Applications of Deep Learning in Industrial Fault Diagnosis: A Survey

Authors: Lu Ziyang

Date: 2024-01-06T00:00:00+00:00

Abstract

In recent years, industrial processes have been developing towards increasing complexity and large-scale operations, and traditional fault diagnosis technologies have encountered a series of challenges when addressing practical industrial process problems. With the excellent performance and unique potential of deep learning in feature extraction and pattern recognition, the application of deep learning technology to fault diagnosis has become a focal point of current research. To this end, this paper introduces several typical fault diagnosis methods based on deep learning. Finally, it discusses the obstacles existing in the application of deep learning to fault diagnosis and prospects for future research directions.

Full Text

A Survey of Deep Learning Applications in Industrial Fault Diagnosis

Abstract

In recent years, industrial processes have been developing toward increasing complexity and large-scale integration, which has posed a series of challenges for traditional fault diagnosis techniques in solving practical industrial problems. With the superior performance and unique potential of deep learning in feature extraction and pattern recognition, the application of deep learning technology to fault diagnosis has become a current research focus. This article introduces several typical fault diagnosis methods based on deep learning. Finally, we discuss the obstacles in applying deep learning to fault diagnosis and prospect future research directions.

Keywords: Deep Learning; Fault Diagnosis; Industrial Processes

With the rapid development of computer technology, communication technology, and sensing technology, modern industrial processes are exhibiting new trends toward complexity, integration, and large-scale expansion. Consequently, system process data are showing trends of massive volume, low value density, and imbalance. Traditional data-driven fault diagnosis methods are ill-adapted to this new data paradigm. Simultaneously, the increasing integration of industrial processes has led to complex coupling between working units, causing systems to exhibit nonlinear characteristics and uncertainty that conventional modeling methods cannot effectively handle.

Generally, a fault refers to the deviation of single or multiple variables in a system from their predetermined trajectories, thereby severely impacting product quality or system performance. Fault detection and diagnosis technology monitors process data to determine whether a system has experienced a fault while also locating and identifying the fault type and position. However, the complexity and large-scale integration of modern industrial processes create intricate dependencies among components, making process data exhibit nonlinear and uncertain relationships. Traditional fault diagnosis methods targeting individual components and units struggle to extract and utilize such nonlinear relationships, resulting in high rates of misdiagnosis and missed detection.

In recent years, deep learning has developed rapidly in both academia and industry, significantly improving recognition accuracy in many traditional identification tasks and demonstrating superior capability in handling complex recognition problems, which has attracted numerous experts and scholars to investigate its theory and applications [1]. Currently, many scholars are attempting to apply deep learning theory to solve domain-specific problems.

Deep learning is a network model developed based on neural networks, characterized by containing multiple hidden layers. Compared with conventional neural network models, the "depth" in deep learning refers to the number of levels of nonlinear operation combinations learned by the network, which corresponds to the number of hidden layers. Deep learning networks possess more hidden layers and thus have better capability to approximate complex functions, offering superior feature extraction ability when facing nonlinear data.

After years of development, deep learning has achieved remarkable success and holds unparalleled advantages in expressing complex functions. Based on a review of domestic and international applications of deep learning in fault diagnosis in recent years, this paper discusses new problems and challenges in fault diagnosis for modern complex industrial processes and prospects the future development of deep learning-based fault diagnosis methods.

1 Deep Learning in Fault Diagnosis Research

The primary role of deep learning is to extract effective features from data and discover the intrinsic structure and essential relationships within datasets. Data representation is crucial for model performance. Good data representation can preserve task-relevant information while eliminating unnecessary interference and noise, thereby improving model performance. In recent years, several widely applied models include Deep Belief Networks (DBN), Stacked Autoencoders (SAE), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). To clearly elaborate on the application of deep learning in fault diagnosis, this paper introduces the principles and methods of fault diagnosis techniques based on these four approaches, along with relevant research.

1.1 Deep Belief Networks in Fault Diagnosis

DBN is a classic deep learning model that can extract feature information from data layer by layer through multiple layers of Boltzmann machines, ultimately forming more abstract high-level representations. This method simulates neural connections in the human brain and can extract essential features from data even with limited training datasets [2].

DBN is a deep neural network model composed of multiple Restricted Boltzmann Machines (RBM), as shown in [Figure 1: see original paper] DBN Model Structure. DBN training employs a greedy pre-training method: first, unsupervised learning algorithms train the network layer by layer to mine feature information from the data, then corresponding classifiers are connected to fine-tune network parameters through supervised learning, optimizing the DBN network's fault classification capability. Its key feature extraction ability is achieved through unsupervised layer-wise training, during which original information is transformed into abstract feature representations through progressive information transfer.

