Research on the Fault Diagnosis Method of a Synchronous Condenser Based on the Multi-Scale Zooming Learning Framework

Baiyun Qian 1, Jinjun Huang 2, Xiaoxun Zhu 3*, Ruijun Wang 3, Xiang Lin 1, Ning Gao 4, Wei Li 1, Lijiang Dong 4 and Wei Liu 1

1 State Grid Xinjiang Company Limited Electric Power Research Institute, Urumqi 830011, China
2 State Grid Corporation of China, Beijing 100031, China
3 Department of Power Engineering, North China Electric Power University, Baoding 071003, China
4 Xinjiang Xinneng Group Co., Ltd., Urumqi Electric Power Construction and Commissioning Institute, Urumqi 830000, China
* Correspondence: zhuxiaoxun@ncepu.edu.cn

Abstract: Under the background of the “strong direct current and weak alternating current” large power grid, the synchronous condenser with dynamic reactive power support capability becomes more important. Due to factors such as manufacturing, installation, and changes in operating conditions, there are many faults associated with the synchronous condenser. This paper studies a fault diagnosis method based on multi-scale zooming learning framework. First, through the energy fully connected (energy FC) layer, the synchronous condenser feature components of the fault signal of the camera are learned, and the transient features of the signal are enhanced. At the same time, the data is adaptively compressed and the effective features are mapped in a distributed manner. The faults are effectively diagnosed and isolated in advance. Secondly, a multi-scale learning framework is constructed to learn the multi-frequency features in the vibration signal. Finally, experiments show that the proposed method has certain advantages over the existing excellent models. The accuracy rate of diagnosis is higher than 99%.

Keywords: state detection; fault diagnosis; multi-scale; zooming learning; signal enhance

1. Introduction

With the rapid development of the ultra-high voltage direct current (UHVDC) transmission system, the weak problem of power transmission at one end of the power system has become increasingly prominent [1]. Once there is a problem with the conveying system, such as a mechanical fault, a commutation fault, or DC blocking, there will be strong reactive power fluctuations in the UHVDC system, resulting in large changes in the grid voltage, and even affecting the transmission capacity of the system, bringing huge security risks to the grid [2]. Therefore, for the safe operation of the UHV transmission system and the power grid, the dynamic reactive power compensation of the lifting system is very important. The synchronous condenser with dynamic reactive power support has a fast adjustment response [3]. Compared with other compensators, it has larger capacity and overload capacity. In particularly, it has a good effect on suppressing the voltage fluctuation caused by DC in the system [4,5]. The synchronous condenser is a synchronous motor without a mechanical load, which is widely used to improve the power factor of the grid and plays an important role in maintaining the voltage level of the grid. Its operating state directly affects the security and economy of the power grid system, so it has certain engineering value to carry out the fault research of the synchronous condenser.

The fault diagnosis method of equipment has always been a hot research trend. The diagnostic methods of equipment are basically divided into three categories from the...
macroscopic aspect, including feature extraction, feature learning, and feature recognition. Because the quality of feature extraction is the basis of fault classification, researchers have conducted many effective quantitative and qualitative analyses of feature extraction methods. Ref. [6] attempts to improve the ensemble empirical mode decomposition (EEMD) algorithm and combine it with a newly proposed denoising process and higher-order spectrum to improve the accuracy and speed of the fault severity and the type detection. In Ref. [7], a symplectic geometric packet decomposition (SGPD) multi-layer decomposition method is proposed. SGPD essentially combines the symplectic geometry theory and the idea concerning the multi-layer decomposition of wavelet packets to decompose the signal into a series of independent components containing the main fault information. The method proposed in Ref. [8] works by extracting the appropriate features of the signal used to determine the condition of the faulty bearing through a multi-output adaptive neuro-fuzzy inference system classifier. However, feature extraction requires a lot of prior knowledge about signal processing technology and diagnostic expertise, which is time-consuming and laborious [9,10].

With the development of artificial intelligence technology, the role of signal parameters in fault classification is weakened. The research of fault diagnosis is turned to pattern recognition based on statistics. The threshold and weight relations between fault features are then analyzed. A variety of classical fault identification methods are proposed and widely used in engineering practice, such as a support vector machine [11], an artificial neural network [12], random forest [13], etc. Although the research of these methods has been relatively mature, they cannot effectively learn the complex nonlinear relations in signals [14]. With the development of deep learning technology, researchers have applied this efficient, automatic, and deep-seated idea to the existing fault diagnosis tasks, such as a convolutional neural network [15], a depth confidence network [16], and an automatic encoder [17]. To complete the fault diagnosis task, we always solve complex problems by a single method. Feature extraction, feature learning, and feature recognition are processes of fault classification, the output of which one can generally be selected as the finally result which depends on the actual diagnosis requirements and diagnosis conditions. It is inadequate to regard one of these processes as a separate identification tool. However, from another perspective, there is the option to combine these three methods to form a brand-new functional module model, depending on whether the diagnostic performance can be effectively improved. Several researchers answered this question. Ref. [18] proposed a planetary gear fault diagnosis method based on variable mode decomposition power spectral entropy (VMD) and deep neural networks (DNNs). Ref. [19], similar to in this paper, proposed a fault diagnosis method based on correlation analysis and deep belief networks. Ref. [20] proposed a new sensor fault diagnosis method for gas leak monitoring based on the naive Bayes classifier (NBC) and the probabilistic neural network (PNN).

