional generative adversarial network, BiGAN) 应用于故障检测,在数据空间与潜在空间中对数据样本进行联合判断,提升了检测效果,缩短了检测用时。为了进一步增强特征提取的有效性,提高模型整体性能,学者们从网络结构和训练方式等角度对 BiGAN 进行了改进。Patil 等<sup>[15]</sup>使用 PCA 方法对数据进行降维和特征提取,并用 BiGAN 模型检测异常网络流量,在测试阶段获得了更好的效果。Kaplan 等<sup>[16]</sup>提出了一种改进的 BiGAN 异常检测方法,通过增加额外的训练步骤,提高了 BiGAN 在异常检测任务中的性能。

上述 BiGAN 模型未考虑时间序列问题,在进行特征提取的时候会将重要的历史信息遗漏。循环神经网络 (recurrent neural network, RNN) 能够充分考虑样本间的时序关联性,增强模型特征提取能力。针对 RNN 训练过程中易出现梯度消失及梯度爆炸的问题, Hochreiter<sup>[17]</sup>提出了长短时记忆网络 (long short term memory, LSTM)。Yu 等<sup>[18]</sup>提出了卷积 LSTM 来描述过程数据的分布,并学习用于故障检测的时间序列数据的有效特征。Park 等<sup>[19]</sup>结合了自编码器和长短时记忆网络来检测罕见故障事件并实现不同类型故障的分类。

综上所述,本文考虑工业数据的时序相关性,并结合双向生成对抗网络的优势,提出一种基于重构误差和长短时编解码的双向生成对抗网络 (LSTMED-BiGAN using residual input, RIL-BiGAN),并将其应用到工业过程的故障检测中,主要贡献如下:

1) 进行数据预处理设计,采用降噪自编码器构建重构误差作为 BiGAN 的输入,以增大正常样本与异常样本的相对差异度,有效地减少统计量计算中正常信息对异常信息的淹没;

2) 将长短时自编码结构引入双向生成对抗网络,提取特征时不仅关注当前时刻样本,还对历史样本加以关注,在博弈过程中提升生成器、编码器以及判别器的性能,从而实现对高维时间序列的特征提取和故障检测。

## 1 BiGAN 的基本原理

生成对抗网络是一种无监督的深度学习模型框架,其结构如图 1(a) 所示。GAN 的原理主要基于博弈论中的零和博弈,该框架由生成器 G 和判别器 D 构成,二者相互竞争并同时优化其目标。将简单的噪声分布,如均匀分布或者高斯分布作为生成器的输入,生成器将其映射到与真实样本相同的数据空间,经过训练,生成与真实样本相似的模拟样本;判别器对模拟样本和真实样本进行标记,将二者区分。GAN 的目标函数公式为

$$\min_G \max_D V(D, G) = E_{x \sim P(r)} [\log(D(x))] + E_{x \sim P(i)} [\log(1 - D(x))]$$

式中:  $P(r)$  是真实数据的分布,  $P(i)$  是模拟数据的分布, D 的目标是最大化识别输入样本来自于真实样本的概率, G 的目标是生成与真实样本无限接近的模拟样本。

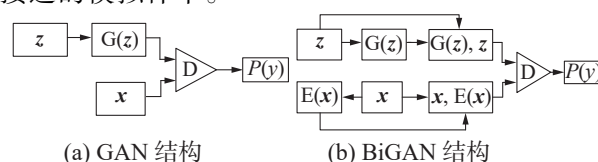


图 1 生成对抗网络结构及其改进

Fig. 1 Generate adversarial network structure and its improvement

BiGAN 模型结构如图 1(b) 所示。除了 GAN 框架中的生成器 G 和判别器 D 外, BiGAN 还包含编码器 E, 它可以将数据  $x$  映射到潜在表示  $z$ 。编码器 E 采用了 Dumoulin<sup>[20]</sup>提出的建模策略进行学习, 即  $E=G^{-1}$ 。BiGAN 中的判别器不仅在数据空间中对样本进行判断, 而且联合样本的潜在表示 (编码器输出和生成器输入) 进行判断。因此, BiGAN 可以避免在故障检测中获得潜在表示的昂贵计算步骤<sup>[14]</sup>。

判别器 D 被训练以最大化来自编码器输出的数据和来自 G 的样本的准确鉴别概率, 生成器 G 被训练以最小化  $1 - \log D(x, z)$ 。因此, BiGAN 训练目标可以被看作极小极大问题:

$$\min_{G, E} \max_D V(D, E, G)$$



$$V(D, E, G) = E_{x \sim P(x)}[\log D(x, E(x))] +$$

$$E_{z \sim P(z)}[\log(1 - D(G(z), z))]$$

$$\log D(x, E(x)) = E_{z \sim P_E(x|z)}[\log D(x, z)]$$

$$\log(1 - D(G(z), z)) = E_{x \sim P_G(x|z)}[\log(1 - D(x, z))]$$

式中:  $P(x)$  是真实数据的分布,  $P(z)$  是潜在表示的分布,  $P_E(x|z)$  和  $P_G(x|z)$  是从编码器和生成器  $G$  获得的分布。

## 2 基于 RIL-BiGAN 模型的故障检测

### 2.1 模型结构及统计量构建

所提 RIL-BiGAN 的模型结构从 2 方面进行改进: 一方面, 设计了一种数据预处理方式, 采用降噪自编码器对样本进行去噪重构, 得到重构误

差并将其作为模型输入, 从而增强对微小故障的检测能力。另一方面, 采用长短时编码器, 并将长短时解码结构引入生成器, 同时采用滑动窗对数据进行序列化采样, 以提高模型对工业过程时序特征的提取能力。本文将重构误差作为 BiGAN 的输入样本  $z$ , 样本  $x$  通过序列化处理后进入 LSTM 自编码器 (LSTM auto-encoder, LSTM-AE) 的编码端得到潜在表示  $E(x)$ , 将简单的噪声分布, 如均匀分布或者高斯分布作为生成器 LSTM-AE 的解码端 (BiGAN 的生成器) 的输入  $z$ , 生成与 BiGAN 的输入样本  $x$  相似数据分布的样本  $G(z)$ 。 $x$  与  $E(x)$ 、 $G(z)$  与  $z$  被送入判别器, 从而对样本正常与否进行联合判断。RIL-BiGAN 模型结构如图 2 所示。

