Hindawi
Science and Technology of Nuclear Installations
Volume 2022, Article ID 2675875, 12 pages
https://doi.org/10.1155/2022/2675875

Research Article

A Data-Driven Fault Prediction Method for Nuclear Power Systems Based on End-to-End Deep Learning Framework

Lu Chao,[1] Chunbing Wang,[1] Shuai Chen,[2] Qizhi Duan,[1] and Hongyun Xie[1]

[1]*State Key Laboratory of Nuclear Power Safety Monitoring Technology and Equipment, China Nuclear Power Engineering Co., Ltd., Shenzhen 518172, China*
[2]*Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei, Anhui 230031, China*

Correspondence should be addressed to Qizhi Duan; qzduan_cgn@163.com

Received 8 July 2022; Revised 20 October 2022; Accepted 5 November 2022; Published 29 November 2022

Academic Editor: Afaque Shams

With the increase in system complexity and operational performance requirements, nuclear energy systems are developing in the direction of intelligence and unmanned, which also requires a higher demand for its safety so that intelligent fault diagnosis and prediction have become a technology that nuclear power plants need to develop at present. At the same time, due to the rapid development of deep learning technology, it has become a meaningful development direction to predict the fault state of nuclear power plants within the framework of supervised deep learning. Usually, the network structure model used in fault diagnosis and prediction requires professional design, which may cost a lot of time and make it difficult to achieve optimal results. For this purpose, we present an end-to-end deep network for nuclear power system prediction (EDN-NPSP), which can automatically mine the transient features of various detection data in the NPS at the current moment through heterogeneous convolution kernels that can increase the receptive field and then predict the feature evolution results of the NPS in the future through a special deep CNN. The results provide an assessment of the future state of NPS. Based on EDN-NPSP presented in this work, we can avoid complicated manual feature extraction and provide the predicted state directly and rapidly. It will provide operators with useful prediction information and enhance the nuclear energy system fault prediction capabilities.

1. Introduction

Fault diagnosis technology is an important part of the operation support system of nuclear power plants, ensuring the safety and reliability of nuclear energy production [1]. Various faults occur in the operation of nuclear power plants, which greatly impact the safety of the reactor. At the same time, with the increase in the complexity of nuclear energy systems and the level of safety requirements, the prediction of system states, especially fault states, has become a capability that nuclear energy systems need to have in the future [2].

From the perspective of diagnostic methods, the fault diagnosis and fault prediction of the nuclear energy system at the current moment have a similar theoretical basis, that is, based on the current and historical data given by using various sensors in the nuclear energy system, the corresponding features are extracted, and the fault classification is carried out based on these features [3]. According to the analysis of Ma and Jiang, the threshold detection method mainly relies on the status data provided by the sensors of the nuclear energy system at the current moment and sets the thresholds of these status data (including hard thresholds or functional thresholds) [2]. Once a certain state data exceeds a threshold, the nuclear energy system is considered to have a fault. Different from the threshold detection method, the expert diagnosis system makes use of the long-term accumulated experience and professional knowledge of domain experts to make reasoning diagnoses and give diagnostic results [4]. As the level of safety requirements of nuclear energy systems increases, the structure of nuclear energy systems becomes more and more complex, which may reduce the effectiveness of relying only on current data for threshold judgment or failure analysis based on historical experience.

Recently, with the rapid progress of machine learning and data mining methods, data-driven fault diagnosis methods for the nuclear power system (NPS) (all abbreviations definitions in this paper are presented in Table1) have been widely developed [5]. The early data-driven diagnostic method typically performs artificial feature extraction from sensor signals, compares with baseline features, and makes a final diagnostic decision [6], whereas traditional signal transformation methods such as Fourier and wavelet, as well as feature dimension reduction methods such as PCA and t-NSE, are effective feature extraction methods [7–9]. However, with the increasing complexity of the nuclear energy system, it is difficult for this method to effectively mine the correlation characteristics between various data in the nuclear energy system. Benefiting from the powerful nonlinear fitting ability of machine learning, such complex correlation characteristics can be effectively extracted in fault detection and diagnosis (FDD) based on machine learning methods, especially deep learning methods [2, 5]. The most widely developed data-driven diagnostic method at present is based on a deep neural network (DNN) [10, 11]; that is, the various equipment or system-level data in the reactor are directly sent to the DNN constructed by various types of network layers, such as convolution, pooling, sparse activation, and FcNN, and then the state judgment result is directly given. In addition to the diagnostic methods, another notable issue in data-driven diagnostics is how to organize system state data. In the early stage, the input of fault diagnosis systems mainly depends on a single sensor time series or the state vector at the current moment. With the development of tensor techniques, tensor-based diagnosis methods have been gradually being applied to the fault diagnosis of mechanical and electrical systems [12, 13]. Recently, tensor technology has been gradually applied in complex industrial systems and can be coupled with machine learning models, for example, the deep learning model based on tensor factorization developed by Luo et al. [14], analog circuit fault diagnosis method based on tensor product wavelet developed by Jin et al. [15], and rotation machinery fault diagnosis method based on supervised second-order tensor developed by Wei et al. [16].

Recently, in addition to the diagnosis process based on current data, we also need to predict the future dynamic behavior of the nuclear energy system to meet the needs of fault prediction [17]. In recent years, many scholars have begun using the physics-informed neutral network (PINN) framework to carry out the transient analysis of various physical parameters of complex dynamic systems [18]. For example, Raissi et al. constructed a PINN framework to calculate the velocity and pressure field in the fluid by using the information in the visualized fluid scatter plot [19]. Schiassi et al. used the PINN method to solve the core dynamics equation of a nuclear reactor [20]. In addition to this, PINN has also been proven to be widely effective in the fields of environmental analysis, economic systems, and material computing [21], so that it can be an effective method for predicting the state of nuclear energy systems and can be coupled with neural networks for fault diagnosis method.

Table 1: List of abbreviations.

Abbreviation	Full name
CNN	Convolutional neural networks
DNN	Deep neural networks
DPU	Distributed processing unit
FDD	Fault detection and diagnosis
IC	Initial condition
I&C	Instrument and control
IPC	Industrial PC
KNN	k-nearest neighbor
MSE	Mean square error
NPP	Nuclear power plant
NPS	Nuclear power system
PCA	Principal component analysis
PINN	Physical-informed neural networks
SVM	Support vector machine
CGN	China General Nuclear Power Group

In this work, we present an end-to-end deep network for nuclear power system prediction (EDN-NPSP) for nuclear power system fault prediction. This network contains three parts: the first part is a simple PINN, which can predict the state of the NPS in the future; the second part is transient features extraction layers, which can automatically mine the transient features of various detection data in the NPS at different time scales through the hole convolution that can increase the receptive field; and the third part is a fault diagnosis network, which can identify the fault state of the NPS in future. Based on a nuclear power plant simulator and the real instrument and control (I&C) system, we test 20 typical transient conditions, and the result demonstrates the effectiveness of EDN-NPSP.

