# Measurement-Driven Source Tracing of Torsional Sub-synchronous Oscillations Caused by Open-Loop Modal Resonance

Wenjuan Du, *Member, IEEE*, Jian Chen, Yang Wang and H. F. Wang, *Senior Member, IEEE*

***Abstract*** **— Open-loop modal resonance is the mechanism that a grid-connected wind farm excites dangerous torsional sub-synchronous oscillations (SSOs) of a synchronous generator. However, the open-loop modal resonance analysis (OMA) has to rely on the parametric model. Whilst, in practice it is often difficult to gain the parametric information of wind farms so as to establish the parametric model for the OMA. Hence, this paper proposes a method of the OMA based on the measurement data, rather than parametric model. The proposed method is particularly for the application to trace the source of torsional SSOs, which is the trouble-making wind farm taking part in the open-loop modal resonance. The proposed method of the OMA is developed by the machine learning to train a model of convolutional neural network (CNN). In order to solve the problem of lack of sufficient and labeled training data, it is proposed to build a simplified simulation system. Then, the CNN model is trained by deep transfer learning algorithm using the training data generated by the simulation system to trace the source of torsional SSOs in a practical power system from limited measurement data. In the paper, the proposed method is demonstrated and evaluated by a test example.**

***Index Terms*** **—Transfer learning, torsional sub-synchronous oscillations, open-loop modal resonance.**

LIST OF ABBREVIATIONS

| | | |
|---|---|---|
| SSO | - | Sub-synchronous oscillations. |
| SG | - | Synchronous generator. |
| IMA | - | Impedance model-based analysis. |
| WTG | - | Wind turbine generator. |
| OMA | - | Open-loop modal resonance analysis. |
| ML | - | Machine learning. |
| TL | - | Transfer learning. |
| CNN | - | Convolutional neural network. |
| JDA | - | Joint distribution adaptation. |
| DFIG | - | Doubly-fed induction generator. |
| PMSG | - | Permanent magnet synchronous generators. |
| SNR | - | Signal noise ratio. |

## I. INTRODUCTION

### *A. Motivation of study*

DURING the summer of 2015, several incidents of sub-synchronous oscillations (SSOs) caused by the grid-connected wind power generation occurred in the power network in Harmi area of western China. At least in one incident, torsional protection of some synchronous generators (SGs) in the power network was triggered, indicating the danger of torsional SSOs induced by the grid-connected wind power [1-2]. The incidents have caused great concerns because at the time the mechanism about why the SSOs happened in the Harmi power network was completely unknown.

Before the Harmi incidents, a great effort had been spent by power system researchers and engineers to apply the impedance model-based analysis (IMA) in order to understand the mechanism about how a sending-end wind turbine generator (WTG) connected to the grid through a series compensated transmission line may cause the SSOs [3-5]. Later, the IMA was used to examine why the SSOs may be induced by the grid-connected wind power in a power network without any series compensated lines, such as that in the Harmi area [6-7]. However, the IMA cannot describe exactly how the grid-connected wind power may interact with the torsional dynamics of the SGs to excite the torsional SSOs [8]. The authors proposed the open-loop modal resonance analysis (OMA) [9] which focused on the impact of dynamic interactions between the wind power and remainder of the power system. It was theoretically proved that when an oscillation mode associated with the wind power is in the proximity of a torsional oscillation mode of a SG on the complex plane, it is likely that the damping of the torsional oscillation mode may degrade, leading to torsional SSOs in the worst case [10-11]. Hence, the open-loop modal resonance reveals the mechanism about why the torsional SSOs may be excited by the grid-connected wind power from the standpoint of modal condition.

However, the OMA proposed by the authors in [9-11] requires to establish the parametric model of the power network integrated with wind power. This has been the main drawback to apply the OMA in practice, because the manufacturers usually do not provide the parametric information of the WTGs for the sake to protect their commercial interests and intellectual property. Consequently, it is extremely difficult or even impossible to establish the parametric model of a practical power network with grid-connected wind power in order to apply the OMA. To overcome this main drawback has motivated the study presented in this paper, which focuses on the following particular problem to apply the OMA based on the measurement data rather than parametric model.

Fig. 1a shows configuration of a practical power system with grid-connected wind farms. Parameters of the WTGs in the

W. Du, Yang Wang (Corresponding author) and H. F. Wang are with the School of Electrical Engineering, Sichuan University, Chengdu, China. J. Chen is with the School of EEEng., North China Electric Power University, Beijing, China.

This work is completed by the Engineering Special Team of Sichuan University on New Energy Power Systems and supported by the Natural Science Foundation of China under Grant 52077144.

wind farms are unknown. It is known that torsional dynamics of a SG in the power system are excited by a grid-connected wind farm (named as the SSO excitor in the paper) under the condition of open-loop modal resonance. Thus, how can the SSO excitor (i.e., source of open-loop modal resonance), which takes part in the open-loop modal resonance with the SG, be detected by using the operational measurement data of the practical power system?

It is of great practical significance to solve the problem outlined above, because tracing the SSO excitor is the basis to develop counteracting measures to eliminate the danger of open-loop modal resonance which causes the torsional SSOs in the practical power system. Otherwise, if the torsional SSOs are excited, the only way to avoid the collapse of entire power system is to shed the wind power by disconnecting ALL the wind farms. This means considerable loss of wind power and cost of extra spinning reserve to cope with the possible loss. So far, the SSO excitor to cause the SSO incidents in Harmi power network has remained unknown. Detection of the SSO excitor shall help significantly to develop measures for the SSO monitoring and suppression in the Harmi power network.

### *B. Literature review and strategy adopted by the study*

Without the parametric model, the possible solution for tracing the SSO excitor may be the measurement data-based methods. So far, there have been two main categories of such methods proposed and evaluated in the literature for other applications in power systems. The first category of methods is based on the traditional numerical algorithms. In this categories, representative innovative works are the energy flow method for tracing the sources of power system low-frequency oscillations [12-14] and the SSO power flow monitoring method to detect the sources of the SSOs [15-16]. Those methods are proposed on the basis of specific theory about the power system oscillations, i.e., the energy dissipation [12-16] and the impedance model-based analysis [17-18]. They are not applicable for the detection of the SSO excitor in the open-loop modal resonance. The second category of methods is the machine learning (ML) [19-21], which have shown great potential to solve the data-driven problems in power systems [22-24]. However, up to date there has been no work published to apply the ML for the SSO source tracing. The reason may be elaborated as follows.

Success of the ML relies on a large amount of labeled data to train certain deep learning models. However, on one hand, the amount of SSO measurement data of a real power system usually is too limited to be used for training a deep learning model. On the other hand, the SSO measurement data usually is not labeled, i.e., the sources are unknown, and hence cannot be used for training the deep learning model. Therefore, in order to solve the problem of lack of training data so as to develop a ML method to trace the SSO excitor, this paper proposes the following strategy.

First, a simulation system is constructed which is obtained by simplifying the real power system of Fig. 1a. Main simplifications include: (1) Displacement of a grid-connected wind farm by a grid-connected WTG or a constant power source if it is known that the wind farm is not the excitor. If the scale of power system with wind farms of Fig. 1a is very large, the displacement can be applied to a wind power collecting node which is connected to multiple wind farms. Thus, the scale of simulation system is kept to be small. (2) Displacement of a SG by a constant power source if it is known the torsional dynamics of the SG are not excited. The simplified simulation system is shown by Fig. 1b.

Second, by varying the load flow conditions, the parameters of the grid-connected WTGs and the torsional dynamics of the SG in the simulation system, training data of the simulation system is generated by running simulation. The generated data can be used for the ML because not only the number of data is sufficient, but also the data is labeled (i.e., the SSO excitor taking part in the open-loop modal resonance is known).

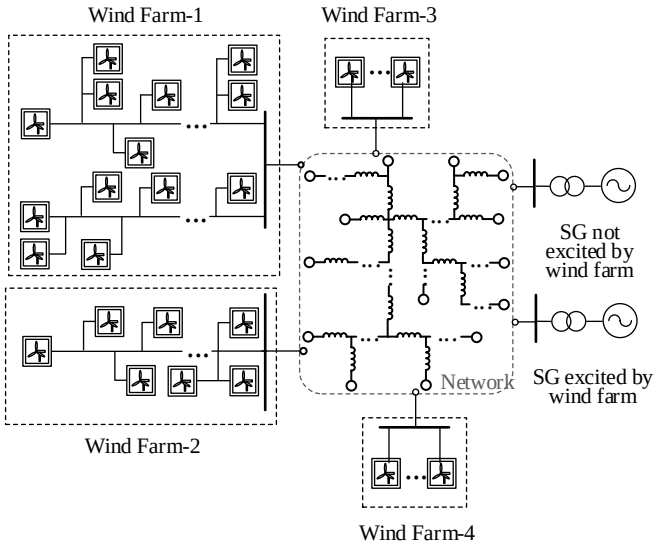



Fig. 1a. Real power system

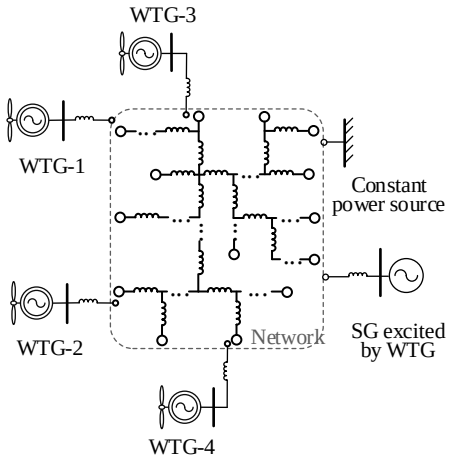



Fig. 1b. Simplified simulation system

Fig.1 Illustration of real power system and simplified simulation system

Third, transfer learning is used to train a deep learning model to trace the SSO excitor in the real power system by using both the labeled data of simulation system and unlabeled measurement data of the real power system. Transfer learning (TL) is one particular type of ML to learn unknown information in the target domain (the real power system) by using the knowledge gained in the source domain (simulation system) [25]. Hence, by use of the deep learning model trained according to the theory of transfer learning, the SSO excitor of open-loop

modal resonance in the real power system is detected.