Although these methods combine the three stages mentioned above, they have are effective in specific tasks. However, feature extraction, feature learning, and feature classification are designed and implemented separately, which will affect the final diagnostic performance. This diagnostic strategy that cannot be optimized at the same time cannot be measured by the same standard. It is difficult to form a good closed-loop diagnosis because of the harsh matching conditions at each stage. Secondly, the operation state of synchronous condenser is complex, and the collected signals usually contain a variety of natural oscillation modes and information unrelated to the fault type, including the multiple coupling effect of multiple moving parts, the operation background noise, etc. Therefore, it is more and more difficult to effectively learn the fault characteristics of multiple frequencies from interference. Finally, the energy of the signal represents the impact information of the equipment components. The impact components and strength of different fault states are different. However, the above research methods did not consider the positive impact of signal energy on diagnostic tasks.

Based on the above motivation, we propose a multi-scale zooming learning framework (MSZLF) as a diagnostic model. Different functional modules in the model undertake dif-
different tasks, and their mapping modes will not interfere with each other. For different fault
tasks, the model focuses on different functions. For the fault identification of equipment, a
one-dimensional signal more fully reflects the fault category information compared to other
dimensional input information. The main diagnostic ideas of this method are as follows.
Firstly, the multi-scale learning mechanism is introduced into the model, with emphasis on
the omni-directional learning idea to express the information from different angles in the
fault data. Secondly, the energy operator is used to estimate the energy in the depth model
framework to increase the probability that important features are efficiently learned by the
subsequent learning unit. Finally, changing the mapping rules and setting the scaling data
mapping process can effectively enhance the display of fault feature.

The main contributions of this paper are as follows. (1) For the one-dimensional fault
signal, MSZLF is proposed. It includes a data acquisition module, a feature learning module,
and a fully connected classification module. It takes into account the multi-scale learning
framework and adopts the feature information scaling function to more comprehensively
reflect the fault information of the synchronous condenser. (2) The data mapping unit
of the energy FC layer is constructed. The Teager energy operator is used to evaluate
the fault feature energy, which can better characterize the transient features and impact
components of the fault information of the synchronous condenser. The full-connection
layer can be used to represent the overall features of the fault in a distributed manner. The
task of “zooming” in the framework completed by this unit can better learn the features
of these different energies so that the proposed MSZLF can efficiently identify fault types
under complex strong interference. (3) For the proposed MSZLF model, a variety of model
validation experiments are carried out. It can be considered that the proposed method has
great advantages over the existing excellent models.

2. Method Display

The deep learning mechanism is mainly used to build the information framework
of multi-layer neurons. By constructing the weight parameters for each piece of neuron
information, the data information on the neuron is encoded layer by layer, and the internal
relationship between the input and the output is preliminarily calculated. The weight pa-
rameters are optimized by using back propagation to gradually and accurately characterize
the nonlinear relationship between the input signal and the recognition task.

The deep learning framework MSZLF proposed in this paper includes three modules,
namely the data acquisition module, the feature learning module, and the fully connected
classification module, as shown in Figure 1. The data acquisition module mainly focuses on
data collection and data set production, and does not participate in learning. The feature
learning module and the fully connected classification module are mainly responsible for
learning and classification. In the deep learning framework, the full-connection operation
and the convolution operation are the two best learning functions. The fully connected
operation tends to observe the overall appearance of the feature at one time, while the
convolution operation focuses on compiling part of the content many times. The character-
istics of fault feature learning and the state recognition of the synchronous condenser
are also considered. The model designed in this paper takes full connection as the primary
mapping area of data, and expands the learning area of fault information. Assuming the
input of the full-connection layer ** layer is $X^{**-1}$, the formula is expressed as follows:

$$X^{**} = F_F(X^{**-1}) = w^{**}X^{**-1} + \beta^{**}$$  \hspace{1cm} (1)

where $F_F(.)$ is the full-connection function, $X^{**}$ is the output of the full-connection $^{**}$
layer, $w^{**}$ is the weight coefficient of the full-connection layer of the $^{**}$ layer, and $\beta^{**}
is the offset of the fully connected layer. The full-connection operation projects the fault
global information of the whole synchronous condenser. The whole data are distributed
and mapped, the feature information is aggregated, and the feature dimension is reduced.
In addition, Teager energy operator technology [21] is introduced in the front of the full-
connection layer, and the structure is shown in Figure 2. It is a nonlinear difference operator
which estimates the energy required by the signal source to generate the dynamic signal by the nonlinear combination of the instantaneous value of the signal and its differences. Based on the energy evaluation of various feature components of the fault signal, the transient features of the signal are enhanced and the characteristics of the impact component are highlighted. This specific feature selection amplification method involves less computation in the operation of the model and has less impact on the computational complexity of the model. At the same time, it provides good learning resources for the model and is an effective feature enhancement method.

\[ X^* = F_F(\phi(X_\epsilon)) \] (3)

