IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE POWER & ENERGY SOCIETY SECTION

Received 17 August 2023, accepted 4 September 2023, date of publication 27 September 2023, date of current version 25 October 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3320051*

RESEARCH ARTICLE

# Transient Stability Assessment Using Deep Transfer Learning

**JONGJU KIM[1], (Member, IEEE), HEUNGSEOK LEE[2], SUNGSHIN KIM[2], (Member, IEEE), SANG-HWA CHUNG[3,4], (Member, IEEE), AND JUNE HO PARK[2,4], (Member, IEEE)**

[1]Korea Southern Power Company Ltd., Busan 48400, South Korea
[2]Department of Electrical Engineering, Pusan National University, Busan 46241, South Korea
[3]Department of Information Convergence Engineering, Pusan National University, Busan 46241, South Korea
[4]Dong-Nam Grand ICT Research and Development Center, Busan 48059, South Korea

Corresponding author: June Ho Park (parkjh@pusan.ac.kr)

This work was supported in part by the National Research Foundation of Korea funded by the Korean Government under Grant NRF-2020R1F1A1076497; and in part by the Ministry of Science and ICT (MSIT), South Korea, under the Innovative Human Resource Development for Local Intellectualization Support Program supervised by the Institute for Information & Communications Technology Planning & Evaluation (IITP) under Grant IITP-2023-2016-0-00318.

**ABSTRACT** This study proposes a deep transfer learning method using a deep convolutional neural network pre-trained with ImageNet for transient stability assessment. The procedure of deep transfer learning, incorporating the role and considerations of the transient stability assessment system, is suggested. The transient assessment system learns the relationship between the severity of disturbances and transient stability in the power system through the proposed training method. The severity of a disturbance based on the physical causal relationship of the angle stability is proposed to train and implement the transient stability assessment model. The power system state variables obtained from the phasor measurement unit are converted into the feature map described by the severity of a disturbance, enabling the training of transient stability characteristics of the power system to the deep convolutional neural network. The training dataset is constructed using the time-domain simulation on IEEE 39 and IEEE 118-bus benchmark power system models configured in MATLAB/Simulink. As a result of deep transfer learning, which involves training with freezing some of the convolutional layers of the pre-trained deep network, the most suitable model for transient stability assessment is selected among the pre-trained deep networks. The effectiveness of the proposed method is compared with other approaches using the confusion matrix, and the robustness against noise interference is also investigated.

**INDEX TERMS** Transient stability assessment, deep transfer learning, pre-trained deep convolutional neural network, VGG, power system stability, deep learning, ImageNet.

## I. INTRODUCTION

In line with the global trend of carbon neutrality, the proportion of renewable power generation in power systems has continuously increased. However, the output fluctuation of renewable power generation dependent on nature adds to difficulties in power generation operation. Additionally, although the continuous increase in the electricity demand causes the need for additional generation reserves, power generation facilities are gradually becoming intensive owing to the difficulty in securing new power generation sites, leading to the concentration of power flow. As changes in the operating environment of power systems accelerate, various uncertainties faced by power systems pose a major threat to stability, and there is a growing need for appropriate countermeasures to overcome the limitations of passive protection and control technologies of existing sequential operations. First of all, real-time analysis of the current state of the power system is required to operate the power system securely. Real-time stability assessment would improve stability by guiding

The associate editor coordinating the review of this manuscript and approving it for publication was Emilio Barocio.

power system operators to perform emergency control and remedial actions at an early stage and resolve inefficiency induced by stability constraints in power system operation. In this regard, although the importance of power system stability assessment has been continuously recognized, the area of real-time assessment remains a great challenge, primarily due to limitations in computing resources.

A power system is a combination of various types of dynamic systems to seek energy balance against continuous changes and disturbances. Stability is the ability to regain a state of operating equilibrium from various disturbances [1], [2]. So as to assess transient stability, it is necessary to interpret the complex dynamic characteristics of the components in a power system. Until now, the most accurate method to verify power system stability is to carry out the time-domain simulation on a case-by-case basis for various disturbances. However, given that requiring an accurate model of the power system, including the control system, the parameter of the component, and the network, and solving high-dimensional nonlinear differential-algebraic equations required considerable computer resources [1], the simulation of all possible disturbances in real-time is unfeasible.

Based on the energy function, an approach to find the stable region of the power system and assess the stability through the state trajectory of the power system after a disturbance was demonstrated [3]. In this regard, an accurate controlling unstable equilibrium point for disturbance was attained for stability analysis in [4] and [5]. State data acquisition using the phasor measurement unit (PMU) was considered [6], and a technique for constructing a lookup table for the real-time application was also suggested [7]. However, when applied to large power systems, energy functions are vulnerable to numerical problems in some areas and are somewhat difficult to use in real-time due to modeling limitations and the unreliability of computational techniques [8].

Deployment of the PMUs provides very high sampling and accurate dynamic state variables with the power system operation system [9]. As a result, machine learning technology began to be used for transient stability assessment (TSA) in earnest. The stability classifications through the neural network [10], support vector machine (SVM) [11], curve fitting [12], and decision tree [13] were proposed. SVMs have been widely used for classification problems with their excellent learning speed and accuracy. In [14], a performance comparison of SVM classifiers using multiple state variables of the generator was conducted, and Wang et al. utilized a core vector machine that can solve the classification problem with a higher dimension than the SVM for transient stability classification [15]. In addition, methods linking the SVM with the Fuzzy analytical hierarchy process [16] and convolutional neural network (CNN) [17] were also conducted. Although SVMs exhibit excellent performance in a normal-size dataset, the scalability of the classification model in the case of the accumulation of datasets remains somewhat insufficient.

Further, Ma et al. suggested a hybrid technique of determining and quantifying stability by comparing with measured values based on simulation results based on the equal area criterion [18]. An approach to predict transient stability based on the maximal Lyapunov exponent estimated by a recursive least squares-based method was proposed [19]. In [20], a nonlinear semiparametric model obtained with the LASSO algorithm was used for transient stability analysis. Besides, Ashraf classified critical machines through one machine infinite bus transformation and carried out rotor angle prediction with the Kalman filter [21], and in [22], stability classification through networked shapelet learning based on the inherent spatial-temporal correlations of power systems was carried out.

Since the perceptron was introduced, approaches using neural networks for the classification problem have been continuously attempted. Research about constructing a model for classifying sequence data by organizing individual classification neural networks in parallel [23], assessing the stability using the stacked sparse autoencoder as a feature extractor [24], a training method of including a dimension of time in the input state variables [25], an assessment model stacked with restricted Boltzmann machines [26], and radial basis function neural network-based model to determine the criticality level of the generators [27] were suggested.

CNN [28] and long short-term memory (LSTM) [29] have demonstrated excellent performance in image and sequence data recognition, respectively, creating an opportunity for the widespread use of artificial intelligence technology. Accordingly, CNN and LSTM have been applied to TSA and achieved high accuracy. Notably, a study on determining transient stability with LSTM specialized in processing sequence data and methods of converting state data into an image format to utilize CNN were presented [30], [31], [32]. In [33], [34], [35], and [36], the performance of classifying transient stability was further enhanced by combining CNN with LSTM.

The significance of the approaches suggested so far is clear. Nonetheless, considering the crucial role of TSA in power systems, inconsistent assessment timing and relatively long response time after the fault clearing should be supplemented. To be considered for practical application, improving the accuracy and response time of TSA is crucial to provide sufficient time for emergency control or remedial action immediately after the assessment. Most importantly, as training data accumulates persistently, designing the model with the potential for advancement is necessary.

Given that time-domain simulation is the most accurate method to assess transient stability of power systems, a deep neural network-based TSA using the time-domain simulation results can be an alternative to satisfy accuracy, feasibility, and responsiveness. However, in order to learn numerous hyperparameters to use a deep neural network, not only a vast amount of dataset is required, but also an appropriate design of the network structure should be premised so that the gradient of the error can be effectively propagated during the process of learning through the backpropagation algorithm [37].

In this study, in order to overcome these limitations and build a high-performance stability assessment system, a pre-trained deep convolutional neural network (PDCNN) is applied to TSA. Deep transfer learning (DTL), which brings the backbone of a neural network model trained with a large-scale dataset and applies it to another field, can overcome the insufficiency of the training dataset and provide significant usefulness in designing neural networks for TSA.

The contributions of this paper are:

1) As the first attempt at DTL of a deep CNN trained with ImageNet for power system stability assessment, the method and criterion of exploiting a PDCNN to TSA are proposed.
2) The TSA system is constructed by learning the proposed disturbance severity index, which possesses the physical characteristics of angle stability. The disturbance severity index synergizes with the excellent classification competence of the deep CNN, resulting in fast response time and high accuracy in the assessment.
3) In this study, an attempt to apply deep CNN pre-trained with completely different tasks to transient stability assessment through fine-tuning proves that DTL is a way to alleviate the need for vast amounts of the training dataset, which is a prerequisite for training deep neural networks.

The remainder of the paper is organized as follows. Section II establishes the primary considerations and proposed approach for constructing the TSA system using the neural network. The background of DTL for TSA is addressed in Section III. Section IV examines the theoretical basis for configuring the training dataset and processing the PMU data. The detailed process of constructing the TSA model through fine-tuning is addressed in Section V. In Section VI, the performance of the proposed method is verified and compared with previous studies.

**TABLE 1. Details of the neural network-based transient stability assessment systems proposed in previous studies.**

| Model | Observation window length | Stability decision threshold | Classification label | Response time[b] | Ref. |
|---|---|---|---|---|---|
| CNN | 5 cycles | 0.9 | 0/1 | 15 cycles | [31] |
| CNN | 15 cycles | N/A | 0/1/2 | 0–9 cycles | [32] |
| CNN | 20 cycles | 0.08 | 0/1 | 13 cycles | [33] |
| LSTM | 5 cycles | 0.4 | 0/1 | 1–9 cycles | [30] |
| GCN[a] LSTM | 20 cycles | 0.9 | 0/1 | 0 | [34] |

[a]Graph convolutional network; [b]The waiting time for assessment with the best accuracy after the fault clearing.

## II. CONSIDERATIONS FOR TRANSIENT STABILITY ASSESSMENT USING NEURAL NETWORKS

In previous research, neural networks of various structures have been applied, and in particular, studies using LSTM and CNN constitute the majority. In implementing TSA using artificial intelligence technology, there are essential options that affect performance and feasibility. Because those options are issues that still need to be resolved and cause various difficulties in adopting the neural network-based TSA in an actual power system, the following must be considered to move forward from previous studies.

### A. OBSERVATION WINDOW AND STABILITY DECISION THRESHOLD

A compromise about the period of the observation window (OW), which is the temporal range of input data, must be found in the trade-off relationship between the accuracy and the timing of the stability decision. Moreover, the OW determines the size of the input data of the neural network; thus, it is a factor that significantly affects the structure of the neural network. The setting of the stability decision threshold (SDT) influences the performance of the TSA system due to acting as a criterion for assessing stability from the inference result of the trained neural network. Notwithstanding these two hyperparameters play a critical role in the neural network-based stability assessment, the OW and SDT, whose value was experimentally set depending on each case study of the power system model and training data, as shown in Table 1, make it challenging to exploit the TSA system to an actual power system.