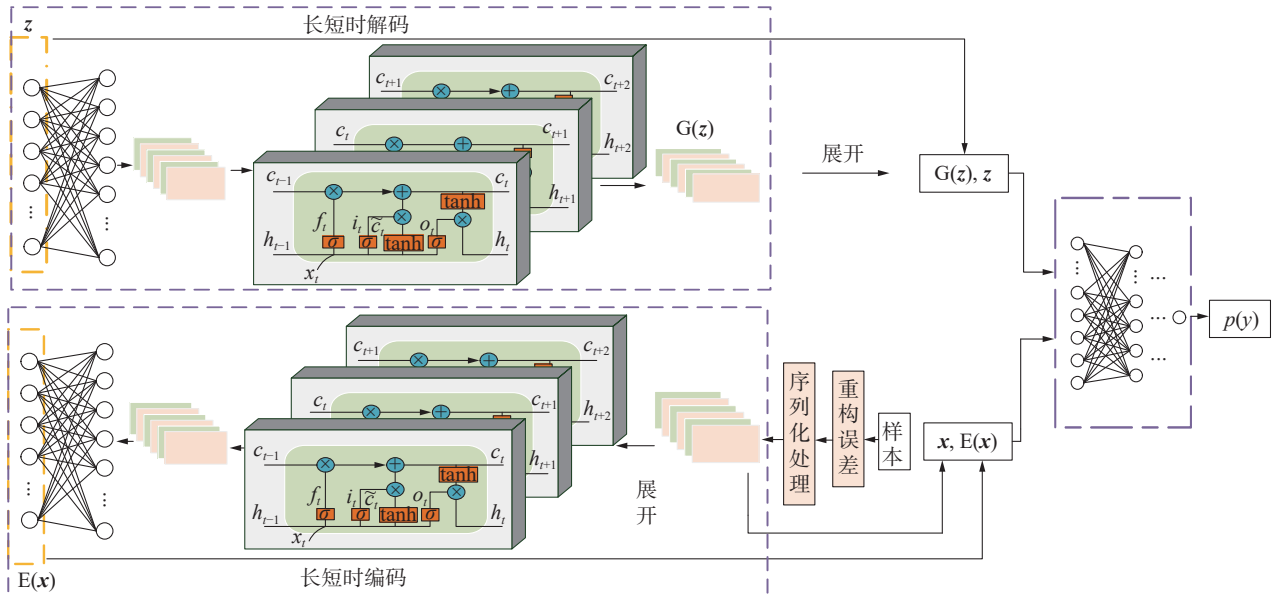


图 2 RIL-BiGAN 模型结构

Fig. 2 RIL-BiGAN model structure

RIL-BiGAN 方法中判别器结合待测样本的过程和时序特征判断样本是否为真实样本, 由于模型训练采用正常工况样本作为真实样本, 故判别器得到的结果可以判断待测样本是否为正常工况样本。判别器由全连接层组成, 各层神经网络个数分别为 256、128、64、10、1。若判别器认为待测样本与正常工况样本服从相同数据分布则判别器输出  $P(y)$  近似为 1, 反之判别器输出  $P(y)$  近似为 0, 其中  $P(y)$  为判别器最后一层神经元的输出, 对应样本正常的概率。结合判别器的上述性质, 其统计量构造公式为

$$D_{\text{score}} = 1 - P(y)$$

同时利用核密度估计公式求取控制限:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right)$$

式中:  $x$  为未知工况的待测样本,  $P(y)$  为样本  $x$  对

应输入时判别器的输出,  $\hat{f}(x)$  为样本  $x$  的预测概率密度;  $n$  为样本点个数,  $K(\cdot)$  为核函数,  $h$  为带宽。本文核函数采用高斯核函数, 带宽通过带交叉验证的网格搜索方法确定。

### 2.2 长短时编解码网络搭建

自编码器可以将原始数据经过编-解码操作, 得到重构后的数据集。重构后的数据集与原始数据集维度相同, 但因为中间的编码过程是将原始数据映射到低维的空间, 再进行解码, 过程中难免产生信息的损失。自编码器的重构效果与编、解码映射息息相关, 若编码器能提取的特征越好, 重构后的数据越能还原真实数据。普通的自编码器只能处理当前时刻输入数据, 但对于时间序列数据, 不仅要关注其当前时刻的数据, 对其前一段时间的数据信息也需要加以利用, 因此本文采用长短时记忆网络来进行编码和解码操作。

LSTM-AE 模型框架如图 3 所示。首先,采用滑动窗对工业过程数据进行序列化处理,以滑动窗宽度设置为 5 举例,以 1 为步长在原始数据  $\mathbf{X}$  上滑动,每次滑动得到新的样本  $\mathbf{x}^*$ ,滑动结束得到序列化采样的样本数据集  $\mathbf{X}^*$ 。在 LSTM-AE 中,输入

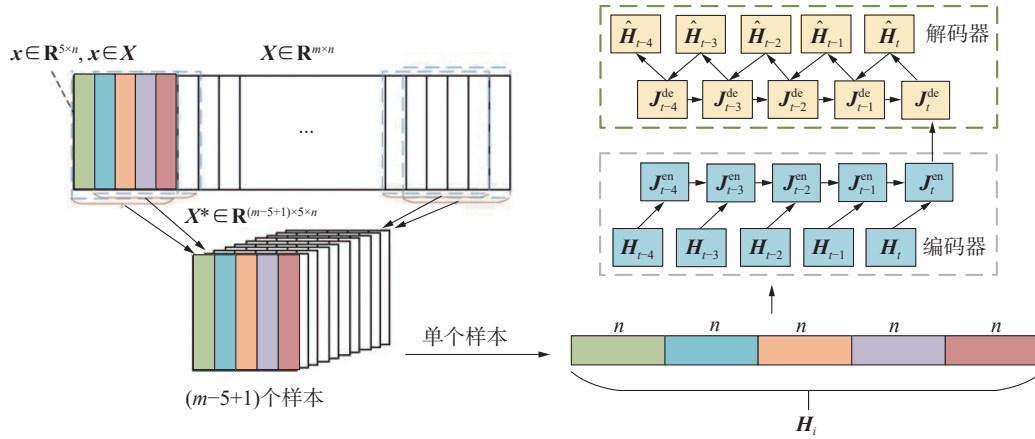


图 3 LSTM-AE 模型框架

Fig. 3 LSTM-AE model framework

从图 3 中可以看出, LSTM-AE 是一个分层模型,主要分为编码和解码 2 部分,两者呈镜像对称,即编码器的输入特征数和输出特征数与解码器的输出特征数和输入特征数分别对应,两者网络层数一致,中间隐藏层的特征数也是相等的。