It can effectively enhance the discrimination information component of the original signal, enhance the extraction of global information, and reduce the feature dimension. Because the sliding receptive field of convolution process cannot extract all the features at once, it can often achieve good results by using convolution to learn the sliding receptive field on lower dimensional data. On the one hand, each feature point is given an expression weight by using the receptive field to slide-map the fault information. On the other hand, the ratio of the low-dimensional data to the receptive field is smaller than the ratio of the original data to the receptive field. Therefore, when learning information, the number of slides becomes less, and each receptive field can capture more features at once. The
information perceived by the convolution kernel reflects the local fault features of the signal. In convolution, two-dimensional convolution is the most famous feature learning process in the field of computer vision. In the field of fault diagnosis, scholars usually convert one-dimensional signals into two-bit images as common image inputs, and use two-dimensional convolutional neural networks for learning. However, due to the dimension transformation of feature points, the original feature ordering is disordered, and the correlation between features is greatly reduced. This transformation method will reduce the highlight of key features and increase the interference of irrelevant information, resulting in a poor recognition effect. Therefore, this paper selects one-dimensional convolution as the main unit of feature learning, and the expression is as follows:

\[ X_{s,j}^* = \sum X_{s}^* \odot \omega_i + b_i \]  

where "\( \odot \)" is a convolution process, \( X_{s}^* \) is the feature of the input convolution layer, \( X_{s,j}^* \) is the \( j \)-th channel of the output feature, \( \omega_i \) is the \( i \)-th convolution kernel, and \( b_i \) is the offset.

Meanwhile, \( \text{Conv}(.) \) is defined as a convolution operation. It points from all feature maps of convolution input to all feature maps of convolution output. In addition, for the spatial and positional relationship of the local features of the fault, by increasing the number of convolution layers and the size of the receptive field of convolution, the learning range of the model is expanded [22]. In other words, in the model, both local features and global features can be widely learned. In order to make the model learn fault information better, we add the \( \text{BN} \) layer and activation function before and after convolution. Together, they form a convolution block \( F_{C} \) layer. The expression is as follows:

\[ F_{C}(\bullet) = \text{ReLU}(\text{Conv}(\text{BN}(\bullet))) = \max(0, \text{Conv}(\text{BN}(\bullet))) \]  

then

\[ X_{s}^{C_s} = F_{C}(X_{s}^{EF,Ps}) \]  

where "\( \text{ReLU} \)" is the rectified linear units, \( \text{BN} \) is the batch normalization operation, \( X_{s}^{EF,Ps} \) represents the features of the possible output before the convolution block, and \( X_{s}^{C_s} \) is the output of the convolution block. When features are extracted by convolution block, useful features and interference information are extracted at the same time, which may affect the efficiency of learning. We introduce the pooling layer to “De redundant” the extracted information, improve the information flow efficiency, and reduce the risk of over-fitting. The output of the pooling layer is:

\[ X_{s}^{Ps} = \text{MaxPool}(X_{s}^{C_s}) \]  

However, as the depth of the model increases, the gradient disappearance problem caused by multiple abstractions becomes more and more obvious. This is because as the number of convolution layers increases, it is difficult for the gradient to fall and converge. The most important thing is that there are many reasons for the fault of the synchronous condenser. For example, stator winding turns to a short circuit, and the winding will continue to be damaged by a large turn in the short-circuit current, which is a common electrical fault of a synchronous motor. At the same time, the stator core may be seriously burnt out due to overheating for a long time, and eventually develop into phase-to-phase short-circuit or grounding faults. In addition, the stator is also prone to turn short-circuit faults due to the variable processing technologies and operating conditions. However, the short-circuit effect is different due to different causes of short-circuit faults. Therefore, we should consider the different forms of data. In addition, the fault signal of the synchronous condenser is a complex signal composed of many different components, which seriously affects the recognition performance. In this paper, the fault data are divided into different input periods by preprocessing, and the information is expanded into a multi-scale mapping scheme. The original fault signal is then preprocessed by standardization. Multiple paths
are then inputted to expand the learning width. The aggregation vector of feature outputs of multiple scales can be expressed as:

\[ X' = [X_1^P, X_2^P, \ldots, X_N^P] \] (8)

\( N \) represents the number of paths. After this step of the network, the information is inputted into the full-connection classification module, which can analyze the extracted features and finally obtain high-precision diagnosis results. The output of the fully connected classification module can be expressed as:

\[ R = \text{Softmax}(F_F(Dropout(F_F(X'), \theta))) \] (9)

where \( \text{Softmax} \) is the probability classification function. \( \text{Dropout}(\cdot, \theta) \) is carried on \( X' \) which is a random inactivation function. \( \theta \) is the inactivation rate.

The back propagation of the model learns the parameters in the network [23]. We can see that this learning process is a continuous iterative process. In other words, these parameter values are constantly updated so that the full-connection learning and convolution learning present the most consistent fault-state information. For different diagnostic tasks, different learning frameworks are set to improve the generalization ability of the model [24]. The advantage of the proposed model MSZLF is that it considers the influence of sample length on the complexity of model training. When the model is inputted, the multi-scale learning effect is considered. In the data learning layer of the model, the Teager energy operator is used to calculate the energy relationship of different features. The key features are highlighted and the useless features are weakened. Then, the most comprehensive neurons are fully connected with the neurons of the next layer by using the full-connection layer, and there is no same-layer or cross-layer connection between neurons [25]. After data filtering, the calibrated data are inputted into a one-dimensional convolutional neural network for feature abstract expression. Finally, the learned feature vectors are merged along the channel axis to expand the multi-dimensional features. The fault of the synchronous condenser is diagnosed in the fully connected classification module. In this process, the scaling learning mechanism is used to enlarge the continuously extracted feature functions, so that the learning results of the model are more clearly. Through the set model learning process, we can see that the flow of learning data in the model is regular and orderly. Therefore, through the continuous self-iteration of the model, the performance of the model is constantly improved.