2. Method

In this section, we will introduce the deep network EDN-NPSP from three aspects, as described in the introduction, including the PINN for the state prediction, the feature extraction network, and the fault identify network. Additionally, a data organization process will also be introduced. The overall architecture of the EDN-NPSP is shown in Figure 1.

2.1. Data Organization. Consider an NPS with n sensors, which can measure various variables, including core power, fuel temperature, coolant temperature, and mass flow. Simultaneously, it is assumed that the sampling period of each sensor is T. The data series presenting the operation state of the NPS is shown in Figure 1, where V_i means the time series data of ith sensor in the NPS. v_i^k means the kth sampling data in V_i. We introduce an observation window [22] with the size of $n \times m$ to cover the time series data cluster $\mathbf{V}_{ii=1\ldots n}$, where n is the number of sensors in the NPS. Then, a dynamic matrix $\vee$ containing the operation data from T_0 to $T_0 + mT$ can be built.

If the data flow direction is defined as shown in the blue arrow in Figure 2, $\mathbb{V}$ will become a tensor. So that, at time $T_0 + mT$, the dynamic tensor $\mathbb{V}_{T_0}$ can be built, which

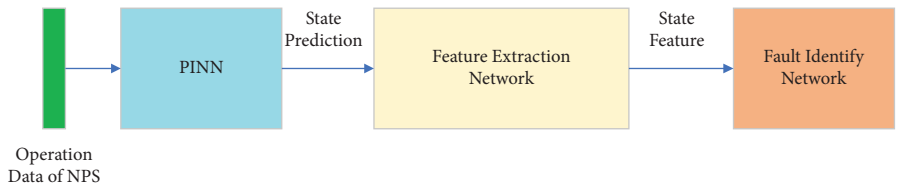

Figure 1: Three components of EDN-NPSP.

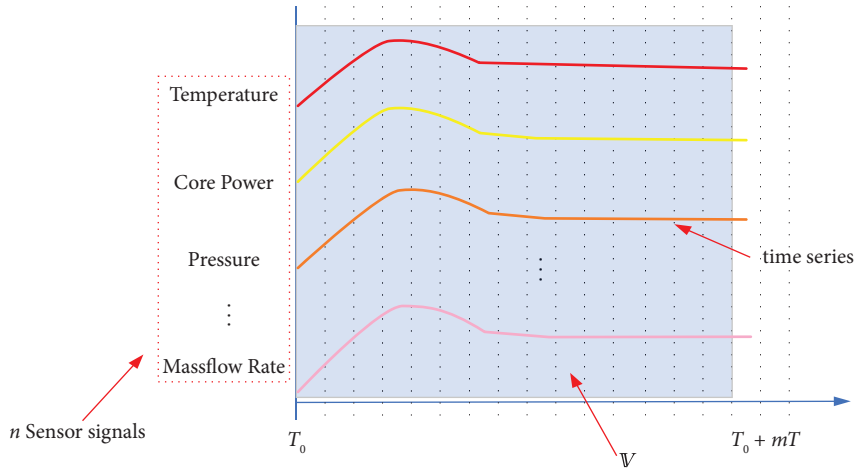

Figure 2: Data series and observation window of NPS with the size of $n \times m$.

contains all the data from T_0 to $T_0 + mT$. The tensor $\mathbb{V}_{T_0}$ is regarded as the analytic target of the fault prediction.

2.2. State Prediction. To provide system states at different moments in the future, we will construct a PINN for state prediction of the NPS, where the tensor $\mathbb{V}_{T_0}$ is the input. We refer to the UNet structure and design a low-dimensional feature mapping scheme with 3 down-sampling. Although UNet was first used in classification tasks, it is also commonly used in regression tasks in recent years due to its good low-dimensional mapping ability and data reconstruction ability, which is the reason why we use UNet as the State prediction network. The network has 4 levels, and each level includes two conv + ReLU modules; the down-sampling is implemented by the maxpool module. To improve the fitting accuracy and training efficiency of the network, we constructed a skip connection at each level. The detailed network structure is shown in Figure 3.

2.3. Feature Extraction and Fault Identify. Faults identified in nuclear power systems at some point in the future can be regarded as a classification problem about $\mathbb{V}_{T_0}$. For this purpose, the classification function of the vibration tensor can be given as follows:

$$\eta_i = Clf\left[Fe\left(\mathbb{V}_T\right)\right] (i = 1, 2, 3 \ldots.), \tag{1}$$

where Clf means the classification function, Fe means the feature extraction function, and η_i means the ith classification of $\mathbb{V}_T$. Actually, equation (1) provides a normal form of the nuclear power system fault prediction based on tensors, that is, as the sample analyzed in this work, $\mathbb{V}_T$ belongs to the sample space; as the classification of system faults in the future, η_i belongs to the classification space. For $\mathbb{V}_T$ of each moment T in future, the fault prediction system presented in this work can calculate η_i, which means the future state of the nuclear power system belongs to the normal or fault state.

According to equation (1), the fault prediction of the nuclear power system can be converted to the classification problem of tensors. Therefore, the primary question is how to extract the key features in the tensor $\mathbb{V}_T$.

For feature extraction of the tensor $\mathbb{V}_T$, we presented a deep convolutional neural network (CNN) as the submodule of the fault prediction system in Figure 3. By observing the data features in the tensor $\mathbb{V}_T$, we can find that in the process of extracting the features of V by this DCNN, the key challenge is how to extract both the time series features of a single variable and the coupling features between variables. So, we propose a strategy for alternately using heterogeneous convolutions, which is shown in Figure 4. The DCNN for feature extraction contains 10 modules with the same structure, each of which consists of conv, batch normalization, and ReLU layers. The ReLU function can be shown as follows:

$$\mathrm{ReLU}(x) = \begin{cases} x, & x > 0 \\ 0, & x < 0 \end{cases}. \tag{2}$$

In the first three modules of this DCNN, we adopt three special convolution kernels, respectively, for extracting different features. The convolution layer in the first module adopts a 3×1 convolution kernel, and its main function is to