采用长短时编解码的方式不仅能够提取样本的过程特征,还能对样本的时序特征加以利用,从而提升特征的有效性,增强模型的科学性。所提 RIL-BiGAN 采用对抗思想,不仅从数据本身中学习相关特征,而且通过学习数据以及其潜在空间的分布生成逼真的虚假样本对判别器产生干扰,从而在训练中增强判别器的判别能力,最终对样本的异常与否作出判断,增强模型的鲁棒性。

### 2.3 数据预处理设计

当生产过程出现异常工况时,通常是偏移量作用到过程变量上,若将这部分偏移量视作导致样本故障的异常信息,故障样本数据即可视作正常工况信息与异常信息的叠加,公式为

$$\mathbf{x}_{\text{train}} = \mathbf{x}_{\text{train}}^r + \mathbf{x}_{\text{train}}^f; \mathbf{x}_{\text{train}}^f = 0$$

$$\mathbf{x}_{\text{test}} = \mathbf{x}_{\text{test}}^r + \mathbf{x}_{\text{test}}^f; \mathbf{x}_{\text{test}}^f \gg \mathbf{x}_{\text{test}}^r$$

式中:训练样本为  $\mathbf{x}_{\text{train}}$ ,测试样本为  $\mathbf{x}_{\text{test}}$ ,上标 f 表示导致样本故障的异常信息,上标 r 表示样本中除异常信息外的信息,记作正常信息。当导致样本故障的异常信息仅发生在局部变量上或表现为微小偏移时,异常信息量级远小于正常信息。此时对原有测试样本  $\mathbf{x}_{\text{test}}$  直接计算统计量,异常信息易被正常信息淹没导致检测效果不理想,为此采用

长度为 5 的时间序列数据,输出也是长度为 5 的时间序列。输入数是向量形式,对于单个样本  $\mathbf{x}^*$ ,对其进行进一步展开,将二维矩阵转化为一维向量,转换后的向量用  $[\mathbf{H}_t, \mathbf{H}_{t-1}, \mathbf{H}_{t-2}, \mathbf{H}_{t-3}, \mathbf{H}_{t-4}]$  表示。对应解码输出为  $[\hat{\mathbf{H}}_t, \hat{\mathbf{H}}_{t-1}, \hat{\mathbf{H}}_{t-2}, \hat{\mathbf{H}}_{t-3}, \hat{\mathbf{H}}_{t-4}]$ 。

降噪自编码器构造的重构误差作为故障检测模型输入样本,其原理如图 4 所示。

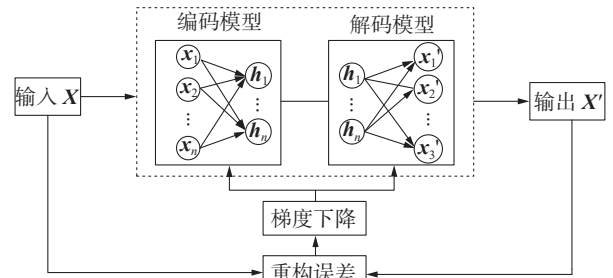


图 4 重构误差原理

Fig. 4 Schematic of reconstruction error

训练集编解码过程公式为

$$\mathbf{l}_{\text{train}} = f_c(\mathbf{w}_c \mathbf{x}_{\text{train}} + b_c); \hat{\mathbf{x}}_{\text{train}} = f_d(\mathbf{w}_d \mathbf{l}_{\text{train}} + b_d) \quad (1)$$

测试集正常信息编解码过程公式为

$$\mathbf{l}_{\text{test}} = f_c(\mathbf{w}_c \mathbf{x}_{\text{test}}^r + b_c); \hat{\mathbf{x}}_{\text{test}}^r = f_d(\mathbf{w}_d \mathbf{l}_{\text{test}} + b_d) \quad (2)$$

自编码器如式 (1) 所示将训练样本  $\mathbf{x}_{\text{train}}$  编码得到  $\mathbf{l}_{\text{train}}$  再经解码得到重构输出  $\hat{\mathbf{x}}_{\text{train}}$ , 又因为自编码器通过最小化损失函数确定模型参数, 因而训练好的自编码器模型对训练样本的重构输出近似于输入, 即  $\hat{\mathbf{x}}_{\text{train}} \approx \mathbf{x}_{\text{train}}$ 。模型对测试样本的正常信息重构输出如式 (2) 所示, 因为  $\mathbf{x}_{\text{test}}^r$  与  $\mathbf{x}_{\text{train}}$  为同一工业过程中的正常样本, 均服从同一数据分布  $\mathbb{D}_1$ , 因而根据式 (1) 和 (2) 可得  $\mathbf{l}_{\text{train}}$  与  $\mathbf{l}_{\text{test}}$  服从相同的数据分布  $\mathbb{D}_2$ ,  $\hat{\mathbf{x}}_{\text{train}}$  与  $\hat{\mathbf{x}}_{\text{test}}^r$  也服从相同的数据分布  $\mathbb{D}_3$ 。又  $\hat{\mathbf{x}}_{\text{train}} \approx \mathbf{x}_{\text{train}}$ , 可知  $\mathbb{D}_1$  与  $\mathbb{D}_3$  为相似的数据分布, 故  $\hat{\mathbf{x}}_{\text{test}}^r \approx \mathbf{x}_{\text{test}}^r$ 。

