Data-Driven Low Frequency Oscillation Mode Identification and Preventive Control Strategy Based on Gradient Descent

Yiwei Fu, Lei Chen
Department of Electrical Engineering
Tsinghua University
Beijing, China
chenlei08@tsinghua.edu.cn

Zhe Yu, Yishen Wang, Di Shi
GEIRI North America
San Jose, USA
zhe.yu@geirina.net

***Abstract*—Accurate mode identification and effective preventive control strategy of low frequency oscillation (LFO) are vital to improve the small signal stability of power system. This paper proposes a novel data-driven method based on Convolutional Neural Network (CNN) to identify the low frequency modes. The application of feature selection and feature fusion makes the CNN model well adapted to the complexity of large-scale power system. The model can predict LFO modes of a power system in operation scenarios with different power injections and topologies. By invoking the trained CNN model, a preventive control method based on gradient descent is developed to increase the damping of critical modes. Case study demonstrates that the proposed method can efficiently identify the oscillation modes and the obtained preventive control strategy can effectively prevent the occurrence of LFO.**

***Index Terms*— low frequency oscillation; mode identification; preventive control; CNN; gradient descent.**

I. INTRODUCTION

With the continuous development of interconnected power system, the low frequency oscillation (LFO) problem caused by weak or negative damping electromechanical modes has become increasingly prominent [1]. Since LFOs can pose a great threat to the stability of power grid and may even induce disastrous consequences such as system splitting and widespread blackouts [2], it is significant to online identify the dominant modes and generate preventive control strategy when there are critical modes.

The online identification methods of LFO dominant modes can be typically classified into two categories [3]: (i) model-based methods, and (ii) measurement-based methods. The former is based on linearized dynamic models of the power system, where the eigenvalue analysis can be processed at the online operating point [4]. However, its applicability to large-scale power system is limited, because the accurate modelling of the large system is nearly impossible and the computation time is too long to satisfy the efficiency requirement of online identification [5]. Whereas the other category is focused on using the measurement data extensively to identify the weak or negative damping modes. With the data obtained from Phasor Measurement Units (PMU) and Wide Area Measurement System (WAMS), the dynamic feature of the power system can be described more correctly through this approach [6]. There are many typical measurement-based methods used in the mode identification of LFO, such as Fast Fourier Transform (FFT) [7], Auto Regressive (AR) [8], Auto Regressive Moving Average (ARMA) [9-10], ESPRIT [11], Prony analysis [12] and Stochastic Subspace Identification (SSI) [13]. The methods based on ring-down signal is not suitable for online monitoring. Therefore, online identification methods should be based on ambient data, and the accuracy is usually limited by noise [5].

The model-based methods and the measurement-based methods both have their own advantages and disadvantages. The results of the model-based methods are restricted by the accuracy of the models and the parameters of the system, while the results of the measurement-based methods are mostly affected by the noises and the performances of the algorithms. These two different mode identification systems are applied simultaneously in many power grids. Their identification results are usually different. Moreover, the identification results are often abandoned and the knowledge behind the data are seldom mined for further analysis. As a data mining method, deep learning has strong ability of feature extraction [14]. But for now, only a few studies have applied deep learning methods to the electromechanical mode identification. As a classic deep learning algorithm, Deep Belief Network (DBN) has been proposed to be combined with Prony analysis to realize mode identification in [15]. The basic idea is to build a DBN model for the order identification of the LFO signal, and then use this order when setting the initial order of the Prony algorithm to get the final identification result. Another study directly uses the DBN model with LFO signal as input to quickly extract a single dominant mode in the regional oscillation scenario and local oscillation scenario, respectively [16]. In addition, there are also a few studies using different shallow neural network

This work is supported by National Key R&D Program of China (2018YFB0904500), SGCC Science and Technology Program (SGJS0000DKJS1801231) and Tsinghua University Tutor Research Fund.

methods to achieve LFO mode identification [17-18]. However, most of the data-driven identification methods mentioned above are applied together with the measurement-based methods, and the characteristics of power system operation are hardly considered when building the data-driven models.

This paper synthesizes the data accumulated by the model-based and measurement-based mode identification systems in the actual power grid as the training set of a CNN-based mode identification model, and then apply the identification results to the stability assessment and preventive control of power system.

When there are critical modes, preventive control strategies should be carried out to increase the damping of the modes. In the existing literatures, the widely used control measures for suppressing LFO encompasses three approaches [19], classified as: (i) changing power grid structure, (ii) strengthening control equipment, and (iii) adjusting operation scenario, mainly generator outputs. In the operation stage, the third one is the only applicable measure. There is a relationship between the small signal stability of power system and the operation scenario of the system. Adjusting operation scenario properly can prevent low frequency oscillation. Most of the existing studies use sensitivity analysis to determine the proper control strategies [20]. The sensitivity of the mode damping to the output power of each generator is first calculated, and then the generator outputs are changed according to the sensitivities to increase the damping of the critical mode [21]. However, the sensitivity-based control method cannot accurately describe the complex relationship between operation scenario and mode damping in large-scale power grid.