Admittedly, deep learning may have the disadvantages of a complex learning process, a long training time, and high computational complexity. However, the signal representation of the synchronous condenser is complex, the signal cycle is not clear, the operation under variable conditions is frequent, and the amount of monitoring data for the equipment is huge in long-term operation. Traditional methods are increasingly difficult to ensure the accuracy of diagnosis. It is meaningful to sacrifice some computational complexity and computing time for more accurate diagnostic results. Moreover, the proposed MSZLF method has universal and flexible structural parameters. The structural parameters of the model can be adjusted according to the exact diagnosis task. The computational complexity and training time are reduced as much as possible while ensuring the diagnostic performance.

3. Experimental Verifications

3.1. Introduction to Experimental Data

The experimental data come from a 300 MVar large-scale synchronous condenser at Tianshan Station in Urumqi, Xinjiang, China. One of the two ends of the equipment is the cranking end, and the cranking gear is driven by the motor for cranking. The other end is the excitation end. The stator is connected to the rotating magnetic field generated by the three-phase current and the rotor is connected to the stable magnetic field formed by the direct current to maintain the rotor rotation. Both ends are supported by sliding bearings. The long-term operation of the equipment has produced rubbing and misaligned
The vibration sensors are arranged as follows: No. 1 is the x direction of the crank end, No. 2 is the y direction of the crank end, No. 3 is the x direction of the excitation end, and No. 4 is the y direction of the excitation end. Collected by a SKVMA vibration monitoring analyzer, the sampling frequency is 6666 Hz. The on-site synchronous condenser equipment is shown in Figure 3 and the structure diagram is shown in Figure 4.

Figure 3. Synchronous condenser experiment device.

Figure 4. Structure diagram.

As a reactive power compensation device, the large-scale synchronous condenser is a synchronous motor running in the state of motor. Due to the uneven rotor material, asymmetric structure, processing conditions, and assembly errors during design, manufacturing, and installation, as well as the scaling, thermal bending, failure of parts, electromagnetic interference force, and other reasons during machine operation, a dynamic imbalance fault is caused. The abnormal vibration caused by the contact and collision between the rotor and the fixed parts due to the serious deformation of the shaft center caused by the bending and misalignment of the rotor, the insufficient clearance, and the bending and deformation of the non-rotating parts is called dynamic and static friction. These two fault states greatly threaten the safe and stable operation of the synchronous condenser. Based on this, the fault modes adopted in this paper include normal, dynamic, and static rubbing, as well as dynamic imbalance. The labels are coded as 0, 1, and 2. Each fault state contains 380 fault samples, and each fault sample is set with a different length to ensure that the input features have different scales. It is worth noting that when the structure of the model is flexibly changed, the length of the sample needs to be analyzed to ensure that the feature information can be smoothly learned by the network during the network operation. After mixing all the samples, random numbers are designed to scramble the data. The training set and the test set are divided in a ratio of 4:1. Table 1 shows the fault description of the synchronous condenser.
Table 1. Fault description.

| Running State          | Labels | Fault Mechanism                                                                 | Spectral Characteristics                                                                 |
|------------------------|--------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|
| Normal operation       | 0      | -                                                                              | Slight fundamental frequency vibration                                                   |
| Dynamic and static rubbing | 1       | 1. The vibration frequency bandwidth includes both the low-frequency part related to the rotational speed frequency and the high-order harmonic component related to the natural frequency. It is also accompanied by abnormal noise, and can be judged according to the vibration spectrum and sound spectrum. 2. Vibration changes over time. Under certain rotational speed and load conditions, the vibration vector changes due to local heating of the contact, and the phase change is opposite to the rotation direction. 3. The moment of contact friction will cause a severe phase jump (more than 10° phase change). When friction is localized, its trajectory has additional loops, both synchronously and asynchronously. 4. During friction, the axis track always processes in the opposite direction, i.e., in the opposite direction to the rotation of the shaft. Self-excited vibration may also occur due to friction, and the self-excited whirl frequency is the first-order natural frequency of the rotor. However, the whirl direction is opposite to the direction of rotor rotation. | 1. The time–domain waveform has the phenomenon of “top clipping”, or waveform distortion. 2. In addition to the rotor power frequency, there are also very rich high-order harmonic components in the frequency spectrum; there are precise frequency division components such as 1/2 frequency, 1/3 frequency, and 1/N frequency. 3. Vibration changes over time. Under certain rotational speed and load conditions, the vibration vector changes due to local heating of the contact, and the phase change is opposite to the rotation direction. 4. According to field experience, the loose parts are characterized by higher harmonics, and the friction is characterized by sub-harmonics. The loosening vibration of the components comes from the unbalanced force, so the loosening vibration obviously changes with the rotation speed, and the rubbing is controlled by the size of the gap, which is not closely related to the rotation speed. |
| Dynamic unbalance      | 2      | 1. The vibration frequency is consistent with the rotational speed frequency, the amplitude of the higher harmonics of the rotational speed frequency is very low, and the time domain waveform is close to a sine wave. 2. The centrifugal force generated by the unbalance of the rigid rotor is proportional to the square of the rotational speed, while the vibration measured in the bearing seat increases with the increase in the rotational speed, but is not necessarily proportional to the square of the rotational speed. This is due to the nonlinearity between the bearing and the rotor. 3. Around the critical speed, the amplitude peaks, and the phase difference is nearly 180° before and after the critical speed. | The waveform is simply harmonic and has less burr. The axis trajectory is a circle or an ellipse. The 1X frequency is the main. The axial vibration is small. The amplitude increases as the rotational speed increases. There are resonance peaks above the critical speed. |