Compared with traditional manual feature extraction techniques, DBN-based feature extraction reduces uncertainty caused by human operations, enabling intelligent fault diagnosis. Additionally, DBN-based feature extraction eliminates dependence on extensive signal processing and empirical screening required by conventional methods, achieving intelligent fault feature extraction. Finally, its capability to handle high-dimensional nonlinear data can effectively avoid insufficient diagnostic performance [3]. Based on these advantages, the DBN model is well-suited for fault diagnosis challenges in the new era of industrial "big data" [4].

Currently, DBN models have been widely applied in signal processing. In speech and audio processing, DBN-HMM models for speech recognition have achieved competitive phone recognition accuracy. Convolutional DBN has been applied to audio and speech data for many tasks, including music artist and genre classification, speaker recognition, speaker gender classification, and phone classification, yielding excellent results. DBN has also found extensive applications in image and video processing, language processing, and information retrieval.

Shao et al. [5], Wang et al. [6], and Chen et al. [7] all employed DBN for rolling bearing and gearbox fault diagnosis, comparing their methods with existing mainstream fault diagnosis algorithms and verifying the robustness and accuracy of their proposed approaches. Subsequently, Li et al. [8] continued research based on previous work, investigating high-dimensional data anomaly detection and identification, rolling bearing fault diagnosis, gearbox deep fault feature extraction and identification, and high-background noise information extraction and fusion using DBN, achieving effects and advantages unmatched by traditional methods. Lei et al. [9] proposed an attention mechanism-based multi-scale DBN model, further enhancing the effectiveness of DBN models in bearing fault diagnosis. Wang et al. [10] proposed an Extended Deep Belief Network (EDBN) to improve DBN's fault classification capability by making full use of useful information in raw data. Applied to the Tennessee Eastman (TE) process for fault classification, results demonstrated that EDBN achieves better feature extraction and fault classification performance than traditional DBN. Wei et al. [11] proposed a DBN-dropout-based nonlinear process fault diagnosis and identification method to reduce DBN model overfitting and improve generalization capability. Tian et al. [12] addressed common issues in high-dimensional data identification using deep neural networks, such as insufficient data, missing data, measurement noise, redundant variables, and high data coupling, by proposing a feature-based DBN method. Application in the TE process showed that compared with traditional fault identification algorithms, this scheme achieves fast convergence and high accuracy in identifying chemical process abnormalities. With the development of DBN models, their application fields have expanded from sensor health diagnosis to compressors, gearboxes, rolling bearings, and beyond. However, research in process fault diagnosis remains limited, representing a significant area for future investigation.

DBN-based fault diagnosis methods can be broadly divided into the following steps: (1) Acquire system process data through sensors and perform preprocessing such as normalization and denoising; (2) Divide the collected signals into batches and split them into training and testing datasets; (3) Determine the number of input nodes for the DBN model based on the input signal format, then establish a multi-hidden-layer DBN model; (4) Adjust network parameters layer by layer through greedy pre-training, then connect corresponding classifiers to fine-tune the network parameters; (5) Finally, validate network effectiveness using the testing dataset.

1.2 Stacked Autoencoders in Fault Diagnosis

Stacked Autoencoders (SAE) can effectively extract features from raw data and constitute a deep learning model formed by stacking multiple autoencoders (AE). An autoencoder consists of an encoder and decoder, as shown in [Figure 2: see original paper] Autoencoder Structure. Its purpose is to find optimal parameters (W, b) that enable the decoded output y to reconstruct the input x as accurately as possible. A stacked autoencoder is a deep learning model formed by serially connecting the encoders of multiple autoencoders, as shown in [Figure 3: see original paper] Stacked Autoencoder Structure [4].

The training of autoencoder models still employs gradient descent to minimize the objective loss function. SAE training, like DBN networks, uses greedy pre-training. During the pre-training phase, each AE model is trained individually. When the lower-level model is trained, its encoder output serves as input to the next AE model, which is then trained separately. After training all AE layers, only the encoder portions are retained and concatenated to form the pre-model of the stacked autoencoder. During global fine-tuning, data labels are used as supervision signals to calculate network error, and backpropagation algorithms adjust parameters in the pre-trained model to reduce error. Stacked autoencoders cascade multiple AEs to build multi-layer neural networks, and their output can be regarded as feature representations of input data after multiple dimensionality reductions [13].