Since the identification model of electromechanical modes has been trained by deep learning method, it can be further used in the formulation of preventive control strategies. Thus, with the combination of mode identification and preventive control, a new LFO monitoring and control framework is proposed in this paper. The main contributions are as follows.

1) A multi-input CNN model with grid topology and power flow as inputs is built and well trained to identify the electromechanical modes of large-scale power system under different operation scenarios.

2) A preventive control method combined with the trained CNN model is presented and efficiently solved by gradient descent.

The reminder of this paper is organized as follows. Section II presents the overall monitoring and control framework of LFO. In Section III, each component of the multi-input CNN model is described in detail. Section IV introduces the flowchart and specific steps of the preventive control method. The application of the proposed method in the actual power grid is given in Section V. Then, the proposed method is demonstrated in Section VI by using the data of the central China power grid. The conclusions are drawn in Section VII.

II. Overall LFO Monitoring and Control Framework

The overall LFO monitoring and control framework is based on the combination of the data-driven mode identification process and the preventive control process, shown as Figure 1.

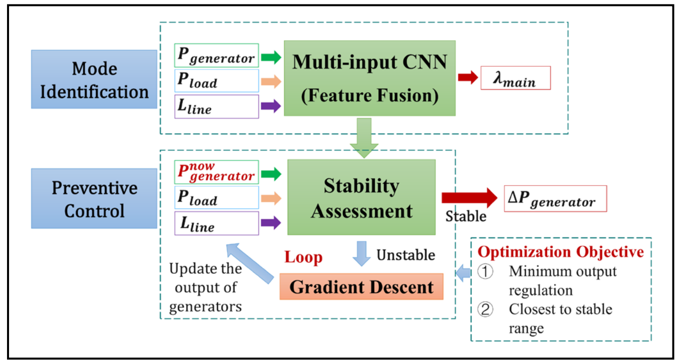

Figure 1. Framework of the proposed method

This framework is actually a unified two-step methodology. The first step is to establish a multi-input CNN model for electromechanical mode identification. After sufficient training with a large sample set, the CNN model can create a mapping between system operation scenario and electromechanical modes, so it can be regarded as a stability assessment module. It is firstly used to identify the electromechanical modes under the current operation scenario and determine whether there is a weak damping mode. If there is, then go to the preventive control step. In the preventive control step, the control strategy will be solved iteratively by gradient descent method. Since the essence of preventive control is to make the output (electromechanical mode) of CNN model within a certain range by adjusting the input (system operation scenario) of it, the trained CNN model needs to be invoked repeatedly to obtain the gradients of the objective function with respect to the input.

III. Data-Driven Electromechanical Mode Identification

In order to accurately identify the electromechanical modes of the power system, the multi-layer CNN model is constructed as depicted in Figure 2. The CNN-based mode identification model is committed to establish the relationship between the system operation scenario and the electromechanical modes. This section will introduce the composition of CNN model in detail.

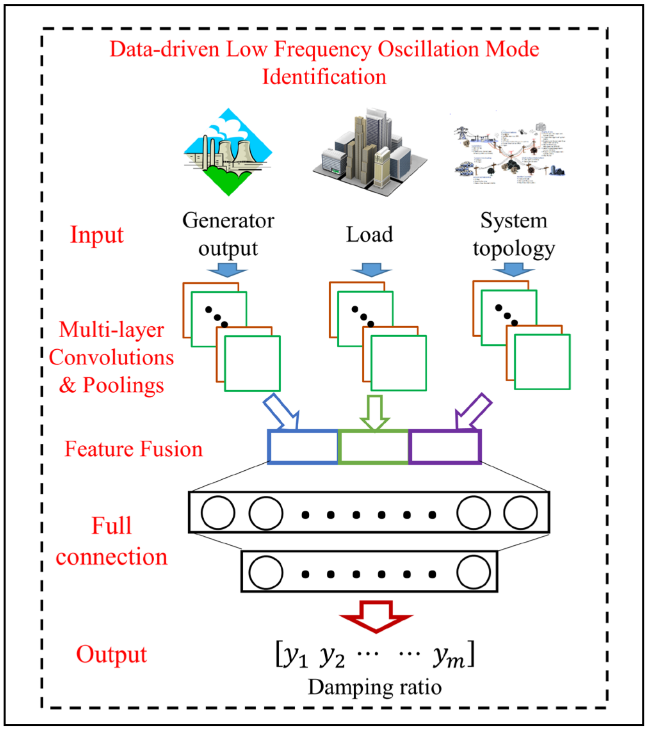

Figure 2. Multi-input CNN model

A. Sample Set

A large sample set, which is accumulated from the mode identification results of the model-based identification system and the measurement-based identification system in the actual power system, is used in the CNN model. Each sample contains the system operation scenario dataset $\boldsymbol{S}$ and the electromechanical mode dataset $[\lambda_1, \lambda_2, \cdots, \lambda_n]$, in which the dataset $\boldsymbol{S}$ includes the generator output vector $\boldsymbol{P}_{gen}$, the load demand vector $\boldsymbol{P}_{load}$, and the system topological connection vector $\boldsymbol{L}$.