测试集编解码过程公式为

$$\hat{\mathbf{x}}_{\text{test}} = f_d(\mathbf{w}_d \mathbf{l}_{\text{test}} + b_d); \mathbf{l}_{\text{test}} = f_c(\mathbf{w}_c \mathbf{x}_{\text{test}} + b_c)$$

重构误差公式为

$$\begin{aligned} \mathbf{x}_{\text{test}}^* &= \mathbf{x}_{\text{test}} - \hat{\mathbf{x}}_{\text{test}} = \mathbf{x}_{\text{test}}^r + \mathbf{x}_{\text{test}}^f - f_d(\mathbf{w}_d \mathbf{l}_{\text{test}} + b_d) = \\ &\mathbf{x}_{\text{test}}^f - [f_d(\mathbf{w}_d \mathbf{l}_{\text{test}} + b_d) - f_d(\mathbf{w}_d \mathbf{l}_{\text{test}}^r + b_d)] = \mathbf{x}_{\text{test}}^f - \hat{\mathbf{x}}_{\text{test}}^f \end{aligned}$$

式中:  $\mathbf{x}_{\text{test}}^r$  为测试集正常信息,  $\hat{\mathbf{x}}_{\text{test}}^f$  为测试集异常信息引起的重构增量, 可以发现重构误差为测试集正常信息与测试集异常信息引起的重构增量的差。由于  $\mathbf{x}_{\text{test}}^r \gg \mathbf{x}_{\text{test}}^f$ , 重构误差的计算过程将测试集正常信息部分进行了抵消, 增大了训练集与测试集的相对差异性, 避免了统计量计算过程中正常信息数据量级远超异常信息数据量级而淹没异常信息数据的问题, 预处理效果在仿真实验得到了进一步验证。

#### 2.4 故障检测流程

本文提出的 RIL-BiGAN 方法分为离线建模和在线检测 2 部分。用正常工况数据建立模型, 得到控制限  $L$ , 通过比较待测样本的统计量  $D_{\text{score}}$  与控制限  $L$  的大小判断样本是否故障, 基于 RIL-BiGAN 的故障检测流程如图 5 所示。

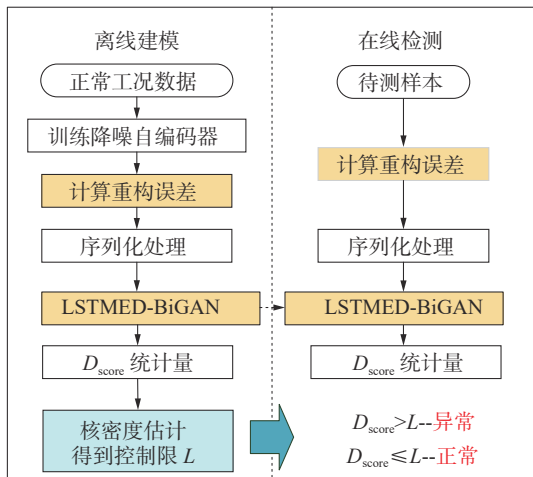


图 5 RIL-BiGAN 故障检测流程图

Fig. 5 RIL-BiGAN fault detection flow chart

离线建模具体过程如下:

- 1) 用正常工况数据集训练降噪自编码器, 得到重构误差;
- 2) 用滑窗对归一化后的重构误差进行序列化处理;
- 3) 用序列化后的数据训练 RIL-BiGAN 模型, 得到  $D_{\text{score}}$  统计量并计算控制限  $L$ ;

在线检测具体过程如下:

- 1) 对于待测样本, 提取重构误差, 用滑窗对归一化后的重构误差进行序列化处理;
- 2) 利用训练好的 RIL-BiGAN 模型得到待测样本的  $D_{\text{score}}$  统计量;
- 3) 通过比较待测样本的  $D_{\text{score}}$  统计量和控制限  $L$  判断样本是否故障。若  $D_{\text{score}} > L$ , 则样本异

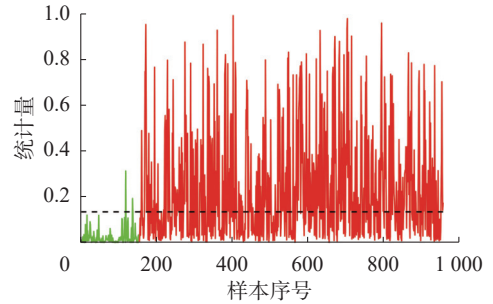
常, 反之样本正常。

### 3 实验结果与分析

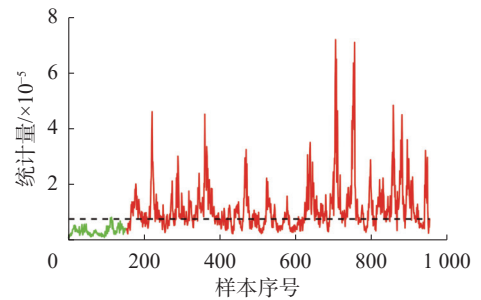
#### 3.1 TE 过程仿真

TE 过程是美国田纳西伊斯曼公司打造的工业模拟器<sup>[21]</sup>, 该过程数据符合高维时间序列的特征, 同时具备强耦合、非线性等过程数据的特点, 广泛应用于工业过程故障检测验证。TE 过程共有 22 个数据集, 本文采用了 1 个正常数据集和 18 个异常数据集, 去除了不易检出的故障 3、9、15 对应的异常数据集, 异常数据集均在 161 个本点引入故障, 建模所用变量包括 22 个过程测量变量和 11 个操作变量<sup>[22]</sup>。

利用 TE 过程数据对重构误差预处理方式的有效性进行仿真验证, 重构误差的预处理方式能够增大训练集与测试集的相对差异性, 从而避免统计量计算过程中正常信息数据量级远超异常信息数据量级而导致异常样本不易被检出的问题, 其中故障 19 报警率提升最大。图 6(a) 为 TE 故障 19 数据在不作预处理情况下正常样本统计量分布区间与正常样本统计量分布情况, 图 6(b) 中为数据在重构误差情况下正常样本统计量分布区间与正常样本统计量分布情况, 对比可知进行重构误差预处理的有效性。