### B. TRADE-OFF BETWEEN ACCURACY AND TIMING

Given that transient stability becomes apparent over time after a power system is affected by a severe disturbance in conformity with the physical nature of synchronous operation between synchronous generators, the temporal feature (the length of the OW) of input data significantly impacts the accuracy and timing of a decision, i.e., tending to be less accurate as it gets shorter. The inference timing, as well as accuracy, is crucial in the TSA system because the prompt assessment can provide sufficient time to perform emergency control in power systems. Previous studies found a compromise between the accuracy and the decision timing through several attempts comparing simulation results. Still, it cannot be asserted that the compromise is an appropriate point that can be used as it is in other data (i.e., other power systems).

### C. NEURAL NETWORK DEPTH AND STRUCTURES

The detailed structure of a network suitable for a specific purpose can only be found by trial and error by referring to various research [38]. Although the general characteristics have been confirmed through a recent study on the correlation of network depth and performance [39], the detailed relationship between hyperparameters of a neural network is like a black box, making it difficult to grasp the exact causal

relationship [40]. The TSA systems composed of limited training data in various studies face challenges in clearly elucidating the detailed structural features of neural networks and explaining why they exhibit good performance, as well as in determining which of the proposed models possesses a more efficient network structure. Moreover, it is very difficult to compare, verify, and select the best network structure that can be generalized because of differences in terms of benchmark models, detailed data acquisition scenarios, and data scales for each research.

### D. TRAINING DATA INSUFFICIENCY AND VARIABILITY

In order to build a TSA system with a neural network, a large number of training data consisting of state data of power systems experiencing various disturbances, which is enough to train a neural network, should be prepared. However, in actual power systems, there are few unstable cases due to tightly designed protection systems, reserve allocation, and operational technologies that prevent the blackout; thus, time-domain simulation results for various disturbances are inevitably used for training a neural network. Moreover, as operation data of the power system will have been accumulated in the database and the amount of the data gradually increases, model performance improvement would be required through additional training. As such, the TSA model should be able to fully generalize transient stability characteristics even with limited training datasets and steadily expand the model through additional training against the continuous accumulation of operational data, although these two aspects may be in conflict.

A neural network model can learn more complex phenomena as it has a large number of hyperparameters, whereas it also requires a large amount of data to be learned sufficiently [38]. In the case of training a simple neural network using the initially obtained small-scale training dataset, the neural network exhibits good performance for the moment, while the performance improvement is inevitably limited when the training dataset is extended. On the other hand, if a deep neural network is indiscriminately used, the neural network cannot be sufficiently trained due to the amount of the training dataset, and thus the expected performance cannot be reached.

### E. PROPOSED APPROACHES FOR ADDRESSING ISSUES

The main issues are summarized as follows:

1) Determination of the OW and the SDT.
2) The compromise between the accuracy and the decision timing of TSA.
3) Design of a suitable neural network structure: A neural network model that can be trained with small-scale data and grow by continuously learning about the increase in the amount of training data.

The methods proposed in this paper to cope with the considerations and issues are as follows:

1) Application of deep CNN through DTL: overcoming the variability of the training dataset and securing a neural network model possessing the finest performance.
2) Construct an adequate training dataset so as to avoid setting experimentally determined hyperparameters such as the OW and the SDT and exert the capability of a deep CNN on TSA.

## III. DEEP TRANSFER LEARNING FOR TRANSIENT STABILITY ASSESSMENT

As a neural network structure deepens, tasks to be learned can be hierarchically decomposed, complex problems can be replaced with simpler ones, and the feature information in data is transmitted hierarchically, contributing to efficient learning [38]. As the layer of CNN deepens, the extracted information becomes more abstract, and the object to which neurons respond changes from simple shape to advanced information, which is the basis for the outstanding performance of a deep CNN [41], [42]. Nevertheless, training a deep neural network requires a vast amount of training data and computing resources. Furthermore, the design of a deep neural network is a significantly complicated task, and if it is designed to possess a deep structure, problems such as vanishing gradient and overfitting would frequently occur during the training process. In this way, designing and verifying a deep neural network by considering various aspects, including training availability, training time, performance, efficiency, and the number of hyper-parameters, requires much effort. Notwithstanding, through DTL, the deep neural networks that are impossible to be trained with only the moderate-scale training dataset (i.e., used in previous studies [30], [31], [32], [33], [34], [35], [36]) can be fine-tuned and applied to TSA. DTL can be an excellent alternative when training data is relatively small, such as in a power system, and available computing resources are limited.

### A. DEEP TRANSFER LEARNING

The ILSVRC (ImageNet Large-Scale Visual Recognition Challenge) is an image recognition competition that classifies ImageNet [43], which comprises 1,200,000 images (see Fig. 1) categorized into 1,000 classes. Because the same data is used to comparatively analyze the overall performance of the models, such as the training speed, the size, parameters, and the top-1 and top-5 accuracy, the best-performing model in the field of image recognition is identified. Since AlexNet [44] won in 2012, CNN-based models have been winning so far, and recently, top-ranked models tend to design the structure of neural networks deeper based on deep learning in earnest.

Transfer learning is motivated by the fact that people can intelligently apply knowledge learned previously to solve new problems faster or with better solutions [45]. It is a valuable method in AI technology to overcome insufficient training data. This method extracts the knowledge from the source task (i.e., ImageNet classification) and applies it to a

**FIGURE 1.** ImageNet dataset image samples.

target (new) task. Research in which PDCNNs with ImageNet have been successfully deep transfer learned in various fields such as medical image, mechanics, physics, and civil engineering can be found [46]. Still, there has yet to be a case where DTL has been used in power systems. Although some studies have partially confirmed the applicability [35], [47], it was limited to cases where transfer learning was performed within a larger simulation model for the neural network learned in the study. It is not an attempt to transfer-learn the representative neural networks trained with ImageNet, such as AlexNet, GoogLeNet [48], ResNet [49], and VGG [50], and apply them to the field of power system stability. The structures of CNNs for TSA used in previous studies are mostly 2, 3, and 4 convolution layers, in contrast to having 5–81 convolution layers, which are representative deep CNN models (see Table 2 ). While the depth of the layer cannot be the sole criterion for the classification performance of neural network models, it is evident that deep neural networks, in general, are capable of modeling more complex features [38], [51]. DTL is simultaneously capable of solving the insufficiency of training data and the difficulty of designing an appropriate deep network in applying deep networks to TSA. The expected merits of DTL for TSA are summarized as follows:

1) Accurate and fast response assessment by the excellent image recognition ability of the deep CNN.
2) Addressing the lack of training dataset, the need for large-scale computing resources to train deep networks, excessive training time, and concerns about overfitting.
3) Scalability to improve performance through additional learning when the amount of training data enlarges continuously due to the acquisition of training data in the future.
4) The state-of-the-art high-performance deep CNNs that have consistently been developed in the field of image recognition can be applied to assess transient stability through the process proposed in this paper. There is no need for additional consideration of hyperparameters related to network structure. As the image recognition ability of the deep CNN improves, the performance of TSA is also enhanced.

### B. FINE-TUNING

Fine-tuning is re-training a deep neural network pre-trained with a vast dataset (e.g., ImageNet) into a new dataset suitable for other tasks. Given that the PDCNN is used as a backbone and only a few layers are initialized and trained, the SGDM, which generally carries out stable learning, is set at a very small learning rate and used for fine-tuning. The initial layers of a trained deep neural network preserve general features applicable to other tasks, while the deep layers have features that are more specialized for the target tasks [51], [52]. When performing fine-tuning to apply a pre-trained deep network to a new task, these characteristics are the foundation for dividing networks into frozen or learnable layers. As shown in Fig. 2, the depth of the learnable layer of a neural network to be trained is generally determined by comparing the size of the new training data and its similarity to the dataset used for pre-training [53].

**TABLE 2.** Comparison of the stability assessment model structure.

| Other approaches for transient stability assessment | | | | | | | Proposed |
|---|---|---|---|---|---|---|---|
| Model | CNN | CNN | CNN | LSTM | GCN LSTM | CNN LSTM | PDCNNs[c] |
| Depth | 3 | 5 | 3 | 3 | 7 | 6 | 8–82 |
| Details | 2Conv[a], 1FC[b] | 3Conv, 2 FC | 2Conv, 1FC | 2LSTM, 1FC | 3Conv 1LSTM 3FC | 4Conv 1LSTM 1FC | 5–81Conv, 1–3FC |
| Ref. | [31] | [32] | [33] | [30] | [34] | [35] | Table III |

[a]Convolution layer; [b]Fully connected layer (i.e., multilayer perceptron); [c]Pre-trained deep convolutional neural networks.

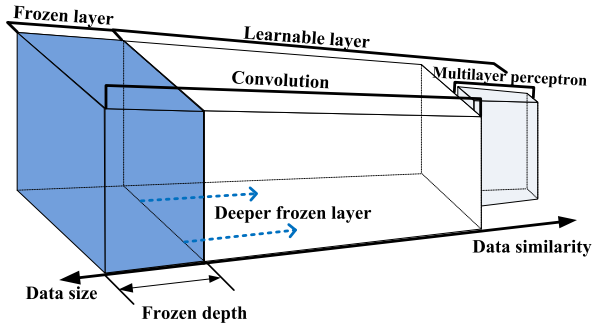


**FIGURE 2.** Relationship between the property of the training dataset and the depth of the frozen layer.

### C. PRE-TRAINED DEEP CONVOLUTIONAL NEURAL NETWORK SELECTION FOR TRANSIENT STABILITY ASSESSMENT

The method of constructing the DTL-based TSA is illustrated in Fig. 3. Candidate models are selected from the PDCNNs on ImageNet (or other large-scale datasets), considering the

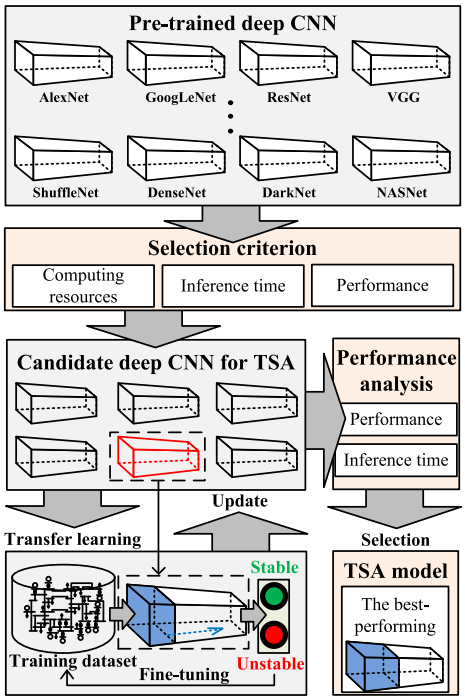


**FIGURE 3.** Deep transfer learning for transient stability assessment.

system environment to implement the TSA system in real-time (inference time within approximately 5ms). The most crucial consideration for applying the PDCNN to TSA is to select a suitable network that requires little time in the inference process, as transient stability should be determined in real-time from power system state variables. In general, the forward propagation of neural networks takes little time, but most of the deep CNN currently proposed for image recognition and classification have very large hyperparameters along with quite deep and complex structures; thus, consideration of inference time should be preceded in selecting networks. In addition to this, the image classification performance, the complexity of the structure, and the number of hyperparameters are considered. The candidate models that meet the priority criteria for real-time application are fine-tuned using the training dataset, which consists of state variables acquired from the power system model to be implemented, and then the most suitable model is selected through performance comparison. The networks arranged in Table 3 are selected as candidates for the real-time TSA model, considering inference speed and detailed specifications among various deep networks trained with ImageNet. All selected models are affirmed to achieve real-time performances from the inference time (i.e., response speed).

