A Novel Method for Missing Data Reconstruction in Smart Grid Using Generative Adversarial Networks

Jiashu Fang, Le Zheng, and Chongru Liu, *Senior Member, IEEE*

***Abstract*—Existing machine-learning research on power grids relies on online measurements without missing data. We propose a missing data reconstruction model based on generative adversarial networks to supplement existing methods. This model fits better spatio-temporal data with several improvements over previous approaches. First, the loss function considers both distribution and value differences, leveraging all available information to minimize differences between original and generated data. Then, a deep-learning architecture incorporating convolutional neural layers and nonlocal blocks is developed to extract the spatial-temporal information in electrical feature maps. The proposed method exhibits enhanced credibility by neglecting invalid consecutive data under phasor measurement unit (PMU) failures (proven by attention maps generated in nonlocal blocks), and higher accuracy than existing models for recovering data under random data missing/PMU failure conditions (proven by numerical results). Finally, the proposed data reconstruction model is effectively applied to an online framework for transient stability assessment.**

***Index Terms*—Generative adversarial networks (GANs), missing data reconstruction, nonlocal blocks, power system.**

I. Introduction

WITH the rapid growth of Internet of Things technologies and smart sensor networks, the modern power system evolves to become more observable and analyzable [1]. Various tasks in power system operation, such as transient stability assessment (TSA) [2] and fault detection [3], have been addressed using data-driven methods. However, missing data are widely witnessed in wide-area measurement systems due to the inevitability of data collection unit malfunction, stochastic packet drops, and communication failure [4], [5]. Specifically, the anomalous/missing data account for approximately 10%–30% of phasor measurement unit (PMU) measurements during the daily operation in China's grids [4], which would significantly degrade the online performance of data-driven methods in the grid [2]. Therefore, in order to maintain the safe operation of modern power grids, there is an urgent need to mitigate the adverse effects of incomplete field data when data-driven methods are applied online.

The integration of ensemble algorithms is the most commonly seen solution to minimize the detrimental effect of incomplete data on data-driven algorithms. Taking the TSA task as an example, the authors in [6] and [7] proposed the ensemble models with different available PMU scenarios to guarantee online performance with online measurements in the presence of missing data, and Tan et al. [2] further developed PMU clusters considering the feature importance. Although the ensemble algorithms make the data-driven methods more tolerable to data loss compared with single-structure models, they undergo high computational burdens, and the limited PMU clusters may fail to cope with widespread missing PMU data. Meanwhile, numerous data-driven models with single structures have been proposed for solving various tasks in power systems. These models are typically built based on a database generated from time-domain simulations that describe the possible operation points in a manner consistent with the actual power system [8], [9], motivating the use of missing data reconstruction methods prior to the existing methods as an alternative to modifying existing algorithms into ensemble structures.

As for missing data reconstruction approaches, the authors in [10] and [11] recovered PMU data loss using advanced matrix optimization techniques based on the low-rank characteristics of spatio-temporal matrices. In [12], missing data were reconstructed via the alternating direction method of multipliers, addressing the completion problem in PMU matrices. However, the performance of these methods is limited as the minimal requirement for obtained measurements may not be satisfied in practice [10]. Besides, the gated recurrent unit (GRU) [13] and the long short-term memory (LSTM) [14] demonstrate high performance to predict the missing data; however, these approaches rely on the former sequence information, which may be unavailable in practical scenarios.

In order to eliminate the abovementioned issues, vanilla generative adversarial networks (GANs) [15], [16] are trained to

Manuscript received 22 November 2022; revised 5 April 2023 and 23 May 2023; accepted 7 August 2023. Date of publication 25 October 2023; date of current version 23 February 2024. This work was supported by National Key R&D Program of China under Grant 2022YFB2402704. Paper no. TII-22-4792. *(Corresponding author: Chongru Liu.)*

The authors are with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China (e-mail: jiashu.fang@hotmail.com; zhengl20@ncepu.edu.cn; chongru.liu@ncepu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2023.3306366.

Digital Object Identifier 10.1109/TII.2023.3306366

reconstruct the missing data by reducing the distribution error between generated and true data. Nonetheless, the vanilla GAN is built based on the multilayer perceptron (MLP) structure, which exhibits limitations when applied to power systems, as the simple MLP layers are less effective in extracting the spatial-temporal electrical features. Several modified GAN-based models are proposed to enhance the ability to learn the distribution of samples [17], [18], [19], [20], [21]. Specifically, a GAN search algorithm is proposed in [17] to automatically discover the optimal GAN architecture for effectively learning the underlying data distribution. Yan et al. [18] introduced a semantics-guided contrastive network, which improves the transferability of GAN-based methods in handling unseen scenarios by leveraging its understanding of the data distribution. Self-attention GAN (SAGAN) [19] introduced the attention mechanism for long-range dependency modeling in image generation tasks, and Ouyang et al. [20] proposed modification for Wasserstein GAN (WGAN) resulting in accelerated training and higher performance; the objective of these works is to generate plausible fake samples that are hard to distinguish from the real samples. However, the applicability of the aforementioned works is limited for data recovery tasks in the following aspects. First, the value differences between the generated and original data are not effectively exploited during the model training process. Second, a common assumption in these works is that a desirable model can be learned from a database where each measurement in the feature map is available; however, this assumption is hardly satisfied if the historical PMU measurements are used as the training set due to the device malfunctions and communication failures [22]. Third, since the feature maps in TSA tasks indicate the dynamic response of the power system where each pixel represents the electrical measurements, determining the authenticity of the entire sample, as is typical in image processing research, may not be sufficient to effectively extract the spatio-temporal correlations that exist between each pixel in the feature maps. Finally, the black-box characteristics of previous deep learning methods [13], [14], [15], [16], [20] limit the visualization of model parameters, leading to concerns from grid operators about the practical performance of intelligent algorithms.

Considering the existing research gap, in this work, we developed a GAN-based model for PMU missing data recovery, where more finer gained feedback is delivered compared with the vanilla GAN, as the proposed discriminator learns to predict the reality of each pixel within samples rather than to predict the reality of the entire sample. The architecture of the proposed method utilizes convolutional neural networks (CNN), the nonlocal block [23], and batch normalization [24]. Specifically, the spatio-temporal data in electrical power systems share similarities in the data structure with images, motivating us to utilize the CNN, a well-known deep learning method for image analysis, to extract characteristics of the electrical data with spatial and temporal dependencies. The sparse interactions and parameter sharing in convolutional operations allow for the reduction of free parameters learned, thereby mitigating the curse of dimensionality and overfitting challenges posed by fully connected networks [25]; the attention mechanism in nonlocal blocks introduces effective multihop dependency modeling abilities to the proposed algorithm, enabling the model to automatically neglect consecutive invalid data due to PMU failures; batch normalization accelerates network training by reducing internal covariate shift. Experiments on the IEEE-39 system have been used to prove the higher performance of the proposed method than the existing approaches under both random data missing (see Table V) and PMU failure (see Table VI) conditions, which can be ascribed by both its model structure and the training optimization process. Furthermore, this article proposes an online TSA framework to demonstrate the effectiveness of the proposed model for practical applications in power grids.