To apply the TL, it is important that the target domain and the source domain have common factors. Those common factors determine the information in the target domain which the TL attempts to learn from the source domain [26]. In studying the SSO caused by the grid-connected wind power generation in a real power system, often a wind farm with many WTGs is represented by an aggregated WTG for simplifying the study [27]. It has been reckoned that the representation of the wind farm by the aggregated WTG can retain the key features of the wind farm to affect the SSO. Hence, in the simplified simulation system, the wind farms which affect the SSO are represented by the aggregated WTGs. The wind farms and SGs which hardly affect the SSO are represented by the constant power sources. Subsequently, the key factors, that the wind power generation affects the SSO, are kept in the simplified simulation system for ensuring the success of the TL to detect the SSO excitor.

In the simplified simulation model, it is not necessary to work out the aggregated parameters of the WTGs. In fact, the parameters of the WTGs are varied to generate labeled training data, i.e., the data of oscillations caused by the WTGs under the condition of open-loop modal resonance. Subsequently, it is ensured that the open-loop modal resonance is also the key common factor shared by the real power system and the simplified simulation system, because the purpose to apply the TL is to detect the SSO caused by the open-loop modal resonance in the real power system.

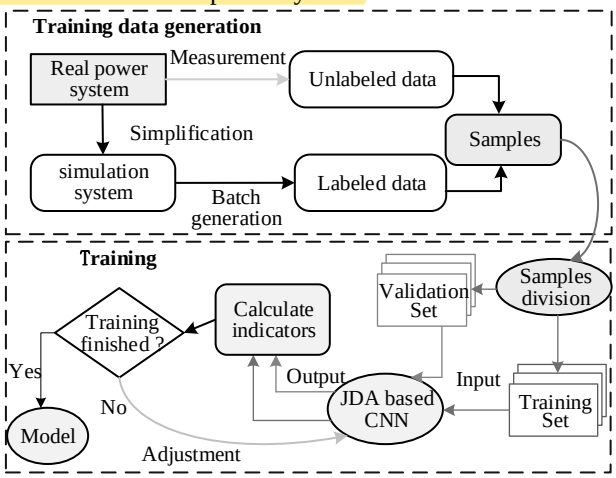



Fig.2 Strategy proposed

The strategy proposed in the paper for tracing the source of open-loop modal resonance to cause the torsional SSOs by the wind power generation, i.e., the SSO excitor, includes two stages, as being illustrated by Fig. 2. At the first stage, simplified simulation system is established firstly. Afterwards, unlabeled training data from the target domain (real power system) are gathered and labeled training data from the source domain (simplified simulation system) are generated. The second stage is to train a convolutional neural network (CNN) by a joint distribution adaptation (JDA) based TL algorithm [28]. For the training, data and the associated labels obtained in the first stage are used. For the convenience of discussion, in the paper, the data and the associated labels are referred to as training samples. In the training, the training samples are divided into two sets, the set of training samples and the set of validation samples. The training is carried out iteratively step by step. In one step of training, one batch of training samples are used to train the CNN. Afterwards, one batch of validation samples are used to test the CNN by calculating the validation indicators. If the performance of the CNN is improved as compared with the CNN trained in the previous step, the CNN trained in the current step is accepted and the training goes on to the next step. Otherwise, the CNN trained in the previous step is taken.

### *C. Organization and main contributions of the paper*

Organization of the paper is as follows.

In the next section, the aim of applying the ML based on measurement to carry out the OMA is introduced firstly. The introduction explains the rationality of the OMA by measurement-based ML. Secondly, the idea to implement the TL to overcome the shortage of training data for the ML is proposed. A simplified simulation system as the source domain to provide rich information of the SSO caused by the open-loop modal resonance is introduced. The strategy to trace the SSO excitor by the TL from the source domain to the target domain is discussed. Thirdly, the details of the JDA based TL algorithm are introduced.

In section III, a test example power system is presented. The test system consists of two SGs and five wind farms with total 120 WTGs, which are either the doubly-fed induction generators (DFIGs) or the permanent magnet synchronous generators (PMSGs). The proposed method of the JDA based TL to trace the source of the SSO excitor in the test system is demonstrated and evaluated. Test results by using the proposed method are compared with those obtained by using the conventional ML method. The comparison demonstrates the effectiveness of the proposed JDA based TL from the source domain (simplified simulation system).

Main contributions of the paper are the proposal and demonstration of a method of the JDA based TL to trace the SSO excitors which cause the open-loop modal resonance in a power system integrated with wind power generation. The paper proposes to construct a simulation system as the source domain in the TL which is simplified from the original real power system integrated with wind power generation. The proposed method of the JDA based TL can effectively train a JDA based CNN by using the source data generated from the simplified simulation system. The trained JDA based CNN can trace the SSO excitor in the original real power system integrated with wind power generation. Hence, the proposed method provides a potential solution to the problem of unbalanced and unlabeled SSO data, which has been the main challenge for applying the measurement-based ML to detect any SSO risk in a real power system.

## II. METHOD FOR TRACING THE SSO EXCITOR

### *A. Aim of the OMA by measurement-based ML*

Denoted $m(\mathbf{A})$ as the parametric model of the real power

system of Fig. 1a, where $\mathbf{A}$ is the state matrix of the model. Denote $\lambda_j$ as a torsional oscillation mode of the *j*th SG in the real power system which is excited by a grid-connected wind farm. Based on $m(\mathbf{A})$, the SSO excitor can be identified by calculating the participation factor of $\lambda_j$ from $\mathbf{A}$ for each of wind farms in the real power system. Let $PF_i$ be the participation factor of the *i*th wind farm and $|PF_i|$ is the maximum among the participation factors of all the wind farms. It is known that the open-loop modal resonance happens between the *i*th wind farm and the torsional dynamics of the *j*th SG. Hence, the *i*th wind farm is identified as the SSO excitor [10-11].

Let $\boldsymbol{x}$ be the vector of measurement signals of the real power system of Fig. 1a. The signals can be the terminal voltage, output current and power of wind farms, etc., which demonstrate the features of the torsional SSOs. Obviously, given $m(\mathbf{A})$, $\boldsymbol{x}$ is determined and vice versa as far as the excitation of torsional SSO under the condition of open-loop modal resonance is concerned. The relations between $\boldsymbol{x}$, $m(\mathbf{A})$ and the *i*th wind farm (SSO excitor) can be simply summarized as

$$\boldsymbol{x} \leftrightarrow m(\mathbf{A}) \rightarrow \lambda_j \rightarrow PF_i \rightarrow i \tag{1}$$

Due to (1), measurement-driven method has been proposed in the literature to identify the participation. For example, in [29], the most effective installing locations of damping controllers are identified by using the measurement data, instead of the parametric model, because the relations shown in (1) give the mapping from the measurable features of the real power system to the participation of damping controllers. In the case of source tracing, the relations given in (1) indicate the existence of the mapping to $i$, which is the label of SSO excitor, i.e., the *i*-th wind farm. The mapping can be expressed as

$$\begin{aligned} &y = f(\boldsymbol{x}); \boldsymbol{x} \in \boldsymbol{X}; y \in \boldsymbol{Y} \\ &\boldsymbol{Y} = \{y | y \text{ is the label of SSO excitor, i.e., the ith wind farm}\} \end{aligned} \tag{2}$$

where $\boldsymbol{X}$ is the feature space and $\boldsymbol{Y}$ the label space of the real power system respectively.

Since system dynamics can be well described by two-dimensional images, i.e., curves of $\boldsymbol{x}$ versus time, the convolutional neural network (CNN) based regression model can be used to learn the complex relationships between the measurable signals and the labels so as to approximate represent the function in (2), i.e., $f_{CNN} \approx f$. Thus, the trained CNN can realize the source tracing of the torsional SSOs, i.e., to find $y$, from the measurement data, i.e., $\boldsymbol{x}$, rather than by using the parametric model $m(\mathbf{A})$.

### *B. Tracing SSO excitor by transfer learning*

The normal CNN training by the deep learning requires a large amount of labeled data, i.e., for each $\boldsymbol{x} \in \boldsymbol{X}$, corresponding label $y \in \boldsymbol{Y}$ (SSO excitors) is known. However, such training data are rarely available from a real power system, as having been discussed in the previous section. Hence, the transfer learning (TL) is adopted to train the CNN for tracing the SSO excitor. The TL is to gain the information in a target domain, about which the knowledge is very limited, by using the information in a source domain, about which the knowledge is rich. The more common factors shared by two domains, the more accurate the transfer learning can be.

Let the SSO excitor tracing in the real power system of Fig. 1a be the target task. A simulation system shown by Fig. 1b is constructed. By varying the parameters of the WTGs and load flow conditions of simulation system, a large amount of labeled data (SSO excitor is known) is generated by running simulation for the CNN training.

The open-loop modal resonance to induce the SSO in a power system is mainly affected by three factors: (1) Dynamics of wind farms and the SGs; (2) Topology and network parameters; (3) Load flow conditions of the power system. The simulation system shares the following common factors of the open-loop modal resonance with the real power system: (1) Parameters of the replacing WTGs in the simulation system are variable such that dynamics of the wind farms are simulated; (2) Topology and network parameters are the same; (3) Variable load flow conditions of simulation system cover the same affecting factor of load flow conditions. Sharing of those common factors is the foundation to ensure the success of the TL from the source domain (simulation system) to the target domain (real power system).

### *C. Training of a JDA based CNN by the TL*

Denote $\boldsymbol{D}_s$ and $\boldsymbol{D}_t$ as the source domain and target domain separately, where subscripts *s* and *t* represent the source and target respectively. Feature space in each domain is $\boldsymbol{X}_s$ and $\boldsymbol{X}_t$. Task in each domain is comprised of its label space $\boldsymbol{Y}$ and a mapping function $f$. Given the labeled source domain and unlabeled target domain, the TL is to learn the mapping function in the target domain to be $f_t$, i.e., $y_t = f_t(\boldsymbol{x}_t); \boldsymbol{x}_t \in \boldsymbol{X}_t; y_t \in \boldsymbol{Y}_t$ [30].

