3D Data Scattergram Image Classification Based Protection for Transmission Line Connecting BESS Using Depth-wise Separable Convolution Based CNN

Yingyu Liang, Yi Ren, Xiaoyang Yang, and Wenting Zha

***Abstract*—The distinctive fault characteristics of battery energy storage stations (BESSs) significantly affect the reliability of conventional protection methods for transmission lines. In this paper, the three-dimensional (3D) data scattergrams are constructed using current data from both sides of the transmission line and their sum. Following a comprehensive analysis of the varying characteristics of 3D data scattergrams under different conditions, a 3D data scattergram image classification based protection method is developed. The depth-wise separable convolution is used to ensure a lightweight convolutional neural network (CNN) structure without compromising performance. In addition, a Bayesian hyperparameter optimization algorithm is used to achieve a hyperparametric search to simplify the training process. Compared with artificial neural networks and CNNs, the depth-wise separable convolution based CNN (DPCNN) achieves a higher recognition accuracy. The 3D data scattergram image classification based protection method using DPCNN can accurately separate internal faults from other disturbances and identify fault phases under different operating states and fault conditions. The proposed protection method also shows first-class tolerability against current transformer (CT) saturation and CT measurement errors.**

***Index Terms*—Convolutional neural network (CNN), battery energy storage station (BESS), depth-wise separable convolution, hyperparameter optimization, fault classification, line protection.**

I. Introduction

With the accelerated development of renewable energy generation technologies, power systems have undergone profound changes. Large-scale renewable energy sources with strong randomness and volatility are being integrated into power grids, which leads to stability problems such as large power fluctuations and voltage and frequency instabilities [1], [2]. To a large extent, the battery energy storage stations (BESSs) can counteract power and energy fluctuations of renewable energy sources and are thus effective in solving the aforementioned challenges [3], [4]. BESSs are currently receiving considerable attention for their ability to enhance the stability and dynamic performance of power systems. However, the integration of BESSs into power grids has certain disadvantages that cannot be ignored. The topology and control system of the BESS and synchronous generator (SG) have marked differences, and thus their fault features differ significantly. Consequently, this causes the performance degeneration and even incorrect operation of conventional protection schemes. It is important to apply a novel protection technique to the special fault features of BESSs for the safety of power grids.

The BESS, wind turbine, photovoltaic power plant, and modular multi-level converter (MMC) based converter stations are converter-interfaced sources, all of which are integrated into the power grid through power electronic converters, and to a certain extent, they exhibit partially similar fault characteristics. References [5]-[8] provide in-depth analyses of the fault characteristics of converter-interfaced sources. The applicability of conventional distance and differential protections in the presence of converter-interfaced sources has also been examined. In [9], a distance protection scheme adapted for a photovoltaic power plant is proposed. However, this scheme may operate incorrectly when the fault resistance is large. In [10], an analytical model based on additional impedance is proposed for an adaptive distance protection scheme, which can reliably and accurately detect fault locations. However, this scheme is only applicable to asymmetrical faults. The distance protection methods proposed in [11] and [12] could perform well under both symmetrical and asymmetrical faults, but they are not phase-segregated and cannot be used as the main protection in transmission networks. Several pilot protection schemes based on similarity metrics are proposed in [13]-[16] and are designed to identify fault phases. However, these schemes are developed in response to the fault features inherent in renewable plants and have the problem of performance degradation in the presence of BESSs. In addition, they are negatively affected by current transformer (CT) saturation. Reference [17] introduces a time-domain differential protection method based on the divergent distribution features of trajectories in the Cartesian plane under different conditions. Reference [18] proposes a protection method based on the internal fault index for transmission lines connecting MMC-based converter stations. This

Manuscript received: December 21, 2023; revised: April 9, 2024; accepted: July 17, 2024. Date of CrossCheck: July 17, 2024. Date of online publication: September 2, 2024.

This work was supported by the Fundamental Research Funds for Central Universities (No. 2024JCCXJD01).

Y. Liang (corresponding author), Y. Ren, X. Yang, and W. Zha are with School of Mechanical Electronic and Electrical Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China (e-mail: liangyingyu2013@163.com; 18434362109@163.com; wfyangxy0106@163.com; wentingzha@126.com).

DOI: 10.35833/MPCE.2023.001008

method functions effectively under different operating modes of MMC and its tolerance to CT saturation has been proven excellent. Reference [19] proposes a current trajectory coefficient based protection for AC transmission line connecting a BESS, which shows an excellent performance. The methods proposed in [17]-[20] perform well in the presence of BESSs, but they require a reasonable setting of key protection parameters, which significantly affects the operating performance of protection and relies heavily on expert knowledge. The protection of a transmission line can be considered a classification problem, the objective of which is to differentiate between internal and external faults. Machine learning is a suitable approach for solving such classification problems.