3.2. Program Operating Environment

In this paper, the verification and evaluation of the model are carried out under the camera test bench. The experimental results of all researches are realized on the PC-side computer on Windows 10 64-bit system. The training and testing of the model are carried out under the deep learning framework of TensorFlow 2.4 in the Python programming language, and the hardware is Intel(R) Core i7-8700K CPU, RTX2080Ti GPU, and 11 GB video memory.

3.3. Hyperparameter Selection and Description

This paper carries out hyperparameter selection experiments. In order to better ensure the stability and convergence of the network, this paper sets the activation function as “ReLU”, the learning rate as 0.0001, the optimizer as “Adam”, the loss function as “Categorical_crossentropy”, and the batch size as 128. In addition, the structural parameters of the network are flexible, and the number of neurons in the fully connected layer is determined according to the length of the input sample and the set of output features. Parameters such as the size of the convolution kernel and the step size of the convolution sliding are determined according to the diagnostic task and signal quality. The depth of
the network is determined based on expected diagnostic results, among other things. In particular, the maximum pooling operation sets the pooling size to 2 and the step size to 2. The training process is gradually set according to the training process to ensure that the network can converge to the best value. In the following experiments, we set the convolution kernel width to $5 \times 1$. The sliding step of convolution kernel is 1. At the same time, BN layer is set before each convolution to ensure the learning quality of features. The edge filling parameter of feature map padding is the "same". In the fully connected classification module, the dropout is set to 0.5 to avoid over-fitting in the decision-making process.

3.4. Structure Exploration

The impact of functional modules in the model on the recognition performance of the model is also verified. In the model, different sample lengths are first divided and inputted into the multi-scale framework as different datasets. Although the model constructed in this paper shows three layers and three scales in the structure diagram, the number of layers of the model and the scale setting of the model are not unique. If the input vibration signal has strong multi-scale characteristics, the scale of the network can be increased. If the input vibration signal has characteristics that are difficult to learn, the number of layers of the network can be increased. Therefore, we can see that the model proposed in this paper is flexible and universal. It can also be applied to many diagnostic tasks. We assume that there are three layers in the network. The set research model is then compared with the single-input, double-input, three-input, and four-input functions, named MSZLF_I, MSZLF_II, MSZLF_III, and MSZLF_IV, respectively. Ten experiments are performed on all networks, as shown in in Table 2. It can be seen from the experimental results that when the scale of the model is increased, the accuracy of the model is gradually improved, and the standard deviation is also gradually reduced. We believe that the performance of the model is gradually enhanced. At the same time, we can also see that the accuracy of the model continues to improve as the input scale of the model increases. However, it is accompanied by an increase in the model complexity and an increase in the model calculation time. When using the model, it is necessary to consider the characteristics of the input characteristics, the coupling degree of the equipment operation, etc., and to observe the noise of the equipment operation environment. Taking these factors into consideration, a model structure that is most suitable for a certain environment to achieve the highest model performance. At the same time, as the network size increases, the structural complexity of the network also increases. In turn, the model computation time increases correspondingly.

Table 2. Comparison results.

| Model Name | Accuracy | Standard Deviation | Train Time | Test Time |
|------------|----------|--------------------|------------|-----------|
| MSZLF_I    | 99.01%   | 1.21               | 3.47 s     | 195 ms    |
| MSZLF_II   | 99.23%   | 1.05               | 4.57 s     | 333 ms    |
| MSZLF_III  | 99.56%   | 0.90               | 5.64 s     | 570 ms    |
| MSZLF_IV   | 99.72%   | 0.79               | 6.91 s     | 859 ms    |

The path of the network can affect the multi-scale learning effect of the network on the vibration signal, and the depth of the network focuses more on the ability of the network to capture useful information in the vibration signal. Assuming the model has three paths, we discuss the influence of the number of model layers on the performance of the model. We set layer 1, layer 2, layer 3, and layer 4 as MSZLF_1, MSZLF_2, MSZLF_3, MSZLF_4, respectively. All networks are tested for 10 times, and the experimental results are shown in Table 3. We can see that the accuracy rate gradually increases with the increase in layers. This is because multi-layer information coding can better capture more discrimination information. MSZLF_2 and MSZLF_3 show the accuracy increases by 0.24% and 0.45%, while MSZLF_4 only shows a 0.08% increase in accuracy. It shows that in the
serial learning structure, too many network layers may lead to excessive abstract mapping, which makes the learning ability of the model too strong, and then leads to model learning information other than fault features, resulting in the over-fitting phenomenon. Meanwhile, the calculation time increases significantly.