The TL to train a CNN, expressed as an abstract function, $f_{CNN}$, to approximately represent the mapping function in the target domain, $f_t$, i.e., $f_{CNN} \approx f_t$. The training is carried out to minimize the following objective function [28]

$$\min_{f_{CNN} \in \mathcal{H}_{\mathbf{K}}} \left[ \sum_{i=1}^{n} L\left( f_{CNN}(\boldsymbol{x}_i), f_t(\boldsymbol{x}_i) \right) + \sigma \| f_{CNN} \|_{\mathbf{K}}^2 + \alpha JDA \right] \tag{3}$$

where $\boldsymbol{x}_i \in \boldsymbol{X}_s \cup \boldsymbol{X}_t$ is the sampled data of both source and target domain; $n$ is the total number of samples; $L(\ )$ is the loss function which measures the fitting error of the CNN to the mapping function $f_t$; Subscript $\mathbf{K}$ indicates that $f_{CNN}$ is written as the kernel function with the base function to be $\phi$, i.e., $\boldsymbol{K}(u,v) = \langle \phi(u), \phi(v) \rangle$; $\| \ \|_{\mathbf{K}}^2$ denotes nuclear norm of the kernel function; $\mathcal{H}$ represents the Hilbert space; $\sigma$ is the coefficient of the structural risk; $\alpha$ is the regularization weight and *JDA* is the joint distribution adaptation to avoid the negative transfer by minimizing the difference of the

probability distribution between the source and target domain [28].

In (3), $L(\ )$ is chosen to be the cross-entropy function, which is a kind of loss function suitable for the problem of classification [31]. Minimizing $L(\ )$ is in order to ensure the smallest error between $f_{CNN}$ and $f_t$. However, the minimization must not be achieved at the cost of complex structure of the CNN, which is ensured by minimizing $\| f_{CNN} \|_{\mathbf{K}}^2$ [32]. Hence, the first two items in (3) is to minimize the structural risk in training the CNN.

In the TL, if the source and target domain are of less common factors, negative transfer may occur which may undermine the classification performance in the target domain [33]. Joint distribution adaptation (JDA) is introduced in [28] to avoid the negative transfer by minimizing the difference of the probability distributions of the source and target domain. When the simulation system of Fig. 1b is simplified from the real power system of Fig. 1a, some factors affecting the open-loop modal resonance of real power system may not exactly be represented by the simulation system. Hence, possible negative transfer may happen and must be avoided to ensure the success of the TL from the simulation system to the real power system. The last item in (3) is to ensure the minimization of difference of probability distribution. Hence, it can have in (3) that [28]

$$JDA = \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \phi(\boldsymbol{x}_{si}) - \frac{1}{n_t} \sum_{j=1}^{n_t} \phi(\boldsymbol{x}_{tj}) \right\|_{\mathcal{H}}^2 + \left\| \frac{1}{n^c} \sum_{x_i \in D_s^c} \phi(\boldsymbol{x}_{si}) - \frac{1}{m^c} \sum_{x_j \in D_t^c} \phi(\boldsymbol{x}_{sj}) \right\|_{\mathcal{H}}^2 \quad (4)$$

where $\boldsymbol{x}_{si} \in \boldsymbol{X}_s$, $\boldsymbol{x}_{ti} \in \boldsymbol{X}_t$; $\boldsymbol{D}_s^c = \{\boldsymbol{X}_i : \boldsymbol{X}_i \in \boldsymbol{D}_s, \boldsymbol{Y}(\boldsymbol{X}_i) = c\}$ is the set of samples with label $c$ in the source domain, $\boldsymbol{D}_t^c = \{\boldsymbol{X}_i : \boldsymbol{X}_i \in \boldsymbol{D}_t, \boldsymbol{Y}^*(\boldsymbol{X}_i) = c\}$ is the set of samples from pre-labeling $c$ in the target domain, $\boldsymbol{Y}(\boldsymbol{X}_i) = c$ and $\boldsymbol{Y}^*(\boldsymbol{X}_i) = c$ are the real label in the source domain and predicted label in the target domain, respectively, $n^c$ and $m^c$ indicate the number of labels in the source and target domain, respectively.

With (3) and (4), the CNN architecture based on the TL is determined by minimizing the structural risk and the difference between two domains based on the regularization theory [34].

Conventional CNN is comprised of one input layer, multiple pairs of convolutional and pooling layers, one fully connected layer, and one Softmax classifier. The CNN can be trained by taking data samples from a single system as the input to produce the labels as the output. In the case of the TL, input data samples are $\boldsymbol{X}_s$ and $\boldsymbol{X}_t$ from two different systems. Hence, a JDA based TL algorithm is proposed to measure the difference between the samples obtained from the simulation system (the source domain) and the real system (the target domain). The structure of the proposed JDA based CNN for tracing the SSO excitor is depicted in Fig. 3. In the proposed structure, the JDA calculation module is added behind the main CNN. The data samples from $\boldsymbol{X}_s$ and $\boldsymbol{X}_t$ are used to train the CNN module in parallel and then to train JDA module for cross calculation.

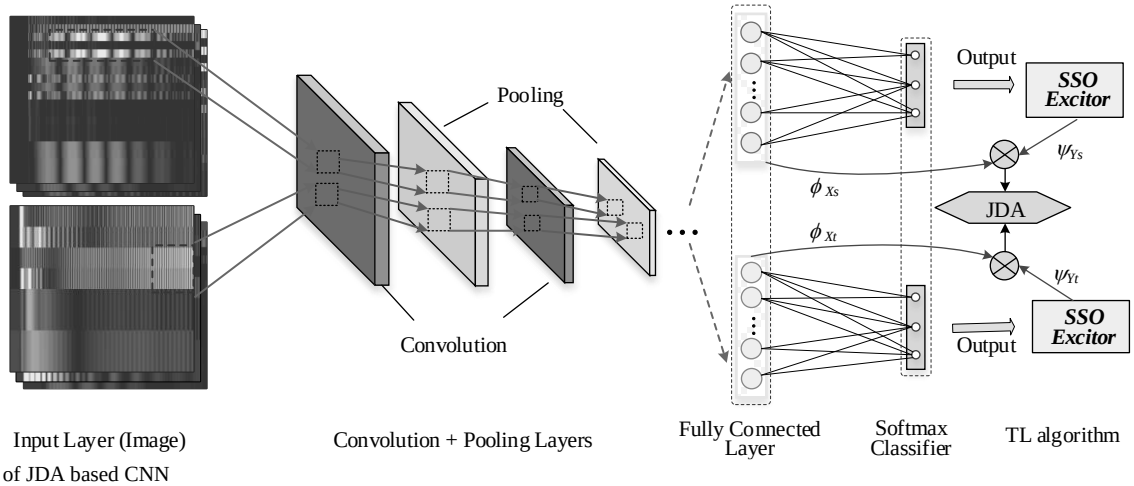



Fig. 3. Structure of JDA based CNN

Based on the proposed structure of the JDA based CNN and by applying the TL algorithm, the minimization problem of (3) can be solved in the PyTorch ML framework [35]. The solution is illustrated by the algorithm displayed in the following table.

**Algorithm:** JDA based TL

**Input:** Labeled training data from the source domain simulation, unlabeled training data from the target domain, $\sigma$ and $\alpha$.

**Output:** Trained CNN for the SSO excitor, *i.e.*, $f_{CNN}$

**Begin:**

*Construct CNN in the Pytorch framework, and use labeled training data from the source domain to pre-train the network*

**For** each epoch **do:**

1. Taking $\boldsymbol{X}_s$ and $\boldsymbol{X}_t$ from training samples of source and target domain as model input, respectively, predict corresponding outputs $\boldsymbol{Y}'_s$ and $\boldsymbol{Y}'_t$
2. Use labeled training data ($\boldsymbol{X}_s$, $\boldsymbol{Y}_s$) in source domain to calculate the structural risk loss of the CNN
3. Calculate the regularization term of joint distribution adaptation according to (4).

4. According to (3), calculate the total loss.
5. Using the total loss, update the model parameters of classifier $f_{CNN}$.

**End for**

**Return:** Trained SSO excitor tracing model $f_{CNN}$

## III. A TEST EXAMPLE POWER SYSTEM

The method proposed in the above section for tracing the SSO excitor by the TL has been tested and evaluated in several example power systems with grid-connected wind farms. Due to the limit of space, in this section only a small-scale power system is presented to demonstrate the procedure to apply the method and the results of the application. The TL was carried out in Anaconda platform. The time domain simulation was conducted in MATLAB. The host computer is equipped with an Intel Xeon Gold-5217 (×2) processor, 128 GB memory, and two NVIDIA Tesla V100 GPUs with 16 GB memory.

### *A. The test example power system*

Configuration of the test example power system is shown by Fig. 4, which is a part of simplified real power system in the western China. $WF_i; i=1,2,\cdots 5$ and $SG_k; k=1,2$ denote respectively the *i*-th wind farm and the *k*-th SG in the example power system. Each of $WF_i; i=1,2,4$ is comprised of 20 PMSGs and each of $WF_i; i=3,5$ consists of 30 DFIGs. The 15th-order model of PMSG given in [36] and 15th-order model of DFIG given in [37] are adopted. The model of SG, including the torsional dynamics, given in [10] is used. Network data are presented in Appendix A.