In recent years, the SAE algorithm has attracted increasing attention from experts and scholars due to its excellent feature extraction capability. Qiu et al. [14] proposed a data fusion algorithm SAEMDA combining stacked SAE with clustering protocols, which fuses similar features and transmits them to sink nodes. Simulation experiments show that SAEMDA can improve data fusion accuracy by 7.5% compared with BPFDA and SOFMDA algorithms while maintaining roughly equivalent network energy consumption. Demetgul et al. [15] addressed the unique nonlinear problems in material handling systems that traditional multi-objective methods cannot solve, proposing a feature extraction algorithm combining diffusion maps (DM), locally linear embedding (LLE), and AE for fault diagnosis based on measurement signals, using Gustafson-Kessel and k-medoids algorithms to classify encoded signals. Results demonstrate that this method improves fault diagnosis accuracy by 90% compared with traditional methods. Zhang [16] proposed a bearing fault diagnosis method based on improved stacked sparse autoencoders to address the problem of traditional fault diagnosis methods relying on feature design and expert experience. The method uses a training set to train a BAS-optimized stacked sparse autoencoder (BAS-SSAE) model and validates the model's fault diagnosis classification capability through a testing set. Then, t-distributed stochastic neighbor embedding (t-SNE) demonstrates that the constructed model can optimize features from raw data and mine more discriminative features. Finally, comparison with random-parameter SSAE, SAE, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) shows the proposed method's superiority in rolling bearing fault diagnosis. Sun [17] demonstrated the use of a sparse autoencoder-based deep neural network method for asynchronous motor fault classification, employing a sparse autoencoder model with unsupervised feature extraction advantages to learn fault features, effectively eliminating interference in feature extraction under the effect of denoising encoding and improving the robustness of feature representation. After SAE extracts features, they are used to train neural networks for identifying asynchronous motor faults. Experiments show this method's unique advantages in asynchronous motor fault diagnosis based on deep learning. Chen et al. [18] proposed a gear fault diagnosis method combining stacked sparse autoencoders (SSAE) with softmax classifiers. Experimental results demonstrate that compared with other shallow learning models in the paper, SSAE can effectively learn the required deep essential features from gear vibration signals and achieve higher recognition accuracy, confirming the method's superiority. The SAE algorithm has also achieved excellent applications in fault diagnosis for robotics, transformers, wind turbine equipment, and nuclear power plants.

Due to its strong feature extraction capability, the SAE algorithm has been continuously used for data feature extraction and fault classification in the fault diagnosis domain. Since both the encoder and decoder of SAE possess good feature extraction and classification capabilities, simple classification algorithms can achieve satisfactory diagnostic results after SAE-based feature extraction.

1.3 Convolutional Neural Networks in Fault Diagnosis

Convolutional Neural Network (CNN) is a very classic feedforward neural network that has been widely applied in image processing. Each CNN layer contains multiple convolution kernels, also called filters, which convolve and pool input signals layer by layer to obtain abstract data representations. Since CNN uses convolution kernels for forward propagation, reducing model parameters, it is highly suitable for processing and learning from massive datasets.

The structure of CNN is shown in [Figure 4: see original paper] Convolutional Neural Network Structure. The front portion consists of alternating convolution and subsampling operations, while the latter portion comprises a multi-layer fully connected neural network for classification, typically using the softmax function. Convolution operations extract certain features from the original signal through convolution kernels. Subsampling layers perform scaling mapping on convolved signals to reduce network complexity and avoid overfitting risks.

The CNN algorithm differs from DBN and SAE algorithms in that it is a supervised learning algorithm. Additionally, due to processing data with convolution kernels, this method reduces network complexity and decreases learning difficulty, making model training simpler. CNN training still uses gradient descent to adjust model parameters. CNN can extract local features from input data and progressively combine them to generate high-level features, effectively achieving fault diagnosis and identification [19].