Reference [21] proposes a decision tree based approach for the fault classification of transmission line. The voltage and current on the protected side of the transmission line are analyzed to calculate the phasors of the odd harmonics. These values are used as inputs to the decision tree for fault classification. This approach has a high accuracy but its generalization ability is poor, and its classification accuracy diminishes for faults under conditions not considered in the training set. Reference [22] proposes a fault location and classification scheme for transmission line based on improved machine learning algorithms. The summation-Gaussian extreme learning machine proposed in [22] performs well in fault classification and location without the need for ad hoc feature extraction. Nevertheless, this scheme does not consider non-ideal conditions such as high-resistance faults and CT saturation. Reference [23] proposes the multiple random forests based fault location scheme, which is used to identify single-phase-to-ground faults on transmission lines connected with wind farms. The amplitude of the zero-sequence current is used to construct the feature set, and a random forest algorithm is used to locate the fault point. However, this scheme is not applicable to symmetrical faults. References [21]-[23] present machine learning based techniques for fault classification and location, but the accuracy of these techniques requires further enhancement, and they exhibit limited generalization. Compared with machine learning, convolutional neural networks (CNNs) have better generalization and feature extraction capabilities and perform better in fault identification. Reference [24] proposes an accelerated CNN based differential protection method, which can be applied to separate the internal faults and inrush currents of power transformers to reduce the risk of false tripping. However, this method is sensitive to the second harmonics and susceptible to their adverse effects. Reference [25] proposes a fault identification method based on CNN with a branched structure. This method feeds voltage and current feature matrices into the CNN to distinguish between DC faults and lightning disturbances. Nevertheless, the recognition accuracy of this method is not high, limiting its use to an auxiliary role in fault recognition. Reference [26] proposes a novel signal-localized CNN for the differential protection of power transformer, which can effectively separate internal faults from inrush currents and external faults even under CT saturation. The datasets in [26] are composed of simulated data and have not been tested with actual engineering field data, which may lead to performance degradation in engineering applications. Reference [27] uses a modified complete ensemble empirical model to convert fault signals into two-dimentional (2D) grey-scale images. The images are fed into the CNN, which mines fault features for the fault classification of distribution network. Similar to [27], the zero-sequence current is transformed using the Grampian angular field to obtain the current feature image, which is then fed to the CNN for recognition [28]. The methods in [27] and [28] involve complex data processing procedures and are computationally intensive, which is impractical for engineering applications. Reference [29] proposes a method for initial fault detection and identification of cables based on current signals measured at one end of the cable using a denoising autoencoder and CNN. This method identifies only initial faults of cables and cannot provide the mainline protection. Reference [30] proposes a line protection scheme based on current trajectory image classification. This scheme can identify internal faults and fault phases with a high identification accuracy. However, the CNN used in this scheme has many parameters, thus requiring a large storage space for the protection device. The summary of the previous studies is presented in Table I.

TABLE I
SUMMARY OF PREVIOUS STUDIES

Method	Description	Feature
Conventional protection methods	Protection methods for analysis of adapt ability problem [5]-[8]	These protection methods may suffer from performance degradation and even rejection in the presence of converter-interfaced sources
	Distance protection methods [9]-[12]	These protection methods are not phase-segregated and cannot be used as the main protection in transmission networks
	Differential protection methods applicable to presence of renewable plants [13]-[16]	These protection methods are developed in response to the fault features inherent in renewable plants and may incur performance degradation in the presence of BESSs. Also, they are negatively affected by CT saturation
	Protection methods for lines connecting BESSs or voltage source converter based high-voltage direct current (VSC-HVDC) stations [17]-[20]	These protection methods are strongly influenced by the rational selection of thresholds and key parameters, which requires expert knowledge
Artificial intelligence based protection methods	Machine learning based protection methods [21]-[23]	The accuracy needs improvement and the generalization ability is poor
	Deep learning [24] based protection methods [30]	The structure of network is complex with a large number of parameters and the computation is intensive

The major contributions of this study are summarized as follows.

1) A theoretical derivation of the trajectory equations of the scattergrams constructed from the currents on both sides is presented. The three-dimensional (3D) data scattergrams are constructed based on the current data, and the differences in the 3D data scattergrams under different fault condi-

tions are analyzed. The depth-wise separable convolution based CNN (DPCNN) is introduced to identify 3D data scattergram images for separating the internal and external faults.

2) The CNN is improved using depth-wise separable (DP) convolution, and the training process is simplified through Bayesian hyperparameter optimization. Compared with artificial neural networks (ANNs) and CNNs, the DPCNN has a higher accuracy rate.

3) The effectiveness of the proposed protection method is assessed using PSCAD and a real-time simulator (RT-Sim) considering various fault conditions, operating states, and non-ideal situations. The results demonstrate that the proposed protection method can be applied to transmission lines connecting a BESS and offers better performance than conventional differential protection methods and cosine-similarity (CS) based pilot protection.

The remainder of this paper is organized as follows. Section II describes how the 3D data scattergrams are constructed in different cases. Section III describes the DPCNN with Bayesian hyperparameter optimization method, and introduces the 3D data scattergram image classification based protection method. Section IV validates the effectiveness of the proposed protection method under different fault conditions, with comparison with other protection methods. Finally, a conclusion is provided in Section V.

II. 3D Data Scattergram

A. Fault Current Characteristics

Figure 1(a) and (b) shows the detailed and simplified models of the modified IEEE 39-bus system with a 150 MW BESS, respectively, where i_b and i_g are the currents on the BESS and grid sides of transmission line 33-19, respectively; and $\dot{I}_b$ and $\dot{I}_g$ are their corresponding phasors. PT is short for potential transformer. The parameters of transmission line 33-19 with a length of 60 km are given as follows. The positive- and zero-sequence impedances are (0.0617+j0.325) Ω/km and (0.227+j0.78)Ω/km, respectively, whereas the positive- and zero-sequence capacitances are 0.0086 μF/km and 0.0061 μF/km, respectively. Other parameters can be found in [31].