In addition to TSA tasks, the data-driven methods based on PMU measurements are widely employed in various power system analysis tasks, such as disturbance identification [26] and voltage stability assessment [27], which also face challenges due to incomplete PMU online measurements. Therefore, the proposed approach, which serves as a precursory online data recovery process, can be extended to these tasks by modifying the database parameters to meet the requirements of specific tasks, such as the range of possible operation points and the construction of feature maps, while the proposed loss function and basic model structure can be retained to extract electrical spatio-temporal information for different tasks. The main contributions of this article include the following.

1) A GAN-based model structure is designed to capture multihop dependency in feature maps; the proposed discriminator is to distinguish the authenticity of each pixel in feature maps, effectively extracting the spatiotemporal relationship in the power system.
2) In addition to the commonly used distribution differences between the generated and real data [28], the proposed loss function introduces the value differences, and allows fully leveraging existing information in feature maps compared with existing methods [29], leading to higher reconstruction accuracy.
3) The complete raw training database (where all measurements in a feature map are available) is no longer mandatory, thus the PMU historical measurements can be directly used as the training database, making the proposed method feasible in real-world applications.
4) The attention map in the proposed algorithm visualizes the model parameters when reconstructing the electrical information. This not only proves the model can adaptively neglect invalid data, but also enhances the credibility of the data-driven method in practice.

The rest of this article is organized as follows. The proposed GAN algorithm is elaborated in Section II. In Section III, we analyze the proposed model in terms of accuracy and attention map visualization, and compare its performance versus existing approaches. The model robustness against field noise is also demonstrated in a large-scale power system. Section IV proposes an online TSA framework where the proposed missing data reconstruction model acts as a predecessor process to the classification algorithm. Finally, Section V concludes this article.

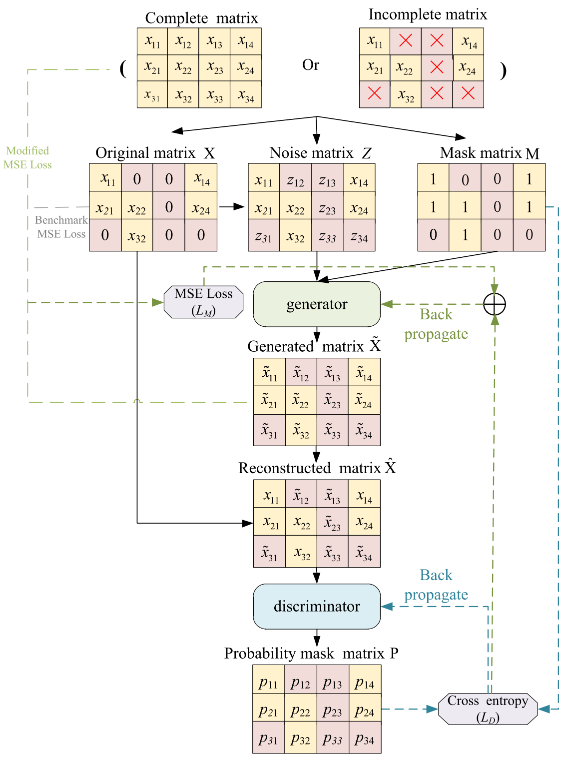

Fig. 1. Training process of the proposed GAN model. The training database has different levels of completeness depending on the source (e.g., it can be either incomplete historical PMU records with large amounts of missing data, or complete raw data generated by simulation [22]).

II. Proposed Data Reconstruction Approach

In this section, we propose a GAN-based model to reconstruct the missing data in power systems. Compared with the benchmark model, [29], the proposed method is more advanced in terms of both model structure and optimization process, making it more applicable to power systems. Moreover, attention maps are generated to visualize the model internal parameters under the PMU failure scenarios.

A. Proposed Model Overview

A GAN-based algorithm is constructed with two deep neural networks, denoted as the generator and the discriminator, to learn spatio-temporal distributions. Specifically, the generator is trained to produce data samples aligned with the distributions of the original data, whereas the discriminator distinguishes the generated data from the original data. The overall architecture of the proposed model is shown in Fig. 1.

1) Generator: Unlike the vanilla GAN [15], which learns the sample distribution using random noise vectors as input, the generator proposed in this article, G, takes two inputs, namely, the noise matrix $\mathbf{Z}$ and the mask matrix $\mathbf{M}$. The mask matrix $\mathbf{M}$ is predetermined according to the missing data within the feature maps. $\mathbf{Z}$ is a prior defined on input noise variables to learn the generator's distribution over $\mathbf{X}$ [15]. The corresponding mask $\mathbf{M} \in \{0,1\}^{d\times T}$ indicates which entries in $\mathbf{X}$ are observed, and its entry $m_i = 1$ if x_i is observed. It should be noted that when the training database is a complete matrix where each measurement in a feature map is known, the original matrix containing missing values fed to the generator is created based on a mask matrix with an artificially set percentage of missing values. The generator would output $\widetilde{\mathbf{X}}$ that has the same dimension as the original matrix $\mathbf{X}$, which includes data in the observed position

$$\widetilde{\mathbf{X}} = G(\mathbf{Z}, \mathbf{M}). \tag{1}$$

2) Discriminator: The discriminator, D, takes mask matrix $\mathbf{M}$ and reconstructed matrix $\hat{\mathbf{X}}$ as input. The reconstructed matrix is a combination of the generated data in the missing data positions and the originally observed data, as shown in the following, where $\odot$ denotes the elementwise multiplication:

$$\hat{\mathbf{X}} = \mathbf{M} \odot \mathbf{X} + (1 - \mathbf{M}) \odot \widetilde{\mathbf{X}}. \tag{2}$$

The objective of the discriminator is to distinguish the observed (true) data from the generated (fake) data, that is, to correctly predict $\mathbf{M}$. The discriminator can be seen as a function D where the ith component of the function output $D(\hat{\mathbf{X}})$ corresponds to the probability that the ith component is predicted as observed data. In contrast, the discriminator in the vanilla GAN is designed to predict whether the output of the generator is completely real or fake.