Zhang et al. [20] proposed an enhanced deep convolutional neural network model to address the problems of mismatched positive/negative value calculations and parameter redundancy caused by low information flow efficiency in deep convolutional neural network activation processes. This method introduces a new activation mechanism based on the Maximum Smoothing Unit (MSF) function to overcome traditional activation function limitations and incorporates attention mechanisms combined with Gated Recurrent Units (GRU) to improve DCNN information flow efficiency and overcome parameter redundancy, comprehensively enhancing traditional DCNN model fault diagnosis performance. The enhanced deep convolutional neural network (EDCNN) model demonstrates significantly improved performance, validated through applications in industrial actuator control systems and industrial acid gas absorption processes. Chen et al. [21] proposed a deep learning algorithm implementing CNN for gearbox fault identification and classification, showing optimal performance in gearbox fault diagnosis compared with peer algorithms. Yan et al. [22] proposed a transfer learning gear fault diagnosis method based on Transformer and CNN to address the problem of insufficient sample data caused by complex operating environments for some gears. Compared with CNN without Transformer, multi-scale CNN, and 2D CNN, Transformer-CNN achieves higher average accuracy in gear fault diagnosis, reaching 99.64%.

After nearly a decade of development, fault diagnosis methods based on CNN have achieved significant breakthroughs and innovations. Currently, CNN-based fault diagnosis technology is often combined with attention mechanisms.

1.4 Recurrent Neural Networks in Fault Diagnosis

Recurrent Neural Networks (RNN) possess both internal feedback and forward feedback. Their internal feedback can transmit state information from previous nodes, providing the network with a memory mechanism where network outputs are no longer limited to current state inputs but also relate to the network's previous internal states. The RNN network structure is shown in [Figure 5: see original paper] Recurrent Neural Network Structure.

RNN has two structures: Jordan and Elman. In Elman RNN, there exists a context layer that combines the hidden state from the previous moment with the current moment's input as hidden layer input. Since Jordan model uses output feedback containing less information, its dynamic representation capability is weaker than the Elman model. When input-output sequences are too long, RNN models cannot effectively preserve information from distant nodes and suffer from gradient vanishing problems. To address this issue, the improved LSTM (Long Short-Term Memory) algorithm was proposed and has become the primary model currently applied in fault diagnosis.

The LSTM network contains forget gates, memory gates, and output gates. The forget gate's integration vector corresponds to cleaning outdated information in neuron hidden states. The memory gate's integration vector adds current moment state information to hidden states. The output gate's integration vector combines information to obtain model output. The difference between RNN and LSTM algorithms and DBN, SAE, CNN algorithms lies in the connections between RNN hidden layers, enabling the network to utilize correlation information between samples. Therefore, RNN is highly suitable for processing sequential data or data with temporal correlations.

Talebi et al. [23] employed two RNNs to detect and isolate unknown sensor or actuator faults in nonlinear systems with uncertain states and sensors containing interference, applying the method to low Earth orbit satellites, with extensive simulation experiments verifying the method's effectiveness and stability. Shao et al. [24] proposed a multi-channel LSTM-CNN method for chemical process fault diagnosis, using the Tennessee Eastman (TE) chemical process for experimental analysis and comparing the MCLSTM-CNN model with LSTM-CNN, LSTM, CNN, RF, and KPCA+SVM models. Experimental results demonstrate that the MCLSTM-CNN model achieves higher diagnostic accuracy with superior fault classification results compared with other models. Mahmoud [25] proposed a machine learning method AE-LSTM for intrusion detection, achieving the highest accuracy among detection methods (Dos, Probe, R2L, U2R, Normal). Zheng et al. [26] proposed a novel unsupervised data mining method for chemical process fault diagnosis based on stacked autoencoders, mainly comprising three steps: convolutional stacked autoencoder feature extraction, t-distributed stochastic neighbor embedding (t-SNE) algorithm feature visualization, and clustering, with experimental results demonstrating the method's effectiveness.

Due to its ability to utilize information between input samples, recurrent neural networks possess unparalleled advantages in prediction. In today's era of large models, RNN is receiving increasing attention in deep learning and will play an increasingly important role in industrial process fault diagnosis.

2. Obstacles in Deep Learning Applications for Fault Diagnosis

As industrial objects become larger and more complex, obtaining accurate and complete fault knowledge through mechanism analysis has become increasingly difficult. Deep learning-based fault diagnosis technology establishes deep neural networks that simulate human brain working principles to learn, understand, and analyze massive process data, adjusting network weights according to input data to enable the network to acquire the ability to interpret fault information from process data. Therefore, the key to deep learning-based fault diagnosis lies in the ability to thoroughly extract knowledge information from process data.

2.1 Characteristics of Large-Scale Complex Industrial Processes

Large-scale industrial systems exhibit specific characteristics due to their numerous functional units and complex internal connections between each unit:

(1) **Massive Data Characteristics:** With the development of sensing and storage technologies, large-scale industrial systems are equipped with numerous sensors, expanding data volume in spatial dimensions. From a temporal perspective, industrial system process data are continuously generated and accumulated, creating the massive data characteristics of complex industrial processes.