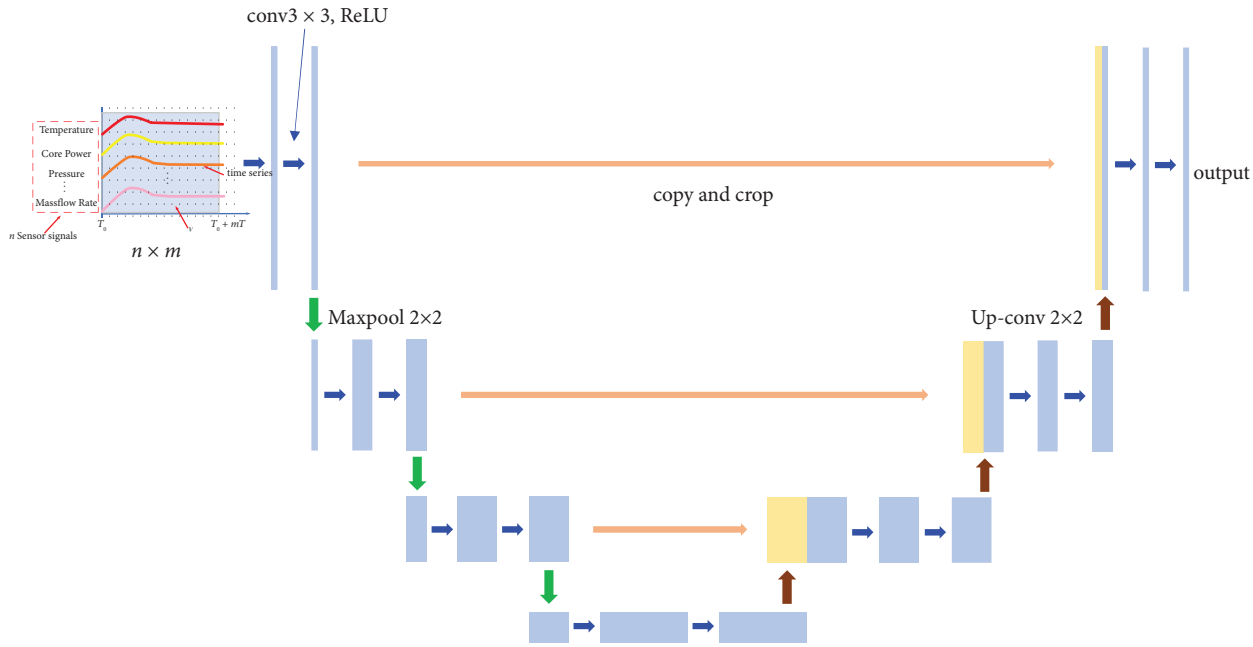

Figure 3: Architecture of state prediction network.

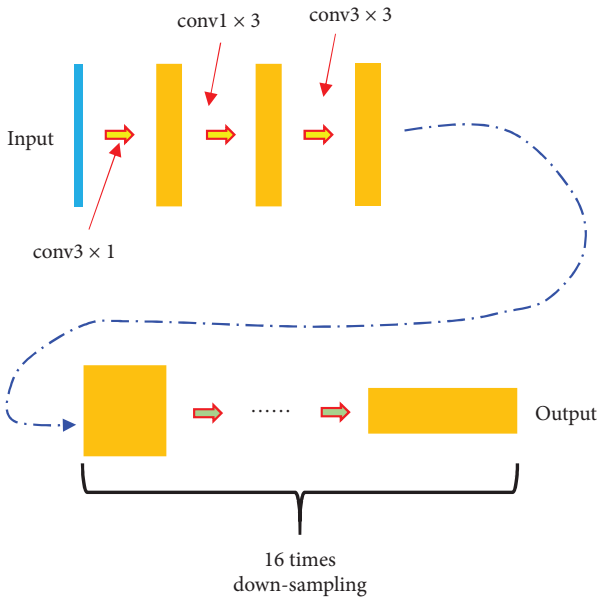

Figure 4: Architecture of feature extraction network.

extract the timing features of various sensor signals in the nuclear power system along the time axis. The convolution layer in the second module adopts a 1×3 convolution kernel, and its main function is to extract the coupling characteristics between various sensor signals in the nuclear power system. To ensure that the input and output dimensions are equal, the padding in the above two convolutional layers is set to 1. The convolution layer in the third module adopts a 3×3 dilated convolution kernel, and the dilation rate is set to 2. Similarly, to ensure that the input and output dimensions are equal, the padding in the above two convolutional layers is set to 2. Expect for the three special convolution layers above, the subnetwork composed of the remaining modules has been down-sampled 16 times, and the channel has been increased to 64. The purpose of these subsampling modules is to reduce the dimension of the fault-identified submodule by further extracting the state features, to reduce the computational overhead. The tensor $\mathbb{V}_T'$ with size $(n/16) \times (m/16) \times 64$ will be regarded as the input for the fault-identified submodule.

To realize the fault prediction of the nuclear power system, we also need to construct an identification network for tensor $\mathbb{V}_{T_2}$. In essence, tensor $\mathbb{V}_T'$ is the spatiotemporal coupling feature of nuclear power system state at T in future, so we construct a fully connected network (FcN) to identify whether the tensor $\mathbb{V}_T'$ reflects fault information. The FcN has $(n/16) \times (m/16) \times 64$ input cells and k output cells, where $k$means the number of the number of ICs that the FcN can identify. The softmax function is regarded as the activation function in the output layer represented as follows:

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}}, \tag{3}$$

where p_i means the output of ith neural cell in the output layer. $z_i = (n/16) \times (m/16) \times 64$ means the input of ith neural cell in the output layer. Actually, the softmax function reflects the probability of the classification. k neural cells in the output layer indicate that the vibration states need to be classified into k ICs. If the output probability of ith neural cell is the maximum probability of the output probability in the output layer, the current vibration state will be classified into ith IC.

The three subnetworks in Figure 5 are connected end to end to form the fault prediction network EDN-NPSP based on a supervised learning framework. The detailed architecture of EDN-NPSP may cause the training process to be

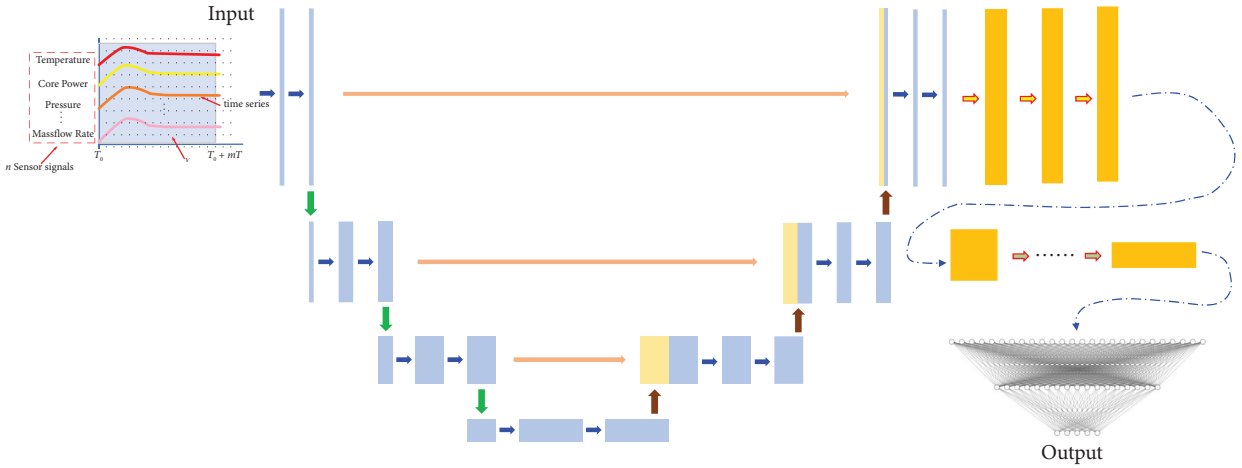

Figure 5: Detailed architecture of EDN-NPSP.