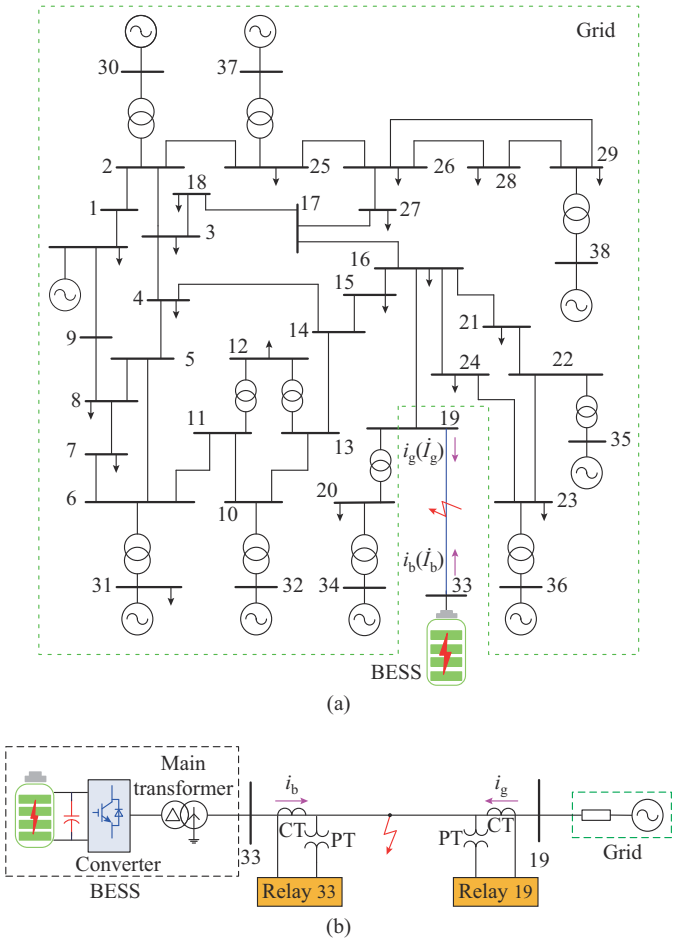

Fig. 1. Modified IEEE-39 bus system with a 150 MW BESS. (a) Detailed model. (b) Simplified model.

Similar to the photovoltaic power plant and VSC-HVDC station, the fault current on the BESS side is controlled by a grid-connected converter, which can be expressed as:

$$i_b = \underbrace{i_d^* \cos\left(\omega_0 t+\varphi_0\right) - i_q^* \sin\left(\omega_0 t+\varphi_0\right)}_{\text{Steady - state component}} + \underbrace{\left(i_d^*-i_{d0}\right)\frac{\mathrm{e}^{-\zeta\omega_n t}\left(\frac{k_p}{k_i}\omega_n \sin\left(\omega_d t\right)-\sin\left(\omega_d t+\beta\right)\right)}{\sqrt{1-\zeta^2}}\cos\left(\omega_0 t+\varphi_0\right) - \left(i_q^*-i_{q0}\right)\frac{\mathrm{e}^{-\zeta\omega_n t}\left(\frac{k_p}{k_i}\omega_n \sin\left(\omega_d t\right)-\sin\left(\omega_d t+\beta\right)\right)}{\sqrt{1-\zeta^2}}\sin\left(\omega_0 t+\varphi_0\right)}_{\text{Transient component}} \quad t\geq 0 \tag{1}$$

where ω_0 is the fundamental frequency of power grid; φ_0 is the initial angle of fault; k_p and k_i are the control parameters of the current loop; ζ is the damping ratio; ω_n is the natural oscillation frequency; ω_d is the damped oscillation frequency; β=arccos ζ; i_d^* and i_q^* are the reference values of the active and reactive fault currents, respectively, which can be obtained by referring to the grid code during a fault; and i_{d0} and i_{q0} are the initial values of active and reactive fault currents, respectively, which are determined by the operating state and control strategy. When the BESS operates in the discharging state, i_{d0} is positive; when the BESS operates in the charging state, i_{d0} is negative. Assuming that the converter operates at a unity power factor prior to the fault, $i_{q0}=0$.

From (1), it can be observed that the fault current of BESS contains both steady-state and transient components, and the transient component decays to 0 within a several milliseconds.

The fault current on the grid side can be expressed as:

$$i_g = I_{g,ac}\cos\left(\omega_0 t+\varphi_g\right) - I_{g,dc}\mathrm{e}^{-t/T_a} \tag{2}$$

where $I_{g,ac}$ is the peak of the power frequency current; $I_{g,dc}$ is the initial value of the decaying DC current; φ_g is the initial

phase angle; and T_a is the decay time constant.

The phase angle difference between the fault currents on the two sides may vary over a broad range during an internal fault. The phase difference may be obtuse, particularly when the BESS operates in the charging state [19]. Figure 2 shows the phase angle difference between $\dot{I}_b$ and $\dot{I}_g$ under different fault resistances and operating conditions when three-phase short-circuit faults occur at transmission line 33-19. Figure 2 clearly shows that when the BESS operates in the charging state, the phase angle difference between $\dot{I}_b$ and $\dot{I}_g$ is obtuse or even close to 180°. In this case, the traditional current phase differential protection will incorrectly identify the internal faults as external faults and fail to operate.