Since samples are collected from two systems, two different sets of electromechanical modes $[\lambda_1^a, \lambda_2^a, \cdots, \lambda_n^a]$ and $[\lambda_1^b, \lambda_2^b, \cdots, \lambda_n^b]$ under the same operation scenario $\boldsymbol{S}$ can be obtained from the actual power grid.

The damping ratio ξ of each electromechanical mode $(\lambda=\sigma + jw)$ can be calculated by formula (1), which is often adopted to evaluate the small signal stability of power system.

$$\xi = \frac{-\sigma}{\sqrt{\sigma^2 + w^2}} \tag{1}$$

Based on this, two sets of damping ratios $[\xi_1^a, \xi_2^a, \cdots, \xi_n^a]$ and $[\xi_1^b, \xi_2^b, \cdots, \xi_n^b]$ can be calculated from these two sets of electromechanical modes. And the average set $\vec{\xi}_{avg}$ of them is then computed as the final set of damping ratios under this operation scenario, which is eventually regarded as the label of CNN model.

Thus, each sample consists of the operation scenario vectors $\boldsymbol{P}_{gen}$, $\boldsymbol{P}_{load}$, $\boldsymbol{L}$ and the average damping ratio vector $\vec{\xi}_{avg}$, while vectors $\boldsymbol{P}_{gen}$, $\boldsymbol{P}_{load}$ and $\boldsymbol{L}$ are used as three inputs of CNN model respectively.

Collinearity is a significant feature of this sample set obtained from actual data, which indicates that the generations and loads in different buses have similar patterns of increase and decrease [23]. In the following section, this problem will be handled.

B. Feature Processing and Feature Fusion

In large-scale interconnected power grids, the three inputs $\boldsymbol{P}_{gen}$, $\boldsymbol{P}_{load}$ and $\boldsymbol{L}$ often contain thousands of features. The excessive dimension of features leads to the input size of CNN model being too large, which brings a great burden to the training of CNN network. The collinearity of the sample set is characterized by the redundancy of the input features. Therefore, it is necessary to implement feature processing before applying them to training CNN model.

Feature processing is the core of feature engineering, which can extract features from the original data to the maximum extent for algorithm and model use [23]. The feature processing in this study mainly includes two steps: data preprocessing and feature selection.

1) *Data Preprocessing:* Cleaning and dimensionless process of the raw data.

The dimensionless method used in this paper is interval scaling method, which can normalize the original features to $[-1,1]$ through:

$$x' = \frac{x - (x_{max} + x_{min})/2}{(x_{max} - x_{min})/2} \tag{2}$$

2) *Feature Selection:* Select significant features as input of machine learning algorithm for training [24].

Feature selection is usually considered from two perspectives:

a) Whether a feature diverges or not: If a feature does not diverge (e.g., variance is close to 0), it means that there is basically no difference in this feature among samples, then it should be deleted. Variance threshold method is one of the most typical methods in this aspect.

b) Relevance between features and labels: The features with high relevance to the labels should be selected first. In this category, correlation coefficient method is commonly used in many studies.

After data preprocessing, the variance threshold method and the correlation coefficient method are successively adopted for feature selection in this study.

First, by comparing the variance of different features in the sample set, the variance threshold method is used to screen out the features with small variance. About 5% of features whose variance is significantly lower than other features have been deleted by this method.

After that, features with higher relevance to labels are further selected from the remaining features by the correlation coefficient method. The labels of the proposed CNN model are composed of the damping ratio of several dominant modes. Therefore, in order to achieve the purpose of feature selection, this paper selects three labels which are close to the stable threshold, calculates their correlation with all the features and sorts the features from high to low according to the correlation values, and then removes the features which are obviously low correlation-coefficient with those important labels.

Since the model has multiple inputs, feature fusion is an important means to improve prediction performance. Generalized feature fusion includes the processes of feature extraction and feature selection [25]. Feature fusion here is more precisely referring to feature combination.

There are two existing techniques of feature combination: serial and parallel combination [26].

a) Serial combination: Connect the two features directly. If the dimension of two input features α and β is p and q, the output feature is $[\alpha, \beta]$ while its dimension is $p+q$. It is carried out by the "concat" operation.

b) Parallel combination: Combine the two features into a complex feature. Take the absolute value of the complex feature as the final feature. If the dimension of two input features α and β is p, the output feature is $\|\alpha + j\beta\|$ while its dimension is p. It is achieved by the "add" operation.

The serial combination is applied in the proposed CNN model. Different features corresponding to the three inputs are merged by "concat" operation after passing through multiple convolution and pooling layers, and further mining is then carried out.

C. Model Structure

CNN model consists of input layer, hidden layer and output layer, while the hidden layer contains three common structures: convolution layer, pooling layer and fully connected layer.