difficult to converge, so a two-step iterative training strategy, shown as follows, is adopted in this study:

(1) First, backpropagation training based on stochastic gradient descent is performed for PINN, where MSE is used as its loss function

(2) Then, the PINN obtained in the previous step is substituted into EDN-NPSP to perform backpropagation training based on small gradient descent, where the cross entropy is used as its loss function

3. Experimental Setup

The experiments were performed on a Yangjiang nuclear power plant (NPP) simulation system, which was built by China General Nuclear Power Group (CGN) firstly and then transformed into a semiphysical simulation system. This simulation system is developed based on the timing scheduling and data management platform SIMEXEC, and the core neutron dynamics and thermal-hydraulic models of the NPP are simulated based on REMARK and RELAP5-HD, respectively. To test the response of the real I&C system when performing fault prediction, we configure the OPC-based data interface in the simulation system and conduct data interaction with the real I&C system through this interface. The architecture of the simulation system is given in Figure 6.

3.1. Choice of Hardware. Generally, two kinds of hardware systems can be applied to NPP I&C control systems: one is industrial PCs (IPCs) with high computational performance and the other is embedded computers with little computational performance. Therefore, an appropriate classification method needs to be chosen for different hardware systems.

3.1.1. IPC. IPCs have become a firmly established part of various industrial environments. The advantages of IPCs are as follows: extremely high arithmetic speed, excellent scalability, and flexibility. With associated software, IPCs are at the core of a wide range of control or calculation tasks, such as motion control, processes or logistics systems, networking

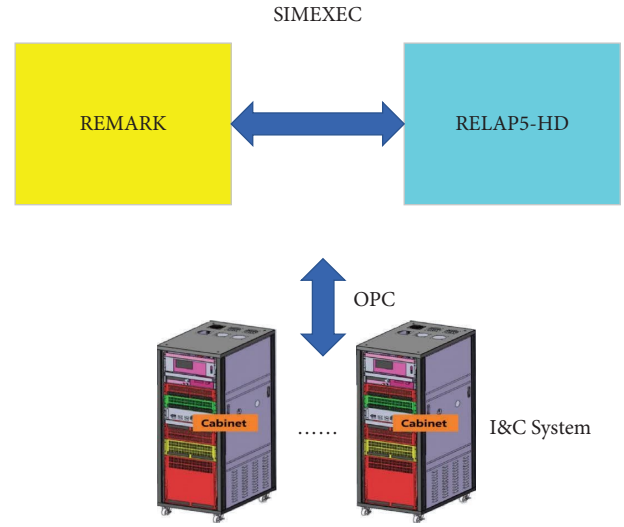

Figure 6: The architecture of the semiphysical simulation system of NPPs.

of system components, data acquisition, or image processing (Figure 7). It is obvious that IPCs also have the ability to diagnose and predict nuclear power system faults based on complex neural networks in real-time.

3.1.2. Embedded Computer. The distributed processing units (DPUs) are a kind of classical control-oriented embedded computer system, which are used for nuclear power system widely (Figure 8). DPUs have lower power consumption, appropriate volume, and lower manufacturing cost while they cannot diagnose system faults based on complex neural networks because of the little computational performance. Conversely, data processing tasks with lower computational costs are more suitable for DPU.

3.2. Hardware Architecture. To test the performance of EDN-NPSP, it is realized based on an additional IPC and embedded DPU, which are shown in Figure 9. The IPC is connected with the I&C system of the Yangjiang power

Figure 7: The industrial PC in the semiphysical NPP simulator.

Figure 8: The embedded computer in the semiphysical NPP simulator.

plant simulation system by a special embedded DPU. The special DPU also communicates with other DPUs in the DCS and acquires sensor signals through a data transmission network based on TCP/IP. The IPC acquires the tensor $\mathbb{V}_{T_0}$ through PROFIBUS-DP. In the IPC, the method based on EDN-NPSP is executed, and the identified results are transmitted to the I&C system backward. Simultaneously, the DPU has also been used to test the performance of EDN-NPSP.

3.3. Data Preparation. EDN-NPSP is an end-to-end network, so no additional preprocessing is required for the data used in this work. 220 samples from 8 working condition classes of the Yangjiang nuclear power plant constitute the training set (shown in Section 4.1). To accelerate the convergence of the network and enhance the stability of the training process, the data in the 220 samples were normalized according to the following:

$$x'_{ij} = \frac{x_{ij} - x_{i\min}}{x_{i\max} - x_{i\min}}, \tag{4}$$

where x_{ij} and x'_{ij} mean the data at $T_0 + mT$ of ith sensor and $x_{i\min}$ and $x_{i\max}$ mean the maximum and minimum values of ith sensor.

3.4. Training of EDN-NPSP. In this section, we also need to briefly introduce the training process of EDN-NPSP. As an end-to-end deep network, the two parts of EDN-NPSP, the state prediction network (Figure 3) and feature extraction network (Figure 4), need to be trained as a whole, where the cross-entropy loss function is considered as follows:

$$\text{Loss}_{ce} = -\frac{1}{N}\sum_{i=1}^{N} y_i \ln x_i, \tag{5}$$

where y_i and x_i mean the correct state class and predicted state class, respectively.

On the other hand, it is difficult for the output and input of the prediction network to be in the same state space after the training process based on equation (5). As a comparison, an additional training process between the input state and prediction state is considered, where the mean square error loss function is considered as follows:

$$\text{Loss}_{\text{mse}} = \left(x'_i - y'_i\right)^2, \tag{6}$$

where x'_i and y'_i mean the correct state and predicted state. So, the loss function of the whole training process can be given as follows:

$$\text{Loss} = -\frac{1}{N}\sum_{i=1}^{N} y_i \ln x_i + \left(x'_i - y'_i\right)^2. \tag{7}$$

So, based on equations (5)–(7), we will evaluate the following two different training processes in the experiment:

(1) End-to-end training only based on Loss_{ce}

(2) End-to-end training based on $\text{Loss}_{ce} + \text{Loss}_{\text{mse}}$

During the training process, the base learning rate was set to 10^{-4} and performed with Python 3.6, PyTorch 1.6.0, and CUDA 10.1 on a computer workstation (Intel-7-—11700K CPU, 16 GB RAM, and RT X3090-24 GB GPU).

