iEnergy
ARTICLE
https://doi.org/10.23919/IEN.2026.0012

A data-driven rapidly predictive identification method for the sub/super-synchronous oscillation dominant frequency of the new energy power system

Chongru Liu[1], Lingran Kong[1], Chenbo Su[1,2] ✉, Jiashu Fang[3], Le Zheng[1] and Aobing Li[1]

ABSTRACT

The sub/super-synchronous oscillation caused by the interaction between renewable energy generation and the power grid seriously affects the stability of the power system. A data-driven rapidly predictive identification method for the dominant oscillation frequency of sub/super-synchronous oscillation is proposed in this paper. The proposed method collects short-term operational data within the oscillation initiation process to predictively identify the dominant frequency when the oscillation presents a relatively stable periodic trajectory, which provides an information basis for the rapid suppression of oscillation. To solve the problem that sub/super-synchronous oscillation data is difficult to obtain, a transfer-learning based approach is developed to transfer knowledge learned in the forced oscillation domain to the sub/super-synchronous oscillation domain. Additionally, residual connection modules, non-local modules, and channel attention modules are introduced into the proposed model architecture, enhancing the model's predictive identification accuracy. Moreover, a weight propagation process is implemented to obtain the weight distribution of electrical channels, enhancing the interpretability of the model. Finally, the proposed method is verified via simulations and measured oscillation data. The results show that the method has high accuracy, strong interpretability, and generality.

KEYWORDS

Artificial intelligence; Dominant frequency; Interpretability; Power system; Sub/super-synchronous oscillation; Transfer learning

With the continuously increasing proportion of renewable energy sources being integrated into the grid, a substantial number of power electronic converters have been connected to the power system, leading to increasingly complex dynamic interactions among various system components. This has resulted in frequent occurrences of sub/super-synchronous oscillations, posing severe challenges to the stable operation of power system[1,2]. The dominant frequency of oscillation is highly important for the selection of oscillation suppression measures. There is an urgent need to establish a fast dominant oscillation frequency identification method to provide a basis for the timely implementation of oscillation control.

Currently, two main types of methods are used to obtain the dominant oscillation frequency: signal processing-based methods and machine learning (ML)-based methods. The signal processing method extracts the signal characteristics and obtains the oscillation frequency through the mathematical transformation method. All these methods exhibit inherent limitations, including the requirement for a long sampling window, the demand for high modeling precision, and the lack of desirable robustness. Specifically, the fast Fourier transform (FFT) is affected by the picket fence effect[3], and a window length of at least 2 s is required to accurately identify the oscillation mode[1]. The robustness of the Prony algorithm[4] highly relies on the accurate model and is susceptible to noise. The empirical mode decomposition (EMD)[5] and the Hilbert–Huang transform (HHT)[6] are usually used in combination, but their effects on higher-frequency oscillation identification are limited.

Compared with the above methods, the ML-based methods can effectively identify the oscillation characteristics, respond quickly, and does not need to accurately model the power grid. In Refs.[7, 8], neural networks are combined with the Prony and FFT methods. The neural network part is used to replace the signal processing process to directly solve the difficult links and improve the accuracy of the algorithm[9]. Some scholars have directly used deep learning algorithms for the extraction of oscillation parameters, such as multi-layer feedforward neural networks[10], convolution neural networks (CNNs)[11,12], long short-term memory (LSTM)[13], graph convolutional networks (GCNs)[14], etc.

Although the ML methods solve the problem of modeling difficulties, the length of the data sampling window of these ML methods at this stage is mostly a few seconds or even more than ten seconds[7–9, 12, 14]and most of them need to identify the frequency based on the oscillation waveform after the oscillation develops to a steady periodic change state, which makes it difficult to meet the needs of rapid identification. With the increasing degree of power electronization of the system, this time lag problem leads to the problem that the suppression measures cannot be put into operation in time, which can easily lead to serious problems such as large-scale unit tripping[15,16].

In this paper, a fast predictive identification method for the dominant oscillation frequency based on transfer learning is proposed. Compared with the FFT method, the proposed method can effectively improve the timeliness of identification. The model has high identification accuracy and can be applied to identify the

[1]State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China; [2]Department of Electrical Engineering, Tsinghua University, Beijing 100084, China; [3]School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China
Address correspondence to Chenbo Su, scb2636@tsinghua.edu.cn

dominant oscillation frequency in other scenarios. The main contributions of this paper include the following:

(1) A fast predictive identification method is proposed. The method leverages the data during the "initial stage of oscillation" to predict the dominant oscillation frequency in the "steady-state stage of oscillation" (when the oscillation presents a relatively stable periodic trajectory), significantly reducing the time required for identification.

(2) A transfer learning method for dominant frequency predictive identification is proposed. It constructs a pre-training process using forced oscillation samples and then transfers the model to the sub/super-synchronous oscillation domain, which effectively solves the problem of the lack of training samples.

(3) An interpretable structure is proposed in this paper.Visualization of the model's internal parameters demonstrates that the model assigns higher weights to key electrical quantities, thereby enhancing its predictive credibility.

The rest of this paper is organized as follows. Section 1 introduces the overall framework of the method proposed in this paper. Sections 2–4 introduce the three parts of the framework in detail, including the transfer learning framework, feature extractor for oscillation parameter identification, and interpretability structure. Section 5 analyses the case study to demonstrate the performance of the constructed model. Section 6 concludes this paper.

1 Overall framework for predictive identification

To achieve rapid predictive identification, we establish a fitting regression model, and the model input and output are shown in Figure 1. As shown in Figure 1, in this paper, we define the stage where the amplitude of the initial oscillation increases as the "initial stage of oscillation", which usually lasts for hundreds of milliseconds. The stage where the amplitude changes relatively little is defined as the "steady-state stage of oscillation". The method proposed in this paper identifies the dominant frequency of the "steady-state stage of oscillation" in advance by using the operation data of the "initial stage of oscillation". It can be regarded as a predictive identification, which has better timeliness.

Figure 2(a) shows the time-varying FFT results of I_a in Figure 1 during the "initial stage of oscillation". Figure 2(b) shows the FFT results of I_a in the "steady-state stage of oscillation". The dominant oscillation frequency described in this paper is the frequency component with the largest amplitude in the FFT results of the "steady-state stage of oscillation", as shown in Figure 2(b). As shown in

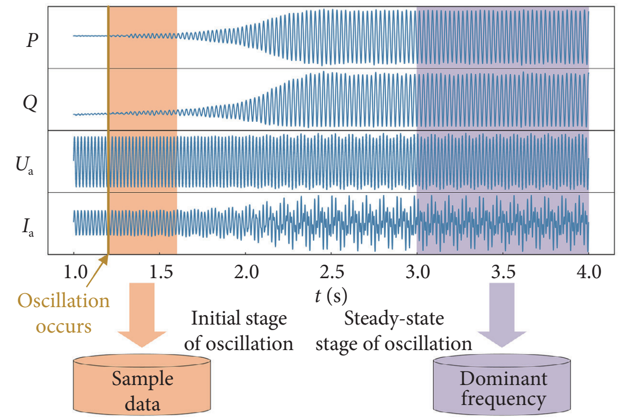


Figure 1 Oscillation predictive identification input and output.