The function of convolution layer is to extract features from input data. It uses convolution kernels to perform an operation similar to continuous saccade on input data, which is mathematically similar to convolution computation. Before input to the convolution layer, the feature sets obtained from the feature processing in the three input vectors ($\boldsymbol{P}_{gen}$, $\boldsymbol{P}_{load}$, $\boldsymbol{L}$) should be reshaped into three matrices ($\boldsymbol{P}_g$, $\boldsymbol{P}_l$, $\boldsymbol{P}_{line}$), in which the system topology is represented by the power flow $\boldsymbol{P}_{line}$ in the transmission lines of the power grid.

The parameters of convolution layer include the kernel size, stride and padding settings, which together determine the output feature size of convolution layer and are the hyper parameters of CNN. In convolution process of the proposed model, for input $\boldsymbol{P}_g$ and $\boldsymbol{P}_l$, set the kernel size and the stride to be 3×3 and 2 respectively, and use the "SAME" padding. Since the size of input $\boldsymbol{P}_{line}$ is much larger than $\boldsymbol{P}_g$ and $\boldsymbol{P}_l$, set the kernel size for $\boldsymbol{P}_{line}$ to be 5×5 and the other two parameters are set the same as those before.

In addition, the role of the pooling layer is to simplify the scale of CNN model, retain the main features, improve the calculation speed and prevent over-fitting. It is usually connected directly behind a convolution layer. In the proposed model, the pooling mode is set as max pooling and a 2×2 filter with the stride set as 2 is selected.

Fully connected layer maps feature space to label space through linear transformation, which can integrate local information with category discrimination in convolution layer or pooling layer. Therefore, it often appears in the last several layers of CNN.

The proposed CNN model consists of 3 convolution layers, 3 pooling layers and 2 fully connected layers, in which the convolution layer and the pooling layer are interlinked into the form like "convolution-pooling".

After 3-layer convolution and pooling, the three inputs are extracted into deep features, and feature fusion is carried out by "concat" operation as mentioned before. The fused features are then processed by 2-layer fully connected layers and the prediction results are finally obtained. The structure of this CNN model can be seen in Figure 2.

The loss function of CNN model is the sum of a square difference function and a regular term in order to avoid over-fitting, as shown in (3).

$$C = \frac{1}{2}\sum_i \left(\xi_i' - \xi_i\right)^2 + \frac{\alpha}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} \tag{3}$$

where, ξ_i' and ξ_i are the predicted and actual value of damping ratio, respectively. α is a parameter of the regular term and $\mathbf{w}$ is a set of weights in the fully connected layer.

From this, a data-driven mode identification based on CNN network is built, and the trained model can be utilized as a stability assessment module in preventive control process.

IV. Preventive Control Based on Gradient Descent

By establishing the CNN-based mode identification model, the correlation between the electromechanical modes and the power change of the generators can be quantitatively analyzed, so as to guide the adjustment of the operation scenario to improve the system damping.

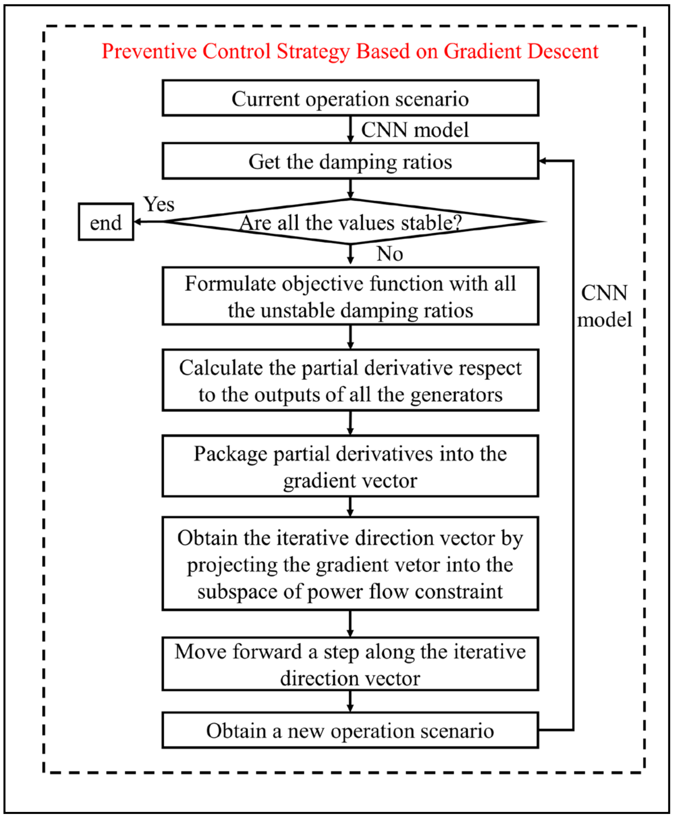

Figure 3. Flowchart of the preventive control process

Since the above CNN model needs to be repeatedly invoked during the solving process of preventive control strategy, the traditional optimization methods are not suitable for this situation. Therefore, the gradient descent method is adopted to solve the problem iteratively as shown in Figure 3.