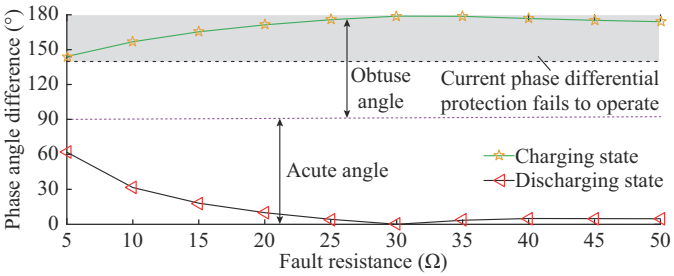

Fig. 2. Phase angle difference between $\dot{I}_b$ and $\dot{I}_g$ under different fault resistances and operating conditions.

When an external fault occurs, we have $i_b \approx -i_g$ according to Kirchhoff's current law and considering that the capacitive current is much less than the fault current of transmission line. This indicates that currents i_b and i_g are nearly the same in amplitude and 180° out of phase. The sum of the currents on both sides $i_{sum}=i_g+i_b\approx 0$.

Table II summarizes the fault current characteristics of i_g and i_b. When the phase angle difference is obtuse, the current phase differential protection, which typically serves as the primary protection for transmission lines, may make an incorrect decision and fail to operate. This severely affects the safety of power systems. Therefore, a new protection method is urgently required.

TABLE II
FAULT CURRENT CHARACTERISTICS OF i_g AND i_b

Condition	Phase angle difference (°)	Sum of i_g and i_b
Internal fault	0-360 (obtuse in charging state)	$i_{sum}=i_g+i_b\neq 0$
External fault/ normal condition	180	$i_{sum}=i_g+i_b\approx 0$

B. Formation of 3D Data Scattergram

To integrate the fault current information, a 3D data scattergram is constructed, where the 3D data are i_b, i_g, and i_{sum}, respectively.

The currents on the BESS and grid sides at each sampling instant are taken as the horizontal and vertical coordinates of the scattergram, respectively. In the following, the dynamic trajectory constructed by (i_b, i_g) is analyzed.

Under the large fault resistance, the decaying DC component of the fault current on the grid side is negligible. The transient component of the fault current of the converter generally decays to zero in a few milliseconds, which is neglected to facilitate the analysis. In this case, the fault currents on both sides can be expressed as:

$$\begin{cases} i_g = I_{g,ac}\cos\left(\omega_0 t+\varphi_g\right) \\ i_b = I_b\cos\left(\omega_0 t+\varphi_b\right) \end{cases} \tag{3}$$

where I_b and φ_b are the amplitude and phase angle of the current on the BESS side, respectively.

To obtain the expression for the trajectory of (i_b, i_g), the derivation process is as follows [32].

According to (3), the relationship between i_b and i_g can be derived as:

$$\frac{i_b}{I_b}\cos\varphi_g-\frac{i_g}{I_{g,ac}}\cos\varphi_b=\sin\left(\omega_0 t\right)\sin\left(\varphi_g-\varphi_b\right) \tag{4}$$

$$\frac{i_b}{I_b}\sin\varphi_g-\frac{i_g}{I_{g,ac}}\sin\varphi_b=\cos\left(\omega_0 t\right)\sin\left(\varphi_g-\varphi_b\right) \tag{5}$$

The sum of the squares of (4) and (5) is:

$$\frac{i_b^2}{I_b^2}+\frac{i_g^2}{I_{g,ac}^2}-\frac{2i_b i_g}{I_b I_{g,ac}}\cos\left(\varphi_g-\varphi_b\right)=\sin^2\left(\varphi_g-\varphi_b\right) \tag{6}$$

From (6), the trajectory of (i_b, i_g) is an ellipse, and the shape of the ellipse depends on the differences in the amplitude and phase angle between i_b and i_g.

When $\varphi_g-\varphi_b=0°$ or 180°, (7) can be obtained, and the trajectory of (i_b, i_g) is a straight line.

$$\left(\frac{i_b}{I_b}\mp\frac{i_g}{I_{g,ac}}\right)^2=0\Rightarrow i_g=\pm\frac{I_{g,ac}}{I_b}i_b \tag{7}$$

When $\varphi_g-\varphi_b=\pm 90°$, (8) can be obtained. The trajectory of (i_b, i_g) is a standard ellipse with the coordinate axis as the major axis. When $I_b=I_{g,ac}$, the ellipse becomes a standard circle.

$$\frac{i_b^2}{I_b^2}+\frac{i_g^2}{I_{g,ac}^2}=1 \tag{8}$$

Figure 3 shows a schematic of the current trajectory with various phase angle differences.

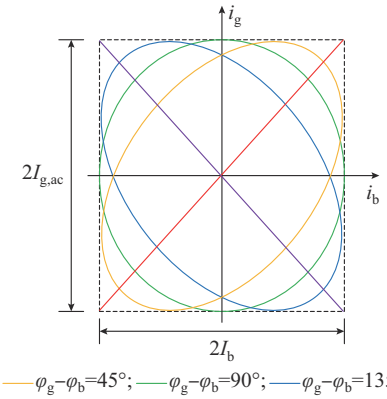

Fig. 3. Schematic of current trajectory with various phase angle differences.

The long axis of the ellipse is in the first and third quadrants when the phase angle difference is acute and is in the second and fourth quadrants when the phase angle difference is obtuse. The dashed box in Fig. 3 indicates the maximum bound-

ing rectangle of the current trajectories, of which the side lengths are $2I_b$ and $2I_{g,ac}$.