Figure 2(a), the FFT method cannot accurately obtain the dominant frequency of oscillation in the "initial stage of oscillation". Therefore, the proposed model can provide more timely dominant oscillation frequency identification than the FFT method.

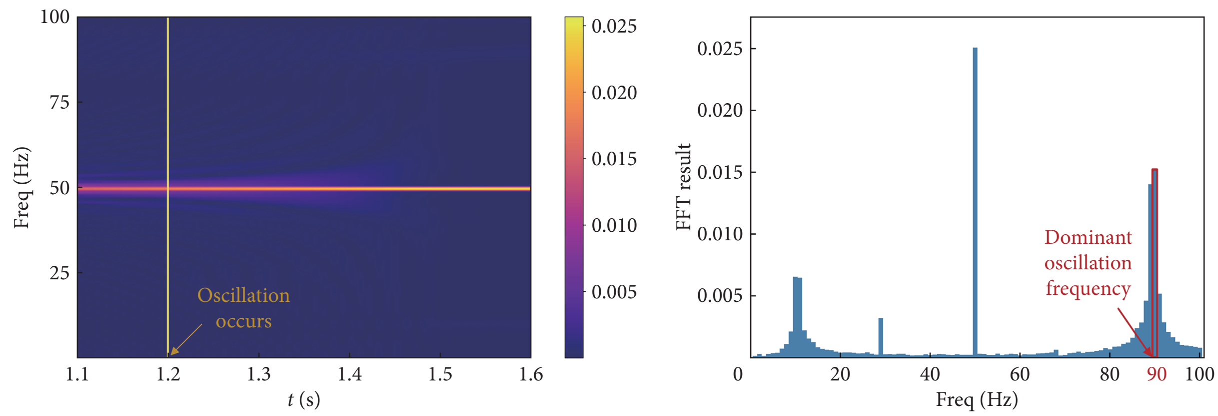
The proposed method adopts an "offline training, online application" approach. During the training phase, the intelligent model is trained via time-series oscillation data and corresponding labels (obtained via FFT). In the online application phase, by inputting the oscillation data into the model, the predicted value of the dominant frequency can be obtained directly, which is earlier than the FFT method. The proposed method in this paper includes three parts: a transfer learning framework, an oscillation predictive identification neural network model, and interpretability analysis, as shown in Figure 3.

2 Data augmentation based on transfer learning

The actual sub/super-synchronous oscillation data is scarce in the field, and relying only on a small amount of oscillation data for model training can easily lead to overfitting. Thus, we propose a transfer learning method[17,18] for oscillation predictive identification. By constructing a large number of forced oscillation samples and generalizing the knowledge learned from them to the actual sub-/super-synchronous oscillation field, the model performance could be effectively improved.

2.1 Sample domain construction

This paper uses a model-based transfer learning method[19,20] to complete the predictive identification task. This method first uses



(a) Initial stage of oscillation (b) Steady-state stage of oscillation

Figure 2 FFT analysis results of I_a within the initial and steady-state stages.

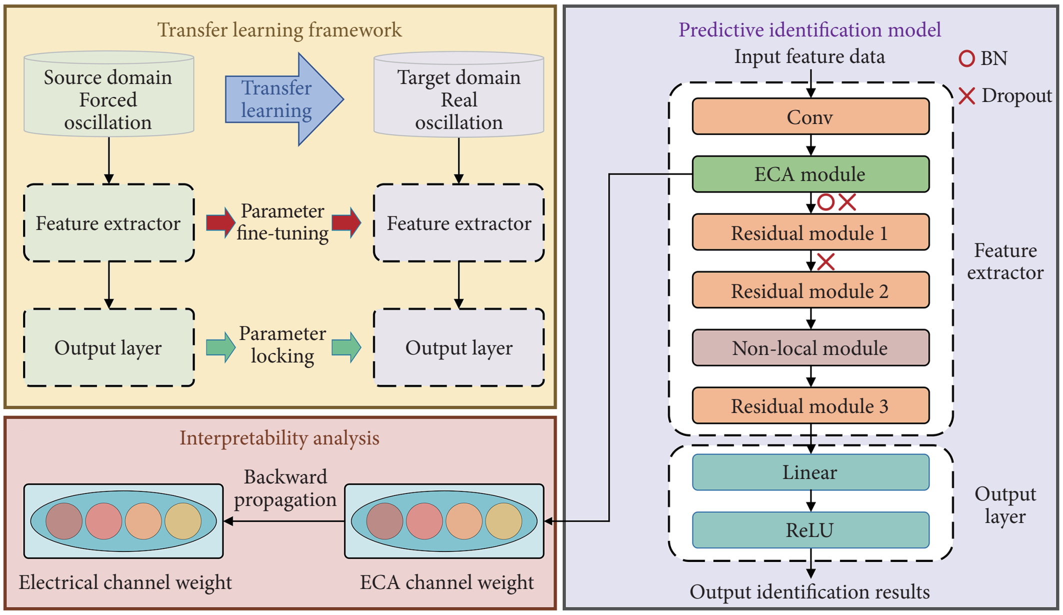


Figure 3 Overall framework of the predictive identification method.

a large number of source domain samples $\mathbb{D}_{\text{source}} = \{X_S, L_S\}$ for pre-training. The training goal is to minimize the loss function, and obtain the corresponding network parameters $W_S = \{\theta_S, \mu_S\}$ and the mapping f_S: $X_S \rightarrow L_S$. Afterwards, a selective transfer is conducted on the network. Some of the network parameters μ_S remain unchanged, and a small number of target domain samples $\mathbb{D}_{\text{target}} = \{X_T, L_T\}$ are used to fine-tune and update the remaining parameters θ_S to obtain the updated parameters. Ultimately, the final network parameters are denoted as $W_T = \{\theta_T, \mu_S\}$, and a mapping relationship f_T: $X_T \rightarrow L_T$ for $\mathbb{D}_{\text{target}}$ is obtained.

In this paper, $\mathbb{D}_{\text{source}}$ and $\mathbb{D}_{\text{target}}$ samples are constructed as follows:

1) $\mathbb{D}_{\text{target}}$ construction

In this work, the sub/super-synchronous oscillation of the new energy power system is selected as $\mathbb{D}_{\text{target}}$. Referring to the data and label selection method shown in Figure 1, the electrical measurement data at the grid connection of new energy power generation during the “initial stage of oscillation” are selected as the characteristic data X_T, and the dominant oscillation frequency of the “steady-state stage of oscillation” (obtained via the FFT method) is selected as the label L_T. Specifically, this paper selects the electrical data, including the active power, reactive power, A-phase voltage, and A-phase current. The reason for this selection is shown in the Appendix.