4. Results and Discussion

4.1. IC Introduction. In this work, 6 fault ICs and 2 normal ICs are considered, including steady-state (IC-1), power changes within normal range from steady-state (IC-2), two main pump rotors are stuck (IC-3), one main pump rotor is stuck (IC-4), two pumps of the secondary circuit are stuck (IC-5), one pump of the secondary circuit is stuck (IC-6), condenser completely loses heat removal capability (IC-7), and control rod withdraws by mistake (IC-8). All of the ICs are shown in Table 2. The corresponding ID of the fault type of each simulation IC is also contained.

The simulation experiment of the NPP is based on 220 different ICs, which are divided into 8 classes. We obtained 11000 tensor samples $\mathbb{V}$ covering the above 220 ICs (in each IC, we obtained 50 tensor samples, and the time interval between tensors is 1 s: 1~5 s, 2~6 s, 3~7 s, ... 50~54 s), where the size of $\mathbb{V}$ is $50 \times 75 \times 4$, which means that we adopt 75 sensor signals in the NPP. The observation window covers the transient process for 5 s, and the number of channels is 4. Figure 10 shows some of the

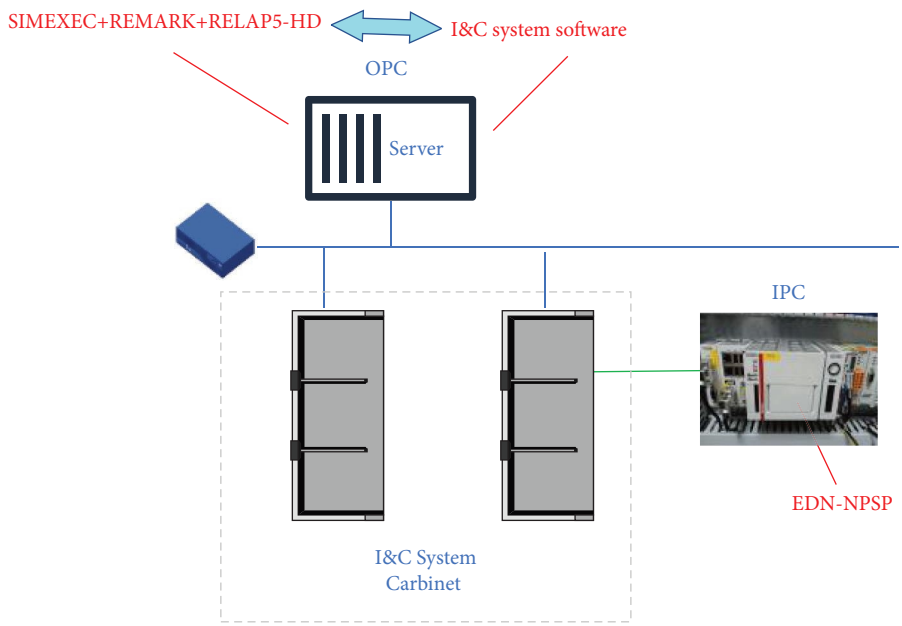

Figure 9: Hardware architecture based on IPC and DPU.

Table 2: Simulation industry conditions.

ID of IC	Class of IC	Sampling period/tensor size
IC1~IC10	(1) Steady-state	0.1 s/ 50 × 75 × 4
IC11~IC40	(2) Power changes within normal range from steady-state	0.1 s/50 × 75 × 4
IC41~IC70	(3) Two main pump rotors are stuck	0.1 s/ 50 × 75 × 4
IC71~IC100	(4) One main pump rotor is stuck	0.1 s/ 50 × 75 × 4
IC101~IC130	(5) Two pumps of the secondary circuit are stuck	0.1 s/ 50 × 75 × 4
IC131~IC160	(6) One pump of the secondary circuit is stuck	0.1 s/ 50 × 75 × 4
IC161~IC190	(7) Condenser completely loses heat removal capability	0.1 s/ 50 × 75 × 4
IC191~IC220	(8) Control rod withdraws by mistake	0.1 s/ 50 × 75 × 4

channels in the typical samples of the 8 ICs. It should be noted that the above samples are all in the early stages of failure, with no sensor signal triggering a threshold alarm or shutdown protection. At the same time, we planned to predict the failure state within the next 30 s. The acquisition process of samples is shown in Figure 10.

4.2. Result. Based on the IPC, the two-step iterative training process of EDN-NPSP is given in Figure 11. As mentioned above, the 1100 samples were divided into two parts: a training set (including 8400 samples of IC2IC8 and 400 samples of IC1) and a testing set (including 2100 samples of IC2-IC8 and 100 samples of IC1).

During the training process, the predicted value is obtained by using forward propagation first. Then, the chain derivative of the backpropagation algorithm is used to calculate the partial derivative of the loss function with respect to each weight. The weight in each neural cell is updated by using the gradient descent method. The training processes are shown in Figure 11. Accordingly, the test results of each epoch are shown in Figure 12, where 200 epochs are implemented during the training process, and two training processes are shown in Figure 11; one is the training based on $Loss_{ce} + Loss_{mse}$ and the other is the training only based on $Loss_{ce}$. Actually, the two training processes can converge after 200 epochs rapidly; this shows that the two training processes make the model have similar diagnostic performance. While because of the existence of $Loss_{mse}$, the training process in Figure 11(a) can make the network model output the prediction results located in the state space. As shown in Figure 12, SSIM is used to evaluate the prediction performance of EDN-NPSP based on the training process in Figure 11(a). The state of the NPP predicted by EDN-NPSP has more than 97.5% similarity with the actual state. Because of this advantage, we only use the training process based on $Loss_{ce} + Loss_{mse}$ in the subsequent experiments.

Alternatively, we also tested the maximum response time of each case and the number of ICs that triggered a shutdown. The maximum response time is defined as the following: the maximum time label in the sample that obtains the correct prediction result. Since all samples considered normal were classified as IC-1 or IC-2, the maximum response time index was not applicable to the identification of IC-1 and IC-2. Additionally, we counted the number of ICs that failed to predict reactor failure before triggering a threshold alarm. Table 3 presents the