The increase in the fault resistance results in a corresponding decrease in the current amplitude on both sides. This implies that the maximum bounding rectangle will also shrink. When the fault resistance is large, the current is small, and the details may be difficult to observe and prone to judgment errors. To facilitate the analysis, we unify the size of the scattergram image formed by points (i_b, i_g) and use adaptive coordinate axes. The values of i_b and i_g have different ranges and are evenly distributed on the horizontal and vertical axes, respectively, according to their maximum and minimum values. Therefore, the actual image maintains equal lengths along the horizontal and vertical axes, although these lengths may correspond to varying ranges. The points in the scattergram are evenly distributed in the square area. In this case, the range of the current trajectory no longer represents the current amplitude information on both sides. Therefore, the sum of the currents on both sides, i.e., i_{sum}, is used as the third dimension of the 3D scattergram, giving a color to each point of the scattergram.

Figure 4 shows the current waveform and 3D data scattergram for the fault phase under a minor internal fault, i. e., with a large fault resistance, when the BESS operates in the charging state.

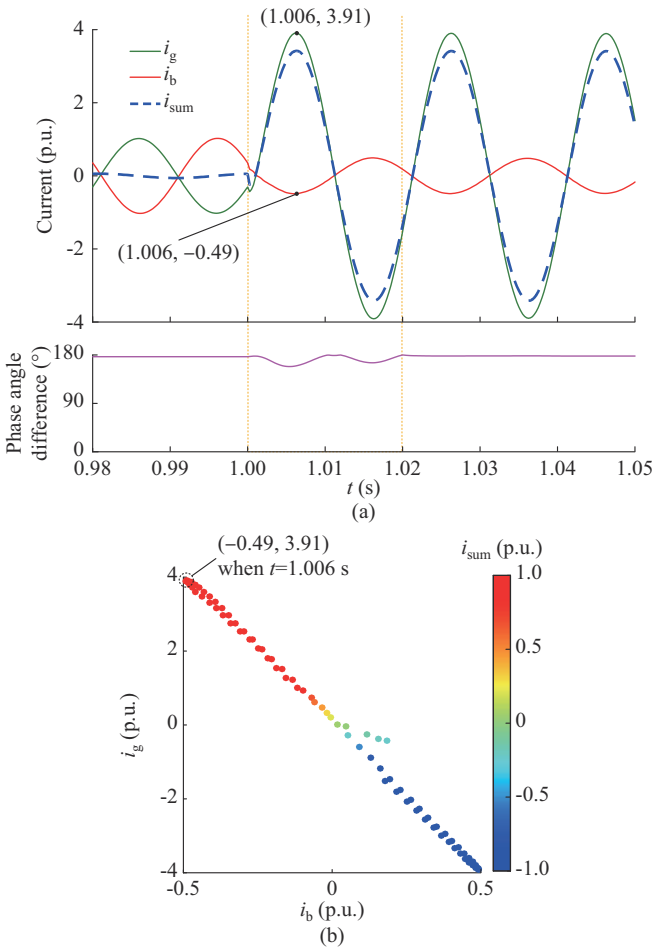

Fig. 4. Formation process of 3D data scattergram under a minor internal fault when BESS operates in charging state. (a) Current waveform. (b) 3D data scattergram for fault phase.

As shown in Fig. 4(b), the 3D data scattergram is constructed using the current sampling data on both ends of the transmission line for a cycle after the fault occurrence (as indicated by the period between the orange dotted lines in Fig. 4(a). The horizontal and vertical coordinates of the 3D scattergram are i_b and i_g, respectively, and the color of each data point is determined by i_{sum} at the current sampling time. The range of the color bar is [−1, 1]. When i_{sum} exceeds the upper limit of the color bar, i.e., $i_{sum} > 1$, the point is red. When i_{sum} is less than the lower limit of the color bar, i.e., $i_{sum} < -1$, the point is blue. As shown in Fig. 4, when $t = 1.006$ s, $i_{sum} = 3.91 - 0.49 = 3.42 > 1$, and this point is red. In addition, the phase angle difference is approximately 180°, and the scattergram is distributed in the second and fourth quadrants.

Figure 5 shows the current waveform and 3D data scattergram for the fault phase under a severe internal fault, i. e., with small fault resistance, when the BESS operates in the charging state.

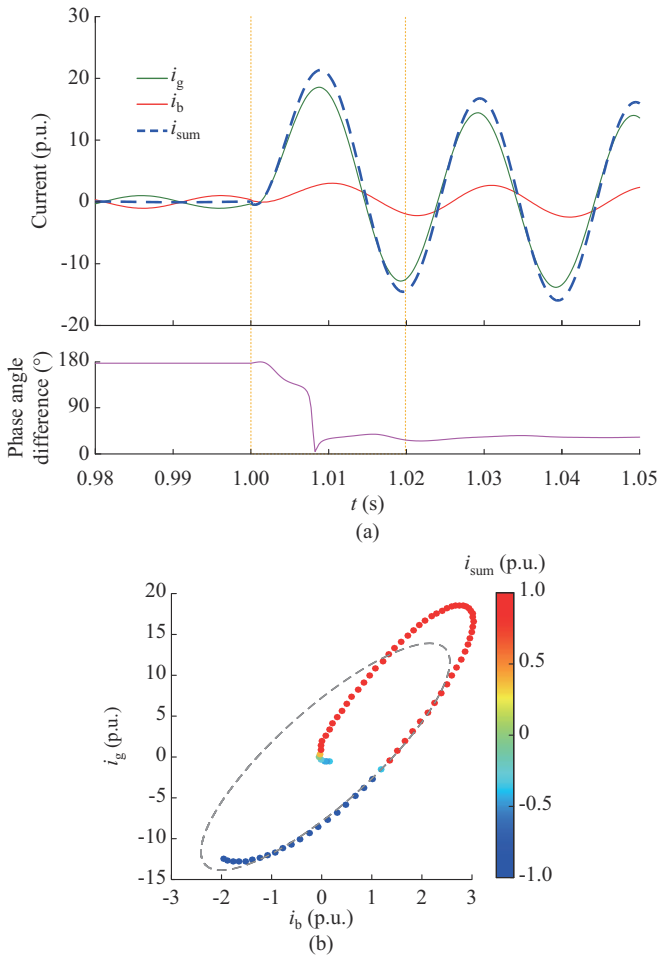

Fig. 5. Formation process of 3D data scattergram under a severe internal fault when BESS operates in charging state. (a) Current waveform. (b) 3D data scattergram for fault phase.