B. Optimization Method of Proposed Model

The loss functions in the generator and discriminator are designed to optimize neural networks' weights. The overall objective of the GAN can be seen as the minimax problem in terms of distribution difference given by

$$\min_G \max_D V(D, G). \tag{3}$$

The value function $V(D, G)$ writes

$$\begin{aligned} V(D,G) = \mathbb{E}_{\hat{\mathbf{X}},\mathbf{M}} \Big[& \mathbf{M}^T \log D(\hat{\mathbf{X}}, \mathbf{M}) \\ & + (\mathbf{1} - \mathbf{M})^T \log(\mathbf{1} - D(\hat{\mathbf{X}}, \mathbf{M})) \Big]. \end{aligned} \tag{4}$$

The proposed model first optimizes the discriminator with fixed generator parameters and then updates the generator with fixed discriminator parameters.

1) Loss Function of Discriminator: The loss function of D is a binary cross entropy function given by

$$\begin{aligned} & \mathcal{L}_D(\mathbf{m}_i, \mathbf{p}_i) \\ & = -\sum_i \left[m_i \log(p_i) + (1 - m_i) \log(1 - p_i) \right] \end{aligned} \tag{5}$$

where m_i is the mask label, and p_i is the probability of mask prediction given by

$$p_i = D(\hat{x}_i, m_i). \tag{6}$$

2) Loss Function of Generator: The loss function of G is divided into two parts considering both the distribution and the value differences. The first loss $\mathcal{L}_G$ evaluates the generated data distribution, thereby encouraging the generator to produce data that closely resemble the real data. Specifically, the use of the logarithmic function in (8) ensures that the generator is heavily penalized for generating highly unrealistic data in unobservable positions, as the logarithm amplifies the weight of less probable events. Besides, the second loss $\mathcal{L}_M$ evaluates the value differences between reconstructed data and original data. Overall, the generator is trained to minimize the weighted sum of the two losses as follows:

$$\mathcal{L}_G(\mathbf{m}_i, \mathbf{p}_i) + \alpha \mathcal{L}_M(\mathbf{x}_i, \hat{\mathbf{x}}_i) \tag{7}$$

where α is a hyperparameter (set to 0.3 in this article); the selection of α is based on cross-validation to achieve high reconstruction accuracy of the missing data, and

$$\mathcal{L}_G(\mathbf{m}_i, \mathbf{p}_i) = -\sum_i (1 - m_i) \log (p_i). \tag{8}$$

The loss function $\mathcal{L}_M$ proposed in this article evaluates the value differences of all existing data, thus maximizing the use of existing knowledge. Specifically, the proposed algorithm (see Fig. 1) can be applied in both situations where complete or incomplete raw datasets are used for the model training process. This gives

$$\mathcal{L}_M (x_i, x_i') = \sum_i (x_i' - x_i)^2 \tag{9}$$

where x_i' and x_i are the predicted and true values of $\widetilde{\mathbf{X}}$, respectively.

In contrast, the loss function $\mathcal{L}_M$ in the benchmark model [29] is

$$\mathcal{L}_M (x_i, x_i') = \sum_i (x_i^{m\prime} - x_i^m)^2 \tag{10}$$

where $x_i^{m\prime}$ and x_i^m are the predicted and true values in the positions where m_i equal to 1, respectively.

As previously discussed, when the training database without any missing data is available, a generated data matrix with missing values is fed to the generator. In such a case, the loss function in (10) does not consider the value differences between the generated data and the true data in the missing data positions during the training process, limiting the model reconstruction ability due to the partial use of the available data.

C. Attention Maps Revealing Model Weights

The attention mechanism is introduced to the proposed algorithm for twofold reasons: on the one hand, attention maps, generated with nonlocal blocks by visualizing model weights, provide an intuitive means to interpret the model behavior and prove the model feasibility in power systems; on the other hand, this mechanism effectively enhances the multihop dependency modeling abilities [23] by neglecting the invalid data, considering that the absence of the whole temporal sequence of failed PMUs (commonly witnessed in practice) more severely hampers the electrical feature extraction with models than the randomly missed data. Especially, the nonlocal block captures long-range dependency by interacting electrical temporal features between any two PMUs in different buses within the power system. With reference to the diagram in Fig. 2, the nonlocal operation can be defined as

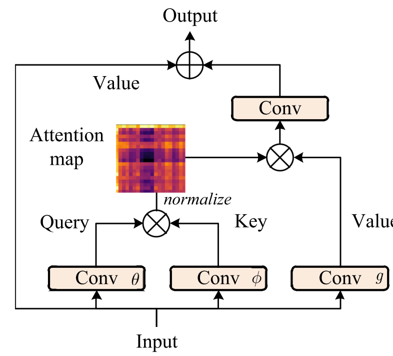

Fig. 2. Nonlocal block for 1-D feature maps. The attention map is generated by the dot-product similarity of the two embedding matrices of the input feature map. "$\otimes$" and "$\oplus$," respectively, represent the matrix multiplication and elementwise sum.

$$\mathbf{y}_i = \frac{1}{d} \sum_{j=1}^{d} f(\mathbf{x}_i, \mathbf{x}_j) g(\mathbf{x}_j) \tag{11}$$

where d denotes the position number (i.e., the number of PMUs located in different buses within the power system) in the feature map, and the corresponding feature maps $\mathbf{X} \in \mathbb{R}^{d \times T}$ indicate measurements covering T moments from various PMUs located in d buses of the power system. i and j of $\mathbf{x}$ and $\mathbf{y}$ are indexes of temporal sequence collected from different PMUs, $g(\mathbf{x}_j) = \mathbf{W}_g \mathbf{x}_j$ and $\mathbf{W}_g$ is a weight matrix, and $\mathbf{y}$ is the output feature map. The pairwise function f, which takes feature maps of two positions as input and produces a scalar, is expressed as

$$f(\mathbf{x}_i, \mathbf{x}_j, H) = \theta(\mathbf{x}_i)^{\mathrm{H}} \phi(\mathbf{x}_j) \tag{12}$$

where $\theta(\mathbf{x}_i) = \mathbf{W}_\theta \mathbf{x}_i$, $\phi(\mathbf{x}_j) = \mathbf{W}_\phi \mathbf{x}_j$, $\mathbf{W}_\theta \in \mathbb{R}^{C' \times C}$, and $\mathbf{W}_\phi \in \mathbb{R}^{C' \times C}$ are learnable weight matrices, and C and C' are the numbers of input and output channels. For simplicity, θ, ϕ, and g are considered in the form of a linear embedding, and are implemented as 1×1 convolutions [23].

$\theta(\mathbf{x}_i)$, $\phi(\mathbf{x}_j)$, and $g(\mathbf{x}_j)$ are named as the *query* of position i, the *key* of position j, and the *value* of position j, respectively. For a given position i, the inner product of its query $\theta(\mathbf{x}_i)$ and the key of each position is calculated to produce the attention distribution. Specifically, normalizing $f(\mathbf{x}_i, \mathbf{x}_j)$ with $\frac{1}{H}$ gives the attention score assigned to $g(\mathbf{x}_j)$ for position i.