2) $\mathbb{D}_{\text{source}}$ construction

Compared with the actual oscillation scenario, the forced oscillation scene is easier to obtain, so this paper selects the forced oscillation sample domain as the transfer learning $\mathbb{D}_{\text{source}}$. In this paper, a wind farm grid-connection model is constructed within the simulation system. By injecting small-signal harmonic current disturbances at the bus of the wind farm outlet, forced oscillation is induced. By continuously changing the frequency, amplitude and other parameters of the injected harmonic current, a large number of $\mathbb{D}_{\text{source}}$ scenarios can be obtained. The selections of input features X_S and label L_S are the same as those in the $\mathbb{D}_{\text{target}}$ selection method.

Owing to the physical characteristic differences between actual oscillation and forced oscillation, we need to use transfer learning. To further improve the transfer learning effect, it is necessary to improve the similarity between the forced oscillation and the sub/super-synchronous oscillation. Therefore, when injecting harmonics, we consider the increase in the oscillation amplitude and multi-frequency coupling to make the forced oscillation closer to the actual oscillation characteristics. The single-frequency harmonic current can be expressed as

$$\begin{cases} I_a = A(t)\sin\left(2\pi ft + \varphi + \dfrac{2}{3}\pi\right) \\ I_b = A(t)\sin\left(2\pi ft + \varphi - \dfrac{2}{3}\pi\right) \\ I_c = A(t)\sin(2\pi ft + \varphi) \end{cases} \tag{1}$$

where f is the set harmonic frequency. To stimulate the coupling effect between more complex power grids, the initial phase angles of the two phases of AC are interchanged, so that the three phases present a negative sequence form. The phase angle φ and amplitude $A(t)$ can be expressed as

$$\varphi = \frac{f}{180} \times \frac{\pi}{50} \tag{2}$$

$$A(t) = \begin{cases} 0, & t < t_0 \\ A_F \dfrac{t - t_0}{t_1 - t_0}, & t_0 \leqslant t \leqslant t_1 \\ A_F, & t > t_1 \end{cases} \tag{3}$$

where A_F is the set steady-state harmonic amplitude, t_0 and t_1 are the set harmonic injection time and harmonic stability time, respectively.

The multi-frequency coupled harmonic waveform used in this paper is shown in Figure 4.

2.2 Data augmentation

The neural network model for oscillation identification established in this paper consists of two parts: a feature extractor and a fitting output layer (as shown in Figure 3). First, the overall model is preliminarily trained by using $\mathbb{D}_{\text{source}}$ samples, so that the model has a certain ability to mine the general features of oscillation data. Then, the weight parameters of the feature extraction part of the

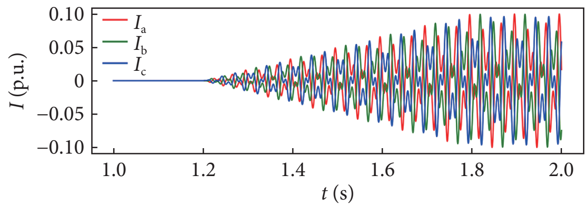


Figure 4 Harmonic current waveform diagram.

pre-trained model are fine-tuned by using the $\mathbb{D}_{\text{source}}$ samples to extract the features of the actual sub/super-synchronous oscillation more effectively.

Compared with training the model with only actual oscillation data, the model can learn more generalized feature representations in the oscillation domain by fully training on large-scale $\mathbb{D}_{\text{source}}$ samples. This "prior knowledge" helps reduce the risk of overfitting when training on $\mathbb{D}_{\text{target}}$, and improves the stability and generalizability ability of the model. The overall transfer learning process can be regarded as a significant enhancement of the quality of sample data.

3 Oscillation predictive identification neural network model

In this section, the proposed neural network oscillation identification model is introduced in detail. Based on the traditional CNN-linear neural network, the proposed model introduces three high-performance modules, namely, residual connection modules[21,22], non-local modules[23,24], and channel attention modules[25,26], to improve the accuracy and interpretability of the model.

3.1 Establishing residual links to preserve complex features

In the oscillation predictive identification problem, the relationship between features and labels is relatively weak. Shallow networks have difficulty in fully exploiting the characteristics of oscillation data. When the power system is modeled through neural networks, an increase in the number of neural network layers can more fully extract complex electrical data features, but the increase in the number of network layers may lead to gradient vanishing, resulting in the degradation of model performance[27]. The residual module[21,22] can effectively solve this problem by introducing a skip connection structure so that the gradient in backpropagation can propagate across one or more layers of the network without disappearing in the layer-by-layer operation. In this paper, the residual connection module is used to replace the deep convolutional network. The relationship between the residual connection module output feature $\boldsymbol{Y}$ and the input feature $\boldsymbol{X}$ can be expressed as

$$\boldsymbol{Y} = f_{\text{main}}(\boldsymbol{X}) + w_s * \boldsymbol{X} \tag{4}$$

where $f_{\text{main}}(\boldsymbol{X})$ represents the main road map and w_s represents the bypass convolution map.

3.2 Introducing a non-local module to achieve global element interaction

The traditional CNN network is limited by the size of the convolution kernel, so that only local adjacent elements can interact with each other and that distant elements cannot interact with each other. For oscillation problems, the electrical characteristics change periodically. Realizing the global interaction of elements can further explore periodic features. To realize the interaction of global elements and better mine global features, a non-local module[23,24] is introduced into the model.

For the input $\boldsymbol{X} \in \mathbb{R}^{C_{\text{in}} \times H \times W}$, where C_{in}, H, W are the number of input channels, the height, and width of the feature data respectively, three independent $1*1$ convolutional layers and shape change operations are carried out to obtain three different feature matrices of $\boldsymbol{\theta}$, $\boldsymbol{\phi}$ and $\boldsymbol{g}$. The three matrices go through a series of multiplication operations to achieve global element interaction. The output matrix $\boldsymbol{Y}$ can be expressed as

$$\boldsymbol{Y} = \sum_{\forall i,j} S(\theta(x_i)\phi(x_j))g(x_j) \tag{5}$$

where S (*) denotes the *softmax* activation operation and x_i and x_j denote the elements at different positions in the input matrix $\boldsymbol{X}$.

The output matrix $\boldsymbol{Y}$ is then restored to its original shape through another convolutional layer and then connected by a residual to form a complete non-local module. The final output matrix $\boldsymbol{Z}$ of the non-local module can be expressed as

$$\boldsymbol{Z} = w\boldsymbol{Y}_{\text{C}} + \boldsymbol{X} \tag{6}$$

where w represents the last layer of the convolutional mapping and $\boldsymbol{Y}_{\text{C}}$ is the matrix after the shape of $\boldsymbol{Y}$ is changed.