When the fault resistance is small, the grid-side fault current clearly contains decaying DC component. The current trajectory generated in this case is somewhat shifted and no longer a standard ellipse, but the overall trajectory is still an elliptical-like arc. As shown in Fig. 5, the phase angle differ-

ence is acute; thus, the 3D scattergram is mainly concentrated in the first and third quadrants.

C. Characteristic Analysis of 3D Data Scattergram Under Various Conditions

1) Internal Fault

The 3D data scattergrams under minor and severe internal faults when BESS operates in the charging state are given in Figs. 4(b) and 5(b).

Figure 6 shows the 3D data scattergrams under severe and minor internal faults when BESS operates in the discharging state. As shown in Fig. 6(a), when the fault resistance is small, the current trajectory is an arc close to an ellipse, the phase angle difference is acute, and the points are mainly concentrated in the first and third quadrants. As shown in Fig. 6(b), when the fault resistance is large, the current trajectory is elliptical, the phase angle difference is acute, and the points are mainly concentrated in the first and third quadrants. Another common feature is that the majority of the values of i_{sum} exceed the range of the color bar under an internal fault, such that the points of the 3D scattergram are mostly red and blue.

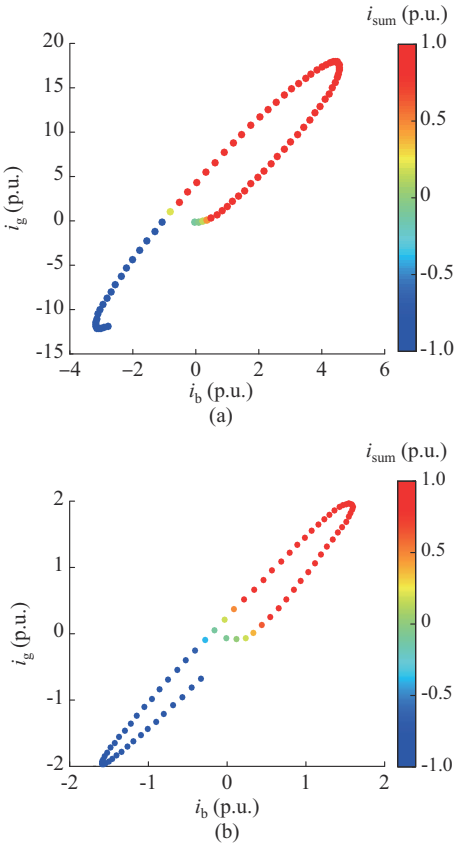

Fig. 6. 3D data scattergram when BESS operates in discharging state. (a) Severe internal fault. (b) Minor internal fault.

Figure 7 shows the current waveform and 3D data scattergram under an internal fault with CT saturation. When the grid-side CT is saturated, the current waveform of i_g is distorted. Accordingly, the 3D data scattergram is distorted, with the distorted points concentrated on the line $i_g=0$, as shown in the green area in Fig. 7(b).

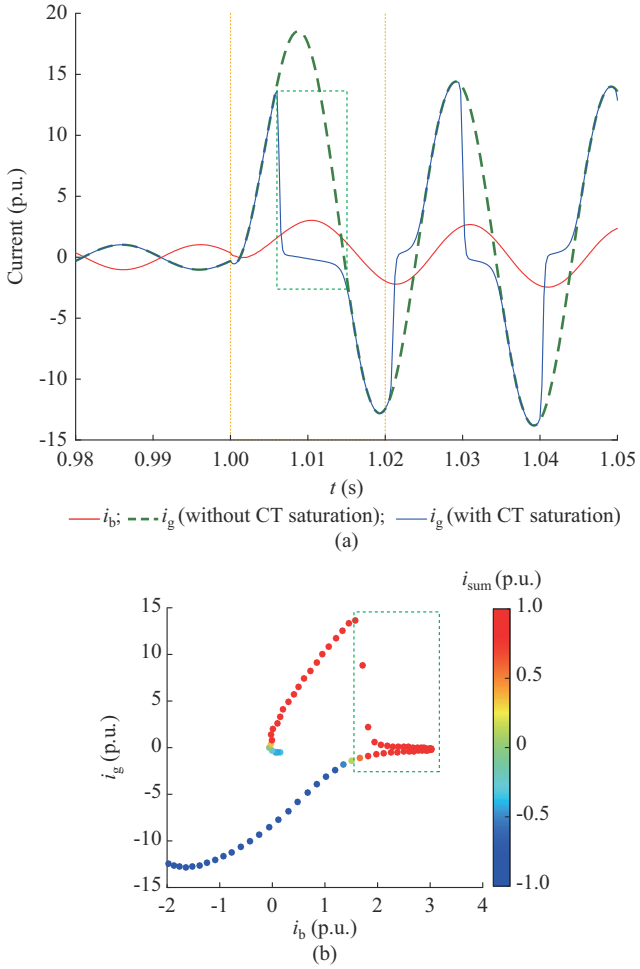

Fig. 7. Current waveform and 3D data scattergram under an internal fault with CT saturation. (a) Current waveform. (b) 3D data scattergram.

