



基于 VAE-DBN 的故障分类方法在化工过程中的应用

张 祥, 崔 哲, 董玉玺, 田文德*

(青岛科技大学化工学院, 山东 青岛 266042)

摘 要: 针对化工过程高维数据的故障特征难以提取的难题, 提出变分自动编码器(VAE)结合深度置信网络(DBN)的混合故障诊断方法。在 VAE 的编码部分对隐变量空间 Z 添加约束, 通过重参数化方法进行反向传播训练, 可无监督地学习不同故障对应的隐变量特征, 其作为 DBN 分类模型的输入特征训练网络, 输入测试集进行故障诊断。田纳西伊斯曼流程(TE)应用结果表明, VAE 能提取原始数据更加抽象有效的特征, VAE-DBN 分类准确。

关键词: 变分自动编码器; 深度置信网络; 故障诊断; 特征提取

中图分类号: TQ018; TP18; TP277

文献标识码: A

文章编号: 1009-606X(2018)03-0590-05

Application of Fault Classification Method Based on VAE-DBN in Chemical Process

Xiang ZHANG, Zhe CUI, Yuxi DONG, Wende TIAN*

(College of Chemical Engineering, Qingdao University of Science & Technology, Qingdao, Shandong 266042, China)

Abstract: To extract the fault feature from a large quantity of high-dimensional data, a variational auto-encoder (VAE) and deep belief network (DBN) combined fault diagnosis method was proposed for chemical process. In the encoding process of VAE, constraints were added to the latent variable space Z , and the backward propagation training was carried out by the re-parameterization method. The latent variables corresponding to different faults could be learned without supervision. Subsequently, the latent variable features learned by VAE were used as input features of the DBN classification model to diagnose the faults. The results showed that VAE could extract more abstract and effective features from the original data, and VAE-DBN had excellent performance in classification accuracy.

Key words: variational auto-encoder; deep belief network; fault diagnosis; feature extraction

1 前 言

特征提取是决定故障诊断准确性和效率的关键步骤。化工装置系统作为一种复杂的非线性动态系统, 组成关系与行为复杂。化工过程产生的多变量故障特征并非每个均能对故障诊断产生相同贡献。与传统的建立机理模型的故障诊断方法^[1-3]不同, 机器学习方法被用于对高维数据降维以提取数据特征。将原高维空间中的数据点映射到低维度空间中, 寻找数据内部的本质结构特征的方法有经典的主元分析(Principal Component Analysis, PCA)、独立成分分析(Independent Component Correlation Algorithm, ICA)、局部线性嵌入算法(Locally Linear Embedding, LLE)、多维标度分析(Multidimensional Scaling, MDS)、t-分布邻域嵌入算法

(t-distributed Stochastic Neighbor Embedding, t-SNE)等^[4-10], 这些方法仍然存在一些限制。PCA 仅限于线性数据, LLE 等非线性方法学习不到数据本质特征, 探索提取故障数据特征的新方法非常有必要。

传统的数据挖掘可通过交互信息和相关系数等特征工程进行数据特征提取和抽取。深度学习通过网络的权值捕捉外界输入模式特征, 并通过网络连接方式组合这些特征提取更高层的特征, 用此方法逐级从大量输入数据中学习得到对输入模式有效的特征, 再将其用于分类、回归和信息检索。

深度学习中自编码器是非常有用的一种无监督学习模型。自编码器由编码器和解码器组成, 前者将原始数据编码成隐层表示, 将数据投射到低维的隐变量空间提取数据的抽象特征; 后者将隐层表示解码成原始表示,

收稿日期: 2017-09-29, 修回日期: 2017-11-24

基金项目: 国家自然科学基金资助项目(编号: 21576143)

作者简介: 张祥(1992-), 男, 山东省济宁市人, 硕士研究生, 化学工程与技术专业; 田文德, 通讯联系人, E-mail: tianwd@qust.edu.cn.

引用格式: 张祥, 崔哲, 董玉玺, 等. 基于 VAE-DBN 的故障分类方法在化工过程中的应用. 过程工程学报, 2018, 18(3): 590-594.

Zhang X, Cui Z, Dong Y X, et al. Application of Fault Classification Method Based on VAE-DBN in Chemical Process (in Chinese). Chin. J. Process Eng., 2018, 18(3): 590-594, DOI: 10.12034/j.issn.1009-606X.217345.