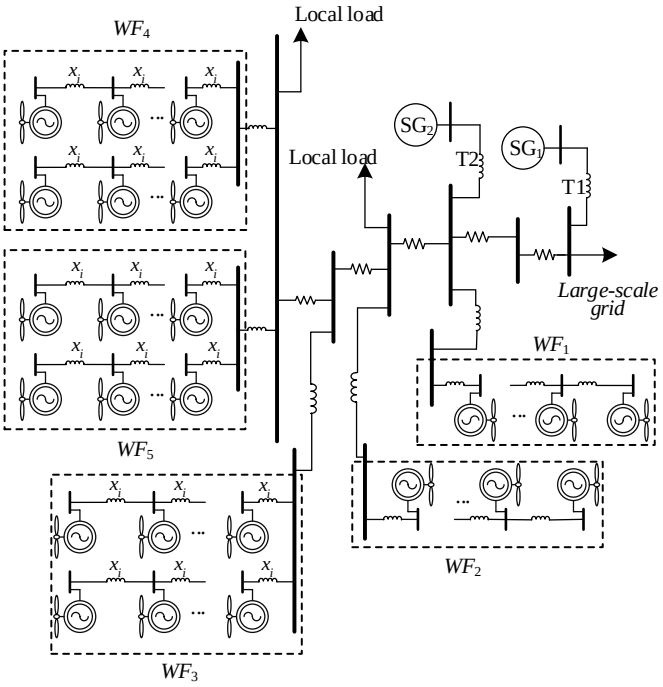
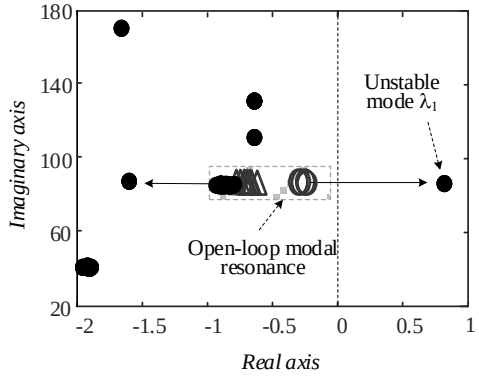
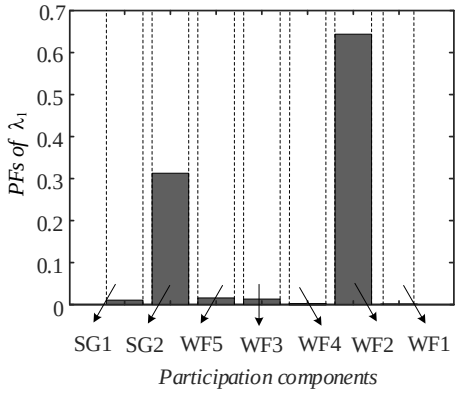



Fig. 4. Configuration of test example power system.

In the test example power system, open-loop modal resonance happened between $SG_2$ and $WF_2$ when one torsional oscillation mode of $SG_2$ and one of oscillation mode of $WF_2$ were in the proximity on the complex plane. This was the case that $WF_2$ excited the torsional dynamics of $SG_2$ to cause growing SSO in the test example power system [9]. Identification of $WF_2$ as the SSO excitor was achieved by using the parametric model of the test example power system. Fig. 5 presents the computational results of open-loop modal resonance and the participation factors of the unstable torsional oscillation mode by using the state-space model of the test example power system. In Fig.5(a), the filled circles represent the oscillation modes of the test example power system, and the red hollow circles and red hollow triangle respectively represent the open-loop oscillation modes of $SG_2$ and $WF_2$.




(a) Results of modal computation



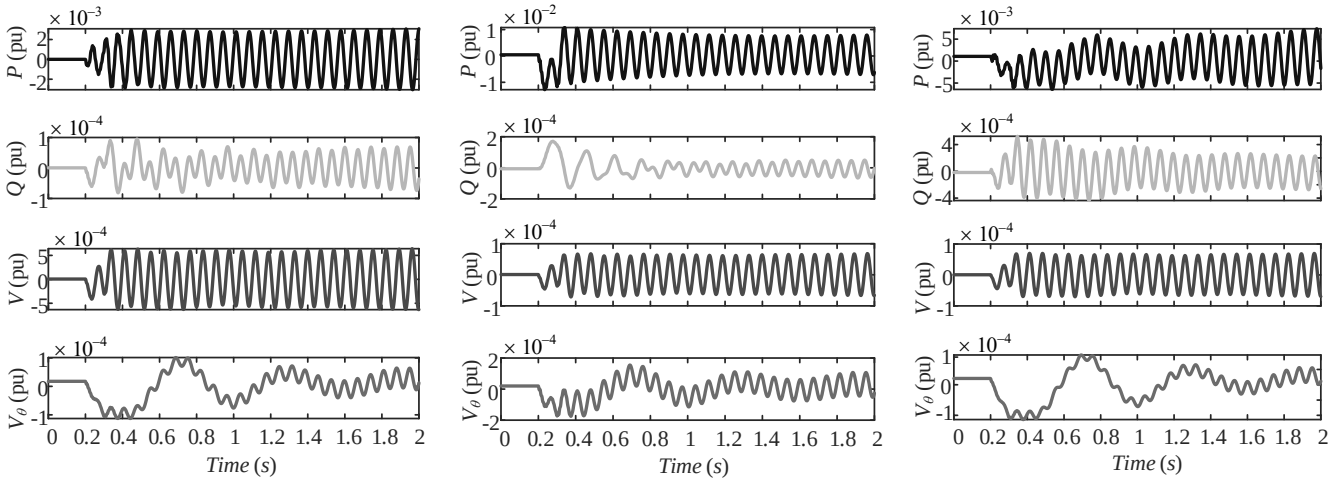

(b) PFs of unstable oscillation modes

Fig. 5 Computational results of open-loop modal resonance based on the parametric model of test power system.

However, in practice the detailed parameters of PMSGs and DFIGs in $WF_i$ ; *i*=1,2,3 of the test example power system normally are unavailable. Hence, computational results obtained by using the parametric model of the test system and presented in Fig. 5 are only for confirming the correctness of method of the TL proposed in the previous section. To apply the proposed method, it is assumed that the only available information is the measurement data when the SSOs occurred. Those data were recorded as $\boldsymbol{D}_t=\{\boldsymbol{X}_t\}$, where $\boldsymbol{X}_t$ contains *N* samples, and the *n*th sample is represented as $\boldsymbol{x}_{tn}=\{\boldsymbol{P}_{w1},\boldsymbol{Q}_{w1},\boldsymbol{V}_{w1},\boldsymbol{V}_{\theta w1},...,\boldsymbol{P}_{wm},\boldsymbol{Q}_{wm},\boldsymbol{V}_{wm},\boldsymbol{V}_{\theta wm}\}$, where $\boldsymbol{P}_{wi}$, $\boldsymbol{Q}_{wi}$, $\boldsymbol{V}_{wi}$, $\boldsymbol{V}_{\theta wi}$ are respectively the active, reactive power output from, the magnitude and phase of terminal voltage of the *i*th wind farm; *i*=1,2,3,4,5; *m*=5.

Ideally, the measurement data of oscillations in the test example power system should be the actual recorded data collected in the field. Unfortunately, those actual data are not available because of the lack of installed measurement device in the field. It is more possible that after the proposed method in the paper is validated by simulation in the laboratory, deployment of measurement device may be considered in a real power system. At present, this study is at the stage of method evaluation by simulation. Hence, for the test example power

system depicted by Fig. 4, data from system simulation are used as the measurement data. Considering the fact that in practice, amount of measurement data of oscillations is very limited, only 500 samples of data were generated to be used as the measurement data for training the proposed JDA based CNN by the TL algorithm. Fig. 6 shows the representative results of torsional SSOs caused by the open-loop modal resonance. It can be seen that the torsional SSOs spread over the whole test example system. Hence, from the measurement data, it was impossible directly to identify the SSO excitor ($WF_2$).




(a) Measurement of $WF_1$ (b) Measurement of $WF_2$ (c) Measurement of $WF_3$

Fig. 6. Representative results of torsional SSOs recorded in the test example power system.

### *B. The simulation system and generation of training data*

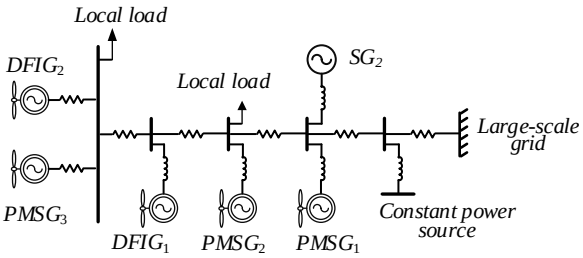



Fig. 7. Configuration of simplified simulation power system.

A simplified simulation system was constructed and shown in Fig. 7, where the topology and network parameters were as same as those of test example power system of Fig. 4. Wind farms were replaced by single WTGs and the unexcited SG was displaced by a constant power source. Loading conditions of the simulation system were changed by varying the local loads such that the active power output from the WTGs and the SGs varied in the range of 90~120% of their rated power. Parameters of converter control systems, including the PLLs, of the WTGs were varied. At each operating condition with each set of parameters of the WTGs, simulation was conducted for 2 second. For the simulation, a disturbance was added at 0.1 second for 100 ms. Results of one simulation were recorded as one sample data in the source domain, for example $\boldsymbol{x}_{sn}=\{\boldsymbol{P}_{w1},\boldsymbol{Q}_{w1},\boldsymbol{V}_{w1},\boldsymbol{V}_{\theta w1},...,\boldsymbol{P}_{wm},\boldsymbol{Q}_{wm},\boldsymbol{V}_{wm},\boldsymbol{V}_{\theta wm}\},m=5$ was the *n*th sample. If the torsional SSOs occurred, participation factors were calculated using the linearized model of simulation system to identify the SSO excitor so as to label the *n*th sample data. If $WF_i$ was identified to be the SSO excitor, the label of $\boldsymbol{x}_{sn}$ was $y_{sn}=i$. If no torsional SSOs occurred, $y_{sn}=0$.