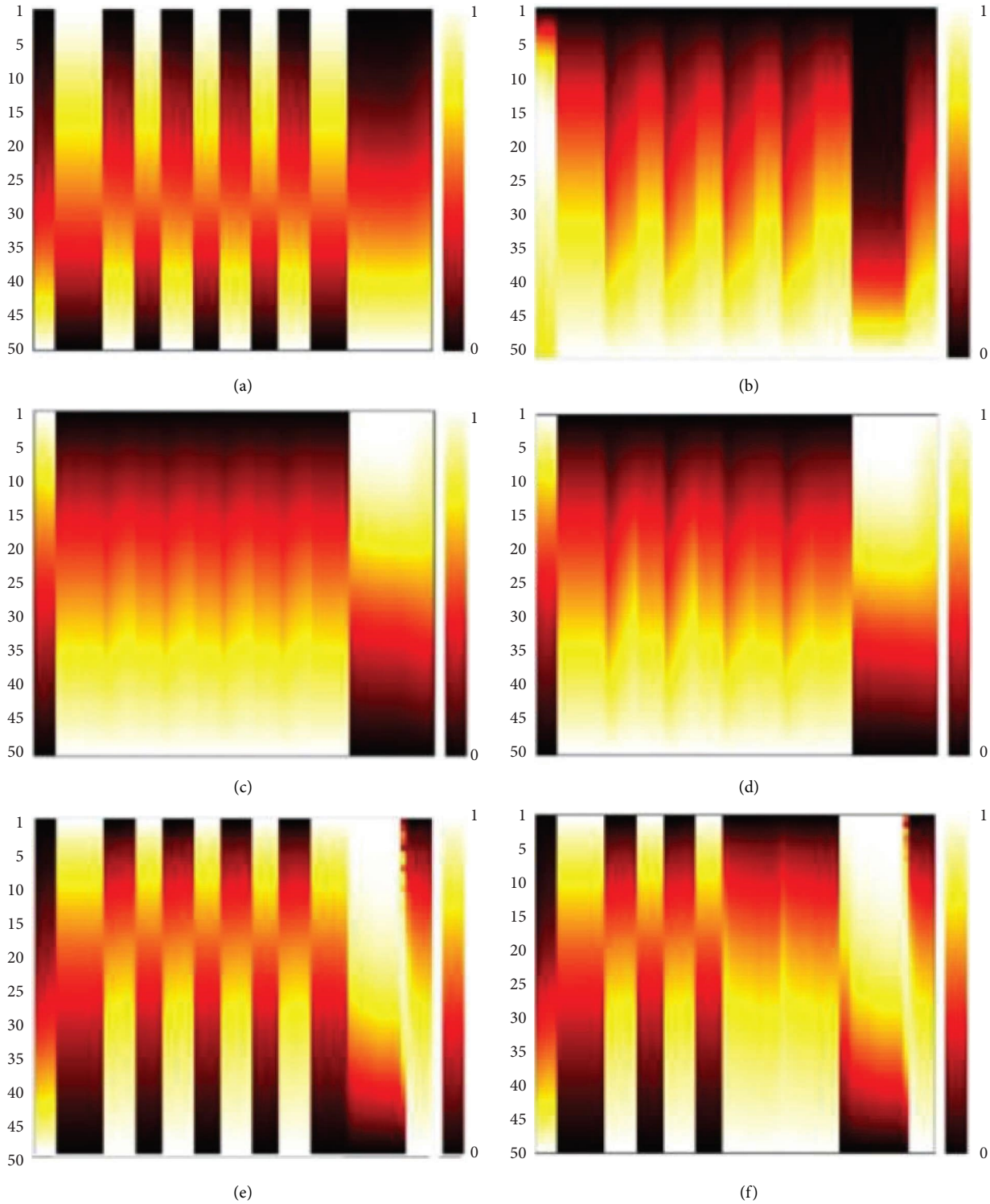

Figure 10: Continued.

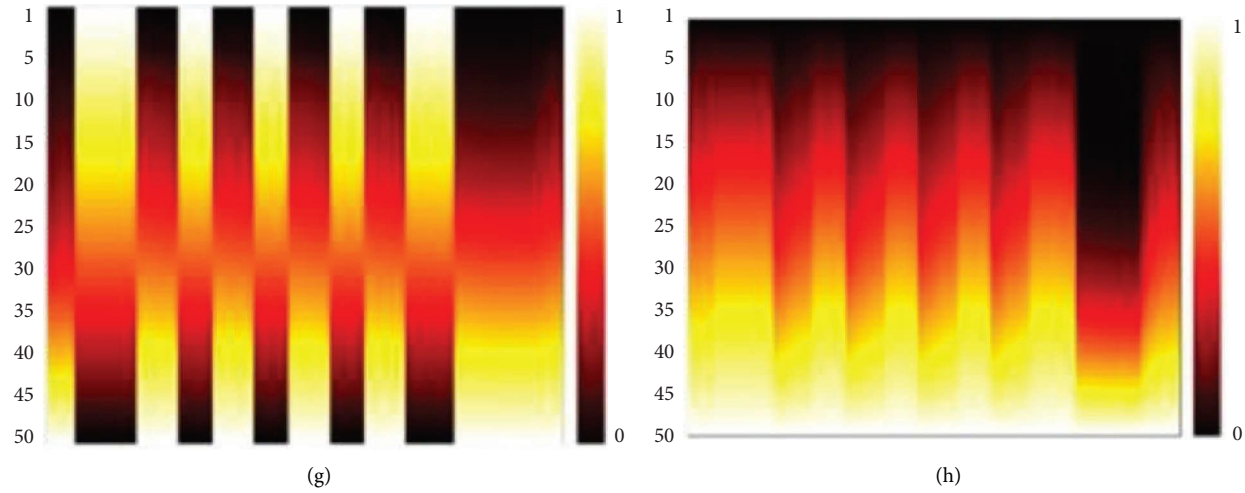

Figure 10: Typical tensor $\mathbb{V}$ consisting of NPP field sensor signals generated by NPP simulation. (a–h) Class 1~8.