(2) **Complexity:** Large-scale industrial process data exhibit strong coupling and nonlinearity. Additionally, some early-stage faults often have very weak dynamic responses that are generally difficult to detect. Some faults are formed by coupling multiple factors with complex transmission paths

and short duration, making them difficult to detect.

(3) **Uncertainty:** The processes of data collection, transmission, and storage often contain substantial noise, making industrial system faults inherently uncertain.

2.2 Obstacles in Deep Learning Fault Diagnosis

(1) **Feature Extraction and Fault Mechanism Mapping:** In solving complex industrial system faults, feature extraction and fault mechanism mapping play crucial roles. Fault mechanisms represent the mapping relationships between equipment fault state signals and equipment system parameters obtained through theoretical or extensive experimental analysis. However, since obtaining complete fault data samples is often impractical, how to combine deep learning with existing fault mechanisms to address "correlation" problems in complex industrial system faults for effective system operating state feature extraction remains a challenging problem without obvious breakthrough progress [4,28-30].

(2) **Complex Fault Diagnosis:** While many effective "observation-inspection-inquiry-measurement" diagnostic methods have been proposed in fault diagnosis research, deficiencies remain in diagnosing early faults, weak faults, compound faults, and system faults, with limited reliable diagnostic methods available. During system operation, inevitable damage and early-stage faults occur with potential characteristics and weak dynamic responses. Compound and system faults, due to multi-factor coupling and complex transmission paths, often render single signal processing methods ineffective for tracing fault causes. The current effective solution for such complex faults is adding sensors to increase monitoring means for detection and diagnosis [4,28-30].

(3) **Small Sample Problem in Industrial Fault Diagnosis:** Massive normal operation data versus small-sample fault state data represent a typical characteristic of industrial big data, yet deep learning requires large amounts of training samples, creating an apparently irreconcilable contradiction. How to design a feature extraction model suitable for multi-source input data with good functionality to solve the complex industrial system "big data" challenge and achieve fault diagnosis is also a significant challenge.

(4) **Interpretability of Deep Learning:** As a black-box model, deep learning often suffers from poor interpretability in fault diagnosis, and its generalization capability is highly questioned. Effectively addressing the interpretability problem of deep learning-based fault diagnosis represents a major challenge.

(5) **Multi-Source Non-Uniform Training Samples:** Although industrial processes generate massive amounts of data, much of this data cannot

be effectively fused and utilized due to different distributions. How to construct deep learning methods that can adapt to and coordinate multi-source data also presents a significant challenge.

This article provides a brief introduction to deep learning applications in fault diagnosis. It first introduces the principles of several relevant deep learning algorithms—DBN, SAE, CNN, and RNN—along with recent domestic and international research on improved algorithms. Finally, it briefly introduces the characteristics of large-scale complex industrial systems, current challenges, and future research directions.

[1] Duan Y J, Lv Y S, Zhang J, et al. Deep learning for control: The state of the art and prospects[J]. Acta Automatica Sinica, 2016, 42(5): 643-654. [2] Dong Yu, Li Deng. Deep learning and its applications to signal and information processing[J]. IEEE Signal Processing Magazine, 2011, 28(1). [3] Haidong Shao, Hongkai Jiang, Xun Zhang, et al. Rolling bearing fault diagnosis using an optimization deep belief network[J]. Measurement Science and Technology, 2015, 26(11): 115002. [4] Ren Hao, Qu Jianfeng, Chai Yi, Tang Qiu, Ye Xin. Research status and challenges of deep learning in the field of fault diagnosis[J]. Control and Decision, 2017, 32(8): 1345-1358. [5] Haidong Shao, Hongkai Jiang, Xun Zhang, et al. Rolling bearing fault diagnosis using an optimization deep belief network[J]. Measurement Science and Technology, 2015, 26(11): 115002. [6] Wang Xingqing, Li Yanfeng, Ting Rui, et al. Bearing fault diagnosis method based on Hilbert envelope spectrum and deep belief network[J]. J of Vibroengineering, 2015, 17(3): 1295-1308. [7] Zhiqiang Chen, Chuan Li, René-Vinicio Sánchez. Multi-layer neural network with deep belief network for gearbox fault diagnosis[J]. J of Vibroengineering, 2015, 17(5): 2379-2392. [8] Chuan Li, René-Vinicio Sánchez, Grover Zurita, et al. Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis[J]. Neurocomputing, 2015, 168(C). [9] L. Xue, L. Ningyun, C. Chuang, H. Tianzhen and J. Bin, Attention mechanism based multi-scale feature extraction of bearing fault diagnosis[J]. Journal of Systems Engineering and Electronics, 2023, 34(5): pp. [10] Yalin Wang, Zhuofu Pan, 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). [11] Wei Y, Weng Z. Research on TE process fault diagnosis method based on DBN and dropout[J]. The Canadian Journal of Chemical Engineering, 2020, 98(6). [12] Wende Tian, Zijian Liu, et al. (PSE)identification of abnormal conditions in high-dimensional chemical process based on feature selection and deep learning[J]. Chinese Journal of Chemical Engineering, 2020. [13] Martin Langkvist, Lars Karlsson, Amy Loutfi. A review of unsupervised feature learning and deep learning for time-series modeling[J]. Pattern Recognition Letters, 2014, 42(1): 11-24. [14] Qiu Lida, Liu Tianjian, Lin Nan, et al. Data fusion algorithm for wireless sensor networks based on deep learning model[J]. Chinese Journal of Sensors and Actuators, 2014, 27(12): 1704-1709. [15] Demetgul M, Yildiz K, Tashin S, et al. Fault diagnosis on material handling system using feature selection and data mining techniques[J]. Measurement, 2014, 55(9): 15-24. [16] Zhang Lei. Research on bearing fault