3.3 Establishing an efficient channel attention module

The input of the proposed network, $\boldsymbol{X} \in \mathbb{R}^{C_e \times L}$, is the operation data of different electrical variables. Where C_e is the number of electrical channels and L is the length of the sampled data. From a physical point of view, the degree of correlation between different electrical data and research problems is different. Guiding the network model to pay more attention to the important information in the channel dimension can further improve the performance of the model.

The efficient channel attention (ECA) module[25,26] is introduced to improve model performance. By adaptively assigning weight coefficients to the features of each channel, the channel features strongly correlated with the oscillation information are enhanced, and the weak correlation features are suppressed. The ECA module obtains the spatial dimensionality reduction characteristics of each channel through the global average pooling (GAP) operation of each channel. Then, a 1-dimensional (1D) convolution operation is performed on the features to capture the interaction information between channels. The original input of each channel is multiplied by the corresponding weights to obtain the final output of the module. The schematic diagram of the ECA module is shown in Figure 5.

To reduce the computational cost while maintaining high performance, the size of the 1D-convolution kernel is determined by the number of input channels, as shown in Eq. (7):

$$\text{kernel} = \left| \frac{\log_2(C_{\text{in}}) + 1}{2} \right|_{\text{odd}} \tag{7}$$

where $|*|_{\text{odd}}$ indicates an odd operation, and the specific method is as follows:

$$\text{kernel} = \begin{cases} \text{kernel}, & \text{kernel is odd} \\ \text{kernel} + 1, & \text{kernel is even} \end{cases} \tag{8}$$

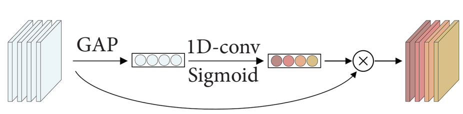


Figure 5 Diagram of the ECA module.

4 Interpretable method based on channel attention

At present, oscillation identification algorithms based on artificial intelligence[7–14] do not consider the interpretability of the model, and the credibility of the model is limited. The problem of poor interpretability caused by black box features is a common defect of deep learning algorithms. The attention mechanism has the advantages of adaptively focusing on key features, revealing the decision-making process of the model, and showing good interpretability[28]. In recent years, mechanisms such as self-attention[29], spatial attention[30], and channel attention[25] have been widely used in power-related problems such as system transient stability assessment[28] and wind power prediction[31], effectively improving the interpretability of the model.

The ECA module allocation channel can be easily presented visually. However, it should be noted that in the proposed model, the input feature of the ECA module represents a weighted combination of the original electrical quantities, which has no practical physical significance. Obtaining only this channel weight distribution does not provide a physically intuitive understanding. Therefore, to explore the focusing ability of specific electrical characteristics, the channel attention distribution of the original electrical quantity is obtained via the weight simplified backwards propagation method. The method presents a more intuitive interpretability. The specific method is as follows.

In the convolution operation, the mapping relationship between the output feature $\boldsymbol{Y}_i$ and the input feature $\boldsymbol{X}_i$ can be represented in the form of a matrix, as shown in Figure 6(b). T represents the transpose operation. The coefficient $w_{i\text{-}j}$ in the convolution mapping matrix is determined by the corresponding convolution kernel parameters. Referring to the forward propagation formula of the feature data, when only the weight relationship of each channel is considered and the specific value is not considered, the channel weight propagation formula can be established, as shown in Figure 6(c). The weight coefficient $k_{i\text{-}j}$ is also determined by the corresponding convolution kernel parameters. To facilitate the calculation, the average value of the 9 positions of the 3 × 3 convolution kernel is selected as the weight transfer coefficient $k_{i\text{-}j}$. When the channel weight distribution of the ECA module (the output channel weight on the right in Figure 6(c)) is obtained, the weight distribution for the original electrical input channel can be extrapolated backwards via matrix operation.

5 Case analysis

In this section, a wind farm system as described in Figure 7 is established through PSCAD/EMTDC to generate data and verify the effectiveness of the proposed method.

5.1 Sample construction

1) $\mathbb{D}_{\text{target}}$ sample construction

Based on the oscillation mechanism, sub/super-synchronous oscillation is triggered. The operating conditions of the farms are constantly adjusted to obtain more oscillation samples. All sub/super-synchronous oscillation is derived from natural resonance. The reason for the oscillation is the energy exchange between the control link or the system capacitance and inductance. The frequency range is 10–26 Hz/74–90 Hz.

In accordance with the method described in Section 2.1, the {P, Q, U_{a}, I_{a}} operation data of a single wind farm outlet is selected as the input feature X_{T}. The sampling window length and sampling frequency are 0.2 s and 2 kHz, respectively. The dominant oscillation frequency of I_{a} during the "steady-state stage of oscillation" is selected as the label L_{T}.

2) $\mathbb{D}_{\text{source}}$ sample construction

In accordance with the method described in Section 2.1, the forced oscillation is triggered by injecting harmonics at the outlet of wind farm 1. The dominant harmonic frequency range is 4–99 Hz, and the additional coupling harmonics are randomly generated.

The numbers of samples of the $\mathbb{D}_{\text{source}}$ and $\mathbb{D}_{\text{target}}$ are 8628 and 994, respectively. The dominant frequency ranges of the two types of samples are 4–99 Hz and 74–90 Hz, respectively. $\mathbb{D}_{\text{source}}$ and