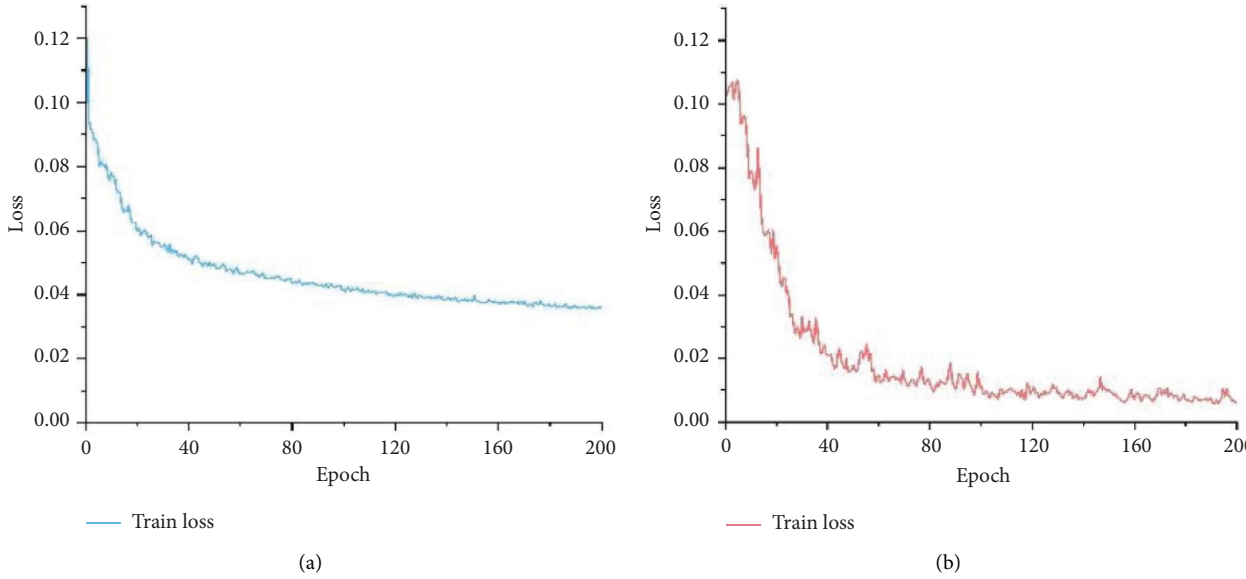

Figure 11: Training process of EDN-NPSP (200 epochs). (a) Training based on $\mathbf{Loss}_{ce} + \mathbf{Loss}_{mse}$. (b) Loss_{ce}.

test results based on the above two indicators. It was found that, except for IC7, the other fault ICs can be predicted faster by EDN-NPSP and respond faster than the threshold alarm. Since the condenser failure occurs in the steam system and its effects propagate slowly in the nuclear reactor, EDN-NPSP may not be able to predict the failure in a short period; at the same time, since the I&C system directly checks the equipment status, the I&C system will give device alarms faster. In the response of END-NPSP to all ICs, we can find that the response time of the secondary circuit fault is longer than that of the fault in the nuclear core and primary loop, which is also due to the longer time for the effect of the secondary circuit fault to propagate to the entire power plant.

4.3. Comparison with Other Machine Learning Methods. A lot of typical machine learning methods can also be used for fault diagnosis of nuclear power systems, so it is necessary to compare the performance of EDN-NPSP and other typical machine learning methods. For ease of analysis, KNN, SVM, CNN10, WaveletCNN, and encoder-decoder (10 layers) are regarded as the candidate classification methods. Expect for CNN10, state features are extracted by the principal component analysis (PCA) method. These comparisons are based on the same test data set, which is generated by the Yangjiang NPP simulator. Similar to the test in Section 4.2, in the comparison of this section, we still test two performance indicators, one is the correct rate of diagnosis or prediction and the other is the response time.

Table 4 shows the prediction or diagnosis accuracy on the 2200 testing samples based on classical machine learning methods compared with EDN-NPSP. For all ICs, the prediction and diagnosis accuracy of EDN-NPSP is similar to the diagnostic accuracy of Wavelet CNN, Encoder-Decoder, and CNN10, better than the traditional

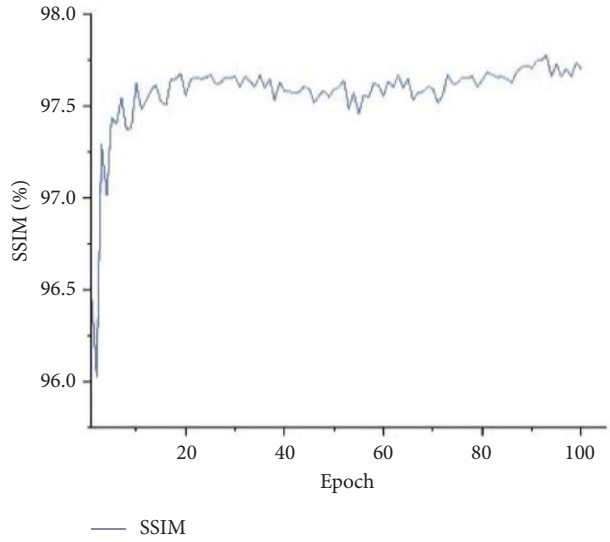

Figure 12: SSIM of state prediction network for test samples (prediction time: 30 s).

Table 3: Simulation industry conditions for EDN-NPSP test.

ID of classes	Number of ICs	Maximum/average response time of EDN-NPSP		Number of ICs that triggered alarm
1	2	—	—	—
2	6	2.0 s	1.2 s	0
3	6	3.0 s	2.7 s	0
4	6	3.0 s	2.8 s	0
5	6	6.0 s	4.3 s	0
6	6	11.0 s	7.5 s	0
7	6	18.0 s	7.5 s	2

Table 4: Accuracy comparison of fault prediction and diagnosis methods.

Class ID	EDN-NPSP (%)	PCA + KNN (%)	PCA + SVM (%)	CNN10 (%)	Wavelet CNN (%)	Encoder-decoder (%)
1	100.00	97.00	100.00	100.00	100.00	100.00
2	100.00	92.00	96.33	100.00	100.00	100.00
3	98.33	93.67	97.33	99.33	98.67	99.00
4	97.67	91.33	93.33	97.33	97.33	96.00
5	95.67	89.67	90.33	95.33	94.67	95.67
6	94.67	88.33	89.00	96.00	93.33	93.67
7	95.00	91.33	90.33	93.67	95.33	94.67
8	96.33	90.66	88.66	99.00	93.67	93.67
Total	96.82	91.27	92.55	96.59	96.09	96.27

machine learning methods PCA + KNN and PCA + SVM. For IC class 1–7, the identification accuracy of EDN-NPSP is the best among all the methods or not much different from the method with the best identification accuracy. Only for IC class 8 (control rod withdraws by mistake), EDN-NPSP is significantly different from CNN10, which has the best identification accuracy of this class. For fault class 8, the main reason for the decreased prediction accuracy of EDN-NPSP is that when the initial power level is low, the reactor state changes caused by the control rod withdrawal are often mistaken for normal transient processes. At the same time, due to the low initial power level and the effect of reactivity feedback in the nuclear reactor core, the mistaken withdrawal of control rods may not cause serious accidents, so this prediction error can be tolerated. At the same time, not only for EDN-NPSP but also for WaveletCNN and encoder-decoder, their prediction accuracy of class 8 has decreased significantly.

We also test the fault response time of EDN-NPSP and other machine learning methods, which is shown in Table 5. It can be found that, compared with the traditional machine learning method and the deep learning methods, EDN-NPSP has the fastest response speed to various NPP faults, mainly because the state prediction network in

Table 5: Time performance of EDN-NPSP and other machine learning methods.

Method	Maximum response time (s)	Average response time (s)	Number of ICs that triggered alarm
EDN-NPSP	18.0	4.3	2
PCA + KNN	31.0	15.1	6
PCA + SVM	23.0	11.3	5
CNN10	21.0	8.1	3
WaveletCNN	20.0	8.7	2
Encoder-decoder	25.0	7.6	3

EDN-NPSP constructs the state of the reactor in the future. However, other machine learning methods may not be able to quickly perceive fault features because features of the fault in initial phases are not obvious. At the same time, in the prediction or diagnosis of various fault conditions, the number of ICs that triggered alarm in the EDN-NPSP testing process is the least, which also proves that EDN-NPSP has a good response speed to failure in the initial stage.