训练目标为最小化重构误差. 自编码器已成功用于故障诊断领域^[11-14], Sun 等^[11]提出了将稀疏自动编码器 (Sparse Autoencoder, SAE) 与深度神经网络 (Deep Neural Network, DNN) 组合进行轴承故障诊断的 SAE-DNN 模型.

SAE 和降噪自编码器 (Denoising Autoencoder, DAE) 学习到的隐变量 (Latent Variable) 是自动学习的, 无任何先验约束. 本工作希望学习产生的表征空间 Z 符合某个分布, 实现无监督表征空间 Z 的约束. 变分自编码器 (Variational Auto-encoder, VAE)^[15] 是一类重要的生成模型 (Generative Model), 通过变分推断和重参数化 (Re-parameterization), 可在编码网络上添加一个约束, 使生成的隐变量 z 大致遵循标准正态分布, 可学习到数据更高层的特征. 将学习到的隐变量 z 用作其它分类算法的输入特征, 可提高分类和聚类操作的准确性. Hinton 等^[16]提出的深度置信网络 (Deep Belief Network, DBN) 用受限玻尔兹曼机 (Restricted Boltzmann Machine, RBM) 进行预训练^[17]. Zhang 等^[18]建立了多 DBN 混合模型, 用于田纳西-伊斯曼 (TE) 化工过程的故障诊断.

本工作选择 DBN 作为分类器, 将 VAE 结合 DBN 用于 TE 化工过程的故障分类, 使用变分自编码器学习原始数据更高层的抽象特征, 将学习到的数据特征输入到一个深层神经网络分类器, 训练分类器识别不同的故障, 在 TE 过程数据重叠较多的故障分类任务中取得了良好的分类效果.

2 基于 VAE-DBN 的故障诊断模型

2.1 变分自动编码器

VAE 是一种生成模型, 也是一种无监督的特征学习方法, 类似传统自编码器的编码和解码过程. VAE 对学习的表征空间添加约束, 用于约束高维空间的隐变量 z 满足高斯分布, 隐变量 z 能提取更加抽象有效的特征, 实现观测样本 x 到 z 的编码过程. VAE 形式上和自动编码器类似, 但与传统的 SAE 和 DAE 不同. VAE 的网络结构模型如图 1 所示, x 由 z 产生, z 到 x 是生成模型 $p(x|z)$, 可认为是自编码器中的解码器, x 到 z 是识别模型, $q(z|x)$ 类似于自编码器的编码器.

根据式(1)对观测样本 $p(x)$ 建模. 式(1)中, 引入条件概率 $p(x|z)$ 近似计算观测样本 x 的分布 $p(x)$. VAE 的核心思想为尝试采样, 得到最有可能产生观测样本 x 的隐变量 z , 基于 z 计算 $p(x)$. 通过构造新函数 $q(z|x)$ 表示输入观测样本 x 条件下, 产生与之相关的隐变量 z 的分布. VAE 采用了变分推理的原理构建 $q(z|x)$.

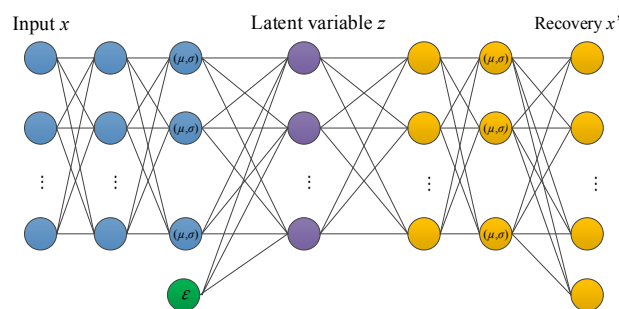


图1 VAE网络结构示意图
Fig.1 Schematic diagram of VAE

$$p(x) = \int_z p(x, z, O) dz = E_z[p(x|z, O)], \quad (1)$$

其中, O 为参数的集合, $p(x)$ 为观测样本 x 的分布, $p(x, z, O)$ 为观测样本 x 与隐变量 z 的联合概率分布, $p(x|z, O)$ 为 x 相对于 z 的条件概率.

目标是最大化似然的期望, 变分推理中优化的目标函数是变分的证据下界 ELBO 函数 $L(q)$:

$$L(q) = E_q\{\ln[p(x|z, O)] - \text{KL}[q(z|x, O) \| p(z, O)]\}, \quad (2)$$

其中, KL 散度用于衡量 2 个分布间的距离, 式(2)右边期望中的第 2 项 $\text{KL}[q(z|x, O) \| p(z, O)]$ 表示将 x 编码到隐变量 z .