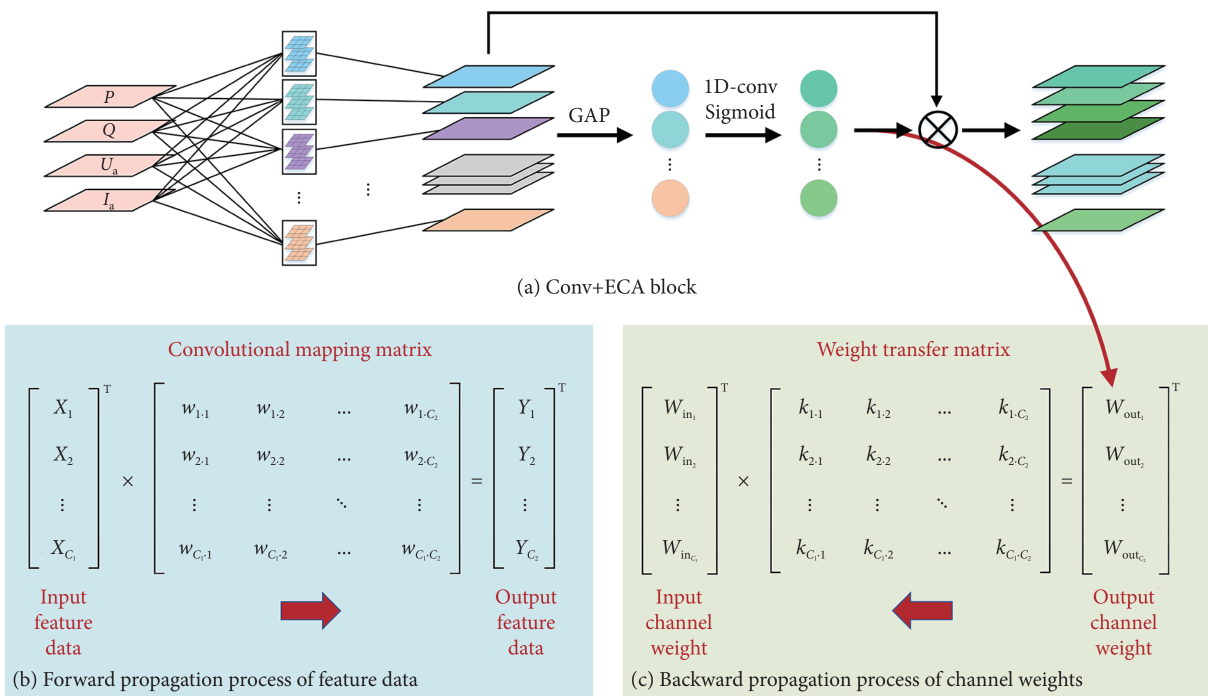


Figure 6 Schematic diagram of the backwards propagation of the channel attention distribution.

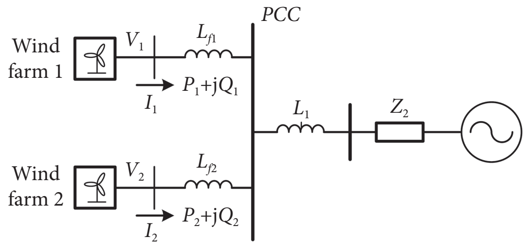


Figure 7 Simulation system model with two wind farms.

$\mathbb{D}_{\text{target}}$ samples are randomly divided into training sets and test sets at a ratio of 4:1.

5.2 Model implementation details

1) Data pre-processing

To eliminate the influence of dimensionality between different channels of different samples, the maximum-minimum normalization of each channel's data is used. Then, we pre-transform the time series data into a [10, 40] two-dimensional matrix format, allowing interaction among more sampled data points at different positions through subsequent convolution operations.

2) Model parameter setting

The predictive identification model is established on the PyTorch framework. Table 1 lists the detailed parameters of the model. In the pre-training stage, the batch size, learning rate, and training epochs are 200, 0.001, and 500, respectively. In the fine-tuning stage, the three values are 200, 1.8e–4, and 200. The model loss function is MSEloss.

The performance of the model is given by the mean absolute percentage error (MAPE) as follows:

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - l_i}{l_i}\right| \times 100\% \tag{9}$$

where y_i and l_i are the corresponding model output predictions and labels, respectively.

5.3 Comparison with the FFT method

The proposed method only needs 0.2 s of data during the "initial stage of oscillation" to identify the dominant frequency of the oscillation. In Section 1, Figure 2(a) shows the FFT analysis results during the "initial stage of oscillation". The FFT method cannot identify the dominant frequency within the initial 0.2 s. The proposed method is timelier.

In addition, Table 2 shows the differences in data acquisition and online application time between the proposed method and the FFT method. The calculation program is based on the Python language and the PyTorch framework. The computer processor is an Intel (R) Core (TM) i7-8700 CPU @ 3.20 GHz. Although the proposed method needs to collect four electrical quantities, these electrical quantities are easy to collect in the system. Because these data is collected locally, the cost of data collection and communication is not significantly increased. The calculation time in this paper is increased compared with that of the FFT method, but it is still within 15 ms. Compared with those of the FFT method, although the data acquisition and calculation times of the proposed method increase, it can still be easily implemented in engineering.

In terms of total time (sampling window length and calculation time), the time used by the proposed method is only approximately 10% that of the FFT method, which obviously has better timeliness.

Table 1 Network parameter settings

Layer	Parameter
Conv1	filter = 16, kernel size = 3, stride = 1
ECA module	channel = 16
Dropout	drop rate = 0.05
Residual module 1	filter = 8, kernel size = 3, stride = 1
Dropout	drop rate = 0.05
Residual module 2	filter = 4, kernel size = 3, stride = 2
Non-local module	channel = 2
Residual module 3	filter = 1, kernel size = 3, stride = 1
Linear	output = 1

Table 2 Comparison of the proposed method and the FFT method

Method	Collected data	Sampling window length (s)	Calculation time (ms)	Total time (s)
FFT	Single electrical quantity	2	4.56	2.00456
The proposed method	Four electrical quantities (P, Q, U_a, I_a)	0.2	14.43	0.21443

5.4 Model performance analysis

After the model pre-training process, the MAPE value of the model on the $\mathbb{D}_{\text{source}}$ test set is 3.71%. After the pre-trained model is fine-tuned by $\mathbb{D}_{\text{target}}$ sample, the MAPE value of $\mathbb{D}_{\text{target}}$ test sets is 4.62%. The above results show that the predictive identification model proposed in this paper has high accuracy and can meet actual needs.

To verify the effectiveness of the transfer learning process in the proposed method, we set up an algorithm that undergoes direct training on the $\mathbb{D}_{\text{target}}$ without pre-training on $\mathbb{D}_{\text{source}}$ as a contrast experiment. The results of the two experiments are shown in Table 3. The comparison of the MAPE values reveals that, compared with the non-transfer learning algorithm, the performance is significantly improved by approximately 3.85%. This result could prove that the proposed transfer process can effectively improve the accuracy of the model.

The triggering mechanism and data characteristics of forced oscillation and sub/super-synchronous oscillation are different. To reveal the differences between the two domains, this paper uses the t-SNE algorithm[32] to visually analyse the distributions of the two domains before and after transfer learning. Figure 8(a) shows the distribution of two domain samples of the original data, and it can be seen that there is no obvious overlapping area. Figure 8(b) shows the distribution of feature extraction results after transfer learning, and the overlapping areas of the two domains are significantly increased. Therefore, the transfer learning process proposed in this paper can effectively reduce the distribution difference between the two types of oscillation samples.