diagnosis method based on improved stacked sparse autoencoder[D]. Hefei University of Technology, 2023. [17] Wenjun Sun, Siyu Shao, Rui Zhao, et al. A sparse auto-encoder-based deep neural network approach for induction motor faults classification[J]. Measurement, 2016, 89: 171-178. [18] Chen Lili, Si Jibing, Dong Shaojiang. Wind turbine bevel gear fault diagnosis based on stacked sparse autoencoder[J]. Machinery Design & Manufacture, 2021(3): 161-164. [19] Peng H-K, Marculescu R. Multi-scale compositionality: Identifying the compositional structures of social dynamics using deep learning[J]. PLOS one, 2015, 10(4):e0118309. [20] Zhang Jiaxin, Zhang Miao, Dai Yiyang, et al. Construction and application of enhanced deep convolutional neural network model for practical chemical process fault diagnosis[J/OL]. Chemical Industry and Engineering Progress, 2024, 1-14. [21] ZhiQiang Chen, Chuan Li, Rene-Vinicio Sanchez. Gearbox fault identification and classification with convolutional neural networks[J]. Shock and Vibration, 2015, 390134. [22] Yan Huiyu, Zhang Chao. Gear fault diagnosis method based on Transformer and convolutional neural network[J/OL]. Mechanical & Electrical Engineering, 2024, (1-11). [23] Talebi H A, Khorasani K, Tafazoli S, et al. A recurrent neural-network-based sensor and actuator fault detection and isolation for nonlinear systems with application to the satellite's attitude control subsystem[J]. IEEE Trans on Neural Network, 2009, 20(1): 45-60. [24] Bilin Shao, Xiaoli Hu, et al. A Multichannel LSTM-CNN Method for Fault Diagnosis of Chemical Process[J]. Mathematical Problems in Engineering. [25] M. Mahmoud, M. Kasem, A. Abdallah and H. S. Kang, AE-LSTM: Autoencoder with LSTM-Based Intrusion Detection in IoT[J], International Telecommunications Conference, 2022,(1-6). [26] Shaodong Zheng, Jinsong Zhao. A new unsupervised data mining method based on the stacked autoencoder for chemical process fault diagnosis[J], Computers & Chemical Engineering, 2020, 135. [27] Matthew J Daigle, Indranil Roychoudhury, Gautam Biswas, et al. A comprehensive diagnosis methodology for complex hybrid systems: A case study on spacecraft power distribution systems[J]. IEEE Trans on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2010, 40(5): 917-931. [28] Baoping Cai, Hanlin Liu, Min Xie. A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks[J]. Mechanical Systems and Signal Processing, 2016, 80: 31-44. [29] Li Deng. A tutorial survey of architectures, algorithms, and applications for deep learning[J]. Apsipa Trans on Signal & Information Processing, 2014, 3: e2. [30] Heshan Fernando, Brian Surgenor. An unsupervised artificial neural network versus a rule-based approach for fault detection and identification in an automated assembly machine[J]. Robotics and Computer-Integrated Manufacturing, 2017, 43: 79-88.

Note: Figure translations are in progress. See original paper for figures.