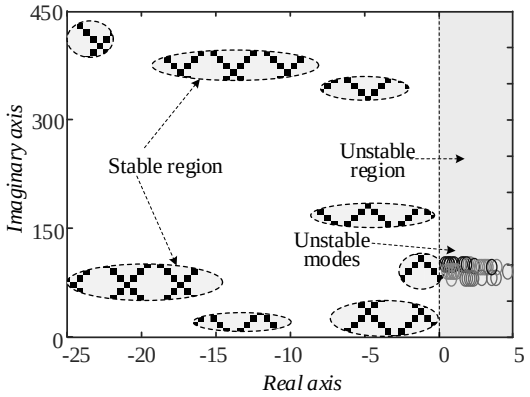
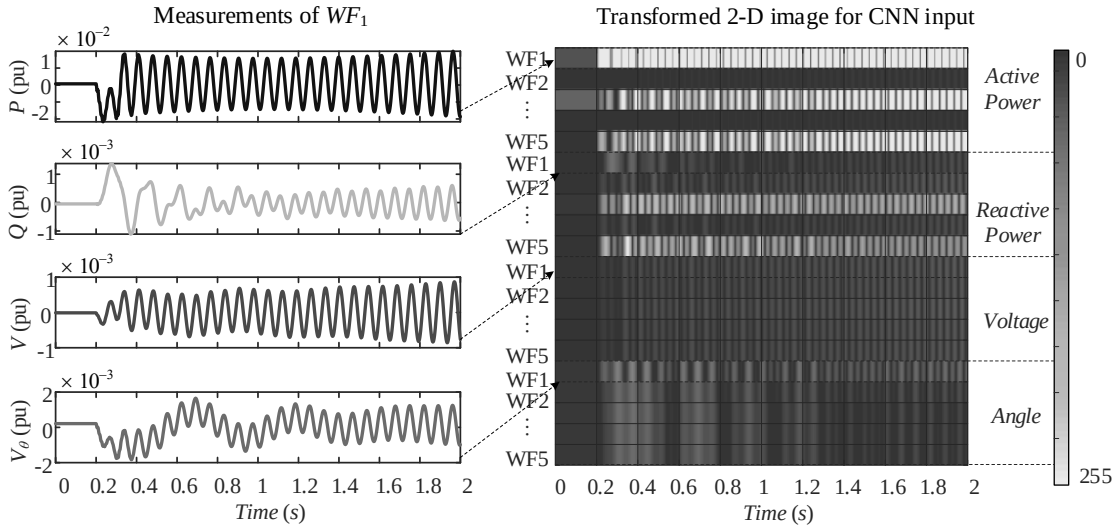
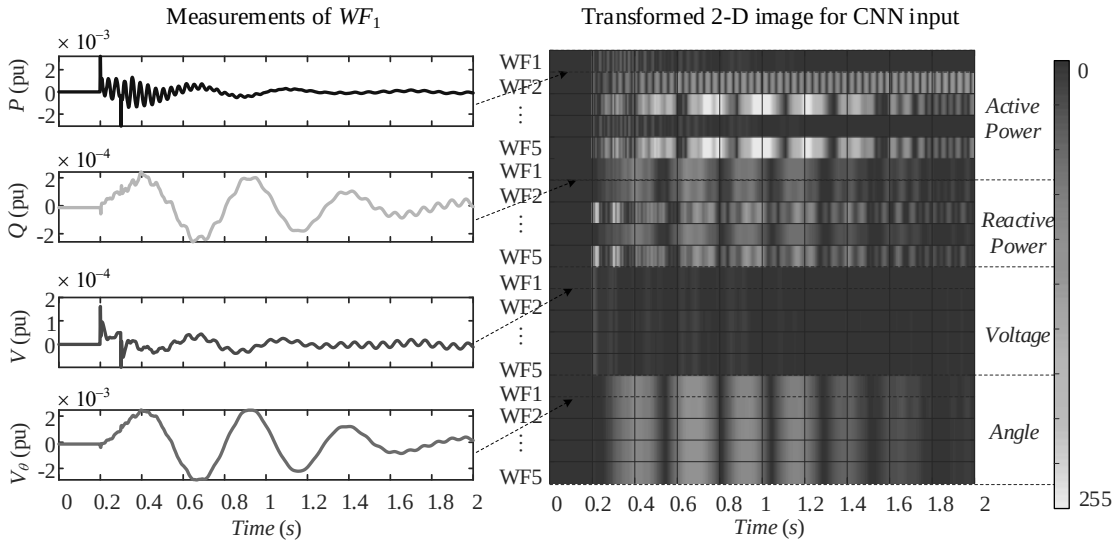



Fig. 8. Positions of dominant oscillation modes of samples in the source domain. Note: different color represents different SSO excitor, × is stable poles.

In total, 11440 simulations were conducted and the disturbance for each simulation was randomly selected as the power or load injection in the range of ±25% nominal value at the terminal of a WTG. Finally, the source domain sample set, denoted as $\boldsymbol{D}_s=\{\boldsymbol{X}_s,\boldsymbol{Y}_s\}$, where $\boldsymbol{X}_s=\{\boldsymbol{x}_{sn}\}$ and $\boldsymbol{Y}_s=\{y_{sn}\}$, was obtained for training the proposed JDA based CNN by the TL algorithm. The total number of sample data was 11440. To demonstrate the features of training data, Fig. 8 presents the positions of dominant oscillation modes of the simplified simulation system in some of the total 11440 simulated cases. Fig. 9 shows two typical samples and their two-dimensional image transformed from the sample of training data as the input to the JDA based CNN to be trained. The 2-D image is obtained by arranging one sample of discrete measurement data from all the wind farms together in the time sequence. The data is standardized with the standardized values varying from 0 to 255 in order to form a clear color map [38].




(a) Representative sample with label of 1




(b) Representative sample with label of 0

Fig. 9. Features of sample training data of simulation system.

### *C. Training of the JDA based CNN and results*

The JDA based CNN model for tracing the SSO excitor was constructed in PyTorch 1.2.0. The JDA based CNN includes the input and output layers, two modules of convolutional and pooling layers for feature extracting, the fully-connected layers and the Softmax classifier, which are the normal parts of a conventional CNN [30]. In addition, a JDA algorithm is added in order to apply the TL. The structure of the JDA based CNN is displayed in Fig. 3. The details of hyperparameter settings for the network are summarized in Appendix B.

A stochastic gradient descent (SGD) algorithm, named the Adam optimizer [39], was employed to train the parameters of the CNN model. Details about the Adam optimizer can be found in [39]. Other hyperparameters are empirically set to make CNN sufficiently trained [31].

A tracing accuracy index (TAI) was proposed to evaluate the model as

$$TAI = \sum_{i=0}^{m} \eta_i / N \tag{5}$$

where $N$ is the total number of samples to be evaluated, $\eta_i$ is the number of samples when the $i$th wind farm is correctly identified as the SSO excitor ($i$=1,2,3,4,5) and $i$=0 is the number of samples when the power system is stable.

In the power system, when using data mining methods for stability analysis, if an unstable sample is incorrectly classified as stable, it will have a serious impact on the system. The classification of a stable sample into an unstable sample is also a misclassification, but its impact on system security is much smaller. Considering the conservative nature of power system operation, it is required that while improving the tracing accuracy, attention should also be paid to reducing the total number of misjudged unstable samples. Hence, confusion matrix [40] was adopted to count the classification number of identified samples, as shown in Table I.

TABLE I
CONFUSION MATRIX

| *Predict label* / *True label* | SSO excitor | Stable |
|---|---|---|
| SSO excitor | *TO* | *FS* |
| Stable | *FO* | *TS* |

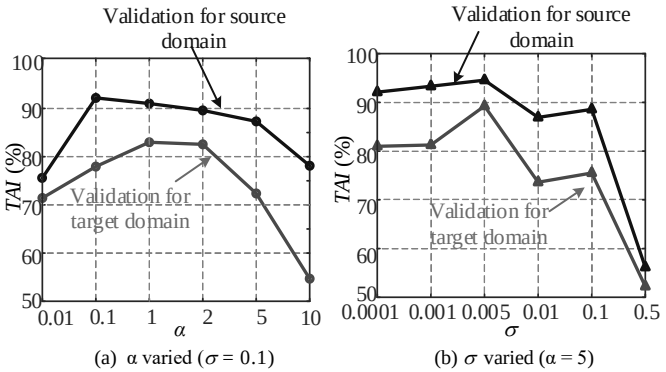
In Table I, *TO* is the number of unstable samples correctly identified (note that *TO* may include the cases of misjudging of SSO excitor from one wind farm to another), *TS* is the number of stable samples correctly identified, *FO* is the number of misjudged unstable samples, *FS* is the number of misjudged stable samples.

Then, precision rate (PR) and the recall rate (RR) are used to evaluate the classification performance of the SSO excitor tracing model. Following classification discrimination index (CDI) proposed in [40] was adopted to comprehensively consider PR and RR.

$$CDI = 2PR \times RR / (PR + RR)$$
$$\begin{cases} PR = TO/(TO + FO) \\ RR = TO/(TO + FS) \end{cases} \quad (6)$$

In training the JDA based CNN for locating the SSO excitor, the TAI and the CDI are calculated. When the indices or training epochs do not satisfy the expected thresholds, parameters of JDA based CNN are updated by using the TL algorithm proposed in Section I.

In training the JDA based CNN for the test example power system with samples from both the source domain and the target domain, 80% of total number of the samples were used for training and the remainder was for validating the model trained. Different values of the regularization weight α and the structure risk coefficient $\sigma$ might possibly affect the training performance of the JDA based CNN. Hence, an initial test training was conducted with $\alpha \in \{0.01, 0.1, 1, 2, 5, 10\}$ and $\sigma \in \{0.0001, 0.001, 0.01, 0.1, 0.5\}$ in 500 training epochs to search the suitable values of α and $\sigma$. Results of initial raining and the validation are presented in Fig. 10.




(a) α varied ($\sigma$ = 0.1) (b) $\sigma$ varied (α = 5)

Fig. 10. Variation of TAI when α and $\sigma$ varied separately.

Results displayed in Fig. 10 indicate that the performance of the JDA based CNN for tracing the SSO excitor is of different performance with values α and $\sigma$ being varied. When $\alpha \in [1.0, 2.0]$ and $\sigma \in [0.001, 0.01]$, performance of the JDA based CNN was good. Hence, with α=1.5 and $\sigma$=0.005, the JDA based CNN was trained by using the TL algorithm proposed in section II.