The overall nonlocal block can be expressed as

$$\mathbf{z}_i = \mathbf{W}_z \mathbf{y}_i + \mathbf{x}_i \tag{13}$$

where $\mathbf{z}_i$ is the output of the block and $\mathbf{W}_z$ transforms $\mathbf{y}_i$, which is consistent with the dimension of the input feature map. The residual operation ($+\mathbf{x}_i$) ensures that local information can also be found in the nonlocal block output.

III. Results of Proposed Method

In this section, the performance of the proposed model is discussed with numerical experiments. Section III-A introduces the database utilized in this article; Section III-B demonstrates model accuracy and discusses the effect of each modified block in the algorithm; Section III-C visualizes the attention maps to reveal the model parameters in case of PMU failures; Section III-D compares the proposed approach with existing approaches; finally, Section III-E evaluates model robustness against field noise in a large-scale power system.

A. Database Generation

Databases are generated by considering different operation conditions and contingencies based on the 50 Hz IEEE-39 power system. A feature map $\mathbf{X} \in \mathbb{R}^{d\times T}$ is utilized to reflect the spatio-temporal dynamics of the power system, representing voltage phase angles observed at various PMUs located in d buses. In this study, we set T to 8 and d to 20, corresponding to the voltage phase angles recorded at 20 buses over eight moments starting from the moment of fault clearance [2]. The voltage phase angles of buses can be directly collected by the PMUs in real-world power grids [30]. A total of 12 000 scenarios were generated, with 4930 scenarios classified as unstable. The dataset was randomly divided into training and testing sets, with an 8:2 ratio.

To ensure accuracy, transient stability simulations were conducted using Power System Toolbox 3.0 [31], which generates the proposed database following an accurate time-domain simulation method. Time-domain simulation solves the differential–algebraic equations that describe power system dynamics under various disturbances while obeying system constraints, and bus phase angles are calculated by solving the nonlinear algebraic network equations. Indeed, the high precision of the power system simulation via the time-domain method has been widely proven in literature [32], [33], [34], [35], making it a reliable method for generating a database that is applicable for TSA tasks in real-world grids.

Existing data-driven TSA classifiers that use incomplete online measurements often experience performance degradation due to missing PMU data [22]. The proposed method is motivated to be used as a precursory data processing process for the TSA model, therefore the sample space of the proposed GAN-based method is aligned with that of the existing TSA classifier model in terms of the operation range of the power system and the baseline operating point. To ensure a significant difference between operating points within the database, the diversity of operation points and various fault conditions are taken into account. Specifically, the loads in power systems are randomly determined with the generation gradually declining from 120% to 80% of the basic operating level [22], with the bus voltage being maintained within 0.95–1.05 p.u. In addition, three-phase-to-ground short-circuit faults are applied to three locations (25%, 50%, and 70% of the length) [36] on four different transmission lines to simulate various fault conditions.

The purpose of the proposed method is to recover the missing data in an incomplete feature map. To enhance the practical applicability of the proposed method, PMU failures have been taken into consideration compared with previous methods that only address randomly missing PMU data [13], [16]. Furthermore, the PMU data used in this study exhibit characteristics, including sampling rate and missing rate, that are consistent with real-world scenarios. Specifically, the sampling rate of the PMU is set to twice per cycle [30]; the data missing rate and PMU failure rate are set to 20% and 10% in the training dataset, respectively, due to the inevitable PMU malfunction and communication loss [4].

TABLE I
DETAILS OF THE GENERATOR

Layer	Description
Conv_1 Non-local_1	[32 filters, size=3, ReLU] $\times$ 2 C'=32
ResBlock_2 Non-local_2	[64 filters, size=3, ReLU] $\times$ 2 C'=64
FC_1 Batch normalization	Dropout=0.5 Sigmoid

TABLE II
DETAILS OF THE DISCRIMINATOR

Layer	Description
Conv_1 Non-local_1	[32 filters, size=3, ReLU] $\times$ 2 C'=32
FC_1 Batch normalization FC_2	Dropout=0.5 softmax

B. Performance of Proposed Model

The model performance is given by mean absolute percentage error (MAPE) and mean square percentage error (MSPE) as

$$\text{MAPE} = \frac{1}{N}\cdot\sum_{i=1}^{N}\left|\frac{\dot{\mathrm{x}}_i - \mathrm{x}_i}{\mathrm{x}_i}\right| \tag{14}$$

$$\text{MSPE} = \frac{1}{N}\cdot\sum_{i=1}^{N}\left(\frac{\dot{\mathrm{x}}_i - \mathrm{x}_i}{\mathrm{x}_i}\right)^2 \tag{15}$$

where N is the total number of reconstructed data in the missing data positions, and $\dot{x}_i$ and x_i are, respectively, the reconstructed and true values.

Details of the proposed generator and discriminator are given in Tables I and II. Several components, including the CNN, the nonlocal block, and batch normalization, are utilized to modify the benchmark model (implemented with two fully connected layers in the generator and discriminator) [29]. In order to investigate the contribution of model components on the overall model performance, the accuracy of different model structures is obtained via the removal of specific components, as given in Table III, where 20% of the data are randomly missing in the test dataset. It is seen that each component enhances the model performance. Specifically, convolutional operations are critical for enhancing the model's ability to extract pertinent electrical data. This is due to the adaptivity of the convolutional kernels, which appropriately balance the contributions of each piece of information within a local receptive field. As a result, the model can proficiently capture the correlations within the electrical data. Besides, the combination of both nonlocal blocks and

TABLE III
PERFORMANCE OF THE PROPOSED MODEL WITH DIFFERENT COMPONENTS

Conv	ResBlock	Non-local	Batch normalization	MAPE (%)	MSPE (%)
✓				3.05	0.38
✓	✓			2.85	0.30
✓	✓	✓		2.23	0.26
	✓	✓	✓	2.18	0.16
✓	✓		✓	2.10	0.15
✓	✓	✓	✓	1.85	0.07

TABLE IV
MODEL PERFORMANCE UNDER DIFFERENT DATA MISSING RATES

Model	MAPE under data missing rate (%)			MSPE under data missing rate (%)		
	10	20	30	10	20	30
Proposed model (proposed loss function)	1.74	1.85	1.99	0.02	0.07	0.09
Proposed model (benchmark loss function)	1.98	2.03	2.20	0.06	0.13	0.21
Benchmark model [28]	2.00	2.46	2.80	0.09	0.16	0.31

batch normalization significantly lifts reconstruction accuracy, as the nonlocal block encourages the model to capture global features [23] and batch normalization accelerates training deep neural nets by reducing internal covariate shift [24]. The learning rate is set to 0.001 based on cross-validation to achieve the highest reproduction accuracy. The basic motivation of architecture design is to realize a relatively simple model structure with fewer parameters and lower computational complexity. Moreover, to ensure the efficacy of our proposed method, we followed the existing algorithmic structures (e.g., 3×3 convolutional kernel in [25]) that have proven to be effective in extracting the spatio-temporal characteristics in similar feature maps of power grid data. To address the challenges of training GAN for data recovery tasks in power systems, our approach leverages a finely tuned discriminator that incorporates spatio-temporal constraints in the power system as prior knowledge during training. In addition, the model architecture is tailored to the characteristics of electrical feature maps with missing data. To prevent overfitting, an early stopping technique, which monitors the performance of the model on the testing set and stops the training process when the MAPE starts to increase, is employed.