The goal of the preventive control is to set all the damping ratios stable as well as minimize the generation power adjustment. The objective function can be formulated as (4).

$$\min \ Obj(\mathbf{P})=\sum_{i}(P_i-P_{i0})^2+M\cdot\sum_{j\in B}\left(\xi_0-f_j(\mathbf{P})\right) \tag{4}$$

where, P_i is the output of generator i under the current operation scenario, P_{i0} is the output of generator i under the original operation scenario, M is a big enough coefficient, ξ_0 is the stability boundary of unstable damping ratios, which is set as 0.03 in this paper. $\mathbf{B}$ is index set of unstable damping ratios, $f_j(\mathbf{P})$ is the jth unstable damping ratio under the current operation scenario via CNN model.

The partial derivative of the objective function $Obj(\mathbf{P})$ respect to P_i is shown as (5):

$$\frac{\partial Obj(\mathbf{P})}{\partial P_i}=\frac{Obj(\mathbf{P}+\mathbf{\Delta_i})-Obj(\mathbf{P})}{\Delta_i} \tag{5}$$

where $\mathbf{\Delta_i}$ is a vector in which $\mathbf{\Delta_i}(i)=\Delta_i$ while others are 0. Then the gradient direction is obtained and a new operation scenario can be created by moving forward a step from the previous operation scenario along the iterative direction which converted by the gradient vector. The process above should be iterated until the weak or negative damping is eliminated.

The specific flowchart of solving the preventive control strategy is given below:

Step 1: Input the current operation scenario into the trained CNN model, and determine whether the system is small-signal stable according to the output damping ratios.

Step 2: If the system is stable, output "No need for preventive control" and exit; if unstable, output "Need for preventive control" and enter the iterative solving process.

Step 3: Calculate the value of the objective function $Obj(\boldsymbol{P})$ with all the unstable damping ratios as (4) under the current operation scenario.

Step 4: Calculate the partial derivative ${\partial Obj(\boldsymbol{P})}/{\partial P_i}$ respect to the outputs of all the generators as (5) and package all the partial derivatives into the gradient vector $\nabla Obj(\boldsymbol{P})$. If the output of generator i can no longer be adjusted along its partial derivative limited by its output constraint, set $\nabla Obj(\boldsymbol{P})(i)=0$.

Step 5: Due to the convergence requirement of power flow, the total change amount of generator output needs to be 0, i.e.

$$\Delta P_1+\Delta P_2+\cdots+\Delta P_n=0 \tag{6}$$

To satisfy the formula (6), the gradient vector $\nabla Obj(\boldsymbol{P})$ should be projected into the subspace V_{sub} represented by (6). It can be easily determined that the dimension of subspace V_{sub} is $n-1$, and a matrix A can be built by its set of base vectors: $A=\left[\vec{b}_1,\vec{b}_2,\cdots,\vec{b}_{n-1}\right]$.

It can be calculated by the formula (6) that:

$$A=\begin{bmatrix}1 & 1 & \cdots & 1\\ -1 & 0 & \cdots & 0\\ 0 & -1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & -1\end{bmatrix} \tag{7}$$

Assuming that vector $\vec{m}$ is the projection of vector $\nabla Obj(\boldsymbol{P})$ on subspace V_{sub}, it can be deduced that:

$$\vec{m}=A\left(A^TA\right)^{-1}A^T\nabla Obj(\boldsymbol{P}) \tag{8}$$

Step 6: Calculate the unit vector $\vec{k}$ (which represents the iterative direction) of $\vec{m}$ through the formula: $\vec{k}=\vec{m}/\|\vec{m}\|$. For each generator i, calculate the maximum move length r_i along vector $\vec{k}$ till its output limit. Then calculate the step length r by $r=\min(r_0,r_1,\cdots,r_n)$ (set $r_i=+\infty$ if $k_i=0$), in which r_0 is the default step length. By using the step length r, the generator output is regulated into $\boldsymbol{P}^{next}$ through the formula: $\boldsymbol{P}^{next} = \boldsymbol{P} + r * \vec{k}$. Thereby move forward a step to another operation scenario.

Step 7: Input the operation scenario obtained by step 6 to the trained CNN model, calculate the damping ratios and determine whether they are in the stable range. If there is no critical mode, then output the preventive control strategy. Otherwise, return to step 3.

The control method proposed above can be applied to the preventive control of power system under normal and unsafe operation scenarios.

V. APPLICATION PROCEDURE

The flowchart of the proposed LFO monitoring and control method applied in actual power systems is shown in Figure 4. The black arrow lines represent the training process of mode identification model, while the red arrow lines describe the process of mode identification and preventive control of the real-time system operation scenario.