2) External Fault

Figure 8(a) shows the 3D data scattergram under an external fault without CT measurement error. When the fault is outside the protection zone, based on Kirchhoff's current law, i_b and i_g have opposite phases and nearly equal amplitudes such that $i_{sum}\approx 0$. As shown in Fig. 8(a), the points are concentrated on the line $i_g=-i_b$, which is shown in green, meaning that i_{sum} approximates 0. Similar to external faults, the currents on both sides are nearly equal in amplitude and opposite in direction under normal conditions. The 3D data scattergram under normal conditions can also be represented by a green line like Fig. 8(a).

Figure 8(b) shows the 3D data scattergram under external fault conditions, where a 10% error is included in the CT measurements. When an external fault occurs with a CT measurement error, the current amplitudes on both sides of the transmission line are no longer equal, and the value of i_{sum} no longer approximates zero. As shown in Fig. 8(b), due to the CT measurement error, the color of the points exhibits a gradient transition from red to blue.

Figure 9 shows the current waveform and 3D data scattergram under an external fault with CT saturation. The current waveform of i_g is distorted when the grid-side CT is saturated, and the distorted part corresponds to the green area in Fig. 9(b).

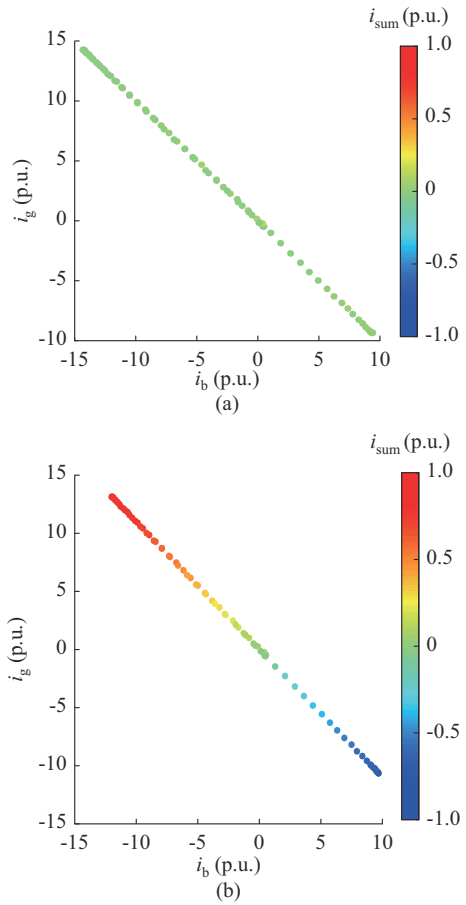

Fig. 8. 3D data scattergram under an external fault. (a) Without CT measurement error. (b) With CT measurement error.

In summary, the 3D data scattergram constructed from [i_b, i_g, i_{sum}] contains a wealth of information, which can be effectively utilized to separate internal faults with fault phases from external faults. The 3D data scattergram has different characteristics under internal and external faults.

1) Without CT saturation and measurement error, the trajectory of the 3D data scattergram under internal faults is close to an elliptical arc, with most of the points colored in red and blue, and the trajectory of the 3D data scattergram under external faults is a straight line in green.

2) With CT saturation, the 3D data scattergram is distorted, and the distorted part is concentrated on the line $i_g = 0$. The parts that are not distorted still show differences in shape and color under internal and external faults.

3) With CT measurement error, the color of points in the 3D data scattergram under an external fault shows a gradient transition.

III. 3D Data Scattergram Image Based Protection Method Using DPCNN

Distinguishing the 3D data scattergram images under internal and external faults is a standard classification problem. CNN is particularly apt for this type of classification problem due to its excellent nonlinear fitting and feature extraction capabilities.

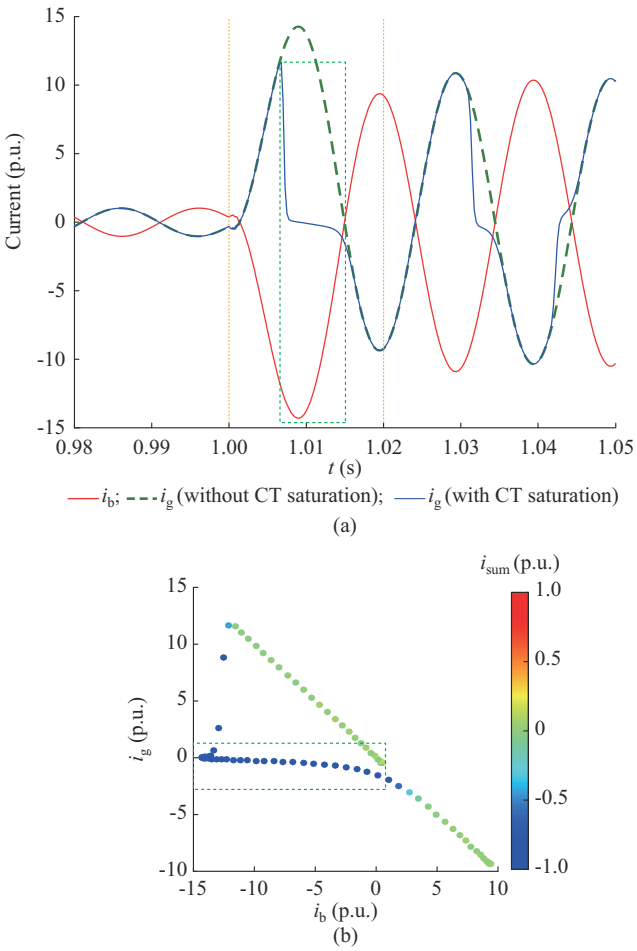

Fig. 9. Current waveform and 3D data scattergram under an external fault with CT saturation. (a) Current waveform. (b) 3D data scattergram.