### Table 3. Comparison results.

| Model Name | Accuracy | Standard Deviation | Train Time | Test Time |
|------------|----------|--------------------|------------|-----------|
| MSZLF_1    | 98.87%   | 1.49               | 3.22 s     | 177 ms    |
| MSZLF_2    | 99.11%   | 1.17               | 4.69 s     | 315 ms    |
| MSZLF_3    | 99.56%   | 0.90               | 5.64 s     | 570 ms    |
| MSZLF_4    | 99.64%   | 1.03               | 7.01 s     | 882 ms    |

#### 3.5. Fault Diagnosis Based on MSZLF

Next, this paper mainly explores the diagnostic performance of MSZLF. The confusion matrix of the proposed MSZLF of the three-scale and three-layer model structure under the original data is shown in the figure. In Figure 5, rows and columns represent predicted tags and real tags, respectively. Diagonal cells represent the accuracy of the three failure modes. The accuracy of MSZLF in the normal category is 100%, which indicates that the model can still distinguish normal samples and fault samples with high accuracy in the actual operating environment. However, the overall diagnosis accuracy rate is 99.56%, which indicates that MSZLF can accurately diagnose different fault states in the actual operating environment.

![Confusion matrix](image1)

Figure 5. Confusion matrix.

The T-SNE manifold learning algorithm is very different from the isometric feature mapping (Isomap) algorithm and the local linear embedding (LLE) algorithm. It is not to use the similarity between samples to keep the same output results before and after dimensionality reduction, and must ensure that the linear relationship between the samples (before and after dimensionality reduction) remains unchanged. It is a nonlinear learning algorithm, which mainly reflects the local structure of data and tends to extract local clusters. It can effectively realize the visualization and dimension reduction in high-dimensional data, and the effect is outstanding.

Figure 6 shows a two-dimensional clustering diagram of the original vibration signal after MSZLF. Because the original vibration fault signal of the synchronous condenser is a mixture of many characteristics, the visual distribution of the signal is irregular and complex, and clustering is difficult, after passing through the full-connection layer of MSZLF. The two-dimensional effect map clearly shows the results of feature clustering. The same feature samples are gathered, and three kinds of fault signals are clearly distinguished, which shows that the classification results of this network are satisfactory.

![T-SNE two-dimensional clustering diagram](image2)
### 3.6. Ablation Study

This part focuses on the impact of MSZLF function modules on the performance of the whole network. We make three ablation models: MSZLF-NF, MSZLF-NE, and MSZLF-NEF. MSZLF-NEF is a network model for MSZLF used to remove energy FC blocks, and MSZLF-NE is a network model for MSZLF used to remove the energy operator layer. MSZLF-NF is a network model of the full-connection layer after MSZLF removes the energy operator layer. Finally, MSZLF is the proposed network. We choose three-scale and three-layer model structures as the structure of MSZLF in the comparative experiment. Ten experiments are performed on all networks. The specific comparative experimental results are shown in Table 4.

| Model Name | Accuracy  | Standard Deviation | Train Time | Test Time |
|------------|-----------|--------------------|------------|-----------|
| MSZLF      | 99.56%    | 0.90               | 5.64 s     | 570 ms    |
| MSZLF-NF   | 97.49%    | 1.27               | 4.72 s     | 325 ms    |
| MSZLF-NE   | 97.01%    | 1.12               | 5.31 s     | 506 ms    |
| MSZLF-NEF  | 95.98%    | 1.52               | 4.28 s     | 279 ms    |

It can be seen from Table 4 that MSZLF without any ablation has the highest accuracy and the lowest standard deviation. It shows that the diagnosis accuracy and stability of the proposed network reach a high level. The MSZLF-NEF network with the energy FC block removed is inferior to other networks in diagnosis accuracy and stability. This is because the energy FC function can effectively capture the global information of the input signal and the components with discrimination features. The global observation process of MSZLF-NEF is difficult to characterize the overall appearance of the input information, and the lack of energy extraction function makes it difficult to effectively compile the useful energy information of the fault. The accuracy of MSZLF-NE is 2.45% lower than that of MSZLF. This is because of the lack of the function of the energy operator to enhance the transient characteristics of the signal and features of the impact component, resulting in the model not extracting the complete discrimination information of the signal when learning. Finally, MSZLF-NF lacking a globally distributed mapping strategy makes it difficult to capture the complete scale information of the signal. In short, any functional module of MSZLF has a specific effect on the learning of fault features. MSZLF has excellent diagnostic ability for the fault signal of the synchronous condenser.

### 3.7. Comparative Experiment

The model selects multi-scale inputs to maximize the application of fault data, which can effectively stimulate the generalization potential of the model. In addition, the performance of the MSZLF model is verified by the retention method. We choose three scales and three layers of the MSZLF model structure as the structure of MSZLF in the comparative experiment. Ten experiments are performed on all networks. The specific comparative experimental results are shown in Table 4.