Table IV further compares the performance between the proposed and benchmark method [29], where the proposed method comprises two main modifications to the benchmark model. First, we designed a model structure suitable for handling spatial-temporal electrical data, as given in Table III. Second, the loss function is modified to improve the model's ability to utilize all known information in the training database, as demonstrated by (9) and (10). For better comparison, the complete training set, where all elements in feature maps are known, is fed to models. The proposed model achieves significantly higher recovery accuracy than the benchmark model under different data missing rates. Also, the performance of the benchmark model degrades significantly with the increase in the missing rate, suggesting that the simple fully connected networks have limited ability to capture the spatial and temporal dynamics of power grids; conversely, the proposed model is more robust to data missing. The proposed loss function (9) further improves the model performance w.r.t. the benchmark loss function (10) owing to the use of more valid information during the training process, significantly enhancing the model ability to extract the feature distribution.

TABLE V
MODELS UNDER DIFFERENT DATA MISSING RATES

Model	MAPE under data missing rate (%)			MSPE under data missing rate (%)		
	10	20	30	10	20	30
LSTM predictor [13]	1.80	2.53	3.21	0.05	0.17	0.58
GRU predictor [12]	1.78	2.15	2.80	0.05	0.16	0.39
Vanilla GAN [15]	1.93	2.10	2.40	0.05	0.13	0.25
SAGAN [18]	1.81	1.98	2.29	0.03	0.12	0.20
WGAN [19]	1.90	2.05	2.35	0.03	0.14	0.21
FAWGAN-GN [20]	1.83	2.03	2.39	0.03	0.14	0.18
Proposed model	1.74	1.85	1.99	0.02	0.07	0.09

C. Attention Map Visualization of Proposed Model

The attention maps generated visualize the model weights in the final nonlocal layer of the generator, revealing the reason behind the high reconstruction ability of the proposed model. When different signals, including invalid and valid information, are adopted as inputs to the model, the valid features are adaptively recognized by model optimization. In order to simulate the PMU failure scenarios, we set all the measurement data to zero for the malfunctioning PMUs.

Attention maps [see Fig. 3(c) and (d)] visualize the attention distribution in two samples with different PMU failures, whose corresponding feature maps are illustrated in Fig. 3(a) and (b). When the missing data in a partially observed sample are reconstructed, the malfunctioning PMU sequences should be neglected, because they cannot provide any useful information for data reconstruction. It is seen from Fig. 3(c) and (d) that the proposed algorithm can adaptively neglect the malfunctioning PMU sequences and focus on the valid data. According to (11), the attention distribution is generated via the direct interaction of each element in feature maps, so the proposed model has global receptive fields to adaptively focus on the valid information.

D. Comparison Versus Existing Methods

In this section, the proposed model is compared versus existing methods, including LSTM [14], GRU [13], vanilla GAN [16], SAGAN [19], WGAN [20], and full attention WGAN with gradient normalization (FAWGAN-GN) [21], which rely on complete training databases where each measurement in a feature map is known. As illustrated in Fig. 3(a), random data missing and PMU failures are, respectively, considered for comparison.

Table V compares the model performance in the case of different data missing rates. When the missing rate is 10%, GRU and LSTM predictors demonstrate higher performance than the vanilla GAN in terms of MAPE, as the specific chain-like

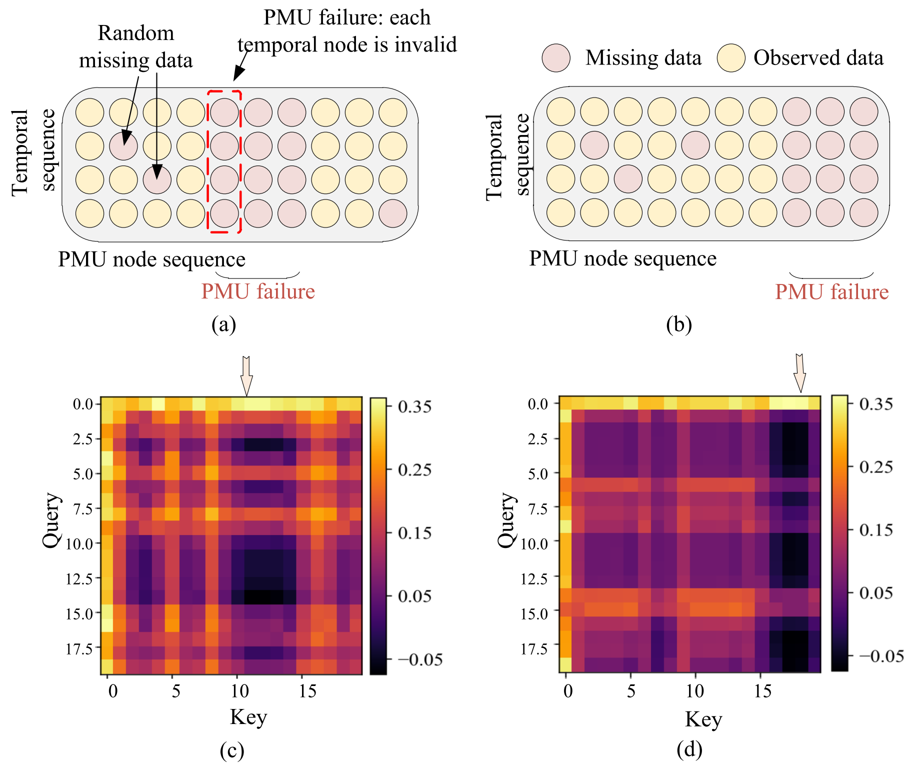

Fig. 3. Attention maps generated in nonlocal block and the corresponding feature maps in two samples. Note each node in spatio-temporal views represents two neighboring nodes for simplicity of drawing, namely, the feature map consists of 20 PMU nodes and eight temporal sections. (a) Spatio-temporal feature map with No. 9–14 PMU failure. (b) Spatio-temporal feature map with No. 15–20 PMU failure. (c) Attention map in scenario (a). (d) Attention map in scenario (b).