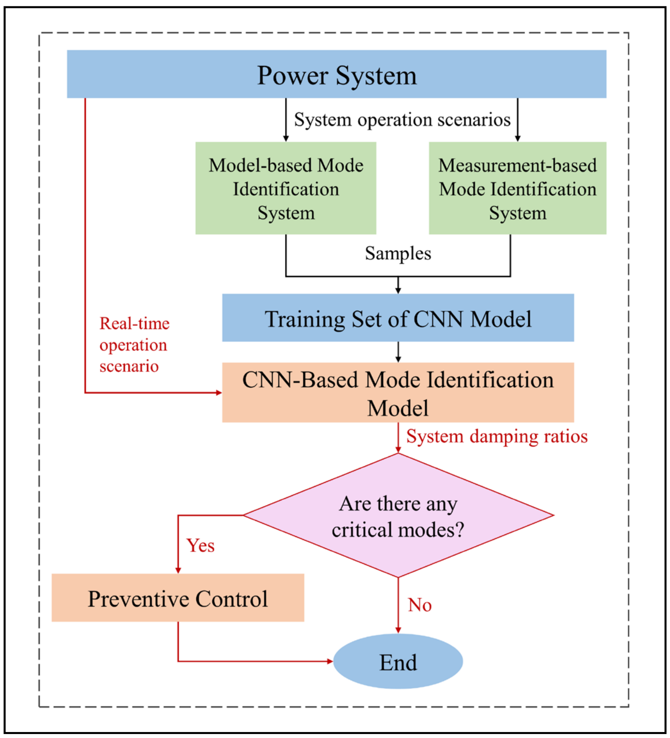

Figure 4. Flowchart of the application in actual system

It can be seen that different identification results under different operation scenarios are obtained from the model-based system and the measurement-based system equipped in the actual power grid. The mode identification result in the training set of CNN model is the combination of these two results. Then the training set is fully trained to get another CNN-based mode identification model. After that, the real-time operation scenario is input into the trained CNN model, and the output damping ratio is used to judge whether the system is small-signal stable or not. If it is not stable, then enters the preventive control process and finally end up with the generator control strategy.

VI. CASE STUDY

In this section, the validity of the proposed mode identification and preventive control methods is tested and verified in the central China power grid. The system consists of 371 generators, 2317 loads and 3305 lines. Therefore, the input dimension of the original data is 5993. After the multi-step feature processing introduced in Section III, the 342-dimensional generator output vector, the 1554-dimensional load power vector, and the 2758-dimensional line power vector are finally selected as the three inputs. The three vectors are respectively reshaped into three feature maps of size 19×18 , 40×39 and 60×46 , in which the dimension deficiency is padding with 0. The sample set is composed of 30,000 samples collected and processed from the actual power system.

A. Analysis of mode identification results

The 30,000 samples are randomly divided into a training set consist of 27,000 samples and a testing set consists of 3,000 samples. After training, the accuracy rate of the multi-layer CNN model on training set is 99.35%, and the average accuracy rate on the testing set is 97.92%.

More specifically, the accuracy of each single sample is depicted as Figure 5. As can be seen, although the accuracy of most test samples is concentrated in 90%-100%, there are still a few samples whose accuracy is less than 90%, and even some individual samples whose accuracy is only about 20%.

Further statistics show that the accuracy of mode identification in the testing set is 97.77% above 90%, 93.17% above 98%, and only 0.83% below 60%, which indicates that a small amount of system operation scenarios cannot be correctly identified by CNN model.

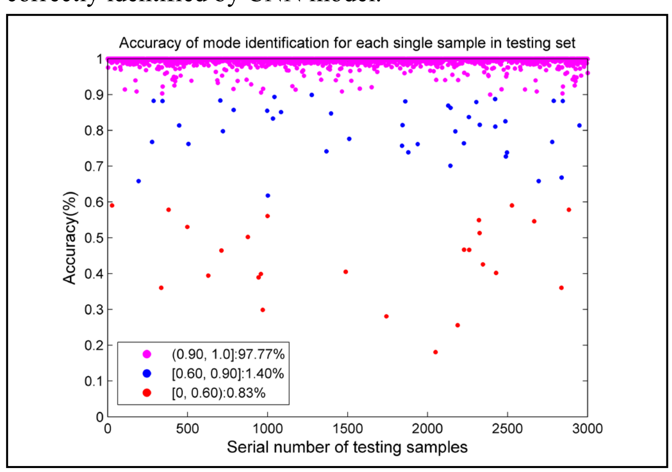

Figure 5. Accuracy of mode identification for each single sample

There are two possible reasons for this result:

1) The effective data produced by the repeatitive operation of power system are quite limited [22], which

indicates that some effective operation scenarios may be missed in the process of sample generation.