(a) 不作预处理的统计量分布



(b) 经过预处理的统计量分布

图 6 训练集与测试集的相对差异性比较

Fig. 6 Comparison of relative differences between training and testing sets

用全连接层编码的 BiGAN 模型与长短时编码的 BiGAN 模型对 TE 过程测试样本进行测试,



编码器隐变量可视化效果如图 7 所示, 其中红色代表正常样本的隐变量, 蓝色代表异常样本的隐变量。标准编码器本质为多层感知机, 其编码得到的正常样本隐变量与异常样本的隐变量重合度较大, 不利于对样本的故障与否进行判别。而采用长短时编码的编码器, 能够增强对时序特征的提取能力, 提高特征提取的有效性, 增大正常样本隐变量与异常样本的隐变量的差异度, 在博弈过程中提高生成器生成数据的逼真程度, 进而联合判别器在数据空间与隐变量空间对样本类型进行判别。

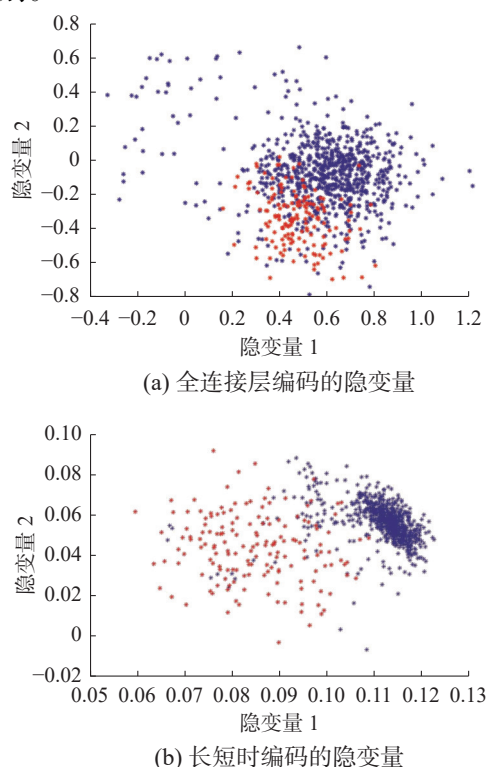


图 7 测试样本编码器隐变量散点图

Fig. 7 Scatter plot of hidden variables for testing sample

采用 TE 过程数据对 RIL-BiGAN 的故障检测性能进行验证, 首先用训练集训练降噪自编码器, 得到训练集的重构误差作为长短时编解码 BiGAN 模型的输入; 然后训练长短时编解码 BiGAN 模型, 生成器中随机噪声连接感知神经元个数为 100, 长短时网络输出维度为 33, 时序化处理时滑动窗口宽度设置为 8, 编码器结构参数与生成器相同, 判别器由多层感知机和 Dropout 组成, 输出层采用 Sigmoid 激活函数, 其余均采用 LeakyReLU 激活函数; 进一步求得训练集的统计量和控制限。对测试样本进行测试, 效果如图 8 所示, 其中绿色实线代表正常样本的统计量, 红色实线代表异常样本的统计量, 黑色虚线为控制限, 所提 RIL-BiGAN 方法能够有效地区分正常样本与异常样本, 从而进行故障检测。

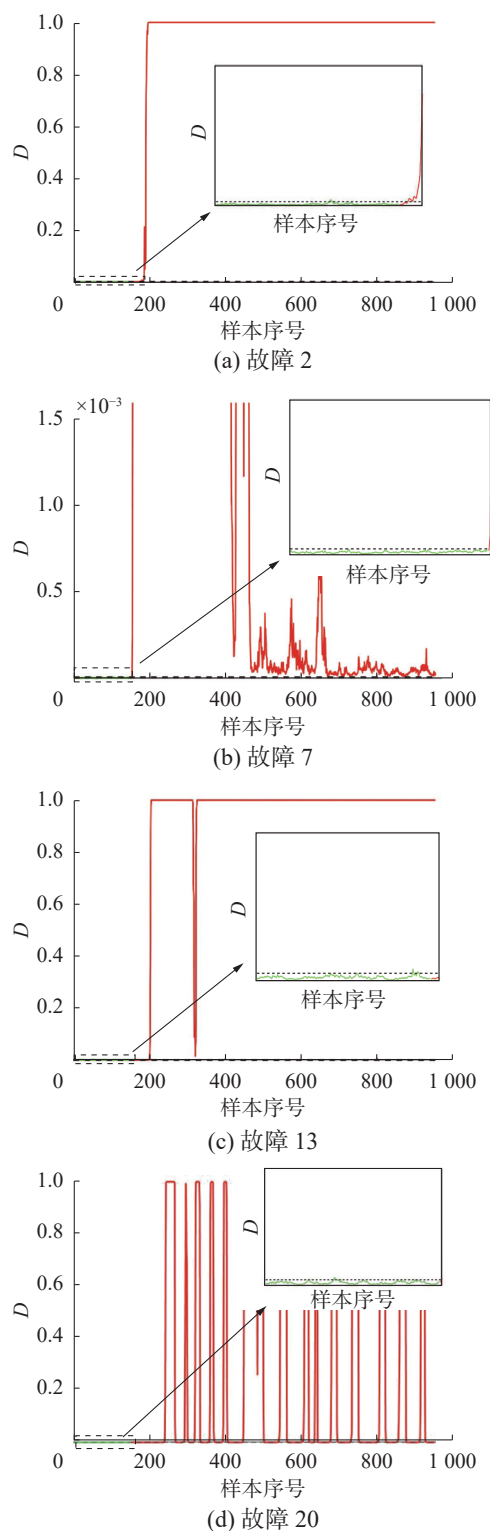


图 8 RIL-BiGAN 故障检测效果

Fig. 8 RIL-BiGAN fault detection effect chart

在保证低误报率的前提下, 将所提方法与 GAN<sup>[23]</sup>、EIGAN<sup>[24]</sup>、BiGAN<sup>[14]</sup>、LSTM-AE<sup>[18]</sup>、对抗训练的 AE<sup>[25]</sup> 方法进行对比, 检测结果如表 1 所示。与传统的 GAN 相比, 本文所提方法在报警率上提升显著。EIGAN 方法将训练集降维后的隐变量信息作为生成器的输入, 解决了生成器输入使用随机噪声, 不包含训练集有效信息的问题。