structure in recurrent networks can effectively extract temporal correlation in the power system. However, with the increase in data missing rates, the performance of methods based on GRU and LSTM degrades significantly; conversely, the models based on the GAN structure, including the proposed method and the model in [16], [19], [20] and [21], maintain high reconstruction accuracy. For all the cases, the proposed model achieves the best performance owing to several reasons. Unlike the discriminator in [16], [19], [20] and [21] that labels the entire reconstructed sample as real or not, the proposed discriminator distinguishes the observed (true) data from generated (fake) data in each partially observed sample. Since the feature maps in each sample indicate the dynamic response of the power system in terms of spatial and temporal relationship, with each pixel representing the electrical measurements of a certain moment in a bus (as described in Section III-A), the proposed model is able to capture finer grained information within the feature maps, leading to improved performance compared with other approaches. Besides, the value differences (9) are introduced to the model optimization process, further enhancing the data reconstruction ability.

Table VI gives the model performance under different PMU failure rates, where 10% data points are randomly missing from the test set. Compared with Table V, the models exhibit more significant performance degradation, suggesting that reconstructing entire sequences is much more difficult than reconstructing random missing data. Besides, even though the additional prior temporal data were provided to GRU and LSTM predictors (as required to predict the following sequence [13], [14]), these models show inferior performance to the GAN structure, and their prediction error increases with sequence due to the accumulated prediction error. Moreover, the GAN-based models with attention computation, including the proposed method, [19], and [21], typically exhibit superior performance compared with other models, as given in Table VI. This suggests that the global receptive field introduced by attention computation enables the model to focus on the most relevant features within the partially observed samples. Among all the models, the proposed model maintains better performance owing to its model structure and optimization process that effectively capture the spatial correlation. The model weights are visualized by attention maps, as

TABLE VI
MODELS UNDER DIFFERENT PMU FAILURE RATES

Model	MAPE under PMU failure rate (%)			MSPE under PMU failure rate (%)		
	10	15	20	10	15	20
LSTM predictor [13]	2.59	3.09	3.68	0.19	0.45	0.61
GRU predictor [12]	2.26	2.51	3.12	0.08	0.25	0.49
Vanilla GAN [15]	2.16	2.41	2.56	0.07	0.23	0.37
SAGAN [18]	2.01	2.15	2.41	0.06	0.17	0.33
WGAN [19]	2.18	2.42	2.61	0.08	0.20	0.39
FAWGAN-GN [20]	2.03	2.13	2.39	0.07	0.17	0.36
Proposed model	1.88	2.00	2.21	0.04	0.15	0.26

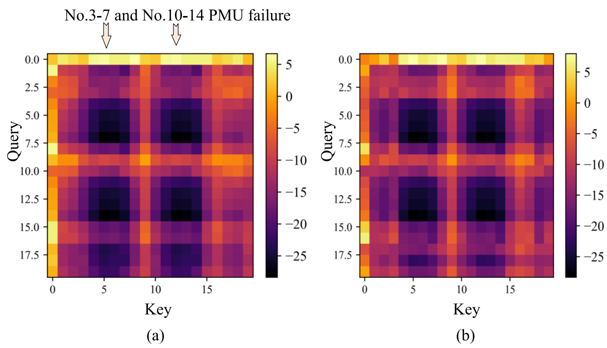

Fig. 4. Attention maps generated in nonlocal block with and without PMU measurements noise. (a) Attention map without PMU measurements noise. (b) Attention map with PMU measurements noise.

shown in Fig. 3, suggesting the proposed model can adaptively neglect the distraction of the invalid sequence and focus on the valid information to reconstruct data.

Besides, the proposed model avoids the dependence on the previous data compared with recurrent networks [13], [14], and is characterized by a unique visualization ability of model parameters (owing to the attention mechanism introduced by nonlocal blocks) compared with other approaches. These benefits enhance the credibility of the proposed model in practice.

E. Performance of Proposed Method on Large-Scale Power System

To evaluate the robustness of our proposed method against field measurement noise and unseen severe scenarios in a large-scale power system, we conducted experiments on the Northeast Power Coordinating Council 48-machine 140-bus system that simulates the Northeastern United States and Canada power grids [37]. Following the database generation structure in Section III-A, 20 PMUs from buses with generators connected are randomly chosen, since the electrical characteristics of generator buses are critical to reflect the power synchronization. In total, 20 000 samples are generated to construct a larger and more diverse dataset responding to the large-scale power system. In order to test the robustness of the data recovery method against unseen scenarios, we set the PMU failure rate to 30% in the training set, but increased it to 50% in the test set, and 20% randomly missing data are presented in both training and test sets. Furthermore, in order to test the model robustness against field noise that is commonly seen in PMU measurements, field noise has been added to the test set in accordance with the IEEE Standard [38]. The measurement error of a vector cannot exceed 1% of its real value for all PMUs, and the approach in [39] is followed to generate realistic noise distribution.

As demonstrated by Fig. 4, the proposed method can adaptively ignore all invalid PMU regions, even in the unseen scenarios where the PMU failure rate is obviously higher than the training scenario, highlighting the robustness of our approach for data recovery. In addition, the attention map in Fig. 4(b) exhibits less clear attention distribution than in Fig. 4(a) due to the presence of measurement noise. Nevertheless, even with measurement noise, our proposed method can still effectively filter out invalid PMU regions. This finding further demonstrates the robustness and efficacy of our approach in field measurements.

TABLE VII
Model Performance Under Different PMU Failure Rates

Accuracy	MAPE under PMU failure rate (%)			MSPE under PMU failure rate (%)		
	30	40	50	30	40	50
without PMU measurements noise	1.98	2.23	2.56	0.12	0.32	0.46
with PMU measurements noise	1.99	2.25	2.59	0.12	0.33	0.48

Table VII demonstrates that the proposed method achieves high accuracy in data recovery even under severe PMU failure rates. Furthermore, Table VII gives that the presence of PMU measurement noise does not significantly affect the model performance. This result is also supported by Fig. 4, where the attention maps generated by the proposed method with and without PMU measurement noise are visually similar. These results highlight the proposed method's capability to extract the underlying electrical characteristics of a large-scale power system and to recover missing data for unseen scenarios, demonstrating its potential for practical applications in power systems.