5. Conclusion

In this work, an end-to-end deep network for NPS fault prediction EDN-NPSP was developed. In EDN-NPSP, dynamic tensors presenting the operation state of NPSs are constructed first. Different from classical data-driven fault diagnosis methods, the tensors can strengthen the association among the monitoring data in the spatial domain and time domain. On this basis, we designed three submodules: the state prediction network, which is used to generate the operation tensor at a certain time in the future; the state extraction network, which is used to extract the features of the generated tensor; and the fault identification network, which is used to identify the fault class. Based on the simulator of the Yangjiang nuclear power plant (containing the I&C system hardware), the prediction accuracy and response speed were evaluated. The results show that, for most fault conditions, EDN-NPSP has similar identification accuracy to some traditional deep learning methods, and the identification performance is better than traditional machine learning methods. On the other hand, due to the role of the state prediction network, EDN-NPSP can detect faults earlier than other machine learning methods and avoid triggering alarms and shutdowns. These experiments demonstrate the potential of EDN-NPSP to conduct fault prediction of NPSs. While some common problems among the deep learning methods, such as how to quickly tell the difference between the normal power change and the power change when the control rod withdraws by mistake, still need to be solved in the future, and it requires EDN-NPSP to further explore the potential of state prediction.

In future work, we will continue to improve the prediction and identification accuracy of fault ICs and further develop a neural network architecture suitable for compound fault prediction.

Data Availability

The original data are stored in the simulator system of China Nuclear Power Engineering Co., Ltd. and can be obtained after approval.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Guangdong Basic and Applied Basic Research Foundation (Project No. 2019B1515120060) and in part by the Open Funds of the State Key Laboratory of Nuclear Power Safety Monitoring Technology and Equipment.

References

[1] J. Li and M. Lin, "Ensemble learning with diversified base models for fault diagnosis in nuclear power plants," *Annals of Nuclear Energy*, vol. 158, Article ID 108265, Aug 2021.

[2] J. P. Ma and J. Jiang, "Applications of fault detection and diagnosis methods in nuclear power plants: a review," *Progress in Nuclear Energy*, vol. 53, no. 3, pp. 255–266, Apr 2011.

[3] S. N. Ahsan and S. A. Hassan, "Machine learning based fault prediction system for the primary heat transport system of CANDU type pressurized heavy water reactor," in *Proceedings of the 2013 International Conference on Open Source Systems and Technologies*, pp. 68–74, Lahore, Pakistan, December 2013.

[4] W. Li and J. Jiang, "Isolation of parametric faults in continuous-time multivariable systems: a sampled data-based approach," *International Journal of Control*, vol. 77, no. 2, pp. 173–187, Feb 2004.

[5] H. Wang, M. Peng, R. Xu, A. Ayodeji, and H. Xia, "Remaining useful life prediction based on improved temporal convolutional network for nuclear power plant valves," *Frontiers in Energy Research*, vol. 8, Article ID 584463, Nov 2020.

[6] V. Venkatasubramanian, R. Rengaswamy, and S. N. Kavuri, "A review of process fault detection and diagnosis. Part II: quanlitative models and search strategies," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 313–326, Mar 2003.

[7] J. Burriel-Valencia, R. Puche-Panadero, J. Martinez-Roman, A. Sapena-Bano, and M. Pineda-Sanchez, "Short-frequency fourier transform for fault diagnosis of induction machines working in transient regime," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 3, pp. 432–440, Jan 2017.

[8] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: a review with applications," *Signal Processing*, vol. 96, pp. 1–15, Mar 2014.

[9] J. Gertler and J. Cao, "PCA-based fault diagnosis in the presence of control and dynamics," *AIChE Journal*, vol. 50, no. 2, pp. 388–402, Feb 2004.

[10] H. A. Saeed, M. J. Peng, H. Wang, and Bw Zhang, "Novel fault diagnosis scheme utilizing deep learning networks," *Progress in Nuclear Energy*, vol. 118, Article ID 103066, Jan 2020.

[11] S. Tang, S. Yuan, and Y. Zhu, "Deep learning-based intelligent fault diagnosis methods toward rotating machinery," *IEEE Access*, vol. 8, pp. 9335–9346, 2020.
[12] H. Zhao, W. Zhang, and G. Wang, "Fault diagnosis method for wind turbine rolling bearings based on Hankel tensor decomposition," *IET Renewable Power Generation*, vol. 13, no. 2, pp. 220–226, Feb 2019.
[13] Z Cheng and R. Wang, "Nearest neighbor convex hull tensor classification for gear intelligent fault diagnosis based on multi-sensor signals," *IEEE Access*, vol. 7, pp. 140781–140793, Dec 2019.
[14] L. Luo, L. Xie, and H. Su, "Deep learning with tensor factorization layers for sequential fault diagnosis and industrial process monitoring," *IEEE Access*, vol. 8, pp. 105494–105506, Jul, 2020.
[15] Y. Jin, G. Chen, and H. Liu, "Analog circuit fault diagnosis based on non-tensor productwavelet network," *Chinese Journal of Scientific Instrument*, vol. 29, pp. 1613–1616, Aug 2008.
[16] F. Wei, G. Wang, B. Ren, J. Ge, and Y. Wang, "Multisensor fused fault diagnosis for rotation machinery based on supervised second-order tensor locality preserving projection and weighted-nearest neighbor classifier under assembled matrix distance metric," *Shock and Vibration*, vol. 2016, Article ID 1212457, Dec 2016.
[17] J. She, T. Shi, S. Xue et al., "Diagnosis and prediction for loss of coolant accidents in nuclear power plants using deep learning methods," *Frontiers in Energy Research*, vol. 9, Article ID 665262, Mar 2021.
[18] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nat. Rev. Phys*, vol. 3, no. 6, pp. 422–440, May 2021.
[19] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, Feb 2019.
[20] E. Schiassi, M. De Florio, B. D. Ganapol, P. Picca, and R. Furfaro, "Physics-informed neural networks for the point kinetics equations for nuclear reactor dynamics," *Annals of Nuclear Energy*, vol. 167, Article ID 108833, 2022.
[21] B. Atilim, P. Barak, A. Radul et al., "Automatic differentiation in machine learning: a survey," *Journal of Machine Learning Research*, vol. 18, pp. 1–43, Apr 2018.
[22] Q. Duan, R. J. Lu, H. Xie et al., "fault diagnosis of air compressor in nuclear power plant based on vibration observation window," *IEEE Access*, vol. 8, pp. 222274–222284, 2020.