题, 提升了生成样本的逼真程度, 增强了故障检测能力, 而 BiGAN 方法将编码器与生成器进行联合训练, 并从数据与潜在空间中进行联合判断, 其效果优于 EIGAN 方法。与 BiGAN 相比, 本文所提方法一方面增加了对时序特征的提取, 另一方面通过对样本的预处理, 使得正常样本与异常样本数据分布的相对差异性增大, 在故障 10、16、19、20、21 上的报警率均有较大提高。与 LSTM-

AE 相比, 本文方法通过生成数据训练模型, 增强了模型的鲁棒性, 在故障 4、10、16、19、20、21 上均有着较好的表现。GAN 与对抗训练的 AE 都采用对抗训练的思想, 对抗训练的 AE 方法相比于传统的 AE 方法和 GAN 方法, 性能得到了提升, 在特征提取方面弱于 EIGAN、BiGAN 和 RIL-BiGAN 方法。通过对比, 可见 RIL-BiGAN 方法能够有效检测 TE 过程的故障。

表 1 TE 过程报警率对比  
Table 1 Comparison of TE process alarm rates

故障编号	GAN <sup>[23]</sup>	EIGAN <sup>[24]</sup>	BiGAN <sup>[14]</sup>	LSTM-AE <sup>[18]</sup>	对抗训练的AE <sup>[25]</sup>	RIL-BiGAN
1	0.98	0.99	0.99	<b>1.00</b>	<b>1.00</b>	0.99
2	0.97	0.98	0.98	0.98	0.99	<b>0.99</b>
4	0.22	0.59	<b>0.99</b>	0.78	0.92	<b>0.99</b>
5	0.26	<b>0.90</b>	0.59	0.37	0.33	0.42
6	<b>1.00</b>	<b>1.00</b>	0.99	<b>1.00</b>	<b>1.00</b>	0.99
7	0.91	0.89	0.99	<b>1.00</b>	<b>1.00</b>	0.99
8	0.94	0.91	0.97	<b>0.99</b>	0.98	0.97
10	0.50	0.64	0.55	0.67	0.54	<b>0.79</b>
11	0.44	0.72	0.90	0.64	0.72	<b>0.93</b>
12	0.90	0.94	0.98	<b>1.00</b>	0.99	0.99
13	0.94	0.90	0.94	<b>0.96</b>	0.95	0.95
14	0.93	0.99	0.99	<b>1.00</b>	<b>1.00</b>	0.99
16	0.55	0.65	0.51	0.65	0.42	<b>0.81</b>
17	0.84	0.93	<b>0.96</b>	0.91	0.92	<b>0.96</b>
18	0.89	<b>0.90</b>	<b>0.90</b>	<b>0.90</b>	<b>0.90</b>	<b>0.90</b>
19	0.19	0.42	0.55	0.13	0.20	<b>0.64</b>
20	0.50	0.68	0.66	0.61	0.59	<b>0.77</b>
21	0.34	0.54	0.50	0.50	0.48	<b>0.64</b>
平均报警率	0.69	0.81	0.84	0.78	0.77	<b>0.88</b>

3.2 磨煤机过程仿真

磨煤机过程数据来源于实际工厂, 磨煤机作为火力发电厂的核心设施之一, 其性能的稳定与否直接影响到电力系统能否可靠运行。在此背景下, 通过对磨煤机进行准确的故障检测, 对于保障电力系统的稳定运行具有至关重要的意义。该研究采用的磨煤机数据集分为训练集和测试集 2 部分, 其中训练集包含 3 500 个正常样本, 测试集则包含前 500 个正常样本和后 500 个异常样本。数据集涵盖了 46 个过程变量, 并已在表 2 中详细描述。

表 2 磨煤机过程变量描述  
Table 2 Description of process variables for coal mills

变量序号	变量描述	变量序号	变量描述
1 ~ 4	进气流量	26	主热气挡板阀开度
5	进气压力	27	分离器电流

续表 2

变量序号	变量描述	变量序号	变量描述
6 ~ 8	进气温度	28	分离器频率
9 ~ 13	煤原料	29	分离器速度
14 ~ 15	输出煤粉压力	30 ~ 35	电机定子温度
16 ~ 22	输出煤粉温度	36 ~ 37	电机轴承温度
23	主冷气挡板控制	38 ~ 40	转子轴承油温
24	主冷气挡板阀开度	41 ~ 42	推力轴承温度
25	主热气挡板控制	43 ~ 46	变速箱油池温度

将 GAN、EIGAN、LSTM-AE、RIL-BiGAN 方法检测效果进行比较, 故障检测效果以及报警率分别如图 9、表 3 所示。所提 RIL-BiGAN 方法无论在报警率还是误报率上均明显好于 GAN 方法, EIGAN 方法在 GAN 的基础上, 将训练集有效信息作为生成器的输入, 效果明显好于 GAN 方法, 报警率与 RIL-BiGAN 方法相差不大, 但误报



率较差。在 TE 过程中异常数目较多, 所提方法性能优于 LSTM-AE, 但对于磨煤机过程, 异常数目较少, 尽管 RIL-BiGAN 方法的报警率与误报率略差于 LSTM-AE 方法, 但其在误报率和报警率上也均有着较好的表现。综合来看, RIL-BiGAN 方法的泛化性能较好。