IV. Proposed Online Framework for TSA Tasks

In order to prove the necessity and effectiveness of the proposed data reconstruction method in practice, an online TSA framework with the proposed data reconstruction model is further designed for IEEE-39 system, as shown in Fig. 5, where the incomplete online measurements are first reconstructed by the trained generator, and then fed into the TSA classifier. Specifically, Conv denotes the convolutional neural layer, Resblock denotes the Residual block, FC denotes the fully connected layer, and BN denotes batch normalization. It should be noted that only the generator is adopted in the process, as the discriminator is to improve the reconstruction accuracy of the generator during the training process. Without loss of generality, in this section, the single CNN-based classifier based on [22] is taken as an example, but can be replaced by other classifier structures. Note that the parameters in both the generator and the classifier are fixed, and the classifier requires no modification to the previous generator model, which maintains the same dimension of the input feature maps.

Considering the 30% data missing rate in test sets, the online classification accuracy increases from 90.2% to 98.2% when the missing data are reconstructed with the proposed method, which has a reconstruction accuracy of 97.12%. Fig. 6 shows the feature representations at the final fully connected layer of the classifier by t-SNE [40]. As seen in Fig. 6(a), the massive overlap of samples with different labels indicates that the classifier cannot effectively distinguish between the stable and unstable samples; however, in Fig. 6(b), the gap between different types of data is increased after the missing data reconstruction. This suggests the proposed framework with the data reconstruction process effectively enhances the overall model performance.

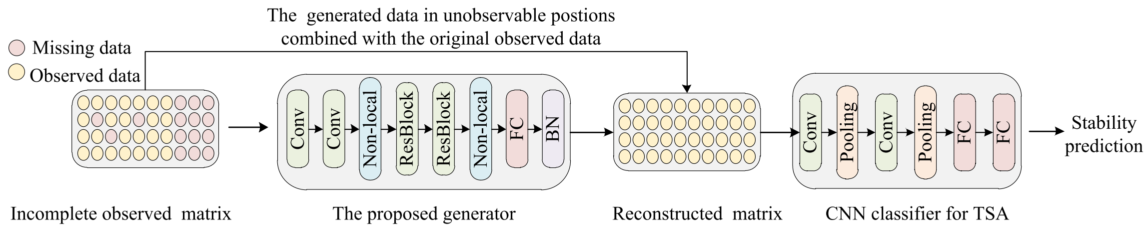

Fig. 5. Online framework for TSA task with the proposed missing data reconstruction model.

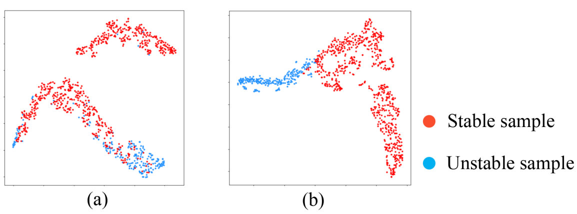

Fig. 6. Visualization of features extracted by the classifier with different databases. (a) TSNE visualization with incomplete database. (b) TSNE visualization with reconstructed database.

Besides, the proposed framework can be readily migrated to other tasks based on the proposed data reconstruction model and the following classifier, such as disturbance identification [26] and system identification [41].

V. Conclusion

In this article, we propose a GAN-based model for missing data reconstruction in smart grids. The effectiveness of the proposed model structure and loss function is verified with ablation experiments, and the comparison with existing approaches shows that the proposed model has superior performance under both random data missing and PMU failure conditions. Besides, the visualization of activation maps within the model's internal layers demonstrates that the model can adaptively focus on valid data and further enhances the credibility of the proposed method in practical applications. Moreover, an online framework for the TSA task is proposed, proving the effectiveness and necessity of the proposed model in the TSA task. Finally, the proposed data reconstruction method can be applied to similar power system applications by adjusting the database parameters to meet specific task requirements. For instance, the method can be employed in constructing a database based on power flow information for voltage stability assessment tasks that also encounter challenges due to incomplete PMU online measurements.

The proposed method, which has been demonstrated to effectively correct missing PMU data in spatio-temporally correlated scenarios, has the potential to be extended to address other missing data correction issues in smart grids. One such issue is online measurements that combine data from PMUs and supervisory control and data acquisition (SCADA) systems due to the high costs associated with PMU deployment. This integration introduces additional challenges with data being collected at different sampling rates; indeed, the SCADA systems typically acquire data at a lower rate (every 2–4 s). This article focuses on restoring samples composed of PMU data with a uniform sampling rate using GAN networks. Future work could explore the extension of the proposed data recovery method to address missing data in online measurements collected jointly by PMUs and SCADA systems, considering the challenge of data collected at different sampling rates. In addition, research efforts can be dedicated to identifying the optimal GAN-based architecture for power systems of varying scales, as the dataset size for power system operating points increases rapidly with the expansion of the power grid. Furthermore, in future research, it is crucial to explore methods for establishing GAN networks with enhanced transferability, ensuring their effectiveness in data recovery when applied to previously unseen power system operating points.

References

[1] K. Chen, Z. He, S. X. Wang, J. Hu, L. Li, and J. He, "Learning-based data analytics: Moving towards transparent power grids," *CSEE J. Power Energy Syst.*, vol. 4, no. 1, pp. 67–82, 2018.

[2] B. Tan, J. Yang, T. Zhou, X. Zhan, and C. Luo, "Spatial-temporal adaptive transient stability assessment for power system under missing data," *Int. J. Elect. Power Energy Syst.*, vol. 123, 2020, Art. no. 106237.

[3] K. Chen, J. Hu, and J. He, "Detection and classification of transmission line faults based on unsupervised feature learning and convolutional sparse autoencoder," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1748–1758, May 2018.

[4] L. Zhu and D. J. Hill, "Cost-effective bad synchrophasor data detection based on unsupervised time-series data analytic," *IEEE Internet Things J.*, vol. 8, no. 3, pp. 2027–2039, Feb. 2021.

[5] W. Liao, S. Salinas, M. Li, P. Li, and K. A. Loparo, "Cascading failure attacks in the power system: A stochastic game perspective," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2247–2259, Dec. 2017.

[6] Y. Zhang, Y. Xu, and Z. Y. Dong, "Robust ensemble data analytics for incomplete PMU measurements-based power system stability assessment," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1124–1126, Jan. 2018.

[7] Y. Zhang, Y. Xu, and Z. Y. Dong, "Robust classification model for PMU-based on-line power system DSA with missing data," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 18, pp. 4484–4491, 2017.

[8] 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.

[9] X. Wang, Q. Zhou, C. Huang, S. Liao, Y. Jiang, and Y. Ge, "A transient stability assessment method using LSTM network with attention mechanism," in *Proc. Proc. IEEE 8th Int. Conf. Adv. Power Syst. Automat. Protection*, 2019, pp. 120–124.

[10] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1006–1013, Mar. 2016.

[11] Y. Hao, M. Wang, and J. H. Chow, "Modeless streaming synchrophasor data recovery in nonlinear systems," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1166–1177, Mar. 2020.