## IV. PROPOSED TRANSIENT STABILITY ASSESSMENT SYSTEM

With the spread of PMUs, an environment where voltage phasors can be acquired in 1 cycle with an error rate of less than 1% has been established [55]. The PMU provides

**TABLE 3.** Pre-Trained deep convolutional neural networks: candidates for deep transfer learning.

| Deep CNN model | Inference time[a] (ms) | Accuracy rank[b] | Parameters (millions) | Depth | Size (MB) |
|---|---|---|---|---|---|
| AlexNet | 1.28 | 10 | 61.0 | 8 | 227 |
| SqueezeNet | 1.60 | 9 | 1.24 | 18 | 5.2 |
| ResNet-18 | 1.79 | 4 | 11.7 | 18 | 44 |
| MobileNet | 3.34 | 3 | 3.5 | 53 | 13 |
| GoogLeNet | 4.54 | 7 | 7.0 | 22 | 27 |
| ResNet-50 | 5.10 | 1 | 25.6 | 50 | 96 |
| VGG-16 | 5.17 | 6 | 138 | 16 | 515 |
| ShuffleNet | 5.40 | 8 | 1.4 | 50 | 5.4 |
| VGG-19 | 5.50 | 5 | 144 | 19 | 535 |
| DarkNet-19 | 5.1–5.17 | 2 | 20.8 | 19 | 78 |

[a]Average per-image inference time over 10 runs reported in [54]; [b]Relative rank with the top-1 accuracy of ImageNet validation among the models in this table.

ease of real-time monitoring and data accumulation of the power system and is a suitable signal for machine learning-based real-time TSA. In this study, only these two types of state variables are considered as inputs of the neural network model, assuming that the PMU can acquire the magnitude and phase of each bus voltage. In a neural network-based model, the training dataset, data property, and pre-processing of the input variable are essential options that determine the performance of the neural network associated with the characteristics of the task and type of the neural network. In this section, the establishment of the training dataset that maximizes the utility of the PDCNN model is proposed.

### A. TRAINING DATASET

In supervised learning of a neural network, the composition of input and its label, that is, the training dataset, is a process of determining the expected inference results through the model learned from the dataset and specifying the task of the model. In order to interpret and learn an intricate system in which various systems are organically combined, such as a power system, the training dataset should be constructed circumspectly based on physical relevance in particular. Likewise, given that the performance of the TSA system should be improved through enduringly additional learning from the operation data accumulated from the actual power system and the time-domain simulation, the critical consideration for building the neural network-based TSA system is that the training dataset should be able to be constructed from the state data of the actual power system as well. Considering these aspects, PMU-based state variables that can be easily accumulated and instantaneously acquired with high accuracy from the power system are most appropriate as input data, and the transient stability index (TSI) (1) can classify the stability from the input data and label whether stable or not [23].

$$\mathrm{TSI} = \frac{360^\circ - |\delta_{max}|}{360^\circ + |\delta_{max}|} \times 100 \tag{1}$$

where $\delta_{max}$ is the maximum phase angle difference between any synchronous generators. Transient stability can be classified (i.e., labeled) as stable in TSI $> 0$, and its class label is tagged as 1; otherwise, as unstable in TSI $\leq 0$, its class label is tagged as -1.

$$\mathcal{Y} = \begin{cases} 1 \text{ (stable)}, & \text{TSI} > 0 \\ -1 \text{ (unstable)}, & \text{TSI} \leq 0 \end{cases} \tag{2}$$

### 1) POWER SYSTEM TRANSIENT STABILITY

Mathematically, transient stability of a power system is a problem to solve the following nonlinear differential equation [1], [8]:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{V}, t) \tag{3}$$

$$0 = \mathbf{g}(\mathbf{x}, \mathbf{V}, t)$$

$$\mathbf{x} = \{x_i | \, i = 1, 2, \cdots, n\}, \quad \mathbf{x}(t_0) = \mathbf{x}_0$$

$$\mathbf{V} = \{V_b | \, b = 1, 2, \cdots, m\}, \quad V_b = [|V_b| \, \theta_b]^{\mathrm{T}} \tag{4}$$

where $\mathbf{x} \in \mathbb{R}^{\mathbf{n}}$ are the state variables, $\mathbf{x}_0$ is the initial values of state variables, $t \in [t_0, T]$ is time, $\mathbf{V} \in \mathbb{R}^{\mathbf{m}}$ are the algebraic variables (i.e., bus voltages), $\mathbf{f}(\mathrm{x})$ is a nonlinear differential equation, $\mathbf{g}(\mathrm{x})$ is a nonlinear algebraic equation, and the number of the generator and bus are $n$ and $m$, respectively. Transient stability is identified from the maximum phase angle difference, $\delta_{max}$, obtained from the state variable $\mathbf{x}$. This is because the state variable $\mathbf{x}$ and the algebraic variable $\mathbf{V}$ are obtained through the solution of the nonlinear differential-algebraic equation:

$$\mathbf{x}(t_0 + \Delta t) = \mathbf{x}_0 + \int_{t0}^{t_0 + \Delta t} \mathbf{f}(\mathbf{x}, \mathbf{V}, t) dt \tag{5}$$

$$0 = \mathbf{g}(\mathbf{x}(t_0 + \Delta t), \mathbf{V}(t_0 + \Delta t), t) \tag{6}$$

$$|\delta_{max}| = \max_{t_0 \leq t < T} \left| \delta_i(t) - \delta_j(t) \right| i, j \in n \tag{7}$$

where $|\delta_{max}|$ represents the maximum phase angle difference between any two generators. Connecting this behavior with the equal area criterion (see Fig. 4), widely known as a helpful method for an intuitive grasp of transient stability, the imbalance energy (area abcd in Fig. 4) stored during the fault period $[t_F, t_C]$ appears as an increase in power angle ($\delta_{max}$), i.e., a release of the stored energy (area defg in Fig. 4). On the other hand, if the fault is cleared at $t_C\prime$, the stability limit is exceeded owing to the energy accumulated during the difference of the fault clearing time ($t_C\prime - t_C$). That is, the imbalance energy stored in the power system during the fault period (i.e., $[t_F, t_C]$ or $[t_F, t_C\prime]$) appears as a change in the state variable $\mathbf{x}$ in the differential algebraic equation, and the trajectory of the algebraic variable $\mathbf{V}$ is obtained by solving the nonlinear algebraic equation. Collectively, concerning the physical aspect of energy accumulation and dissipation through the equal area criterion and the relationship between $\mathbf{V}$ and $\mathbf{x}$ of the nonlinear algebraic equation, the training dataset is constructed to model the correlation between the imbalance energy accumulated during the fault

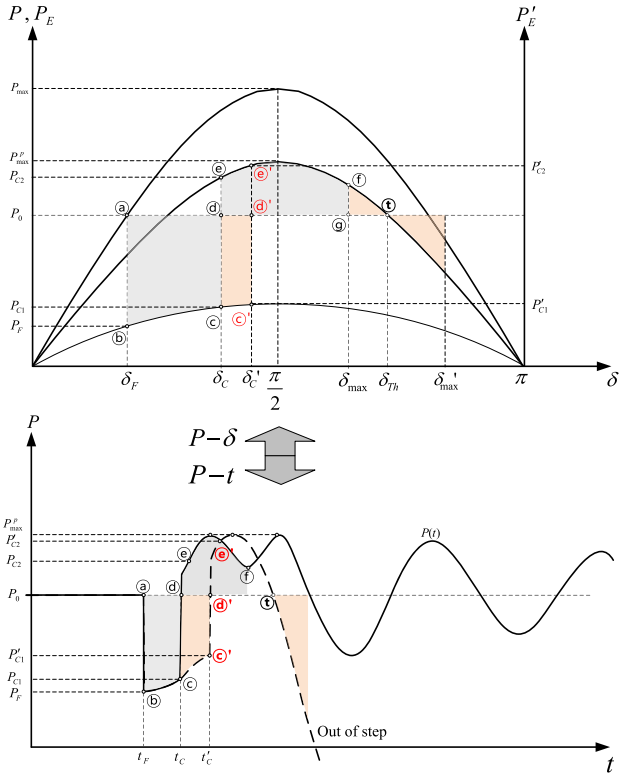


**FIGURE 4.** Equal area criterion for transient stability assessment.

period expressed as $\mathbf{V}$ and the transient stability. The proposed training dataset is

$$\mathcal{D}_{TSA} \triangleq \{(\mathcal{X}_{case}, \mathcal{Y}_{case}) | \, case = 1, 2, \cdots, N_D\}$$

$$\begin{cases} \boldsymbol{\mathcal{X}} \text{ (input)} : \text{accumulated imbalance energy} \\ \boldsymbol{\mathcal{Y}} \text{ (label)} : \text{transient stability index} \end{cases} \tag{8}$$

where $N_D$ is the number of the fault case, $\mathcal{X}_{case}$ is the input of the training dataset, and $\mathcal{Y}_{case}$ is the classification label of the dataset. In conclusion, through the training dataset for constructing the TSA model that exerts the assessment with input data solely during the fault period, the model learns how severe the disturbance occurs in the power system and the effects on generators (i.e., the angle stability). This model targets to assess transient stability immediately after the fault clearing in the power system, as provided in Fig. 5. As depicted by the blue solid line box in Fig. 5, the size of the observation window proposed in this study varies depending on the duration of the fault. The TSA model is trained using only the state data within the fault period, allowing it to assess transient stability immediately after the fault clears. In this method, it is possible to secure control time for the power system operator to perform additional countermeasures compared to other approaches due to the fast response time being made at $T_{C1}$ and $T_{C2}$ for TSA. On the other hand, in other approaches, as summarized in Table 1, fixed observation windows of 5 to 20 cycles (i.e., the yellow solid and dashed lines in Fig. 5) are used, and the inference