假设观测样本 x 服从 $N[\mu(x, O), d(x, O)]$ 的高斯分布, 隐变量 z 服从 $N(0, 1)$ 的高斯分布, 式(2)中最右边的 KL 距离可简化表示为

$$\text{KL}\{N[\mu(x, O), d(x, O)] \| N(0, 1)\} = \frac{1}{2} \text{tr}[d(x)] + \left\{ \mu(x)^T [\mu(x) - k - \log \det d(x)] \right\}. \quad (3)$$

当求解高纬度或复杂分布时, 直接计算后验概率的积分不可行, 只有通过采样的方法近似求解似然函数. 本工作采用 MCMC 采样方法近似, VAE 的网络结构如图 2(a) 所示. 重参数化处理采样操作无法进行梯度反向传播, 将采样操作移动到输出层操作, 可从 $N[u(x), d(x)]$ 中结合 $e \sim N(0, 1)$ 采样得到 $q(z|x)$ 的均值 $u(x)$ 和协方差矩阵 $d(x)$, 计算 $z = u(x) + d^{1/2}(x) * e$. 目标函数转换为式(4), 采样方式换成相乘再求和. 式(4)的梯度在反向传播时可导, VAE 模型如图 2(b) 所示.

$$L(q) = E_q \left\{ \ln \left[p(x|z) = u(x) + d^{1/2}(x) * e \right] - \text{KL}[q(z|x, O) \| p(z, O)] \right\}, \quad (4)$$

经典的 VAE 模型利用 Decoder 部分生成新样本, 而本工作利用 Encoder 部分, 通过反向传播算法进行梯度下降, 训练整个网络. 根据 x 求解 z 的均值方差, 对

隐变量空间进行重构, 提取出原始数据更高层的特征,

避免了维度灾难和计算量大的问题.

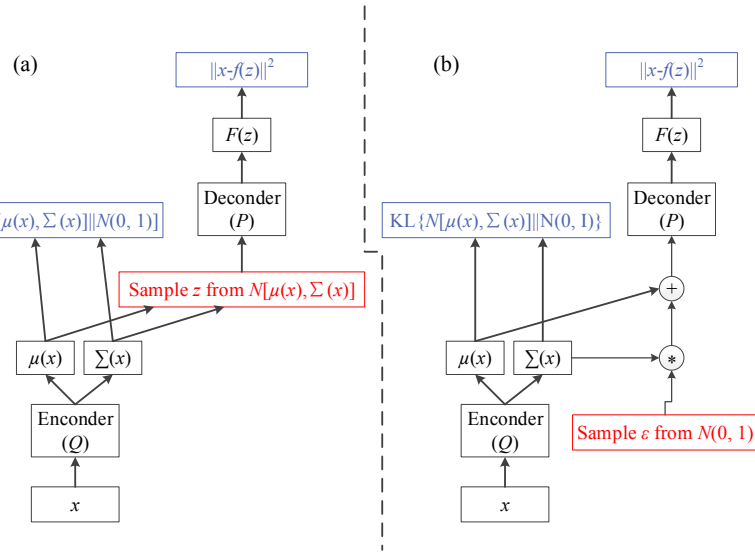


图2 VAE 训练结构

Fig.2 Network of training-time VAE

2.2 受限玻尔兹曼机

DBN 使用 RBM 进行预训练, 预训练过程相当于逐层训练每个 RBM, 经过预训练的 DBN 可用于模拟训练数据. 为进一步提高网络的判别性能, 利用标签数据通过反向传播算法对网络参数微调, 具体结构如图 3 所示.

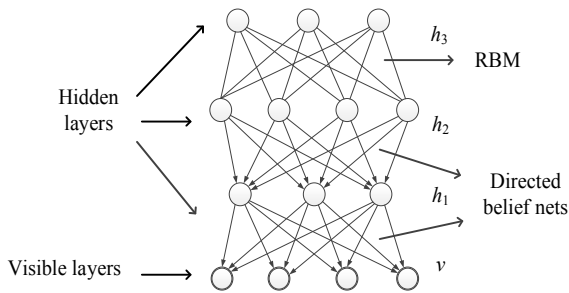


图3 DBN 结构示意图

Fig.3 Schematic diagram of DBN

RBM 可表示成一个二分图模型, 所有可见单元 v 和隐藏单元 h 之间存在着给定状态下的连接值 (v, h) . 假设可见层单元和隐藏层单元均服从伯努利分布, RBM 的能量公式为

$$E(v, h | \theta) = -\sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i W_{ij} h_j, \quad (5)$$

其中, $E(v, h | \theta)$ 为能量函数, a_i 为第 i 个可见单元的偏置阈值, b_j 为第 j 个隐单元的偏置阈值, W_{ij} 为第 i 个可见单元与第 j 个隐单元之间的权值, v_i 为第 i 个可见单元, h_j 为第 j 个隐单元, m 为隐单元数, n 为可见单元数.