Finally, using the validating samples (20% of total number of samples which were not used in the training) as the input, the performances of the JDA based CNN for tracing the SSO excitor under different training epochs were evaluated. Statistical results of evaluation are summarized in Table II.

From Table II, it can be seen that the JDA based CNN achieved excellent performances with α=1.5 and $\sigma$= 0.005. Particularly, after training 500 epochs, both the tracing accuracy (TAI) and the classification discrimination (CDI) reached above 90%. Afterwards, the performance continued to improve with the increase of the training epochs. After 2000 training epochs, the training saturation was reached.

TABLE II
MODEL INDICES ACCORDING TO NUMBER OF TRAINING EPOCHS

| No. epochs | TAI (%) | CDI (%) | Training time (s) |
|---|---|---|---|
| 50 | 78.12 | 72.07 | 293.28 |
| 100 | 85.93 | 81.87 | 585.11 |
| 500 | 92.74 | 95.60 | 2918.18 |
| 800 | 98.43 | 98.27 | 4914.62 |
| 1000 | 100 | 100 | 5822.45 |
| 2000 | 100 | 100 | 11621.54 |

### *D. Further test of the trained JDA based CNN*

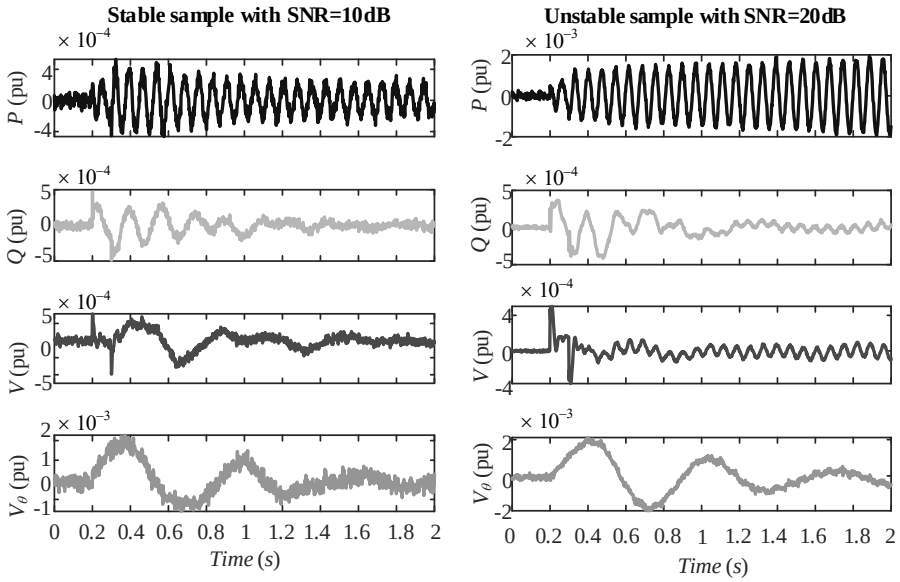
In order to further evaluate the adaptability of the JDA based CNN for tracing the SSO excitor which had been trained, three more scenarios, which were not included in the target domain, were considered. Those three new scenarios were generated from the real power system of Fig.2 as follows.

*Test scenario 1*: The measurement noise was added (Gaussian white noises with SNR=5~20dB). 200 samples of test example power system were generated and recorded.

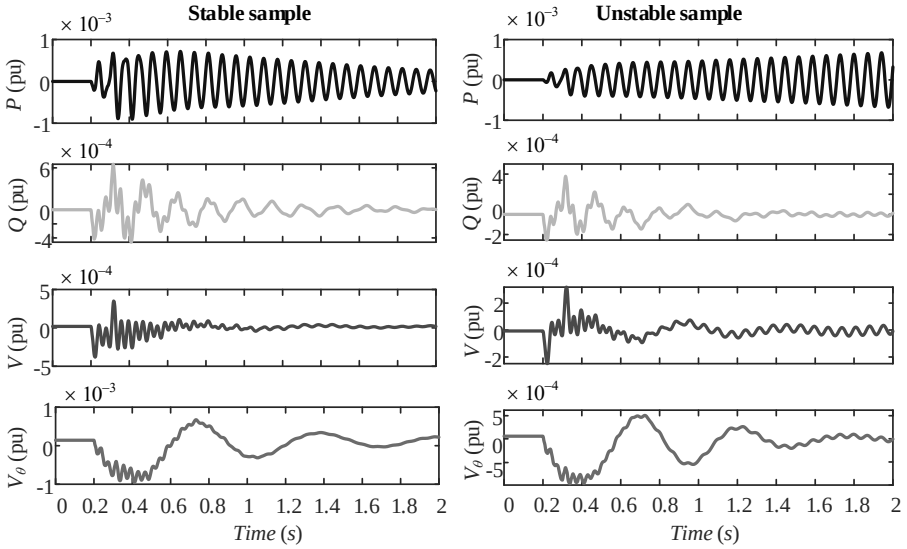
*Test scenario 2*: Previously, control parameters of the WTGs in $WF_2$ of test example power system were fixed when the measurement data were recorded. In this new test scenario, the control parameters were set with different values. Afterwards, 200 new samples of test example power system were generated and recorded.

*Test scenario 3*: By tuning the control parameters of the WTGs in $WF_1$ and $WF_2$, the open-loop modal resonance happened between $SG_2$ and $WF_1$ instead of between $SG_2$ and $WF_2$. Thus, the SSO excitor was $WF_1$ in this new scenario. Then, 200 test samples were generated and recorded.

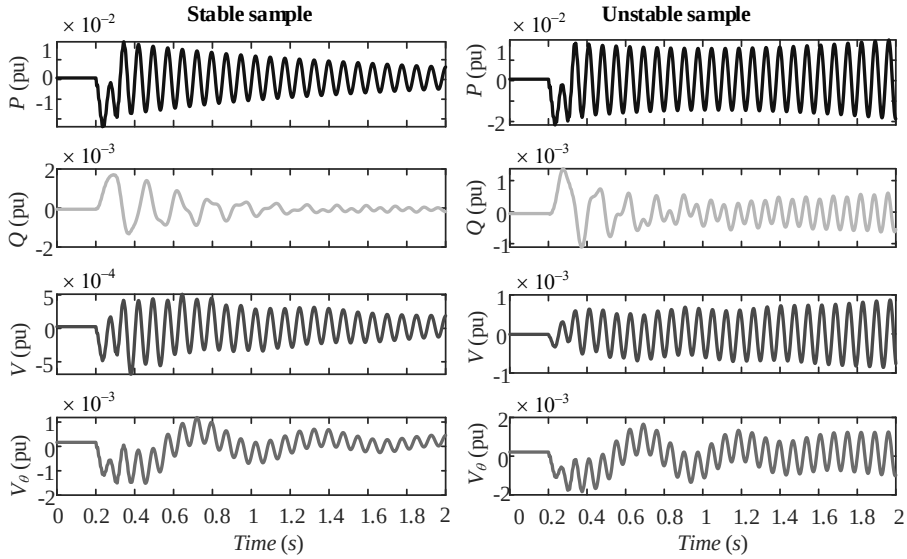
Representative samples recorded for those three scenarios are shown in Figs.11 -13.




(a) Measurements of $WF_2$ of stable and unstable sample in test scenario 1.




(b) Measurements of $WF_2$ of stable and unstable sample in test scenario 2.




(c) Measurements of $WF_1$ of stable and unstable sample in test scenario 3.

Fig. 11. Measurements display of a randomly chosen case of three test scenarios.

TABLE III
RESULTS OF JDA BASED CNN

| Source of test sample | TAI (%) | CDI (%) | PR (%) | RR (%) | Test time (s) |
|---|---|---|---|---|---|
| From validation source domain | 100 | 100 | 100 | 100 | 0.16 |
| From training target domain | 96.59 | 97.37 | 99.50 | 95.34 | 0.16 |
| Scenario 1 | 97.23 | 96.70 | 98.92 | 94.59 | 0.18 |
| Scenario 2 | 98.50 | 98.11 | 100.00 | 96.30 | 0.16 |
| Scenario 3 | 93.75 | 95.72 | 94.91 | 96.55 | 0.19 |

Afterwards, 600 test samples recorded in the three scenarios were transformed in to 2-D image and taken as the input to the JDA based CNN trained previously. Results of evaluating the JDA based CNN to trace the SSO excitor is summarized in Table III.

Results in Table III indicate that good performance of the JDA based CNN model trained previously was achieved to trace the SSO excitor. In addition, the trained JDA based CNN was of strong resistance to measurement noise. This merit is due to the CNN architecture as it inherently can filter the noise [41].

In terms of the computational complexities, the computational time consumed by the offline training seems to be high as it can be seen from the last column of Table II. However, this should be affordable to take approximately 2 hours to complete 1000 epochs training with 5 GB samples. After completing the training, it takes about 1 ms. per sample to locate the SSO excitor as being demonstrated in the last column of Table III. This demonstrates the great potential in on-line application for quick disconnection of the SSO excitor to eliminate the torsional SSOs occurred in real-time operation.

TABLE IV
RESULTS OF THE JDA BASED CNN TRAINED WITH POLLUTED MEASUREMENT DATA

| Source of test samples | TAI (%) | CDI (%) | PR (%) | RR (%) | Test time (s) |
|---|---|---|---|---|---|
| From validation source domain | 100 | 100 | 100 | 100 | 0.0619 |
| From training target domain | 94.00 | 92.59 | 90.36 | 94.94 | 0.0748 |
| Scenario 1 | 97.50 | 94.95 | 94.00 | 95.92 | 0.0668 |
| Scenario 2 | 96.00 | 92.00 | 90.20 | 93.88 | 0.0628 |
| Scenario 3 | 92.50 | 90.57 | 88.89 | 92.31 | 0.0738 |

In the training of the JDA based CNN in the previous subsection and the tests presented above, the measurement data of the real power system of Fig. 4 are obtained in the laboratory by conducting simulation. In application, those measurement data normally are polluted with the noise. To examine the specific aspect of impact of measurement noise, the training and the tests were re-conducted with the polluted measurement data from the real power system. The polluted measurement data were obtained by adding Gaussian white noises (SNR=5) to the results of laboratory simulation. Performance of the trained JDA based CNN and the test results are presented in Table IV. From Table IV, it can be seen that although the measurement data were very badly polluted (the level of measurement noise in real power systems normally is much lower than the Gaussian white noises with SNR=5), the trained JDA based CNN is still of good performance to trace the SSO excitor.