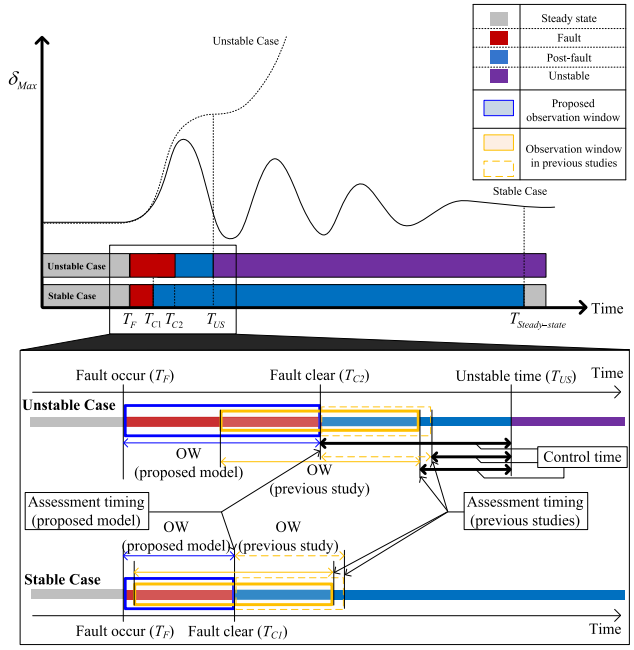


**FIGURE 5.** Comparison of the observation window.

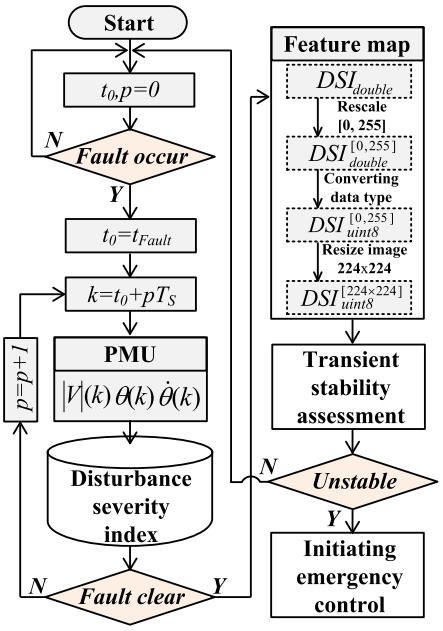


**FIGURE 6.** Flow chart of processing the state variable to the feature map in the proposed system.

results of the model are delayed until the SDT is met. Other approaches using CNN and LSTM require different waiting periods (i.e., observation window) for each case to enhance the accuracy of the assessment after fault clearing, leading to delays and uncertainty in the assessment timing, making it challenging to use the assessment result as an initiation trigger for emergency control of the power system. Contrary to the response time of previous studies (see Table 1), the TSA model learned with the proposed dataset is built to infer transient stability as soon as the fault clears in all cases and is suitable for application in power system operation.

### 2) DISTURBANCE SEVERITY INDEX

So as to construct the training dataset, it is necessary to quantify the effect of disturbance on transient stability of the power system. Taken together, the disturbance severity index (DSI), which includes the location of faults, the distribution of power flows, acceleration energy, and the fault duration, is defined. It can be obtained from time-series status variables from the PMU. In (3) and (4), the algebraic variable $\mathbf{V}$, the voltage phasor acquired in real-time from the PMU, is

$$\begin{aligned}\mathbf{V}\left(t_p\right) &= \left[|\boldsymbol{V}|\left(t_p\right)\boldsymbol{\theta}\left(t_p\right)\right]^{\mathrm{T}} \\ &\begin{cases}|\boldsymbol{V}|\left(t_p\right) = \{|V_b||\, b=1,2,\cdots,m\} \\ \boldsymbol{\theta}\left(t_p\right) = \{\theta_b|\, b=1,2,\cdots,m\} \\ \dot{\boldsymbol{\theta}}\left(t_p\right) = \left\{\dot{\theta}_b\middle|\, b=1,2,\cdots,m\right\}\end{cases} \\ \dot{\theta}_b\left(t_p\right) &\triangleq \frac{\theta_b\left(t_p\right)-\theta_b\left(t_p-\Delta T_S\right)}{\Delta T_S}\end{aligned} \tag{9}$$

where $t_p$ is the sampling time, $\Delta T_S$ is the sampling interval, the subscript $b$ is the bus number, $m$ is the total number of buses, $|\boldsymbol{V}|$ is the magnitude of the bus voltage, $\theta$ is the phase of the bus voltage, and $\dot{\theta}$ is the phase angle change with respect to the sampling interval. Because the voltage distribution of the entire power system during the fault period is determined by factors such as the location of the voltage source, line admittance, fault impedance, and fault location, it can be observed that the voltage magnitude of all buses in the PMU conveys the information about the fault and network properties. Also, the power flow is the function of voltages and phase angles, so the trajectories of voltages, phase angles, and derivatives of phase angles during the fault period are caused by unbalanced energy under the influence of the fault; thus, these variables stand for the disturbance severity influencing the transient stability. From (9), the DSI is defined as

$$\begin{aligned}&\textbf{\textit{Disturbance severity index}}: \textbf{\textit{DSI}} \\ &\triangleq \left\{\left[|\boldsymbol{V}|\left(k\right)\boldsymbol{\theta}\left(k\right)\dot{\boldsymbol{\theta}}\left(k\right)\right]^{\mathrm{T}}\middle|\, k=t_F, t_F+\Delta T_S,\cdots,t_C\right. \\ &\quad \left.-\Delta T_S, t_C\right\}\end{aligned} \tag{10}$$

where $k$ is the sampling time, $t_F$ is the fault occurrence time, and $t_C$ is the fault clearing time. Collectively, as shown in Fig. 5, the DSI is the trajectory of the variables of the fault period, and the temporal range of data varies depending on the fault clearing time.

### 3) FEATURE MAP GENERATION

The power system time-series data acquired from the PMU should be converted into the appropriate type for the CNN-based model so that the PDCNN effectively recognizes the correlation between the trajectory of power system state data and the transient stability. As shown in Fig. 6, the DSI defined in (10) is three-dimensional time series data

sampled only during the fault period, which is instantaneously converted into the feature map after the fault clearing and is imparted to the TSA model. To convert the DSI into the feature map, the state variables defined in (9) are sampled only during the fault period and assigned to each dimension in the form of a 3-dimensional tensor, as described in (10). The RGB image data format consists of 3-dimensional data with red, green, and blue components. Each dimension is represented by an 8-bit unsigned integer, ranging from 0 to 255, indicating the brightness or intensity of the respective color channel. Therefore, by linearly transforming the data distribution of each dimension in (10) into 8-bit unsigned integers ranging from 0 to 255, the data acquires the same attributes as an RGB image. In MATLAB, to perform the described process, as shown in Fig. 6, each state variable is transformed to have a data scale from 0 to 255 using "*rescale*", and then converted to an 8-bit unsigned integer data class using "*uint8*". These transformed state variables are then combined into a 3-dimensional tensor using "*cat*", and "*imresize*" is used to adjust the size of the input image to fit the CNN based model.

### B. FINE-TUNING FOR TRANSIENT STABILITY ASSESSMENT

The PDCNN models are defined as (11).

$$\begin{aligned} &H^{\mathcal{B}}\left(\boldsymbol{W}_{\mathcal{F}}, \mathcal{C}, \mathcal{X}_i\right) = \hat{\mathcal{Y}}_i \\ &H^{\mathcal{B}} \in \left\{ \begin{array}{l} \text{AlexNet, SqueezeNet, ResNet18, } \ldots, \\ \text{ResNet50, MobileNet, GoogLeNet, } \ldots, \\ \text{VGG16, VGG19, ShuffleNet, DarkNet19} \end{array} \right\} \\ &\boldsymbol{W}_{\mathcal{F}} \in \{10\%, 30\%, 50\%, 70\%, 90\%\} \\ &\mathcal{C} = \{\textit{Stable}, \textit{Unstable}\} \end{aligned} \tag{11}$$

$H^{\mathcal{B}}$ is the deep CNN model pre-trained with ImageNet to be the backbone, $\boldsymbol{W}_{\mathcal{F}}$ is the ratio of layers whose learning rate is set to 0 (i.e., the frozen layer), $\hat{\mathcal{Y}}_i$ is the inference result for input $\mathcal{X}_i$, and $\mathcal{C}$ is the category (i.e., class labels). The cross-entropy loss function with the regularization term [38] for training the PDCNN (11) is

$$E_L = -\frac{1}{N_D}\sum_{i=1}^{N_D}\left\{\mathcal{Y}_i \ln \hat{\mathcal{Y}}_i + (1-\mathcal{Y}_i)\ln\left(1-\hat{\mathcal{Y}}_i\right)\right\} + \lambda\frac{1}{2}\|\boldsymbol{W}\| \tag{12}$$

where $N_D$ is the number of samples, the output of the model for input $\mathcal{X}_i$ is $\hat{\mathcal{Y}}_i$ (the inference result), the class of input $\mathcal{X}_i$ is $\mathcal{Y}_i$ (the class label), $\boldsymbol{W}$ is the weight vector, and $\lambda$ is the regularization coefficient to reduce overfitting. The optimizer to minimize the loss function (12), the stochastic gradient descent with momentum (SGDM) [40], is

$$\boldsymbol{W}^{j+1} = \boldsymbol{W}^{j} - \alpha\nabla E_L\left(\boldsymbol{W}^{j}\right) + \gamma\left(\boldsymbol{W}^{j} - \boldsymbol{W}^{j-1}\right) \tag{13}$$

where $j$ is the iteration number, $\alpha$ is the learning rate, $\boldsymbol{W}$ is the weight vector, $\nabla E_L\left(\boldsymbol{W}^{j}\right)$ is the gradient of the loss function (12) with respect to $\boldsymbol{W}$ at iteration $j$, and $\gamma$ is the momentum which determines the contribution of the previous gradient step to the current iteration. In the training

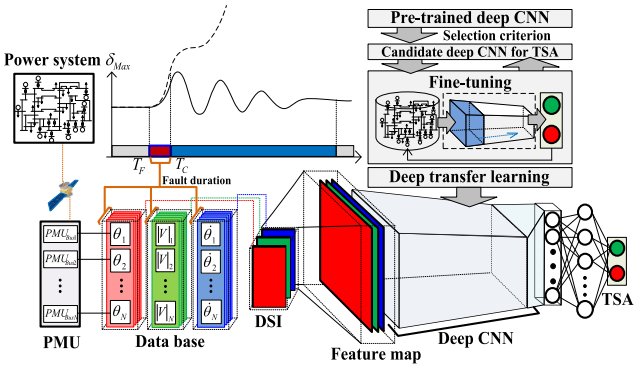


**FIGURE 7. An overall framework of the proposed transient stability assessment model.**

dataset (8), the feature map converted from the DSI (10) and the transient stability (2) for each case are assigned to the input $\mathcal{X}_i$ and label $\mathcal{Y}_i$, respectively. The training dataset is constructed accordingly, as (14).

$$\mathcal{D}_{TSA} = \{(\mathcal{X}_i, \mathcal{Y}_i) | \, i = 1, 2, \cdots, N_D\} \tag{14}$$

Using the training dataset (14), the PDCNNs with 10–90% depth of the frozen layer ($\boldsymbol{W}_{\mathcal{F}}$) are trained by the SGDM (13) to minimize the loss function (12) and, consequently, form the TSA model.