[12] M. Liao, D. Shi, Z. Yu, Z. Yi, Z. Wang, and Y. Xiang, "An alternating direction method of multipliers based approach for PMU data recovery," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4554–4565, Jul. 2019.

[13] J. J. Q. Yu, A. Y. S. Lam, D. J. Hill, Y. Hou, and V. O. K. Li, "Delay aware power system synchrophasor recovery and prediction framework," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3732–3742, Jul. 2019.

[14] X. Guo, S. Zhu, Z. Yang, H. Liu, and T. Bi, "Consecutive missing data recovery method based on long-short term memory network," in *Proc. 3rd Asia Energy Elect. Eng. Symp.*, 2021, pp. 988–992.

[15] I. Goodfellow et al., "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[16] C. Ren and Y. Xu, "A fully data-driven method based on generative adversarial networks for power system dynamic security assessment with missing data," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 5044–5052, Nov. 2019.

[17] C. Yan et al., "Zeronas: Differentiable generative adversarial networks search for zero-shot learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 12, pp. 9733–9740, Dec. 2022.

[18] C. Yan, X. Chang, M. Luo, H. Liu, X. Zhang, and Q. Zheng, "Semantics-guided contrastive network for zero-shot object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Jan. 04, 2022, doi: 10.1109/TPAMI.2021.3140070.

[19] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-attention generative adversarial networks," 2018. [Online]. Available: https://arxiv.org/abs/1805.08318

[20] X. Ouyang, Y. Chen, and G. Agam, "Accelerated WGAN update strategy with loss change rate balancing," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2021, pp. 2545–2554.

[21] J. Fan, X. Yuan, Z. Miao, Z. Sun, X. Mei, and F. Zhou, "Full attention Wasserstein GAN with gradient normalization for fault diagnosis under imbalanced data," *IEEE Trans. Instrum. Meas.*, vol. 71, no. 1, Jul. 2022, Art. no. 3517516.

[22] Y. Zhou, Q. Guo, H. Sun, Z. Yu, J. Wu, and L. Hao, "A novel data-driven approach for transient stability prediction of power systems considering the operational variability," *Int. J. Elect. Power Energy Syst.*, vol. 107, pp. 379–394, 2019.

[23] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7794–7803.

[24] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn.*, 2015, pp. 448–456.

[25] Z. Shi et al., "Convolutional neural network-based power system transient stability assessment and instability mode prediction," *Appl. Energy*, vol. 263, 2020, Art. no. 114586. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261920300982

[26] Z. Li, H. Liu, J. Zhao, T. Bi, and Q. Yang, "A power system disturbance classification method robust to PMU data quality issues," *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 130–142, Jan. 2022.

[27] H. Cai and D. J. Hill, "A real-time continuous monitoring system for long-term voltage stability with sliding 3D convolutional neural network," *Int. J. Elect. Power Energy Syst.*, vol. 134, 2022, Art. no. 107378. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061521006177

[28] C. Zhang, S. R. Kuppannagari, R. Kannan, and V. K. Prasanna, "Generative adversarial network for synthetic time series data generation in smart grids," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids*, 2018, pp. 1–6.

[29] J. Yoon, J. Jordon, and M. van der Schaar, "GAIN: Missing data imputation using generative adversarial nets," in *Proc. 35th Int. Conf. Mach. Learn.*, 2018, pp. 5689–5698.

[30] P. M. Joshi and H. Verma, "Synchrophasor measurement applications and optimal PMU placement: A review," *Electric Power Syst. Res.*, vol. 199, 2021, Art. no. 107428. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779621004090

[31] J. H. Chow and K. W. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE Trans. Power Syst.*, vol. 7, no. 4, pp. 1559–1564, Nov. 2002.

[32] D. Fabozzi and T. Van Cutsem, "On angle references in long-term time-domain simulations," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 483–484, Feb. 2011.

[33] P. S. Kundur and O. P. Malik, *Power System Stability and Control*. Hoboken, NJ, USA: McGraw-Hill, 2022.

[34] S. Fan and H. Ding, "Time domain transformation method for accelerating EMTP simulation of power system dynamics," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1778–1787, Nov. 2012.

[35] Y. Liu, K. Sun, R. Yao, and B. Wang, "Power system time domain simulation using a differential transformation method," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3739–3748, Sep. 2019.

[36] A. Gupta, G. Gurrala, and P. S. Sastry, "Instability prediction in power systems using recurrent neural networks," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, 2017, pp. 1795–1801.

[37] D. Osipov and K. Sun, "Adaptive nonlinear model reduction for fast power system simulation," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6746–6754, Nov. 2018.

[38] *IEEE Standard for Synchrophasor Data Transfer for Power Systems*, IEEE Std C37.118.2-2011 (Revision of IEEE Std C37.118-2005), Dec. 2011.

[39] M. He, V. Vittal, and J. Zhang, "Online dynamic security assessment with missing PMU measurements: A data mining approach," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1969–1977, May 2013.

[40] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, 2008.

[41] I. Kamwa and L. Gerin-Lajoie, "State-space system identification-toward MIMO models for modal analysis and optimization of bulk power systems," *IEEE Trans. Power Syst.*, vol. 15, no. 1, pp. 326–335, Feb. 2000.

Jiashu Fang received the B.S. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2015, and the M.S. degree in electrical power system from the University of Manchester, Manchester, U.K., in 2016. She is currently working toward the Ph.D. degree in electrical engineering with North China Electric Power University.

Her research interests include applications of machine learning and data science in power systems.

Le Zheng received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2011 and 2017, respectively.

From 2017 to 2019, he had been a Postdoctoral Research Fellow with Stanford University, Stanford, CA, USA. He is currently a Faculty Member of the North China Electric Power University, Beijing. His current research interests include stability and control of power electronics dominated power systems and renewable energy conversion systems, and machine learning applications in power systems.

Chongru Liu (Senior Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2000 and 2006, respectively.

She is currently a Professor, Ph.D. Supervisor, and the Dean of Undergraduate Academic Affairs Office, North China Electric Power University, Beijing. In 2009 and 2012, she was selected for Beijing Science and Technology Nova Talent Support Plan and the Ministry of Education Excellent Talents Support Plan for the New Century.

Dr. Liu is the Standing Director of the Tenth Council of Beijing Electrical Engineering Society, Director of the Committee on Science Popularization and Education, Member of the Fifth and Sixth National Technical Committee on Power System Management and Information Exchange Standardization, Member of the Power System Automation Committee of China Electrical Engineering Society, Member of the Foreign Affairs Committee of China Electrical Engineering Society, and Member of Beijing Science Representatives of the Ninth Congress of the Technical Association.