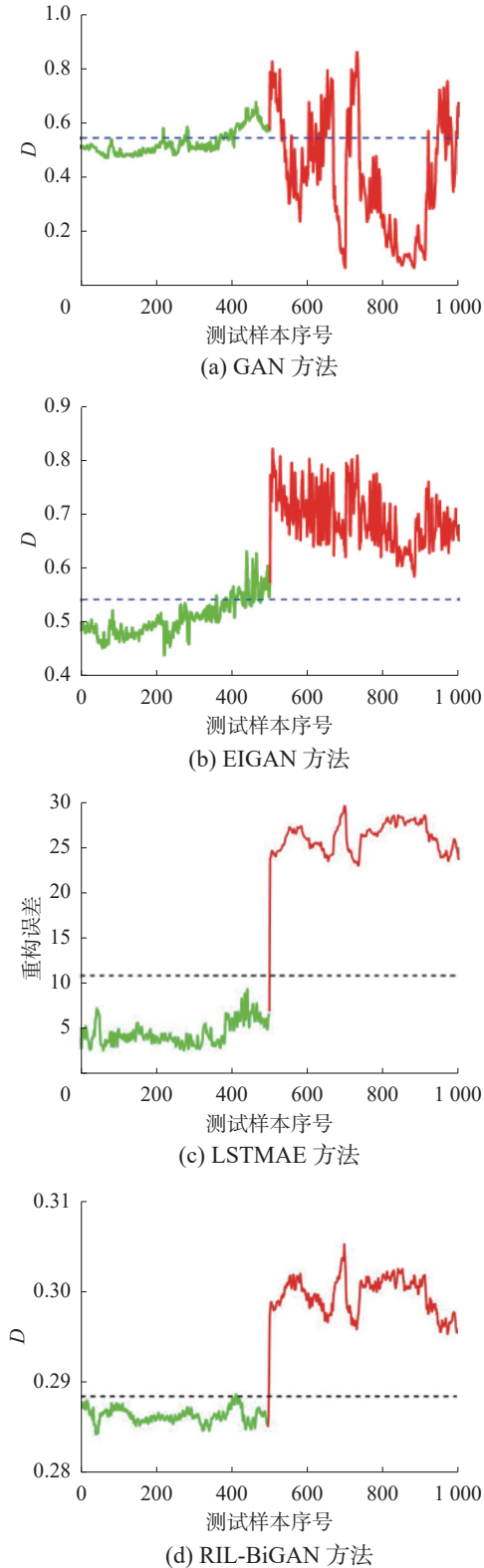


图 9 磨煤机过程故障检测结果

Fig. 9 Process fault detection results of coal mill

表 3 磨煤机过程检测结果对比

Table 3 Comparison of process inspection results for coal mills

指标	GAN	EIGAN	LSTMAE	RIL-BiGAN
误报率	0.152	0.212	0.000	0.006
报警率	0.295	1.000	0.998	0.994

采用贡献度法对磨煤机过程进行故障诊断, 结果如图 10 所示, 其中变量 38 对样本故障贡献度最大, 图 11 中 (a)、(b) 分别为磨煤机测试样本的 38 号变量在重构误差前后的变化趋势以及分布情况。重构误差前, 异常样本与正常样本的变量量级相同; 重构误差后, 异常样本量级远大于正常样本量级, 正常样本与异常样本的 38 号变量的相对差异度更大, 能够有效地避免统计量计算过程中正常信息数据量级远超异常信息数据量级而淹没异常信息数据的问题, 提高正常样本与异常样本的区分度, 增强模型的故障检测性能。

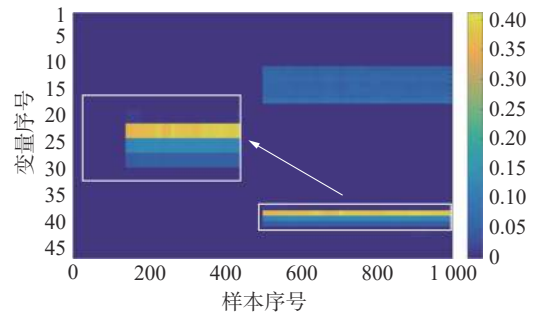


图 10 贡献度法故障诊断结果

Fig. 10 Contribution method fault diagnosis results

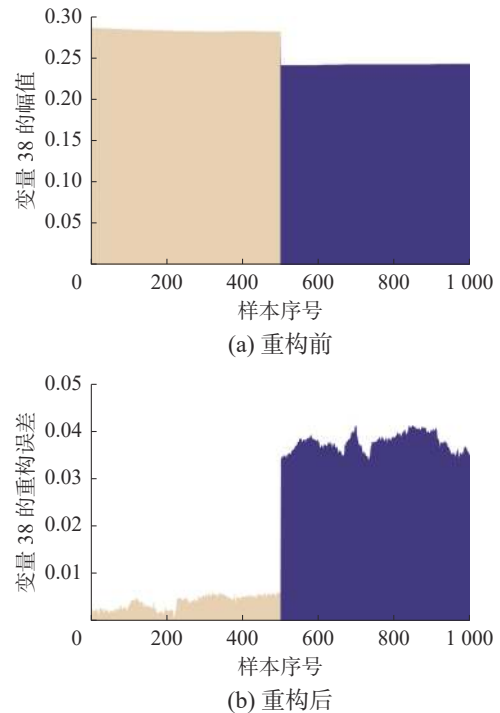


图 11 重构误差前后测试样本变量 38 分布条形图

Fig. 11 Bar chart of 38 distribution of test sample variables before and after reconstruction error

## 4 结束语

本文提出了一种改进双向生成对抗网络的工业过程故障检测方法。一方面, 重新设计样本预处理方法, 以样本与其降噪的重构输出的差作为 BiGAN 的输入, 减少了异常样本正常信息对异常信息的淹没, 增强了模型对异常样本的检出能力。另一方面, 进行 LSTM-AE 模型的编码与解码结构的设计, 在双向生成对抗网络编码阶段, 有效地提高了对数据的特征提取能力, 提升生成器生成数据的逼真程度, 在博弈过程中增强了模型对高维时间序列进行故障检测的性能。目前所提方法无法对过程的故障进行分类, 如何结合分类器实现故障诊断, 是本文需要进一步研究的问题。