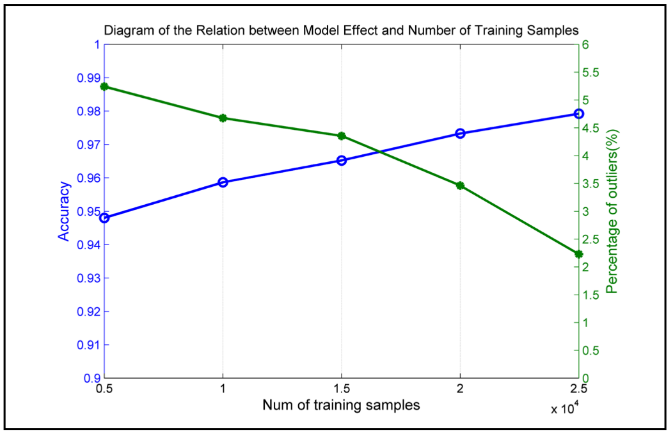

Figure 6. Accuracy and outlier proportion under different training sets

Theoretically, the more samples the training set has, the more likely it is to cover more effective data of the power system. To verify this thesis, the identification accuracy rates of the same testing set under different scales of training sets are compared and analyzed. After randomly selecting 3000 samples as the fixed testing set, 5000, 10000, 15000, 20000 and 25000 samples were respectively selected by stochastic extraction from the remaining 27000 samples to form five training sets. From this, five CNN models are trained and obtained. Outliers are defined as points with identification accuracy less than 60%. The average accuracy rate of the testing set and the proportion of outliers under different training sets are shown as Figure 6.

These results demonstrate:

- With the enlarged scale of training set, the accuracy of mode identification increases gradually and tends to be saturated, which shows that an associated rise in proportion of effective data are indeed realized.
- The proportion of outliers decreases with the increasing size of training set, which also indicates that the effect of mode identification does improve with the enrichment of training set.

Therefore, it is necessary to first explore the required data size when adopting the data-driven method to analyze the power system.

2) The sample sizes under the two kinds of topological changes are relatively smaller, which leads to insufficient mining of the relationship between the operation scenario and the electromechanical modes after the change of the topology.

In other words, although the proposed CNN model has the ability to identify the mode after the change of topology theoretically, it also needs sufficient data to reflect these changes.

B. *Preventive control strategy*

Given a certain operation scenario, the result of small-signal stability calculation shows that the power system has a negative damping mode. According to the process described in Section IV, the preventive control strategy for the generators in this system should be activated. The specific preventive control strategy is expressed in TABLE I. The strategy is reducing the output of three generators, and simultaneously increasing the output of another four generators while maintaining power balance.

Generator Number	*Output before control (pu)*	*Output after control (pu)*	*Amount of adjustment (pu)*
116	6.246	5.560	-0.686
258	1.771	5.642	3.871
264	0.343	1.210	0.867
431	5.088	1.099	-3.989
484	2.760	0.279	-2.481
650	6.979	8.995	2.016
694	2.893	3.295	0.402

TABLE I. THE PREVENTIVE CONTROL STRATEGY

In order to verify the effectiveness of the obtained strategy, the CNN model is invoked to identify the electromechanical mode before and after control. And the identification results are compared with the real mode of the system as shown in TABLE II.

	Damping ratio (real)	*Damping ratio (CNN)*
Before control	-0.0153	-0.0136
After control	0.0447	0.0413

TABLE II. EFFECT OF THE PREVENTIVE CONTROL STRATEGY

From this table, it is notable that the identification results of CNN model before and after control are both relatively close to the real values. Moreover, after the implementation of generation control strategy, the damping ratio of the system has been significantly increased, thus the negative damping mode is eliminated.

VII. CONCLUSION

This paper presents a novel monitoring and control method of LFO based on deep learning technology. The proposed multi-layer CNN model can effectively calculate the system damping ratios under a certain operation scenario, so as to determine whether the system is small-signal stable or not. Through the repeated calls of the well-trained CNN model, a preventive control method based on gradient descent is then designed. This method can provide the generation control strategy to eliminate the weak or negative damping mode of power system. The case study demonstrates that the CNN-based mode identification method has high accuracy under most of operation scenarios and the increase of effective data from the power system can improve the model reliability. It is also verified that the generator control strategy obtained by the proposed preventive control method is practical in the real power system.

References

[1] Tao Jin, Siyi Liu, Rodolfo C.C. Flesch and Wencong Su. “A method for the identification of low frequency oscillation modes in power systems subjected to noise.” *Applied Energy*, vol. 206, pp. 1379-1392, Nov. 2017.

[2] Morteza Khosravi-Charmi and Turaj Amraee. “Wide area damping of electromechanical low frequency oscillations using phasor measurement data.” *International Journal of Electrical Power & Energy Systems*, vol. 99, pp. 183-191, Jul. 2018.

[3] Mario R.Arrieta Paternina, Rajesh Kumar Tripathy, Alejandro Zamora-Mendez and Daniel Dotta. “Identification of electromechanical oscillatory modes based on variational mode decomposition.” *Electric Power Systems Research*, vol. 167, pp. 71-85, Feb 2019.

[4] Tripathy, Praveen, et al. "Estimation of low-frequency modes in power system using robust modified Prony." *IET Generation, Transmission & Distribution*, vol. 10, no. 6, pp. 1401-1409, Apr. 2016.

[5] Joice G. Philip and Trapti Jain. “An improved Stochastic Subspace Identifcation based estimation of low frequency modes in power system using synchrophasors,” *Electrical Power and Energy Systems*, vol. 109, pp. 495-503, Feb. 2019.