Table 4 shows a comparison of the effects of fine-tuning different parts of the network parameters during the transfer process. After the transfer and update of the three high-performance modules added in this paper, the feature extraction ability of the model for

Table 3 Comparison of model accuracy with and without the transfer learning process

Training sample		$\mathbb{D}_{\text{target}}$ MAPE (%)
Pre-training	Transfer fine-tuning	
×	$\mathbb{D}_{\text{target}}$	8.47
$\mathbb{D}_{\text{source}}$	$\mathbb{D}_{\text{target}}$	4.62

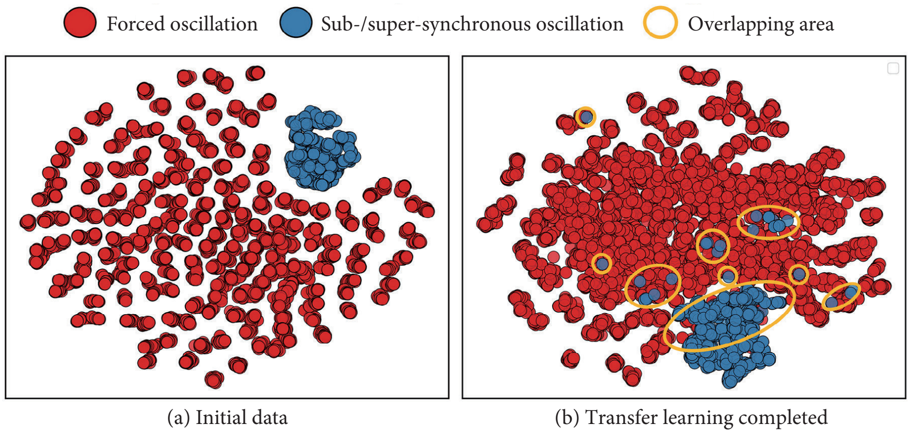


Figure 8 Distribution of forced oscillation and sub/super-synchronous oscillation before and after transfer learning.

Table 4 Performance comparison of fine-tuning different parts of the network

Model transfer section	$\mathbb{D}_{\text{target}}$ MAPE (%)
Output layer	9.78
Feature extractor	4.62

the actual oscillation domain can be better enhanced. Therefore, the accuracy of the model is improved.

To study the contributions of model components to model performance, ablation experiments are performed to remove specific components. The results of the ablation experiments (as shown in Table 5) indicate that the residual module, non-local module, and ECA module added in this paper have positive effects on model accuracy and that the combination of the three can further improve model accuracy.

To demonstrate the performance of the proposed method, the current network models commonly used in the field of power system oscillation (GCN, CNN-GRU, and CNN-LSTM) are constructed as comparative experiments, and the comparison results of the models are shown in Table 6. The results show that the proposed model is superior to the above traditional models in terms of the identification accuracy of $\mathbb{D}_{\text{target}}$.

Table 5 Performance of various models in ablation experiments

Model module				$\mathbb{D}_{\text{target}}$ MAPE (%)
Conv	Residual	Non-local	ECA	
√				6.31
√	√			5.21
√	√	√		5.10
√	√		√	5.09
√	√	√	√	4.62

Table 6 Comparison of the identification accuracies of different network models

Model	$\mathbb{D}_{\text{target}}$ MAPE (%)
GCN	7.70
CNN-GRU	6.75
CNN-LSTM	5.57
The proposed model	4.62

5.5 Model interpretability analysis

The attention weight distribution map visually shows how much attention the deep learning model pays to different parts of the input data. In this section, the channel weight distribution is analysed, and it is proven that the oscillation predictive identification model can achieve automatic focusing on strongly correlated electrical features.

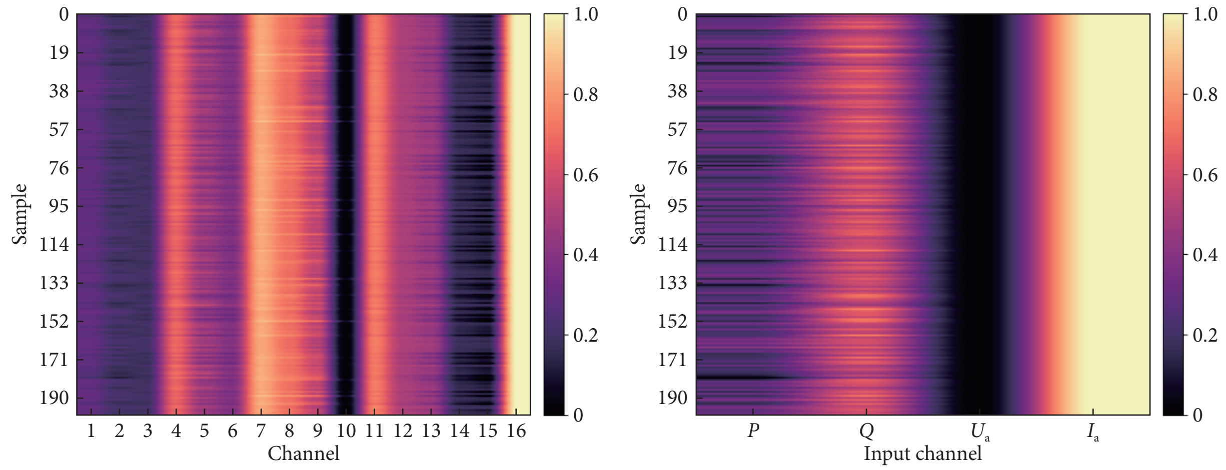
Figure 9(a) shows the distribution of the weights of each channel in the ECA module in the form of a heatmap. The vertical axis corresponds to different samples, and the horizontal axis represents different channels. As shown in the figure, among the 16 channels, the ECA module assigns higher weight coefficients to the 4th, 7th, 8th, 11th, and 16th channels, and the characteristics of these channels have a higher decisive effect on the output results.

In the matrix relation in Figure 6(c), the output channel weight on the right side of the equation is the channel weight of the ECA module shown in Figure 9(a). The weight matrix on the left side of the equation can also be obtained by the network parameters, so the weight distribution of the initial input channel can be obtained by reversing backwards through the matrix operation. Figure 9(b) shows the weight distribution of the initial electrical input channel obtained from the backwards extrapolation of each sample. In this example, the forced oscillation is excited by the harmonic current injected into the outlet of the wind farm, and the system oscillation disappears after the harmonic injection stops. Therefore, the corresponding electrical quantity I_a has the highest correlation with the system oscillation. According to Figure 9(b), among the 4 channels of the original input, the attention weight of the I_a channel is the highest. The results show that the proposed model can adaptively focus on the characteristics of the strongly correlated electrical quantity of the oscillation.

5.6 Generalization analysis of the model

In practical engineering, the short-circuit ratio of the system may change. To verify the applicability of the proposed model for unanticipated conditions, we change the initial value of the impedance of the grid line of the simulation system to simulate the change in the short-circuit ratio. Refer to $\mathbb{D}_{\text{target}}$ acquisition method, another sub/super-synchronous oscillation sample set is obtained, which is called the test domain $\mathbb{D}_{\text{test}}$. $\mathbb{D}_{\text{test}}$ contains 494 samples. The frequency range is 5–22 Hz/78–95 Hz, and the frequency range exceeds $\mathbb{D}_{\text{target}}$. Table 7 shows the identification accuracy of different algorithms in $\mathbb{D}_{\text{test}}$. The proposed model still has higher accuracy than the other models do.