## 参考文献:

- [1] YIN Shen, YANG Xuebo, KARIMI H R. Data-driven adaptive observer for fault diagnosis[J]. *Mathematical problems in engineering*, 2012(1): 832–836.
- [2] GE Zhiqiang. Review on data-driven modeling and monitoring for plant-wide industrial processes[J]. *Chemometrics and intelligent laboratory systems*, 2017, 171: 16–25.
- [3] JIANG Qingchao, YAN Xuefeng, HUANG Biao. Performance-driven distributed PCA process monitoring based on fault-relevant variable selection and Bayesian inference[J]. *IEEE transactions on industrial electronics*, 2016, 63(1): 377–386.
- [4] YU Gang. Fault feature extraction using independent component analysis with reference and its application on fault diagnosis of rotating machinery[J]. *Neural computing and applications*, 2015, 26(1): 187–198.
- [5] ZHONG Bin, WANG Jing, ZHOU Jinglin, et al. Quality-related statistical process monitoring method based on global and local partial least-squares projection[J]. *Industrial & engineering chemistry research*, 2016, 55(6): 1609–1622.
- [6] LEE J M, YOO C, CHOI S W, et al. Nonlinear process monitoring using kernel principal component analysis[J]. *Chemical engineering science*, 2004, 59(1): 223–234.
- [7] WANG Kai, CHEN J, SONG Zhihuan. Performance analysis of dynamic PCA for closed-loop process monitoring and its improvement by output oversampling scheme[J]. *IEEE transactions on control systems technology*, 2019, 27(1): 378–385.
- [8] DEWI Y N, RIANA D, MANTORO T. Improving Naïve Bayes performance in single image pap smear using weighted principal component analysis[C]//2017 International Conference on Computing, Engineering, and Design. Kuala Lumpur: IEEE, 2017: 1–5.
- [9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84–90.
- [10] SARATH C A P, LAULY S, LAROCHELLE H, et al. An autoencoder approach to learning bilingual word representations[J]. *Advances in neural information processing systems*, 2014, 3: 1853–1861.
- [11] GOODFELLOW I J, POUGET J, MIRZA M, et al. Generative adversarial networks[J]. *Advances in neural information processing systems*, 2014, 3: 2672–2680.
- [12] 朱晓荣, 张佩佩. 基于 GAN 的异构无线网络故障检测与诊断算法 [J]. *通信学报*, 2020, 41(8): 110–119.  
ZHU Xiaorong, ZHANG Peipei. Fault detection and diagnosis method for heterogeneous wireless network based on GAN[J]. *Journal on communications*, 2020, 41(8): 110–119.
- [13] 郭俊锋, 王森生, 孙磊, 等. 基于生成对抗网络的滚动轴承不平衡数据集故障诊断新方法 [J]. *计算机集成制造系统*, 2022, 28(9): 2825–2835.  
GUO Junfeng, WANG Miaosheng, SUN Lei, et al. New method of fault diagnosis for rolling bearing imbalance data set based on generative adversarial network[J]. *Computer integrated manufacturing systems*, 2022, 28(9): 2825–2835.
- [14] YANG Xin, FENG Dajun. Generative adversarial network based anomaly detection on the benchmark Tennessee Eastman process[C]//2019 5th International Conference on Control, Automation and Robotics. Beijing: IEEE, 2019: 644–648.
- [15] PATIL R, BIRADAR R, RAVI V, et al. Network traffic anomaly detection using PCA and BiGAN[J]. *Internet technology letters*, 2022, 5(1): 1–6.
- [16] KAPLAN M O, ALPTEKIN S E. An improved BiGAN based approach for anomaly detection[J]. *Procedia computer science*, 2020, 176: 185–194.
- [17] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural computation*, 1997, 9(8): 1735–1780.
- [18] YU Jianbo, LIU Xing, YE L. Convolutional long short-term memory autoencoder-based feature learning for fault detection in industrial processes[J]. *IEEE transactions on instrumentation and measurement*, 2021, 70: 1–15.
- [19] PARK P, DI MARCO P, SHIN H, et al. Fault detection and diagnosis using combined autoencoder and long short-term memory network[J]. *Sensors*, 2019, 19(21): 4612.
- [20] VINCENT D, ISHMAEL B, BEN P, et al. Adversarially learned inference. (2016–06–02)[2023–06–07]. <https://arxiv.org/abs/1606.00924>.

- [org/abs/1606.00704](https://arxiv.org/abs/1606.00704).
- [21] 潘丙翱. 基于自编码器的高维时间序列异常检测 [D]. 杭州: 浙江大学, 2021.  
PAN Bing'ao. Journal of chemical engineering anomaly detection of high-dimensional time series based on self-coder[D]. Hangzhou: Zhejiang University, 2021.
- [22] GE Zhiqiang, SONG Zhihuan. Distributed PCA model for plant-wide process monitoring[J]. Industrial & engineering chemistry research, 2013, 52(5): 1947–1957.
- [23] 宋励嘉, 童楚东. 基于缺失变量估计误差的工业过程监测方法 [J]. 高校化学工程学报, 2019, 33(1): 167–173.  
SONG Lijia, TONG Chudong. Industrial process monitoring based on estimation error of missing variables[J]. Journal of chemical engineering of Chinese universities, 2019, 33(1): 167–173.
- [24] 吴晓东, 熊伟丽. 采用编码输入的生成对抗网络故障检测方法及应用 [J]. 智能系统学报, 2022, 17(3): 496–505.  
WU Xiaodong, XIONG Weili. Fault detection method and its application using GAN with an encoded input[J]. CAAI transactions on intelligent systems, 2022, 17(3): 496–505.
- [25] AUDIBERT J, MICHIARDI P, GUYARD F, et al. USAD: UnSupervised anomaly detection on multivariate time series[C]//Proceedings of the 26th ACM SIGKDD

International Conference on Knowledge Discovery & Data Mining. Virtual Event: ACM, 2020: 23–27.

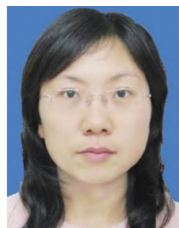
#### 作者简介:



牟建鹏, 硕士研究生, 主要研究方向为复杂工业过程建模和故障诊断。  
E-mail: [moujianpeng123@163.com](mailto:moujianpeng123@163.com)。



刘文韬, 博士研究生, 主要研究方向为系统辨识、复杂工业过程建模。  
E-mail: [wliu@126.com](mailto:wliu@126.com)。



熊伟丽, 教授, 博士生导师, 主要研究方向为复杂工业过程建模与监控、智能软测量技术。获得江苏省科学技术二等奖 1 项, 授权发明专利 30 项, 发表学术论文近百篇。E-mail: [greenpre@163.com](mailto:greenpre@163.com)。