同一层神经元具有独立性, 概率密度满足独立性:

$$p(h | v) = \prod_{i=1}^n p(h_i | v), \quad (6)$$

$$p(v | h) = \prod_{i=1}^m p(v_i | h), \quad (7)$$

其中, $p(h_i | v)$ 为隐层神经元 h_i 被显层神经元激活的概率, $p(v_i | h)$ 为显层神经元 v_i 被隐层神经元激活的概率.

参数求解采用似然函数求导方法, 已知联合概率分布 $p(v, h)$, 通过对隐藏层节点集合的所有状态求和, 可得到可见层节点集合的边缘分布 $P(v)$:

$$P(v, h | \theta) = \frac{e^{-E(v, h | \theta)}}{Z(\theta)}, \quad (8)$$

其中, $Z(\theta)$ 为归一化因子或配分函数, 表示对可见层和隐藏层节点集合的所有可能状态求和.

由于 RBM 模型具有特殊的层间连接、层内无连接的结构, 在给定可见单元的状态时, 各隐藏层单元的激活状态之间条件独立, 第 j 个隐单元的激活概率为

$$P(h_j = 1 | v) = \sigma(b_j + \sum_{i=1}^n W_{ij} v_i), \quad (9)$$

其中, σ 为 sigmoid 激活函数.

由于是双向连接, 显层神经元同样能被隐层神经元激活, 激活概率

$$P(v_i = 1 | h) = \sigma(a_i + \sum_{j=1}^m w_{ij} h_j), \quad (10)$$

2.3 VAE-DBN 故障诊断模型

无监督特征学习能从大量未标记数据中学习到的有效的数据特征. 在故障诊断领域, 标记故障耗时费力, 需要具体而详细的实验设置. 本工作提出了一种结合无

监督特征学习算法 VAE 的故障诊断方法，采用 VAE 将原始数据映射到低维的隐变量空间 Z，将学习到的隐变量输入到神经网络分类器中，VAE 和 DBN 的混合模型结构如图 4 所示。