In order to verify whether the proposed method is applicable to



(a) ECA channel attention distribution (b) Initial electrical input channel attention distribution

Figure 9 ECA and initial electrical channel attention distributions.

Table 7 Comparison of the identification accuracies of different algorithms in $\mathbb{D}_{\text{test}}$

Model	$\mathbb{D}_{\text{test}}$ MAPE (%)
GCN	8.62
CNN-GRU	8.18
CNN-LSTM	8.56
The proposed model	7.96

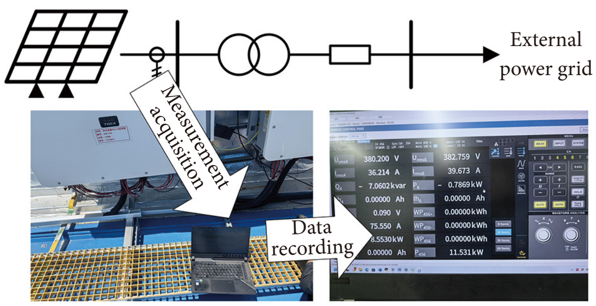


Figure 10 Actual measurement data equipment diagram.

different oscillation scenarinos, the measured oscillation data from the outlet of a photovoltaic power converter in a distribution system (as illustrated in Figure 10) in Shandong Province, China, are used as the model input to verify the universality of the trained model for on-site actual oscillation. The oscillation frequency of the actual oscillation data is 28/72 Hz, and the sub-synchronous oscillation component dominates. The oscillation initiation mechanism is the change of the PV output. The oscillation frequency and the generation mechanism are different from those of the above simulation oscillation data. After testing, the model correctly outputs a predicted frequency of 28 Hz. This demonstrates that method proposed in this paper is universal for other scenarios.

6 Conclusions

To accurately and quickly identify the oscillation frequency in a short time and provide a basis for the rapid suppression of oscillation, this paper proposes a predictive identification method for the dominant oscillation frequency and introduces the transfer learning process and attention mechanism, which improve the practicability and credibility of the proposed method. The experimental results demonstrate that the proposed model achieves high accuracy and is generalized for dominant frequency identification of oscillation in other scenarios. The main contributions of this paper are as follows:

(1) The proposed method realizes frequency identification based on early oscillation data by establishing an effective feature extractor for electrical data. Thus, the proposed method significantly reduces the time required for identification.
(2) The proposed framework reduces the problem of insufficient ML training accuracy caused by the lack of sub/super-synchronous oscillation data, establishing feature transfer learning between force and sub/super-synchronous oscillation.
(3) Weight backwards extrapolation is used to visualize the parameter weight distribution, making the proposed method interpretable. Therefore, the proposed model can adaptively focus on correlated electrical channel characteristics that trigger oscillation.

The proposed idea of predicting the dominant frequency of the "steady-state stage of oscillation" during the "initial stage of oscillation" can be further applied to predict other oscillation parameters, such as the amplitude and damping ratio, to achieve more comprehensive oscillation information acquisition.

Appendix

To select effective data features, it is necessary to study the relationship between the electrical quantity and the oscillation parameters and select the electrical features related to the oscillation frequency.

The whole new energy generation grid-connected system can be considered to be composed of two parts: the new energy generation subsystem and the remaining subsystem. The system state space equation of the new energy generation subsystem can be expressed as

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\Delta \boldsymbol{X}_1 = \boldsymbol{A}_1\Delta \boldsymbol{X}_1 + \boldsymbol{B}_1\Delta \boldsymbol{Z}_1 \\ \Delta \boldsymbol{Y}_1 = \boldsymbol{C}_1\Delta \boldsymbol{X}_1 + \boldsymbol{D}_1\Delta \boldsymbol{Z}_1 \end{cases} \tag{A1}$$

where $\Delta \boldsymbol{Y}_1 = \{\Delta P_1, \Delta Q_1\}$ represents the output variable of the new energy generation subsystem, which is composed of the power sent by the new energy generation; $\Delta \boldsymbol{Z}_1 = \{\Delta V_1, \Delta I_1\}$ represents the input variable of this subsystem, which is composed of the voltage and current at the grid connection between the new energy generation subsystem and the remaining subsystem. $\Delta\boldsymbol{X}_1$ is the state variable of the new energy generation; $\boldsymbol{A}_1$, $\boldsymbol{B}_1$, $\boldsymbol{C}_1$ and $\boldsymbol{D}_1$ are the coefficient matrices.

The system state space equation for the remaining subsystems is as follows:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\Delta\boldsymbol{X}_2 = \boldsymbol{A}_2\Delta\boldsymbol{X}_2 + \boldsymbol{B}_2\Delta\boldsymbol{Y}_1 \\ \Delta\boldsymbol{Z}_1 = \boldsymbol{C}_2\Delta\boldsymbol{X}_2 + \boldsymbol{D}_2\Delta\boldsymbol{Y}_1 \end{cases} \tag{A2}$$

where $\Delta\boldsymbol{X}_2$ is the remaining system state variable. $\boldsymbol{A}_2$, $\boldsymbol{B}_2$, $\boldsymbol{C}_2$ and $\boldsymbol{D}_2$ are the coefficient matrices.

The equation of state of the interconnected system can be deduced as follows:

$$\begin{bmatrix} \dfrac{\mathrm{d}}{\mathrm{d}t}\Delta\boldsymbol{X}_1 \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\Delta\boldsymbol{X}_2 \end{bmatrix} = \boldsymbol{A}\begin{bmatrix} \Delta\boldsymbol{X}_1 \\ \Delta\boldsymbol{X}_2 \end{bmatrix} \tag{A3}$$

where $\boldsymbol{A}$ is the characteristic matrix of the system. The oscillation frequency can be further obtained by solving for the eigenvalues of the $\boldsymbol{A}$ matrix.

The above formula shows that the system matrix $\boldsymbol{A}$ is related to the operating state of the new energy generation $\Delta\boldsymbol{S}_1 = \{\Delta\boldsymbol{Y}_1, \Delta\boldsymbol{Z}_1\} = \{\Delta P_1, \Delta Q_1, \Delta V_1, \Delta I_1\}$. Therefore, the oscillation frequency is related to the operating state of the new energy generation $\Delta\boldsymbol{S}_1 = \{\Delta P_1, \Delta Q_1, \Delta V_1, \Delta I_1\}$. Therefore, the active power, reactive power, A-phase voltage and A-phase current at the outlet of the new energy generation are selected as the input features.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No. U23B20126) and the National Natural Science Foundation of China (No. 52407117).

Article history

Received: 10 June 2025; Revised: 31 July 2025; Accepted: 19 August 2025

Additional information

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] Zhang, X., Zhang, F., Shi, J. (2025). Research on sub/supersynchronous oscillation parameter identification based on synchrophasor: The potential opportunities and challenges of non-spectral analytical techniques. *Proceedings of the CSEE*, 45(12): 4578–4593. (in Chinese).