![Figure 6. T-SNE two-dimensional clustering diagram.](image-url)
experiment. Seven commonly used networks are selected for comparison. For instance, the k-nearest neighbor algorithm (KNN) sets the number of nearest neighbors to 3 and DBN is stacked by a three-layer restricted Boltzmann machine. The feature extraction part of 1-DCNN consists of three layers of convolution. There are three hidden layers of the back propagation (BP) neural network (BPN): SVM uses the linear kernel function, naive Bayes (NB) adopts the Gaussian naive Bayes classifier (Gaussian NB), and random forest (RF) has 10 decision trees. After 10 experiments, the recognition results of the eight network models are shown in Table 5. MSZLF has obvious advantages. However, the computational time is more long than other methods. This is inevitable for the deep learning model. Moreover, the magnitude of computing time is at the ms level, which can still meet the requirements of real-time monitoring.

| Model Name | Average Accuracy | Standard Deviation | Test Time |
|------------|------------------|--------------------|-----------|
| MSSLF      | 99.56%           | 0.90               | 570.0 ms  |
| KNN        | 70.21%           | 1.91               | 40.8 ms   |
| DBN        | 87.98%           | 1.25               | 205.1 ms  |
| 1-DCNN     | 92.70%           | 0.98               | 171.7 ms  |
| BP-DNN     | 88.75%           | 2.62               | 154.3 ms  |
| SVM        | 86.29%           | 2.13               | 99.4 ms   |
| NBC        | 65.61%           | 3.26               | 101.9 ms  |
| RF         | 85.74%           | 1.79               | 137.2 ms  |

4. Conclusions

Through the analysis of the rotor vibration signal of the synchronous condenser, this paper explores a fault diagnosis method of the synchronous condenser and carries out functional experimental verification. The specific advantages are as follows:

1. This paper constructs an energy FC block, a convolution feature extraction module, and a fully connected classification module. For the synchronous condenser rotor vibration signal, they can effectively and adaptively learn and classify the feature components. The classification accuracy reaches 99.56%.

2. To verify the performance of different functional modules in the model, the ablation study is set up in this paper. The rationality of module construction is explored, and the effect of these functional modules on the fault identification of the synchronous condenser is expressed.

3. In this paper, a structural analysis experiment and a comparative experiment are set up. The structural analysis experiment aims to verify the performance of the model under different structures. The comparison experiment focuses on the performance comparison between the model and the existing network. The experiments show that the model has superior diagnostic performance. Compared with other models, it had certain advantages.

In future work, we will carry out a joint detection of multiple signal sources for the fault of the synchronous condenser to further improve the application effect in the actual project.

Author Contributions: Conceptualization, X.Z. and B.Q.; methodology, J.H.; software, X.Z.; validation, X.L., N.G. and W.L. (Wei Li); formal analysis, L.D.; investigation, R.W.; resources, X.Z.; data curation, W.L. (Wei Liu); writing—original draft preparation, B.Q.; writing—review and editing, B.Q.; visualization, X.Z.; supervision, J.H.; funding acquisition, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Hebei Province Outstanding Youth Fund (grant number E2019502080).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.
**Data Availability Statement:** The data used to support the findings of this study are available from the corresponding author upon request.

**Acknowledgments:** This paper is supported by the State Grid Xin-jiang Electric Power Co., Ltd. “State Grid Xinjiang Electric Power Research Institute 2022 300 MVar synchro-nous condenser rotor fault diagnosis technology service”.

**Conflicts of Interest:** We declare that there are no actual or potential conflicts of interest between us and any other person or organization, including financial, personal, or other relationships with others or organizations, which may improperly affect or believe to affect their work.

**References**

1. Liu, Y.; Xia, C.; Yu, G.; Cao, J.; Ma, X. Synchronous Condenser Fault Diagnosis Method Based on RBF Neutral Network. In Proceedings of the 2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2), Changsha, China, 8–10 November 2019; pp. 2193–2196. [CrossRef]

2. Wei, C.; Sun, L.; Wang, W.; Lin, Y.; Tian, W.; Cheng, M. A Fault Diagnosis Strategy for Rotor Windings Inter-turn Short Circuit of Synchronous Condenser. In Proceedings of the 2022 22nd International Conference on Electrical Machines and Systems (ICEMS), Harbin, China, 11–14 August 2019; pp. 1–5. [CrossRef]

3. Zhou, J.; Ruan, L.; Zhou, L.; Wang, J.; Ma, Y.; Tao, Q.; Ling, Z. A Method of Evaluating the Stator Inter-Turn Short Circuit Fault of Synchronous Condenser. In Proceedings of the 2018 21st International Conference on Electrical Machines and Systems (ICEMS), Jeju, Korea, 7–10 October 2018; pp. 520–524. [CrossRef]

4. Sun, C.; Zhang, Y.; Xu, S.; Tian, W.; Gu, B.; Cheng, M.; Chao, W. Equivalent analysis of Stator Inter-Turn Fault in Different Winding Topologies of Synchronous Condenser. In Proceedings of the 2022 IEEE 5th International Electrical and Energy Conference (CIEEC), Nangjing, China, 27–29 May 2022; pp. 4994–4998. [CrossRef]

5. Wang, C.; Li, D.; Hu, J.; Ren, Z.; Ye, L. Calculation on Electromagnetic Characteristics of Synchronous Condenser Under Rotor Eccentricity. In Proceedings of the 2018 21st International Conference on Electrical Machines and Systems (ICEMS), Jeju, Korea, 7–10 October 2018; pp. 2633–2637. [CrossRef]