## V. CONSTRUCTING PROPOSED ASSESSMENT SYSTEM

Training and verification datasets are constructed through time-domain simulation in power system models of different scales and characteristics to verify the proposed TSA system. The assessment model is selected by comparing the performance of models trained by varying the depth of the frozen layer. The design of the neural network model, data processing, training, and DTL for assessment model construction are performed using MATLAB R2022b with Deep Learning Toolbox and Image Processing Toolbox. All the tests are fulfilled on a computer with Intel Core i9-7900X 3.30GHz CPU, 16GB RAM, and RTX 3090Ti GPU.

### A. TRANSIENT STABILITY ASSESSMENT MODEL

The DTL-based TSA system proposed in this study is illustrated in Fig. 7. The state variables of the power system are measured by PMUs and stored in the database, from which the DSI is reconstructed. Immediately after the fault clearing, the DSI is converted into the feature map (see Fig. 6). The feature map represents the power system state variable during the fault period in the type of image data, which is then fed into the deep CNN model. The TSA model is constructed by fine-tuning PDCNNs, as shown in Fig. 3, using the training dataset obtained from the power system model to be implemented.

### B. TRAINING DATA GENERATION

#### 1) SIMULATION MODEL

To verify the performance and feasibility of the proposed system, time-domain simulation models of the IEEE 39-bus

**TABLE 4.** Parameters for the benchmark power system.

| Component (model) | Parameter | Setting range |
|---|---|---|
| Synchronous generator (sixth-order model) | $X_d$; $X_d'$<br>$T_{do}'$; $T_{do}''$; H | 1–2.3; 0.15–0.4<br>4.5; 0.02–0.03; 3.5–10 |
| Prime-mover and governor (tandem compound) | $T_2$; $T_3$; $T_4$; $T_5$<br>$F_2$; $F_3$; $F_4$; $F_5$<br>$K_p$; $R_p$; $T_{sr}$; $T_{sm}$ | 0; 10; 3.3; 0.5<br>0; 0.36; 0.36; 0.28<br>1; 0.05; 1e-3; 0.15 |
| Exciter (IEEE-Type 1) | $K_a$; $T_a$; $E_{fmn}$; $E_{fmx}$ | 200; 1e-3; 0; 5 |
| Power system stabilizer (PSS-4B) | $K_S$; $F_L$; $K_L$<br>$F_I$; $K_I$; $F_H$; $K_H$ | 1; 0.2; 30<br>1.25; 40; 12; 160 |

**TABLE 5.** Configuration of the benchmark system.

| Benchmark model | Machine | Load | Bus | Line |
|---|---|---|---|---|
| IEEE 39-bus | 10 | 19 | 39 | 46 |
| IEEE 118-bus | 54 | 91 | 118 | 177 |

and the IEEE 118-bus system, which are widely used as the benchmark system, are constructed using MATLAB/Simulink. The parameters and configuration of each model are set up as Tables 4 and 5.

#### 2) DATABASE GENERATION

The training dataset of the TSA model is acquired through time-domain simulation for a three-phase short circuit fault in transmission lines and buses, as summarized in Table 6. The fault duration is set using a uniform distribution ranging from 200ms to 600ms, and the simulation time is set between 5 and 10 seconds. Case-1 involves conducting fault simulations on all transmission lines and buses for various load scenarios on the IEEE 39-bus system, resulting in a sufficient training dataset. On the other hand, case-2 considers only bus faults on the IEEE 118-bus system, which is larger and more complex than case-1. Because of the limited and insufficient fault simulation on the larger benchmark system, case-2 presents a challenging dataset for neural network models. The case studies are composed of two types: one with sufficient training data (case-1), and the other with insufficient training data, referred to as case-2. Case-2 is used to verify the generalization ability of the TSA model within the limitations of the training dataset. In actual power systems, it is impossible to obtain all potential fault cases that may occur, so the capability of generalizing the characteristics of transient stability from the training dataset needs to be verified.

### C. PRETRAINED DEEP CONVOLUTIONAL NEURAL NETWORK SELECTION

In order to apply the deep CNN pre-trained with ImageNet to TSA, the PDCNNs (11) listed in Table 3 are fine-tuned

**TABLE 6.** Configuration of training dataset for the case study.

| Condition | Details | |
|---|---|---|
| Fault type | Three-phase short circuit fault | |
| Data location | All buses | |
| State variable | Voltage phasor | |
| Sampling rate | 60Hz | |
| Fault duration | 0.2–0.6s (uniform distribution) | |
| Case study | Case-1 (C1) | Case-2 (C2) |
| Benchmark system | IEEE 39-bus | IEEE 118-bus |
| Fault location | Line and bus | Bus |
| Fault case (stable/unstable) | 3,528 (2,801/727) | 2,476 (1,831/645) |
| Dataset split ratio (training/validation) | 70% (2,470/1,058) | 70% (1,733/743) |
| Remarks | Relatively small benchmark system and sufficient training dataset: easy case | Relatively large-scale benchmark system and insufficient training dataset: hard case |

**TABLE 7.** Optimization parameters for fine-tuning.

| Training option | Setting |
|---|---|
| Optimizer | SGDM[a] |
| $L_2$ regularization | 1e-4 |
| Batch normalization statistics | population |
| Gradient threshold | infinite |
| Gradient threshold method | $L_2$norm |
| Initial learning rate | 1e-4 (learnable layer)<br>1e-3 (fully connected layer) |
| Momentum | 0.9 |
| Maximum epochs | 40 |
| Mini batch size | 4; 8; 16; 32; 48; 64 |
| Frozen layer depth(%) | 10; 30; 50; 70; 90 |

[a]Stochastic gradient descent with momentum.

using the optimization parameter in Table 7 to compose the best-performing TSA system. The performance comparison is conducted through case studies by varying the depth of the frozen layer to select the most suitable model for TSA.

In Fig. 8(a), the overall good performance of DTL models is verified. Because the difference between generality and specificity learned by the neural network from the former task depends on the depth of the layer, performance variations are observed for an increase in the frozen layer. Fig. 8(b) represents a magnified figure expanding the dotted line box in Fig. 8(a). Each model demonstrates its best performance with different depths of the frozen layer due to the property of each model. Comparing the transient stability feature map (i.e., the DSI) and ImageNet from the perspective of Fig. 2, these datasets (see Figs. 1 and 9) are dissimilar visually and differ significantly in scale; thus, excellent performance is observed at the proportion of the frozen layer 10–50% due to the low similarity and relatively small training dataset. From the comparison of the accuracy of the trained PDCNN models shown in Fig. 8(c), it is evident that the VGG-19 based

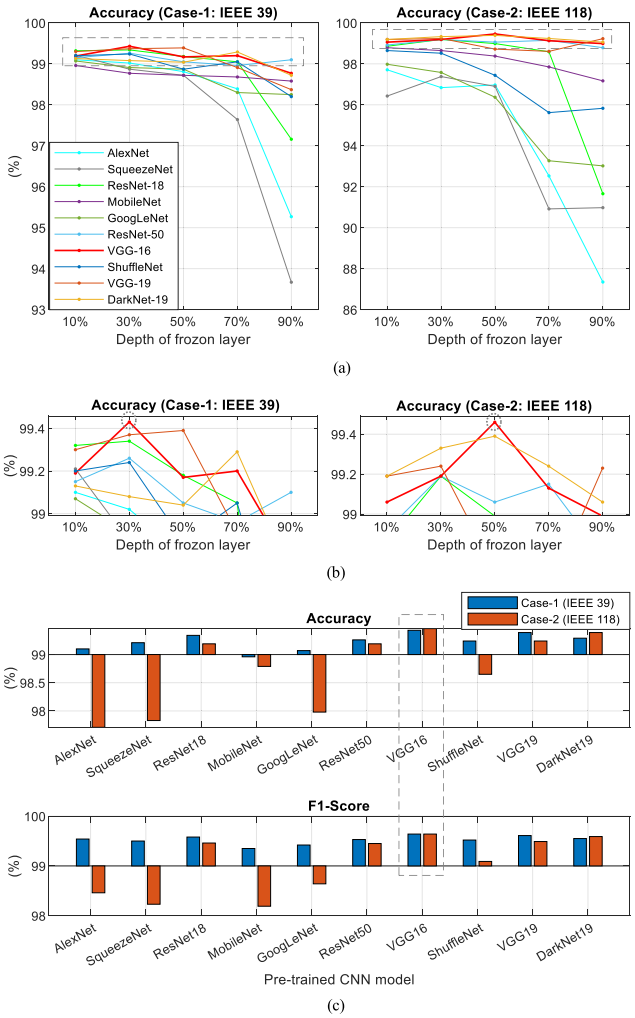


**FIGURE 8.** Comparison of the training result for selecting a pre-trained deep convolutional neural network. (a) Training results of the models depending on the depth of the frozen layer. (b) Training results (zoomed in the gray dotted box of (a)). (c) Best performances of each model at the optimal frozen layer depth.

model exhibits the best performance for both case-1 and case-2. Relatively deeper and more complex models exhibit rather disappointing classification results. This indicates that the training dataset used for TSA in this study is somewhat insufficient to train massive networks. As a result, relatively simple models among candidate networks reveal better performance. While most of the models show excellent performance, VGG-16, of note, proves the best classification competence. In the training process, not only the effect of the mini-batch size but also the depth of the frozen layer on the performance appears to have the slightest difference. As a result, in the case studies of this research, it has been confirmed that VGG-16 is the most suitable model for DTL in TSA, considering the size of the training dataset in this study. As summarized in Table 3, VGG-16 can determine transient stability from a feature map within 5.1ms, so it has performance sufficient for real-time application and is relatively small-scale compared to other PDCNNs, which

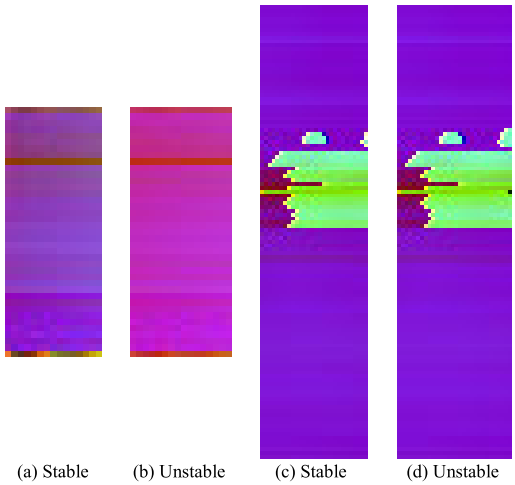


**FIGURE 9.** The proposed feature map and classification label. (a) Stable case sample: 265ms transmission line (#9bus to #39bus) short circuit fault on IEEE 39-bus model. (b) Unstable case sample: 280ms transmission line (#9bus to #39bus) short circuit fault on IEEE 39-bus model. (c) Stable case sample: 480ms #49bus short circuit fault on IEEE 118-bus model. (d) Unstable case sample: 500ms #49bus short circuit fault on IEEE 118-bus model.

has advantages in terms of saving training time and computing resources. Starting from the next section of this paper, VGG-16, which serves as the backbone of the assessment model, is referred to as the DTL model.