[2] Hao, Q. (2024). Research on identification and location methods of sub-/super-synchronous oscillation in wind power systems. Master's Thesis, North China Electric Power University, Beijing, China. (in Chinese)

[3] Shao, Y., Yao, Y., Liu, H., Lv, R., Zan, P. (2021). Power harmonic detection method based on dual HSMW window FFT/apFFT comprehensive phase difference. In: Proceedings of the 2021 40th Chinese Control Conference (CCC), Shanghai, China.

[4] Yalcin, N. A., Vatansever, F. (2021). A new hybrid method for signal estimation based on Haar transform and Prony analysis. *IEEE Transactions on Instrumentation and Measurement*, 70: 6501409.

[5] Zhao, Y., Su, Y. (2020). The extraction of micro-doppler signal with EMD algorithm for radar-based small UAVs' detection. *IEEE Transactions on Instrumentation and Measurement*, 69: 929–940.

[6] He, Z., Zhu, J., Lu, C., Li, Y., Huang, L., Chen, Y. (2023). Parameter identification of DFIG model based on HHT marginal spectrum. In: Proceedings of the 2023 3rd New Energy and Energy Storage System Control Summit Forum (NEESSC), Mianyang, China.

[7] Zhu, W., Tang, Y., Zhou, Y., Zeng, Z. (2009). Identification of power system low frequency oscillation mode based on improved Prony algorithm. *Power System Technology*, 33: 44–47+53. (in Chinese)

[8] Zhu, W, Ma, J., Zeng, Z., Zhou, Y. (2012). Low frequency oscillation mode recognition based on segmental Fourier neural network algorithm. *Power System Protection and Control*, 40: 40–45. (in Chinese)

[9] Ding, R., Shen, Z. (2020). Power system low frequency oscillation mode identification based on exact mode order-exponentially damped sinusoids neural network. *Automation of Electric Power Systems*, 44: 122–131. (in Chinese)

[10] Gupta, A. K., Verma, K. (2016). PMU-ANN based real time monitoring of power system electromechanical oscillations. In: Proceedings of the 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India.

[11] Fu, Y., Chen, L., Yu, Z., Wang, Y., Shi, D. (2020). Data-driven low frequency oscillation mode identification and preventive control strategy based on gradient descent. *Electric Power Systems Research*, 189: 106544.

[12] Chen, J., Chen, Y., Luo, F., Li, N. (2024). Intelligent identification method of sub/super-synchronous oscillation parameters based on CNN and synchrophasor. In: Proceedings of the 2024 9th Asia Conference on Power and Electrical Engineering (ACPEE), Shanghai, China.

[13] Ye, S., Wei, J., Xu, Z., Zhang, W., Liu, X., Zhang, C. (2020). LSTM-based rapid identification of dominant low frequency oscillation modal features in power system. In: Proceedings of the 2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2), Wuhan, China.

[14] Zhang, J., An, H., Wu, N. (2020). Low frequency oscillation mode estimation using synchrophasor data. *IEEE Access*, 8: 59444–59455.

[15] Feng, S., Cui, H., Chen, J., Tang, Y., Lei, J. (2021). Applications and challenges of artificial intelligence in power system wide-band oscillations. *Proceedings of the CSEE*, 41: 7889–7905. (in Chinese)

[16] Zhao, Y., Sun, S., Liu, X., Nie, Y. (2022). Identification of power system wide-band oscillation based on improved CNN-LSTM. *Smart Power*, 50: 48–54+96. (in Chinese)

[17] Hao, Q., Liu, C., Wang, J., Wang, X., Su, C., Zheng, L. (2023). Location of sub-/super-synchronous oscillation source for direct-drive wind turbines based on deep subdomain adaptation. *Automation of Electric Power Systems*, 47: 27–37. (in Chinese)

[18] Hao, Q., Wang, J., Wang, X., Li, J., Su, C., Liu, C. (2023). Localization of sub/super-synchronous oscillation sources in wind power systems based on deep adaptation network. In: Proceedings of the 2023 4th International Conference on Advanced Electrical and Energy Systems (AEES), Shanghai, China.

[19] Xu, H., Yang, L., Long, X. (2022). Underwater sonar image classification with small samples based on parameter-based transfer learning and deep learning. In: Proceedings of the 2022 Global Conference on Robotics, Artificial Intelligence and Information Technology (GCRAIT), Chicago, IL, USA.

[20] Zhao, B., Liu, X., Zhang, L., Chen, Z., Zhang, L., Xie, C. (2024). Fault diagnosis of proton exchange membrane fuel cell integrated system based on GoogleNet and transfer learning. *Proceedings of the CSEE*, 44: 5147–5158. (in Chinese)

[21] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for

image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA.

[22] Jafar, A., Lee, M. (2020). Hyperparameter optimization for deep residual learning in image classification. In: Proceedings of the 2020 IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C), Washington, DC, USA.

[23] Fang, J., Zheng, L., Liu, C. (2024). A novel method for missing data reconstruction in smart grid using generative adversarial networks. *IEEE Transactions on Industrial Informatics*, 20: 4408–4417.

[24] Fang, J., Chen, K., Liu, C., He, J. (2023). An explainable and robust method for fault classification and location on transmission lines. *IEEE Transactions on Industrial Informatics*, 19: 10182–10191.

[25] Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q. (2020). ECA-net: Efficient channel attention for deep convolutional neural networks. In: Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA.

[26] Xu, Y., Zou, Z., Liu, Y., Zeng, Z., Wen, Y, Jin, T. (2025). Information fusion diagnosis of switching tube open-circuit fault in V2G charging piles based on multi-scale convolutional neural network and dual-attention mechanism. *Proceedings of the CSEE*, 45: 2992–3003. (in Chinese)

[27] Wang, W., Feng, B., Huang, G., Liu, Z., Ji, W., Guo, C. (2023). Short-term net load forecasting based on self-attention encoder and deep neural network. *Proceedings of the CSEE*, 43: 9072–9083. (in Chinese)

[28] Fang, J., Liu, C., Su, C., Lin, H., Zheng, L. (2023). Multi-stage transient stability assessment of power system based on self-attention transformer encoder. *Proceedings of the CSEE*, 43: 5745–5759. (in Chinese)

[29] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I. (2017). Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA.

[30] Zhang, X., Wang, Z. (2023). Spatial proximity feature selection with residual spatial–spectral attention network for hyperspectral image classification. *IEEE Access*, 11: 23268–23281.

[31] Cui, Y., Wang, Y., Huang, Y., Wang, Z., Wang, M. (2023). Closed-loop wind power ultra-short-term forecasting strategy based on multi-attention framework and guided supervised learning. *Proceedings of the CSEE*, 43:1334–1347. (in Chinese)

[32] Van der Maaten, L., Hinton, G. (2008). Visualizing data using tSNE. *Journal of Machine Learning Research*, 9: 2579–2605.