6. Hoseinzadeh, M.S.; Khadem, S.E.; Sadooghi, M.S. Quantitative diagnosis for bearing faults by improving ensemble empirical mode decomposition. *ISA Trans.*, 2018, 83, 261–275. [CrossRef] [PubMed]

7. Cheng, J.; Yang, Y.; Li, X.; Cheng, J. Symplectic geometry packet decomposition and its applications to gear fault diagnosis. *Mech. Syst. Signal Process.* 2022, 174, 109096. [CrossRef]

8. Rajabi, S.; Azari, M.S.; Santini, S.; Flammini, F. Fault diagnosis in industrial rotating equipment based on permutation entropy, signal processing and multi-output neuro-fuzzy classifier. *Expert Syst. Appl.* 2022, 206, 117754. [CrossRef]

9. Fan, Z.; Xu, X.; Wang, R.; Wang, H. Fan Fault Diagnosis Based on Lightwave Multiscale Multitattentiation Feature Fusion Network. *IEEE Trans. Ind. Inform.* 2021, 18, 4542–4554. [CrossRef]

10. Fan, Z.; Xu, X.; Wang, R.; Wang, H. CF-HSACNN: A joint anti-noise learning framework for centrifugal fan state recognition. *Measurement* 2022, 202, 111902. [CrossRef]

11. Liu, C.; Zhu, Y.; Zhan, K.; Guo, S.; Liu, K. Prediction of bolted joint looseness based on feature weighted-multiple kernel support vector machine. *J. Mech. Sci. Technol.* 2021, 35, 2407–2418. [CrossRef]

12. Gunerker, R.S.; Jalan, A.K.; Belgamwar, S.U. Fault diagnosis of rolling element bearing based on artificial neural network. *J. Mech. Sci. Technol.* 2019, 33, 505–511. [CrossRef]

13. Yang, B.-S.; Di, X.; Han, T. Random forests classifier for machine fault diagnosis. *J. Mech. Sci. Technol.* 2008, 22, 1716–1725. [CrossRef]

14. Zhu, X.; Wang, R.; Fan, Z.; Xia, D.; Liu, Z.; Li, Z. Gearbox fault identification based on lightweight multivariate multidirectional induction network. *Measurement* 2022, 193, 110977. [CrossRef]

15. Wang, M.-H.; Lu, S.-D.; Hsieh, C.-C.; Hung, C.-C. Fault Detection of Wind Turbine Blades Using Multi-variate CNN. *Sustainability* 2022, 14, 1781. [CrossRef]

16. Tang, J.; Wu, J.; Hu, B.; Liu, J. Towards a fault diagnosis method for rolling bearing with Bi-directional deep belief network. *Appl. Acoust.* 2022, 192, 108727. [CrossRef]

17. Aljemely, A.H.; Xuan, J.; Jawad, F.K.; Al-Azzawi, O.; Alhumaima, A.S. A novel unsupervised learning method for intelligent fault diagnosis of rolling element bearings based on deep functional auto-encoder. *J. Mech. Sci. Technol.* 2020, 34, 4367–4381. [CrossRef]

18. Li, Y.; Cheng, G.; Liu, C.; Chen, X. Study on planetary gear fault diagnosis based on variational mode decomposition and deep neural networks. *Measurement* 2018, 130, 94–104. [CrossRef]

19. Peng, B.-S.; Xia, H.; Liu, Y.-K.; Yang, B.; Guo, D.; Zhu, S.-M. Research on intelligent fault diagnosis method for nuclear power plant based on correlation analysis and deep belief network. *Prog. Nucl. Energy* 2018, 108, 419–427. [CrossRef]

20. Tan, Q.; Mu, X.; Fu, M.; Yuan, H.; Sun, J.; Liang, G.; Sun, L. A new sensor fault diagnosis method for gas leakage monitoring based on the naïve Bayes and probabilistic neural network classifier. *Measurement* 2022, 194, 111037. [CrossRef]

21. Bozchalooi, I.S.; Liang, M. Teager energy operator for multi-modulation extraction and its application for gearbox fault detection. *Smart Mater. Struct.* 2010, 19, 075008. [CrossRef]
22. Sainath, T.N.; Kingsbury, B.; Saon, G.; Soltan, H.; Mohamed, A.-R.; Dahl, G.; Ramabhadran, B. Deep convolutional neural networks for large scale speech tasks. *Neural Netw.* **2015**, *64*, 39–48. [CrossRef] [PubMed]

23. Kim, J.; Kim, K.-I. Partial Discharge Online Detection for Long-Term Operational Sustainability of On-Site Low Voltage Distribution Network Using CNN Transfer Learning. *Sustainability* **2021**, *13*, 4692. [CrossRef]

24. Liu, C.; Jia, G. Industrial Big Data and Computational Sustainability: Multi-Method Comparison Driven by High-Dimensional Data for Improving Reliability and Sustainability of Complex Systems. *Sustainability* **2019**, *11*, 4557. [CrossRef]

25. Lewenstein, M.; Nowak, A. Fully connected neural networks with self-control of noise levels. *Phys. Rev. Lett.* **1989**, *62*, 225. [CrossRef] [PubMed]