### *E. Comparison between the proposed JDA based CNN with the TL algorithm and the conventional CNN without the TL*

Common factors shared by the target domain and the source domain are important for the successful TL. In the proposed method for tracing the SSO excitor, the simplified simulation system is proposed as the source domain because the data from the real power system (target domain) are too limited to be used to train a CNN. The simplified simulation system shares two important common factors with the real power system, where the SSO excitor is to be traced. The first common factor is that the SSO is caused by the open-loop modal resonance. The second common factor is the same configuration and parameters of power system network. A wind farm in the real power system is represented by a WTG so as to simplify the real power system. This is the usual strategy to represent the wind farm as the aggregated WTG, which can retain the key characteristics of the wind farm to affect the SSO [27]. Hence, the simplified simulation system (source domain) is of the same key factors to affect the SSO. This ensures the TL to train the JDA based CNN by mainly using the data from the simplified simulation system to trace the SSO excitor in the real power system.

In addition to the common factors shared by the target domain and the source domain, of course, as being discussed above, the TL algorithm is the most important factor to determine the success to trace the SSO excitor by the proposed JDA based CNN. In order to demonstrate the decisive role of the TL algorithm with the JAD, in this subsection, tests are carried out to compare the performance between the proposed JDA based CNN with the TL and the conventional CNN without the TL. Test results of comparison are as follows.

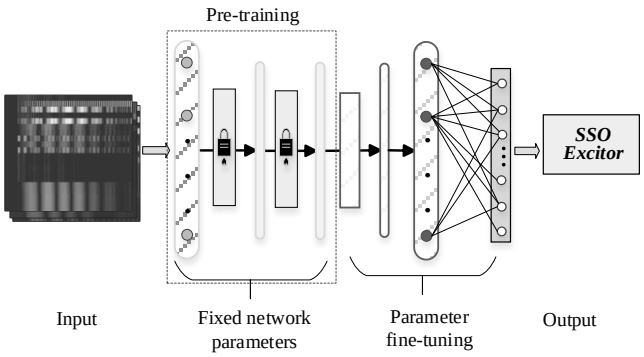



Fig. 12. Structure of conventional CNN with pre-training

Structure of the conventional CNN is displayed in Fig. 12. Two types of commonly used algorithms [31,42] are used to train the CNN for the performance comparison. The first one has no pre-training and uses the data from the simplified simulation system to train the CNN because the data from the real power system are not sufficient for the training [42]. The second uses the data from the real power system to conduct a pre-training.

Afterwards, the data from the simplified simulation system are used to train the CNN [31]. Performance and test results of trained conventional CNN without and with the pre-training are presented in Table V and VI.

TABLE V
RESULTS OF CONVENTIONAL CNN WITHOUT PRE-TRAIN

| Source of test sample | TAI (%) | CDI (%) | PR (%) | RR (%) | Test time (s) |
|---|---|---|---|---|---|
| From target domain | 75.50 | 78.67 | 65.56 | 97.37 | 0.21 |
| From validation source domain | 99.00 | 97.95 | 96.00 | 100 | 0.21 |
| Scenario 1 | 79.68 | 87.71 | 79.36 | 98.04 | 0.22 |
| Scenario 2 | 80.00 | 72.97 | 58.06 | 98.18 | 0.22 |
| Scenario 3 | 76.50 | 86.45 | 6.00 | 100.00 | 0.20 |

TABLE VI
RESULTS OF CONVENTIONAL CNN WITH PRE-TRAINING

| Source of test sample | TAI (%) | CDI (%) | PR (%) | RR (%) | Test time (s) |
|---|---|---|---|---|---|
| From training target domain | 94.50 | 93.74 | 90.00 | 97.82 | 0.16 |
| From training source domain | 89.87 | 89.39 | 55.56 | 99.14 | 0.21 |
| Scenario 1 | 90.50 | 90.06 | 85.15 | 95.56 | 0.18 |
| Scenario 2 | 89.50 | 87.07 | 82.05 | 92.75 | 0.17 |
| Scenario 3 | 87.05 | 70.70 | 61.40 | 83.33 | 0.20 |

By comparing the results presented in Table V and VI with those in Table III and IV, it can be seen that the performance of the JDA based CNN proposed in the paper is better than that of conventional CNN. Particularly, test results of the proposed JDA based CNN is of the better performance to handle with the cases which are not included in the training (Scenario 2 and 3). This indicates the better general applicability of the proposed JDA based CNN and thus validates the effectiveness of the JDA based TL algorithm adopted in the paper. Indicator PR is the ratio to wrongly judge the unstable cases as the stable cases. Low values of the PR in Table V and VI indicate that the conventional CNN will not be accepted in practical application.

## IV. CONCLUSIONS

The open-loop modal resonance analysis (OMA) can trace the source of torsional SSOs, which is a grid-connected wind farm taking part in the open-loop modal resonance to excite the torsional dynamics of a SG (hence, the wind farm is the torsional SSO excitor) in a power system. Main drawback of the OMA is the requirement to establish the parametric model in order to conduct the OMA, because in practice, it is difficult to obtain the parametric information of wind farms. Hence, this paper proposes a method of the OMA based on the measurement data of a practice power system with the grid-connected wind farms. The core of the proposed method is the deep transfer learning (TL) algorithm to train a JDA based CNN for tracing the SSO excitor.

In practice, the measurement data of the power system are neither sufficient nor labeled and hence not suitable for training the CNN. The paper proposes to build a simulation system by simplifying the practical power system with the wind farms. A large amount of labeled training data generated by the simplified simulation system together with the limited measurement data from the practical power system are used to train the JDA based CNN by using the TL algorithm. Applicability of the JDA based CNN to trace the SSO excitor in the practical power system is ensured by the principle of the TL.

Proposed method of the OMA based on the measurement data has been evaluated in several test example power systems with grid-connected wind farms. In the paper, a small-scale test example power system is presented. The proposed JDA based CNN trained by the TL algorithm is compared with the conventional CNN without using the TL. The comparison demonstrates the better performance of the JDA based CNN with the TL algorithm to trace the SSO excitor.

Hence, main contributions of the paper are summarized as follows.

1) A measurement-data driven method of machine learning (ML) to apply the OMA for tracing the SSO excitor in a real power system with grid-connected wind power generation is proposed and evaluated. The method does not need to establish the parametric model of the real power system. This avoids the difficulty to gain parametric information of wind power generation in practice.

2) Data of oscillations in real power systems are too limited to be used in the ML. In order to solve this problem of data unbalance, the JAD based CNN is proposed to apply the TL to use the data from the source domain for the SSO excitor tracing in the target domain (the real power system with the grid-connected wind power generation).

3) Simplified simulation system is proposed as the source domain in the TL. The simplified simulation system shares the common factors to affect the SSO with the real power system. This ensures that a good performance of the TL is achieved.

The proposed method is yet to be evaluated by using the actual oscillation data from a real-world power system with the grid-connected wind power generation. This is the ultimate target to be pursued by this study in the near future to convince the grid company to install the measurement device to collect the data for applying the OMA.

## REFERENCES

[1] A. E. Leon and J. A. Solsona, "Subsynchronous interaction damping control for DFIG wind turbines," *IEEE Trans. Power Syst.*, vol.30, no.1, pp. 419–428. Jan. 2015.

[2] L.Chen, W. Zhao, F. Wang, et al., "An Interharmonic Phasor and Frequency Estimator for Subsynchronous Oscillation Identification and Monitoring", *IEEE Trans. Instru & Measu.*, vol.68, no.6, pp.1714-1723, Jun. 2019.

[3] Z. Miao, "Impedance-model-based SSR analysis for type3 wind generator and series-compensated network," *IEEE Trans. Energy Convers*., vol. 27, no. 4, pp. 984–991, Dec. 2012.

[4] L. Wang, X. Xie, Q. Jiang, et al., "Investigation of SSR in Practical DFIG-Based Wind Farms Connected to a Series-Compensated Power System," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2772-2779. 2015.

[5] H. Liu, X. Xie, X. Gao, H. Liu and Y. Li. "Stability Analysis of SSR in Multiple Wind Farms Connected to Series-Compensated Systems using Impedance Network Model," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3118 - 3128. 2018.

[6] Y. Song, X. Wang and F. Blaabjerg, "Impedance-Based High-Frequency Resonance Analysis of DFIG System in Weak Grids," *IEEE Trans. Power Electron.*, vol.32, no.5, pp.3536–3548, May. 2017.

[7] H. Liu, X, Xie, W. Liu, et al., "An oscillatory stability criterion based on the unified dq-frame impedance network model for power systems with high-penetration renewables," *IEEE Trans. Power Syst.*, vol. 33, no.3, pp. 3472-3485, 2018.

[8] W. Du, C. Chen and H. Wang, "Sub-synchronous interactions induced by DFIGs in power systems without series compensated lines," *IEEE Trans. Sustain. Energy*, Vol. 9, No. 3, pp. 1275-1284, 2018.

[9] W. Du, Q. Fu and H. Wang. "Concept of Modal Repulsion for Examining the Subsynchronous Oscillations Caused by Wind Farms in Power Systems," *IEEE Trans. Power Syst.*, vol.34, no.1, pp.518-526, 2019.

[10] W. Du, W. Wang and H. Wang, "Analytical Examination on the Amplifying Effect of Weak Grid Connection for the DFIGs to Induce Torsional Sub-synchronous Oscillations," *IEEE Trans. Power Del.*., vol.35, no.4, pp. 1928-1938, Aug. 2020.

[11] W. Du, B. Ren, H. Wang, et al., "Comparison of methods to examine sub-synchronous oscillations caused by grid-connected wind turbine generators," *IEEE Trans. Power Syst.*, vol.34, no.6, pp.4931–4943, Nov. 2019.