## VI. VERIFICATION

The performance comparison is conducted between the proposed DTL-based model and other approaches, including CNN-based models [31], [32], GINN-based model [33], LSTM-based model [30], and CNN-LSTM-based model [35]. From this point onward in this paper, these other approach models will be denoted as CNN-1, CNN-2, GINN, LSTM, and CLN model, respectively. To achieve objective performance comparison, TSA are constructed in an environment where all conditions for composing the model are exactly the same except for the neural network model. Due to the difference in the training datasets in the case studies, some modifications to the structures of the other approach models are necessary. The structural differences between the proposed model and other approach models can be observed in Table 8. The training is conducted with the same optimizer settings listed in Table 9, and the classification results are summarized as the average value of 10 training results.

### A. PERFORMANCE ANALYSIS

All models for comparing the performance assess transient stability using the state variable only during the fault period accompanied by the composition of the training dataset proposed in this study. The performance of the assessment model is evaluated using accuracy (AC), precision (PR), recall (RC), F1-score (F1S), false positive rate (FPR), and false discovery

**TABLE 8.** Comparison of the structure of assessment models.

| Proposed | Other approaches[a] | | |
|---|---|---|---|
| DTL model (VGG-16[b]) | CNN-1 model [31] | CNN-2 model [32] | GINN model [33] |
| Input[c] | | | |
| Conv[d]-64 | | | |
| Conv-64 | | | |
| Max pool[e] | Input[c] | Input[c] | Input[c] |
| Conv-128 | Conv-32 (64)[f] | Conv-64 | Incep[g]-16 |
| Conv-128 | Conv-16 (32) | Conv-16 | Conv-48 |
| Max pool | FC-32 | Conv-4 | Incep-32 |
| Conv-256 | | FC-64 (128)[f] | Conv-96 |
| Conv-256 | | | FC-128 |
| Conv-256 | | | |
| Max pool | LSTM model [30] | CLN model[h] [35] | |
| Conv-512 | | | |
| Conv-512 | | | |
| Conv-512 | | Input[c] | |
| Max pool | Input[c] | Conv-64 | |
| Conv-512 | LSTM-256 | Conv-16 | |
| Conv-512 | LSTM-128 | Max pool | |
| Conv-512 | FC-32 | LSTM-256 | |
| Max pool | | LSTM-128 | |
| FC-4096 | | FC-32 | |
| FC-4096 | | | |
| Output layer: FC-2 and softmax | | | |

[a]The models proposed in other approaches are trained using the case study dataset in this study; [b]Visual geometry group 16 layers; [c]All input source images are identical but resized to fit the network; [d]Convolution layer; [e]Max pooling; [f]The model structure is changed depending on each case study; [g]Inception block (i.e., CNN variant); [h]The CLN model is arbitrarily constructed as the detail of the network structure is not provided in [35].

rate (FDR) as the performance index (15) [40].

$$\begin{cases} AC = \dfrac{TP + TN}{TP + FP + TN + FN} \\ PR = \dfrac{TP}{TP + FP} \\ RC = \dfrac{TP}{TP + TN} \\ F1S = \dfrac{2 \times PR \times RC}{PR + RC} \\ FPR = \dfrac{FP}{FP + TN} \\ FDR = \dfrac{FP}{TP + FP} \end{cases} \tag{15}$$

As shown in Table 10, TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

A considerable difference in classification ability is confirmed in the performance comparison results (see Fig. 10), even though all conditions are the same except for the neural network model. In the proposed DTL model, the false negative and false positive cases are only 4.75 and 1.25 cases on the IEEE 39-bus system and 1 and 3 cases on the IEEE 118-bus system, respectively. The DTL model, moreover, shows clear superiority in F1-score, used to validate performance from the imbalanced training dataset, such as stable and unstable cases in power systems. The results of the two case studies can be seen in Fig. 11, where most models perform well in case-1. However, a clear performance difference is

**TABLE 9.** Optimization parameters for training comparison models.

| Parameter | Other approaches | | |
|---|---|---|---|
| | CNN-1; CNN-2; GINN model | LSTM model | CLN model |
| Optimizer | SGDM | Adam[a] | |
| Initial learning rate | 1e-2 | 1e-3 | |
| Momentum | 0.9 | N/A | |
| Squared gradient decay factor | N/A | 0.999 | |
| Gradient decay factor | N/A | 0.9 | |
| Epsilon | N/A | 1e−8 | |
| Gradient threshold | N/A | 1 | |
| Mini batch size | 16 | | |
| Maximum epochs | 40 | | |
| $L_2$ regularization | 1e-4 | | |
| Batch normalization statistics | population | | |

[a]Adaptive moment estimation.

**TABLE 10.** Confusion matrix.

| True class | Predicted class | |
|---|---|---|
| | Stable | Unstable |
| Stable | True positive (TP) | False negative (FN) |
| Unstable | False positive (FP) | True negative (TN) |

observed in case-2, where learning about transient stability characteristics is relatively difficult. These results confirm the ability of the proposed DTL model to learn and generalize transient stability in power systems. Because the proposed model has a deep layer structure, it can be inferred that if implemented in an actual power system, the deep neural network model can be upgraded continuously using the training dataset being accumulated from the actual operation data and simulation. Notably, through the procedure proposed in this study transfer learning of superior neural networks developed for other tasks enables the continuous improvement of the DTL model itself. The reason why the performance of other approach models is rather insufficient compared to the results in the original research is the difference in assessment timing between this study and the previous studies. In previous studies, the TSA models delay the response time, i.e., the decision timing, until identifying more apparent clues about transient stability, as shown in Table 1, so as to obtain an accuracy of 99% or higher. This is because, in general, the classification accuracy is upregulated as the response time is delayed owing to the physical characteristics of transient stability becoming apparent as time progresses after the fault clearing. However, in this study, the accuracy of other approaches diminishes as the assessment is set to be taken immediately after the fault clearing. The outstanding modeling capabilities of the deep CNN used in this study and the classification ability remaining in the network pre-learned with over 1 million ImageNet datasets are able to immediately and accurately assess transient stability only using the state variables during the fault period.

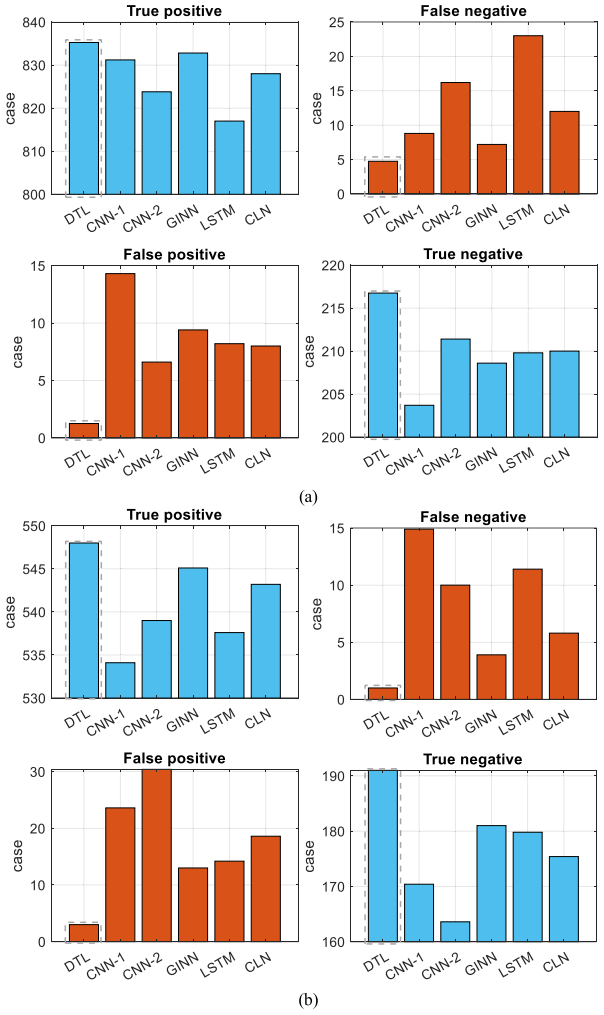


**FIGURE 10.** Comparison of the confusion matrix. (a) Case-1: IEEE 39-bus system. (b) Case-2: IEEE 118-bus system.

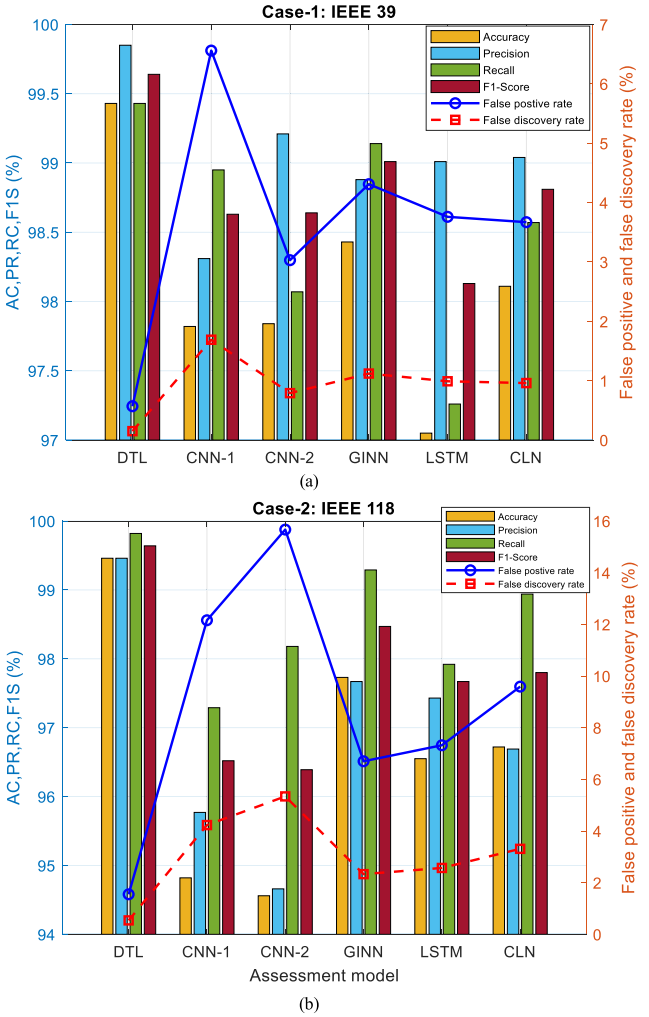


**FIGURE 11.** Comparison of the transient stability assessment performance. (a) Case-1: IEEE 39-bus model. (b) Case-2: IEEE 118-bus model.