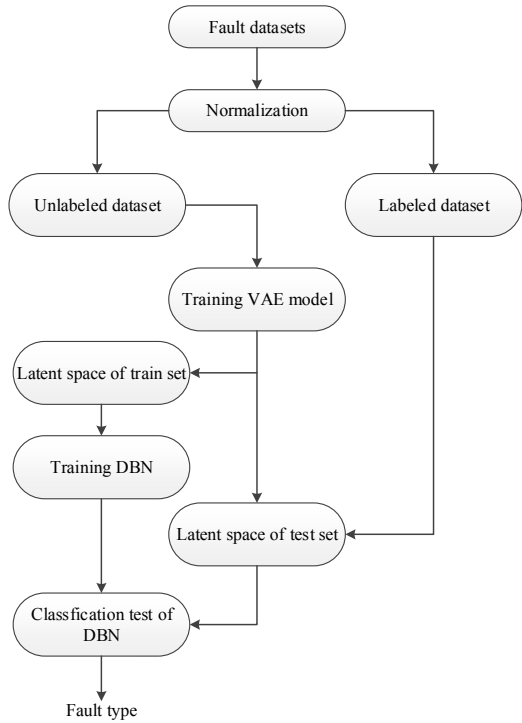


图 4 VAE-DBN 模型故障诊断流程
Fig.4 Procedure for fault diagnosis using VAE and DBN

3 应用案例

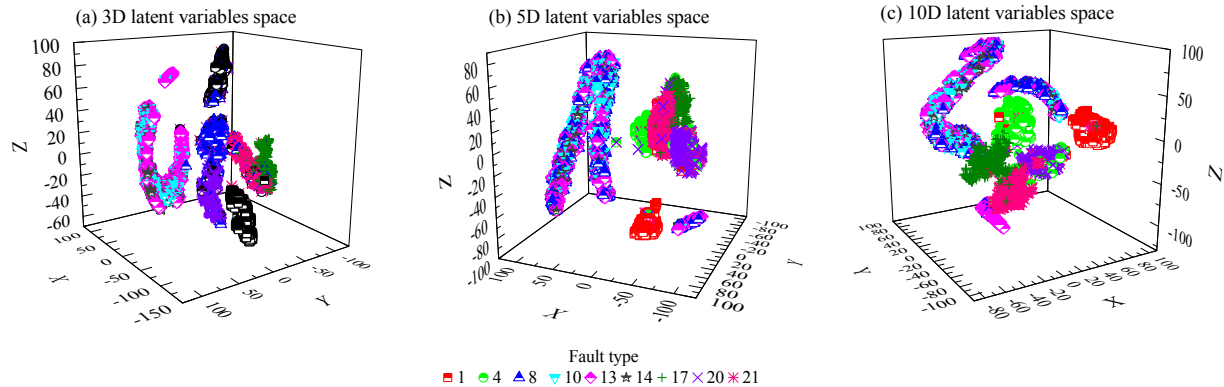


图 5 隐变量特征空间
Fig.5 Features of latent variable

3.2 对比实验

将 9 种故障的原始数据的训练集和测试集统一起来，通过 VAE 学习数据的潜在特征，再重新划分特征训练集和测试集。将特征训练集输入 DBN 进行训练，经过多次重复实验，选择的网络结构参数如表 3 所示。

3.1 特征学习

选取 TE 过程的 9 种类型的故障数据作为故障数据集进行学习，每类故障数据均归一化为[0, 1]之间的数据，具体的故障类型如表 1 所示。每种类型的故障数据集均包含 480 个训练样本和 800 个测试样本，VAE 的参数设置如表 2 所示。

用 t-SNE 的数据可视化技术展示隐变量特征，表明所提方法可有效地学习故障数据的潜在特征，具体效果如图 5 所示。相比于 Isomap 等无参方法，t-SNE 的参数可调节，结构更灵活，低维空间数据分布更好。从图可看出，VAE 可学习到原始故障数据的潜在特征，变量维度为 10 时故障能最大限度地被识别。

表 1 故障类型
Table1 Type of faults

Fault type	Variable name	Occurrence type
1	Change of APC feed ratio	Jump
4	Cooling water temperature of reactor	Jump
8	Changes of A, B & C components in feed 4	Random
10	Temperature variation of feed 4	Random
13	Variation of reaction kinetic characteristics	Drift
14	Unknown	Unknown
17	Unknown	Unknown
20	Unknown	Unknown
21	Unknown	Unknown

表 2 VAE 参数设置
Table 2 Parameters setting up of VAE

Input node	52
Encoder node	50
Decoder node	50
Learning rate	0.001
Dimension of latent variable	3/5/10

对同样的特征集合，VAE-DBN 与 SAE 的分类效果对比如图 6 所示。SAE 采用与 DBN 类似的预训练方式，通过逐层叠加的方式学习数据特征，最后输出分类结果。由图可知，所提方法对大部分故障分类的准确度均最高。VAE-DBN 比 SAE 分类准确度有较大提高，特别是对于

故障 20, VAE-DBN 的分类准确度提高了约 20%。对于故障 14, SAE 的分类准确度高。故障 4 和 14 是反应器冷却水温度变化导致的故障, 数据重叠较多, 两种方法的分类准确度均较低。

表 3 参数设置

Table 3 Parameters setting up of DBN & SAE		
Parameter	DBN	SAE
Input layer node	52	52
Hidden layer node	70/70/70	200
Output layer node	9	52
Momentum	0.3	0.4
Pre training learning rate	0.001	1
Sparsity ratio		0.4
Learning rate		1

VAE 可从无标注数据中学习到的较好的数据特征, 同时降低特征的维度, DBN 作为分类器使用可使分类错误率尽可能小, 两种方法结合可得到更好的数据特征, 使故障分类准确率更高。

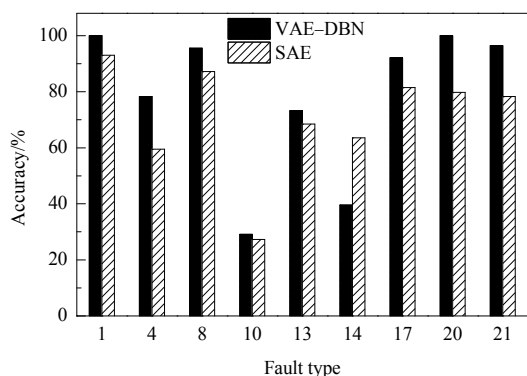


图 6 VAE-DBN 与 SAE 分类效果的对比

Fig.6 Comparison of classification performance of VAE-DBN and SAE

4 结 论

用无监督特征学习算法变分自动编码器(VAE)找到故障各自对应的潜在特征, 通过 t-SNE 将其映射到三维空间实现可视化, 并选取不同分类器进行对比, 得到以下结论:

(1) 通过对比潜在空间隐变量的维度, VAE 能得到最能反映原始数据特征的隐变量。

(2) 田纳西-伊斯曼(TE)应用结果表明, 将隐变量空间的潜在特征集合输入到分类器中进行故障分类, 比其它分类算法的精度更高, 有效提高了故障分类的准确性。

参考文献:

[1] Tian W D, Hu M G, Li C K. Fault Prediction Based on Dynamic

Model and Grey Time Series Model in Chemical Processes [J]. Chin. J. Chem. Eng., 2014, 22(6): 643–650.

- [2] Tian W D, Sun S L, Guo Q J. Fault Detection and Diagnosis for Distillation Column Using Two-tier Model [J]. Can. J. Chem. Eng., 2013, 91(10): 1671–1685.
- [3] 毛海涛, 田文德, 梁慧婷. 基于双层机器学习的动态精馏过程故障检测与分离 [J]. 过程工程学报, 2017, 17(2): 351–356.
- [4] Mao H T, Tian W D, Liang H T. Fault Detection and Isolation of Dynamic Distillation Process Using Two-tier Machine Learning [J]. Chin. J. Process Eng., 2017, 17(2): 351–356.
- [5] Tenenbaum J B, De Silva V, Langford J C. A Global Geometric Framework for Nonlinear Dimensionality Reduction [J]. Science, 2000, 290(5500): 2319–2323.
- [6] Roweis S T, Saul L K. Nonlinear Dimensionality Reduction by Locally Linear Embedding [J]. Science, 2000, 290(5500): 2323–2326.
- [7] Belkin M, Niyogi P. Laplacian Eigenmaps for Dimensionality Reduction and Data Representation [J]. Neural Comput., 2003, 15(6): 1373–1396.
- [8] Silva V D, Tenenbaum J B. Global Versus Local Methods in Nonlinear Dimensionality Reduction [J]. Advances in Neural Information Processing Systems, 2003, 15: 1959–1966.
- [9] Hinton G E. Visualizing High-dimensional Data Using t-SNE [J]. Vigiliae Christianae, 2008, 9(2): 2579–2605.
- [10] Hinton G, Roweis S. Stochastic Neighbor Embedding [J]. Advances in Neural Information Processing Systems, 2003, 41(4): 833–840.
- [11] Laurens V D M. Accelerating t-SNE Using Tree-based Algorithms [J]. Journal of Machine Learning Research, 2014, 15(1): 3221–3245.
- [12] Sun W, Shao S, Zhao R, et al. A Sparse Auto-encoder-based Deep Neural Network Approach for Induction Motor Faults Classification [J]. Measurement, 2016, 89: 171–178.
- [13] Lu C, Wang Z Y, Qin W L, et al. Fault Diagnosis of Rotary Machinery Components Using a Stacked Denoising Autoencoder Based Health State Identification [J]. Signal Process., 2017, 130(C): 377–388.
- [14] Shao H, Jiang H, Zhao H, et al. A Novel Deep Autoencoder Feature Learning Method for Rotating Machinery Fault Diagnosis [J]. Mech. Syst. Sig. Process., 2017, 95: 187–204.
- [15] Cheriadat A M. Unsupervised Feature Learning for Aerial Scene Classification [J]. IEEE Trans. Geosci. Remote Sens., 2013, 52(1): 439–451.
- [16] Na J, Jeon K, Lee W B. Toxic Gas Release Modeling for Real-time Analysis Using Variational Autoencoder with Convolutional Neural Networks [J]. Chem. Eng. Sci., 2018, 181: 68–78.
- [17] Hinton G E, Osindero S, Teh Y W. A Fast Learning Algorithm for Deep Belief Nets [J]. Neural Comput., 2006, 18(7): 1527–1554.
- [18] Hinton G E, Salakhutdinov R R. Reducing the Dimensionality of Data with Neural Networks [J]. Science, 2006, 313(5786): 504–507.
- [19] Zhang Z, Zhao J S. A Deep Belief Network Based Fault Diagnosis Model for Complex Chemical Processes [J]. Comput. Chem. Eng., 2017, 107: 395–407.