[12] Y. Shu, X. Zhou and W. Li. "Analysis of Low Frequency Oscillation and Source Location in Power Systems," *CSEE Journal of Power and Energy Systems*, vol.4, no.1, 2018, pp.58-66.

[13] L. Chen, Y. Min and W. Hu, "An energy-based method for location of power system oscillation source," *IEEE Trans. Power Syst.*, vol.28, no.2, pp. 828 - 836, May. 2013.

[14] S. Maslennikov, B. Wang, and E. Litvinov, "Dissipating energy flow method for locating the source of sustained oscillations," *Int. J. Electr. Power Energy Syst*., vol.88, pp. 55–62, Jun. 2017.

[15] Y. Ren, X. Wang, L. Chen, et al., "Component damping evaluation in sub-synchronous oscillation based on transient energy flow method," *IET Gener. Transm. Distrib*., vol.14, no.3, pp. 460–469, Nov.2019.

[16] X. Xie, Y. Zhan, J. Shair, et al., "Identifying the source of subsynchronous control interaction via wide-area monitoring of sub/super-synchronous power flows," *IEEE Trans. Power Del*., vol.35, no.5, pp.2177-2185, Dec. 2019.

[17] S. C. Chevalier, P. Vorobev, and K. Turitsyn, "Using Effective Generator Impedance for Forced Oscillation Source Location," *IEEE Trans. Power Syst*., vol.33, no.6, pp. 828 - 836, May. 2018.

[18] X. Xie, Y. Zhan, H. Liu, et al., "Wide-area monitoring and early-warning of subsynchronous oscillation in power systems with high-penetration of renewables," *Int. J. Electr. Power Energy Syst*., vol.108, pp. 31–39, Jun. 2019.

[19] B. Wang, and K. Sun, "Location methods of oscillation sources in power systems: a survey," *J. Mod. Power Syst. Clean Energy*, vol.5, no.2, pp.151-159, Mar. 2017.

[20] Z. Ping, X. Li, W. He, T. Yang, and Y. Yuan, "Sparse learning of network-reduced models for locating low frequency oscillations in power systems," *Applied Energy*, vol.262, Mar. 2020.

[21] Y. Meng, Z. Yu, D. Shi, D. Bian, and Z. Wang, "Time Series Classification for Locating Forced Oscillation Sources," *IEEE Trans. Smart Grid*, vol.12, no.2, pp.1712-1721, Mar. 2021.

[22] J. Mayo-Maldonado, J. Valdez-Resendiz, D. Guillen, et al., "Data-Driven Framework to Model Identification, Event Detection, and Topology Change Location Using D-PMUs", *IEEE Trans. Instru & Measu.*, vol.69, no.9, pp.6921-6933, Sep. 2020.

[23] H. Banna, S. Solanki, and J. Solanki, "Data-driven disturbance source identification for power system oscillations using credibility search ensemble learning," *IET Smart Grid*, vol.2, no.2 pp.293-300, 2019.

[24] S. Maslennikov, and E. Litvinov, "ISO New England Experience in Locating the Source of Oscillations Online," *IEEE Trans. Power Syst.*, vol.36, no.1, pp. 495 - 503, Jan. 2021.

[25] S. J. Pan, J. T. Kwok, and Q. Yang, "Transfer learning via dimensionality reduction," in *Proceedings of the 23rd AAAI Conference on Artificial Intelligence*, Chicago, Illinois, USA, July 2008, pp. 677–682.

[26] J. Yosinski, J. Clune, Y. Bengio, "How transferable are features in deep neural networks?" in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, Montreal Canada: ACM, December 2014, pp. 3320–3328.

[27] W. Du, W. Dong and H. Wang, "A Method of Reduced-Order Modal Computation for Planning Grid Connection of a Large-Scale Wind Farm," *IEEE Transactions on Sustainable Energy*, vol.11, no.3, pp. 1185-1198, Jun. 2019.

[28] M. Long, J. Wang, G. Ding, J. Sun and P. S. Yu, "Transfer Feature Learning with Joint Distribution Adaptation," in *Proc. 14th IEEE Int. Conf. Computer Vision (ICCV)*, 2013.

[29] A. Heniche and I.Kamwa, "Assessment of two methods to select wide area signals for power system damping control," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 572–581, May. 2008.

[30] S. J. Pan and Q. Yang, "A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.*, vol. 22, no.10, pp. 1345–1359. Oct. 2010.

[31] A. Krizhevsky, I. Sutskever, G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proc. 25th Int. Conf. Neural Information Processing Systems (NIPS)*, 2012.

[32] W. Liu, Y. Wen, Z. Yu, et al., "Large-Marg in Softmax Loss for Convolutional Neural Networks," in *Proc. 33rd Int. Conf. Machine Learning (ICML)*, 2016, pp.507-516

[33] Z. Wang, Z. Dai, B. Póczos, and J. Carbonell. "Characterizing and avoiding negative transfer," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[34] A. Celletti, "Regularization theory," *Springer Berlin Heidelberg*, 2010.

[35] A. Paszke, S. Gross, F. Massa, et al, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," *Neural Information Processing Systems(NeurIPS)*, 2019.

[36] S. Li, T.A. Haskew, R.P. Swatloski, et al, "Optimal and Direct-Current Vector Control of Direct-Driven PMSG Wind Turbines," *IEEE Trans. Power Electron.*, vol.27, no.5, pp.2325-2337, 2012.

[37] L. Fan, R. Kavasseri, Z. Miao, et al, "Modeling of DFIG-based wind farms for SSR analysis," *IEEE Trans. Power Del*., vol.25, no.4, pp.2073-2082, 2010.

[38] A. Gupta, G. Gurrala and P. S. Sastry, "An Online Power System Stability Monitoring System Using Convolutional Neural Networks," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 864-872, March 2019.

[39] D.Kinga and J.B.Adam, "A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, vol. 5, 2015, pp. 1–15.

[40] D.M.W. Powers. "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation". *Journal of Machine Learning Technologies*, vol.2, no.1, pp.37–63, 2011.

[41] L. Zhu, D.J. Hill and C. Lu, "Hierarchical Deep Learning Machine for Power System Online Transient Stability Prediction," *IEEE Trans. Power Syst.*, vol.35, no.3, pp. 2399 -2411, Mar. 2020.

[42] D. Erhan, A. C. Courville, Y. Bengio, et al, "Why Does Unsupervised Pre-training Help Deep Learning?" *Proc Aistats*, vol.11, no.19, pp.625−660, 2010.

## Appendix A

Parameters of the test example power system:

Parameters of transmission lines: $x_1$=0.012, $x_2$=0.004, $x_3$=0.04, $x_4$=0.07, $x_5$= $x_6$=0.01, $x_7$= $x_9$= $x_{10}$= $x_{11}$=0.03, $x_8$=0.0045, $x_{12}$=0.01, $r_1$= $r_5$= $r_6$= $r_{11}$= $r_{12}$=0.001, $r_2$= $r_3$=0.0002, $r_4$=0.0007, $r_7$=0.003, $r_8$=0.0005, $r_9$= $r_{10}$=0.002.

Parameters of *PMSG*s in WF1, WF2, WF4: $X_{pd}$=0.25, $X_{pq}$=0.15, $\psi_{pm}$=1.1, $J_p$=3s, $C_p$=30, $K_{p\_MSC\_pq}$=5, $K_{i\_MSC\_pq}$=20, $K_{p\_MSC\_iq}$=1, $K_{i\_MSC\_iq}$=100, $K_{p\_MSC\_id}$=1, $K_{i\_MSC\_id}$=100, $K_{p\_GSC\_Udc1,2,4}$=1.2,0.07,1.5, $K_{i\_GSC\_Udc1,2,4}$=2900,820, 3300, $K_{p\_GSC\_iq}$=0.1, $K_{i\_GSC\_iq}$=10, $K_{p\_GSC\_qd}$=0.006, $K_{i\_GSC\_qd}$=60, $K_{p\_GSC\_id}$=0.006, $K_{i\_GSC\_id}$=40

Parameters of *DFIG*s in WF3, WF5: $X_{dss}$=0.18, $X_{dm}$=2.40, $X_{drr}$=0.12, $R_r$=0.015, $J_d$=3s, $C_d$=30, $K_{p\_RSC\_p}$=0.1, $K_{i\_RSC\_p}$=35, $K_{p\_RSC\_q}$=0.1, $K_{i\_RSC\_q}$=2, $K_{p\_RSC\_iq3,5}$=1,1.5, $K_{i\_RSC\_iq3,5=}$=80,90, $K_{p\_RSC\_iq}$=1, $K_{i\_RSC\_iq}$=25, $K_{p\_GSC\_Udc3,5}$=50,30, $K_{i\_GSC\_Udc}$=600, $K_{p\_GSC\_qd}$=2, $K_{i\_GSC\_qd}$=200, $K_{p\_GSC\_id}$=2, $K_{i\_GSC\_id}$=200

Loading condition: $P_{PMSGs}$ in WF1, 2, 4=0.05,0.1,0.15, $P_{DFIGs}$ in WF3, WF5 =0.2,0.2.

When generate test samples in test system, the parameter $K_{i\_GSC\_Udc}$ of *PMSG*s in WF2 is set to 800~850 (excluding 820), the parameter $K_{i\_GSC\_Udc}$ of *PMSG*s in WF1 is set to 760~830.

## Appendix B

Network structure parameters: *Input layer* =20*640, [*Convolutional layer* + *BN layer*+*Pooling layer*] × 2={1 × 64, (convolution kernel=5 × 5), bn-Size=64, (pooling kernel=2 × 2), 64 × 32, (convolution kernel=5 × 5), bn-Size=32, (pooling kernel=2 × 2) }, *Fully connected layer* = 5*160, *Dropout layer*=0.2, *BN layer*=800, *Softmax*=800*(5+1); *Activation function=ReLU.*

Network training parameters: *Learning rate*=1e-2~1e-5, *BatchSize*=64, *training epochs*=2000.