### B. ROBUSTNESS AGAINST PMU DATA WITH NOISE

Considering the standard for the total vector error and noise filtering capability of the PMU specified in IEEE standard C37.118.1 and the field test results under interference conditions [55], [56], [57], it is identified that the PMU can be interfered to around 40–60dB signal-noise ratio (SNR). It is necessary to investigate the overall effect of the reported noise interference on TSA because the proposed model should be built from the state variable acquired from the PMU and be applied in a real-time assessment system. Additive white Gaussian noise (AWGN) is a widely used noise model in communication systems, where the noise is uniformly distributed across all frequency bands, and its amplitude follows a normal probability distribution. It is especially common in modeling satellite communication channels, such as PMU systems. By incorporating AWGN into TSA, the robustness to noise can be validated. Therefore, the power system state data (10) is reconstructed using an AWGN with SNR ranging from 40 to 60 dB to examine the impact of noise on the TSA system.

CNN converts a specific area of the input image into a representative value, thereby reducing the influence of slight variations in the input signal on the inference results [38]. In particular, deep CNNs have convolution layers connected in series and perform multiple convolution and pooling operations repeatedly in each layer, which can impart robustness against noise interference in the PMU. As identified in Fig. 12, the CNN-based models generally show good performance deviations against noise, while the LSTM-based model clearly exhibits performance degradation. The DTL model (the dotted gray box in Fig 12) demonstrates the most noise-resistant characteristics in line with the features of the deep network structure. Furthermore, the performance differences among case studies also exhibit the most marginal variances. Tables 11 and 12 indicate the effects of noise interference on performance.

### C. LIMITATION AND POTENTIAL CHALLENGE

The approach to constructing the TSA system through DTL, as proposed in this study, requires a substantial amount of

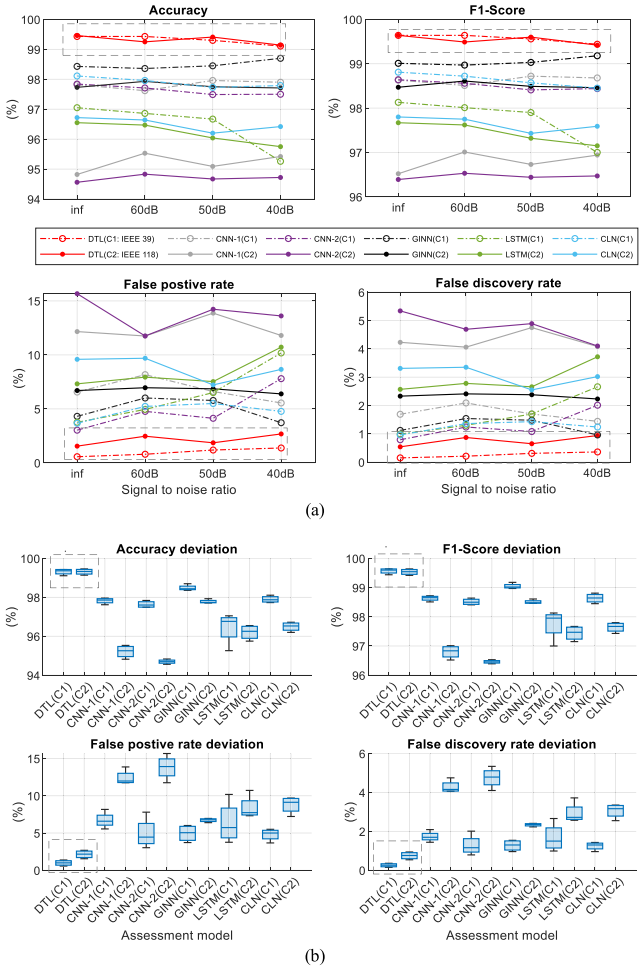


**FIGURE 12.** Comparison of the robustness against noise. (a) Performance variation of models depending on signal-to-noise ratio. (b) Distribution of performance with noise variations.

**TABLE 11.** Performance indexes with different noise levels (Case-1: IEEE 39-Bus).

| SNR | Index | Proposed DTL (%) | Other approaches (%) | | | | |
|---|---|---|---|---|---|---|---|
| | | | CNN-1 | CNN-2 | GINN | LSTM | CLN |
| 40dB | AC | 99.11 | 97.90 | 97.50 | 98.70 | 95.26 | 97.79 |
| | F1S | 99.44 | 98.68 | 98.44 | 99.18 | 97.01 | 98.45 |
| | FPR | 1.38 | 5.55 | 7.8 | 3.72 | 10.18 | 4.77 |
| | FDR | 0.36 | 1.44 | 2.01 | 0.96 | 2.66 | 1.24 |
| 50dB | AC | 99.30 | 97.96 | 97.49 | 98.45 | 96.67 | 97.73 |
| | F1S | 99.56 | 98.72 | 98.41 | 99.03 | 97.90 | 98.57 |
| | FPR | 1.19 | 6.61 | 4.13 | 5.78 | 6.51 | 5.5 |
| | FDR | 0.31 | 1.70 | 1.08 | 1.48 | 1.7 | 1.43 |
| 60dB | AC | 99.43 | 97.62 | 97.71 | 98.36 | 96.86 | 97.96 |
| | F1S | 99.64 | 98.51 | 98.56 | 98.97 | 98.01 | 98.72 |
| | FPR | 0.80 | 8.17 | 4.77 | 6.01 | 4.95 | 5.23 |
| | FDR | 0.21 | 2.09 | 1.24 | 1.54 | 1.3 | 1.36 |
| ∞[a] | AC | 99.43 | 97.82 | 97.84 | 98.43 | 97.05 | 98.11 |
| | F1S | 99.64 | 98.63 | 98.64 | 99.01 | 98.13 | 98.81 |
| | FPR | 0.57 | 6.56 | 3.03 | 4.31 | 3.76 | 3.67 |
| | FDR | 0.15 | 1.69 | 0.79 | 1.12 | 0.99 | 0.96 |

[a]Infinite signal-noise ratio.

**TABLE 12.** Performance indexes with different noise levels (Case-2: IEEE 118-Bus).

| SNR | Index | Proposed DTL (%) | Other approaches (%) | | | | |
|---|---|---|---|---|---|---|---|
| | | | CNN-1 | CNN-2 | GINN | LSTM | CLN |
| 40dB | AC | 99.14 | 95.42 | 94.72 | 97.71 | 95.75 | 96.42 |
| | F1S | 99.42 | 96.94 | 96.47 | 98.46 | 97.15 | 97.59 |
| | FPR | 2.68 | 11.80 | 13.61 | 6.39 | 10.72 | 8.66 |
| | FDR | 0.94 | 4.08 | 4.1 | 2.23 | 3.72 | 3.02 |
| 50dB | AC | 99.41 | 95.09 | 94.67 | 97.75 | 96.04 | 96.20 |
| | F1S | 99.60 | 96.73 | 96.44 | 98.49 | 97.32 | 97.43 |
| | FPR | 1.86 | 13.87 | 14.23 | 6.86 | 7.53 | 7.22 |
| | FDR | 0.65 | 4.75 | 4.89 | 2.38 | 2.66 | 2.55 |
| 60dB | AC | 99.25 | 95.53 | 94.83 | 97.93 | 96.47 | 96.64 |
| | F1S | 99.49 | 97.01 | 96.53 | 98.61 | 97.62 | 97.75 |
| | FPR | 2.47 | 11.75 | 11.75 | 6.96 | 7.94 | 9.69 |
| | FDR | 0.87 | 4.06 | 4.69 | 2.41 | 2.78 | 3.35 |
| ∞ | AC | 99.46 | 94.82 | 94.56 | 97.73 | 96.55 | 96.72 |
| | F1S | 99.64 | 96.52 | 96.39 | 98.47 | 97.67 | 97.80 |
| | FPR | 1.55 | 12.16 | 15.67 | 6.70 | 7.32 | 9.59 |
| | FDR | 0.54 | 4.23 | 5.34 | 2.33 | 2.57 | 3.31 |

training data, which is an inherent drawback of data-driven AI technologies. In actual power systems, acquiring a sufficient number of unstable cases for training is impossible. Hence, accurate modeling of the power system network, machines, and control systems becomes essential as a prerequisite, necessitating the time-consuming process of modeling and simulating the power system to obtain the necessary training dataset. Additionally, from the perspective of power system operation, defining the role of the TSA system in control and operation systems is crucial. The primary purpose of TSA is to enhance the stability of the power system through effective operation and control strategies. Therefore, validation and application in power system operation systems are necessary to ensure its practical implementation. In particular, for effective integration with remedial action schemes (or emergency control), the TSA system must be designed and evaluated considering various constraints, such as PMU data acquisition communication delay, remedial action communication delay, algorithm operation time, circuit breaker interrupting time, and the operation time of the remedial action decision algorithm. The roles and performance criteria required for TSA in these constrained conditions need to be clearly defined, and a seamless relationship between the TSA system and the remedial action scheme should be established.

## VII. CONCLUSION

In this paper, DTL method for composing the TSA system through the excellent learning performance and image classification ability of the deep CNN pre-trained with ImageNet is proposed. The proposed system exhibits high accuracy assessment at the fastest timing, which is challenging to achieve with existing methods. Despite utilizing only the state data during the fault period the PDCNN has effectively learned the characteristics of power system transient stability and thus can assess transient stability immediately after the fault clears. Furthermore, it is verified that DTL is a valuable method for overcoming the restriction of acquiring large-scale datasets and training deep neural networks. Future work will focus on developing the results of this study in connection with power system emergency control to improve power system stability.

## REFERENCES

[1] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. Van Cutsem, and V. Vittal, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1387–1401, Aug. 2004.

[2] N. Hatziargyriou, J. Milanovic, C. Rahmann, V. Ajjarapu, C. Canizares, I. Erlich, D. Hill, I. Hiskens, I. Kamwa, B. Pal, P. Pourbeik, J. Sanchez-Gasca, A. Stankovic, T. Van Cutsem, V. Vittal, and C. Vournas, "Definition and classification of power system stability—Revisited & extended," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3271–3281, Jul. 2021.

[3] A. A. Fouad and S. E. Stanton, "Transient stability of a multi-machine power system. Part II: Critical transient energy," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 7, pp. 3417–3424, Jul. 1981.

[4] A. A. Fouad, V. Vittal, and T. K. Oh, "Critical energy for direct transient stability assessment of a multimachine power system," *IEEE Trans. Power App. Syst.*, vol. PAS-103, no. 8, pp. 2199–2206, Aug. 1984.