[6] Zhang, Shuqing, X. Xie, and J. Wu. "WAMS-based detection and early-warning of low-frequency oscillations in large-scale power systems." *Electric Power Systems Research*, vol. 78, no. 5, pp. 897-906, May 2008.

[7] Hwang, Jin Kwon, and Y. Liu. "Identification of interarea modes from ringdown data by curve-fitting in the frequency domain." *IEEE Transactions on Power Syst.*, vol. 32, no. 2, pp. 842-851, Jan. 2016.

[8] Wies, R. W., A. Balasubramanian, and J. W. Pierre. "Combining least mean squares adaptive filter and auto-regressive block processing techniques for estimating the low-frequency electromechanical modes in power systems." in *Proc. 2006 IEEE Power Engineering Society General Meeting*, pp.1054-1061.

[9] Wies RW, Pierre JW and Trudnowski DJ. “Use of ARMA block processing for estimating stationary low-frequency electromechanical modes of power systems.” *IEEE Trans Power Syst.*, vol. 18, no. 1, pp. 167–73, Feb. 2003.

[10] Chen, G, et al. "A new approach for online identification of low frequency oscillation modes based on auto-regressive moving-average model." *Power System Technology,* vol.34, no.11, pp. 48-54, Nov. 2010.

[11] Joice G. Philip and Trapti Jain. “Analysis of low frequency oscillations in power system using EMO ESPRIT.” *International Journal of Electrical Power & Energy Systems*, vol. 95, pp. 499-506, Feb. 2018.

[12] Annakkage, Udaya D., D. P. Wadduwage and K. Narendra. "Identification of dominant low-frequency modes in ring-down oscillations using multiple Prony models." *IET Generation, Transmission & Distribution*, vol. 9, no. 15, pp. 2206-2214, July 2015.

[13] P. Arunagirinathan and G. K. Venayagamoorthy. “Effects of Data Size on Stochastic Subspace Identification Method for Power System Electromechanical Modes.” *IFAC ParpersOnLine*, vol. 51, no. 13, pp. 668-673, July 2018.

[14] Farias, Gonzalo, et al. "Automatic feature extraction in large fusion databases by using deep learning approach." *Fusion Engineering and Design*, vol. 112, pp. 979-983, Nov. 2016.

[15] Yingkai Wang. “Analysis of low frequency oscillations based on deep learning”, Master dissertation, South China Univ. of Technology, 2017.

[16] Xuejian Wang, “Study on mode identification of low frequency oscillation in power system based on intelligent computing”, Master dissertation, South China Univ. of Technology, 2017.

[17] Tang, Ying Jie , Y. J. Tang , and X. L. Xie . "The new identification method for low frequency oscillation mode in power system based on Prony algorithm and neural network." *Applied Mechanics and Materials,* 321-324(2013):1400-1404.

[18] Shen, Zhongting, and Renjie Ding. "A novel neural network approach for power system low frequency oscillation mode identification." *2019 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE*, 2019.

[19] Chung C Y, Wang Lei, Frederic Howell. “Generation rescheduling methods to improve power transfer capability constrained by small-signal stability.” *IEEE Trans on Power Syst.*, vol. 19, no. 1, pp. 524-530, Feb. 2004.

[20] Zhang, Junbo, et al. "Locally Weighted Ridge Regression for Power System Online Sensitivity Identification Considering Data Collinearity." *IEEE Transactions on Power Syst.*, vol. 33, no. 2, pp. 1624-1634, March 2018.

[21] Junbo Zhang, Chao Lu, Chi Yung Chung, Kun Men and Liang Tu. "Online re-dispatching of power systems based on modal sensitivity identification." *IET Generation, Transmission & Distribution*, vol. 9, no. 12, pp. 1352-1360, Feb. 2015.

[22] Zhang, Junbo, Xiangtian Zheng, Zejing Wang, Lin Guan and C.Y.Chung. "Power System Sensitivity Identification – Inherent System Properties and Data Quality." *IEEE Transactions on Power Syst.*, vol 32, no.4, pp. 2756-2766, July 2017.

[23] Guoqiang Zhong, Li-Na Wang, Xiao Ling and Junyu Dong. “An overview on data representation learning: From traditional feature learning to recent deep learning.” *The Journal of Finance and Data Science*, vol. 2, no. 4, pp. 265-278, Dec. 2016.

[24] Newton Spolaôr, et al. "A Systematic Review of Multi-Label Feature Selection and a New Method Based on Label Construction." *Neurocomputing,* vol. 180, pp. 3-15, Mar. 2016.

[25] Francisco J. Pulgar, Francisco Charte, Antonio J. Rivera and María J. del Jesus. “Choosing the proper autoencoder for feature fusion based on data complexity and classifiers: Analysis, tips and guidelines.” *Information Fusion*, vol. 54, pp. 44-60, Feb. 2020.

[26] Utthara Gosa Mangai, Suranjana Samanta, Sukhendu Das and Pinaki Roy Chowdhury. “A Survey of Decision Fusion and Feature Fusion Strategies for Pattern Classification”, *IETE Technical Review*, vol. 27, no. 4, pp. 293-307, Jul. 2010.