[5] H.-D. Chiang, F. F. Wu, and P. P. Varaiya, "A BCU method for direct analysis of power system transient stability," *IEEE Trans. Power Syst.*, vol. 9, no. 3, pp. 1194–1208, Aug. 1994.

[6] J. H. Chow, A. Chakrabortty, M. Arcak, B. Bhargava, and A. Salazar, "Synchronized phasor data based energy function analysis of dominant power transfer paths in large power systems," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 727–734, May 2007.

[7] P. Bhui and N. Senroy, "Real-time prediction and control of transient stability using transient energy function," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 923–934, Mar. 2017.

[8] J. Machowski, Z. Lubosny, J. W. Bialek, and J. R. Bumby, *Power System Dynamics: Stability and Control*. Hoboken, NJ, USA: Wiley, 2020, pp. 253–262.

[9] J. Rasmussen and P. Jorgensen, "Synchronized phasor measurements of a power system event in Eastern Denmark," *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 278–284, Feb. 2006.

[10] N. I. A. Wahab, A. Mohamed, and A. Hussain, "Fast transient stability assessment of large power system using probabilistic neural network with feature reduction techniques," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 11112–11119, Sep. 2011.

[11] F. R. Gomez, A. D. Rajapakse, U. D. Annakkage, and I. T. Fernando, "Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1474–1483, Aug. 2011.

[12] D. Huang, X. Yang, S. Chen, and T. Meng, "Wide-area measurement system-based model-free approach of post-fault rotor angle trajectory prediction for on-line transient instability detection," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 10, pp. 2425–2435, May 2018.

[13] M. Rahmatian, Y. C. Chen, A. Palizban, A. Moshref, and W. G. Dunford, "Transient stability assessment via decision trees and multivariate adaptive regression splines," *Electr. Power Syst. Res.*, vol. 142, pp. 320–328, Jan. 2017.

[14] D. You, K. Wang, L. Ye, J. Wu, and R. Huang, "Transient stability assessment of power system using support vector machine with generator combinatorial trajectories inputs," *Int. J. Electr. Power Energy Syst.*, vol. 44, no. 1, pp. 318–325, Jan. 2013.

[15] B. Wang, B. Fang, Y. Wang, H. Liu, and Y. Liu, "Power system transient stability assessment based on big data and the core vector machine," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2561–2570, Sep. 2016.

[16] H. Hosseini, S. Naderi, and S. Afsharnia, "New approach to transient stability prediction of power systems in wide area measurement systems based on multiple-criteria decision making theory," *IET Gener., Transmiss. Distrib.*, vol. 13, no. 21, pp. 4960–4967, Nov. 2019.

[17] A. B. Mosavi, A. Amiri, and H. Hosseini, "A learning framework for size and type independent transient stability prediction of power system using twin convolutional support vector machine," *IEEE Access*, vol. 6, pp. 69937–69947, 2018.

[18] S. Ma, C. Chen, C. Liu, and Z. Shen, "A measurement-simulation hybrid method for transient stability assessment and control based on the deviation energy," *Int. J. Electr. Power Energy Syst.*, vol. 115, Feb. 2020, Art. no. 105422.

[19] S. Wei, M. Yang, J. Qi, J. Wang, S. Ma, and X. Han, "Model-free MLE estimation for online rotor angle stability assessment with PMU data," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2463–2476, May 2018.

[20] J. Lv, "Transient stability assessment in large-scale power systems based on the sparse single index model," *Electric Power Syst. Res.*, vol. 184, Jul. 2020, Art. no. 106291.

[21] S. M. Ashraf and S. Chakrabarti, "A single machine equivalent-based approach for online tracking of power system transient stability," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 1688–1696, May 2021.

[22] L. Zhu and D. J. Hill, "Networked time series shapelet learning for power system transient stability assessment," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 416–428, Jan. 2022.

[23] R. Zhang, Y. Xu, Z. Y. Dong, and K. P. Wong, "Post-disturbance transient stability assessment of power systems by a self-adaptive intelligent system," *IET Gener., Transmiss. Distrib.*, vol. 9, no. 3, pp. 296–305, Feb. 2015.

[24] M. Mahdi and V. M. I. Genc, "Post-fault prediction of transient instabilities using stacked sparse autoencoder," *Electr. Power Syst. Res.*, vol. 164, pp. 243–252, Nov. 2018.

[25] S. Jafarzadeh and V. M. I. Genc, "Real-time transient stability prediction of power systems based on the energy of signals obtained from PMUs," *Electr. Power Syst. Res.*, vol. 192, Mar. 2021, Art. no. 107005.

[26] S. Wu, L. Zheng, W. Hu, R. Yu, and B. Liu, "Improved deep belief network and model interpretation method for power system transient stability assessment," *J. Modern Power Syst. Clean Energy*, vol. 8, no. 1, pp. 27–37, Jan. 2020.

[27] B. P. Soni, A. Saxena, V. Gupta, and S. L. Surana, "Identification of generator criticality and transient instability by supervising real-time rotor angle trajectories employing RBFNN," *ISA Trans.*, vol. 83, pp. 66–88, Dec. 2018.

[28] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[29] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[30] J. J. Q. Yu, D. J. Hill, A. Y. S. Lam, J. Gu, and V. O. K. Li, "Intelligent time-adaptive transient stability assessment system," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1049–1058, Jan. 2018.

[31] A. Gupta, G. Gurrala, and P. S. Sastry, "An online power system stability monitoring system using convolutional neural networks," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 864–872, Mar. 2019.

[32] Z. Shi, W. Yao, L. Zeng, J. Wen, J. Fang, X. Ai, and J. Wen, "Convolutional neural network-based power system transient stability assessment and instability mode prediction," *Appl. Energy*, vol. 263, Apr. 2020, Art. no. 114586.

[33] S. K. Azman, Y. J. Isbeih, M. S. E. Moursi, and K. Elbassioni, "A unified online deep learning prediction model for small signal and transient stability," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4585–4598, Nov. 2020.

[34] J. Huang, L. Guan, Y. Su, H. Yao, M. Guo, and Z. Zhong, "Recurrent graph convolutional network-based multi-task transient stability assessment framework in power system," *IEEE Access*, vol. 8, pp. 93283–93296, 2020.

[35] J. Xie and W. Sun, "A transfer and deep learning-based method for online frequency stability assessment and control," *IEEE Access*, vol. 9, pp. 75712–75721, 2021.

[36] J. Kim, H. Lee, S. Kim, and J. H. Park, "Real-time power system transient stability prediction using convolutional layer and long short-term memory," *J. Electr. Eng. Technol.*, vol. 18, no. 4, pp. 2723–2735, Jul. 2023.

[37] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.

[38] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[39] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. 36th Int. Conf. Mach. Learn. (PMLR)*, 2019, pp. 6105–6114.

[40] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.

[41] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. 13th Eur. Conf. Comput. Vis. (ECCV)*, Zurich, Switzerland, Aug. 2014, pp. 818–833.

[42] A. Mahendran and A. Vedaldi, "Understanding deep image representations by inverting them," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 5188–5196.

[43] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
[44] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
[45] 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.
[46] M. Iman, H. R. Arabnia, and K. Rasheed, "A review of deep transfer learning and recent advancements," *Technologies*, vol. 11, no. 2, p. 40, Mar. 2023.
[47] H. Cui, Q. Wang, Y. Ye, Y. Tang, and Z. Lin, "A combinational transfer learning framework for online transient stability prediction," *Sustain. Energy, Grids Netw.*, vol. 30, Jun. 2022, Art. no. 100674.
[48] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1–9.
[49] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[50] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[51] M. Shaha and M. Pawar, "Transfer learning for image classification," in *Proc. 2nd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, India, Mar. 2018, pp. 656–660.
[52] B. Neyshabur, H. Sedghi, and C. Zhang, "What is being transferred in transfer learning?" in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 512–523.
[53] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 1–9.
[54] S. Bianco, R. Cadene, L. Celona, and P. Napoletano, "Benchmark analysis of representative deep neural network architectures," *IEEE Access*, vol. 6, pp. 64270–64277, 2018.
[55] *IEEE Standard for Synchrophasor Data Transfer for Power Systems*, IEEE Standard C37.118.2-2011, Revision IEEE Standard C37.118-2005, 2011, pp. 1–53.
[56] *Amendment 1: Modification of Selected Performance Requirements*, IEEE Standard C37.118.1a-2014, 2014.
[57] R. Ghiga, K. Martin, Q. Wu, and A. H. Nielsen, "Phasor measurement unit test under interference conditions," *IEEE Trans. Power Del.*, vol. 33, no. 2, pp. 630–639, Apr. 2018.

**JONGJU KIM** (Member, IEEE) received the B.S. degree in electrical engineering from Konkuk University, Seoul, South Korea, in 2009, and the M.S. and Ph.D. degrees in electrical engineering from Pusan National University, Busan, South Korea, in 2020 and 2023, respectively. Since 2009, he has been with Korea Southern Power Company Ltd. (KOSPO), Busan. He is currently a Registered Professional Engineer Generation Transmission and Distribution in South Korea. His research interests include power system control, protection, stability, and AI applications to power systems.

**HEUNGSEOK LEE** received the B.S. degree in electrical engineering from Ulsan University, Ulsan, South Korea, in 2012, and the M.S. degree in electrical engineering from Pusan National University, Busan, South Korea, in 2014, where he is currently pursuing the Ph.D. degree in electrical engineering. His research interests include power system stability, security, and deep learning applications to power systems.

**SUNGSHIN KIM** (Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from Yonsei University, Seoul, South Korea, in 1984 and 1986, respectively, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 1996. He is currently a Professor with the Department of Electrical Engineering, Pusan National University, Busan, South Korea. His research interests include intelligent control, intelligent robot, hierarchical learning structures, and fault diagnosis and prognosis.

**SANG-HWA CHUNG** (Member, IEEE) received the B.S. degree in electrical engineering from Seoul National University, Seoul, South Korea, in 1985, the M.S. degree in computer engineering from Iowa State University, Ames, IA, USA, in 1988, and the Ph.D. degree in computer engineering from the University of Southern California, Los Angeles, CA, USA, in 1993. From 1993 to 1994, he was an Assistant Professor with the Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL, USA. He is currently a Professor with the Computer Engineering Department, PNU, Busan. Since 2016, he has been the Director of the Dong-Nam Grand ICT Research Center. He has authored over 270 articles and holds 90 patents. His research interests include embedded systems, wireless networks, software-defined networking, and smart factories.

**JUNE HO PARK** (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Seoul National University, Seoul, South Korea, in 1978, 1980, and 1987, respectively. He is currently an emeritus Professor with the Department of Electrical Engineering, Pusan National University, Busan, South Korea. His research interests include power system control, operation, and AI applications to power systems. He has been a member of the IEEE Power Engineering Society.