A. DPCNN with Bayesian Hyperparameter Optimization

In recent years, to improve the accuracy of image classification, the structure of deep neural networks has become deeper and wider [33], [34]. However, this results in a greater number of parameters and increased computation as well as greater demands on the hardware resources for prediction and training. Deep neural network models often require servers with high computational power, placing high demands on the memory and performance of computers.

To reduce the numbers of parameters and computations without sacrificing accuracy, a CNN architecture based on DP convolution, i.e., DPCNN, is introduced [35].

The DP convolution consists of two main parts: depthwise (DW) convolution and point-wise (PW) convolution. Figure 10 shows the structures of the DP convolution and standard convolution.

As shown in Fig. 10(a), a feature map with M channels is convolved with N convolution kernels of size $D_K \times D_K \times M$ to obtain a feature map of size $D_F \times D_F \times N$. The DP convolution is the decomposition of a standard convolution into DW convolution and PW convolution. The standard convolution kernel of size $D_K \times D_K \times M$ is decomposed into M DW convolution kernels of size $D_K \times D_K \times 1$ and N PW convolutions of size $1 \times 1 \times M$, as illustrated in Fig. 10(b).

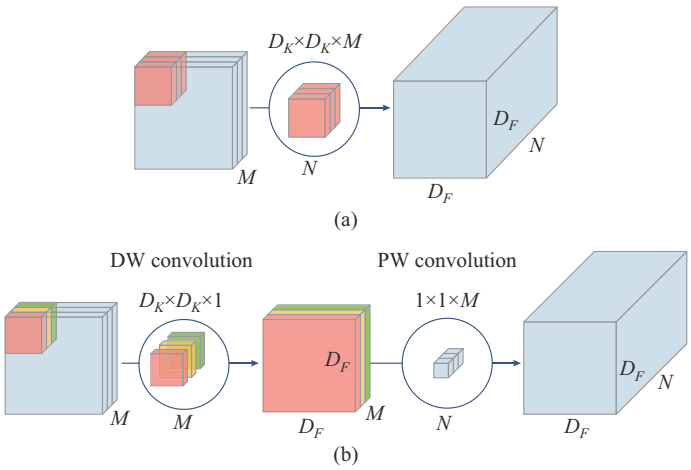

Fig. 10. Structures of convolutions. (a) Standard convolution. (b) DP convolution.

The M DW convolutions correspond to the M channels of the feature map, and each DW convolution kernel performs convolution calculations only in the aspect direction. However, a $1 \times 1 \times M$ PW convolution kernel performs convolution calculations only in the channel direction.

From Fig. 10, the ratio of the number of parameters of the DP convolution (P_D) to that of the standard convolution (P_C) can be calculated as:

$$\frac{P_D}{P_C} = \frac{D_K D_K M + MN}{D_K D_K MN} = \frac{1}{N} + \frac{1}{D_K^2} \tag{9}$$

When $D_K = 3$, P_D is close to 1/9 of P_C.

A comparison of the computational costs of the DP convolution (C_D) and standard convolution (C_C) results in:

$$\frac{C_D}{C_C} = \frac{\left(D_K D_K\right)\left(MD_F D_F\right) + M\left(ND_F D_F\right)}{\left(D_K D_K M\right)\left(ND_F D_F\right)} = \frac{1}{N} + \frac{1}{D_K^2} \tag{10}$$

The computational burden of the DP convolution is substantially reduced and is approximately 1/9 of the standard convolution when $D_K = 3$.

Most deep learning algorithms have several hyperparameters, and the settings of these hyperparameters affect the performance of algorithms. Hyperparameters are generally manually set to tune the performance of the algorithm. However, the manual tuning is often time-consuming and relies heavily on expertise. Therefore, it is imperative to develop automated algorithms to determine the correct hyperparameters.

The hyperparameter search problem can be transformed into an optimization problem, of which the purpose is to determine the optimal parameters of the model and maximize the recognition accuracy of the validation set. The hyperparametric optimization is a black-box problem, in which only the input and output of the function can be obtained during the optimization process, and the expression and gradient of the optimization objective function cannot be obtained. This feature poses difficulties for hyperparametric optimization.

Bayesian optimization can be used to solve extreme value problems for functions with unknown expressions and is perfectly suited for hyperparameter optimization [36], [37]. Bayesian hyperparameter optimization has two core components: a probabilistic regression model and an acquisition function.

1) A probabilistic regression model is used to approximate black-box objective functions. The mean and variance of the objective function are estimated from the function values of the available observations. The most common probabilistic regression models are the Beta-Bernoulli model, Gaussian processes, and random forests.

2) An acquisition function is used to determine the point next to the sample. Common collection functions include probability improvement, expected improvement, and upper confidence bound functions.

The Bayesian hyperparameter optimization is an iterative process that consists of three main steps.

Step 1: select the next most "promising" acquisition point based on maximizing the acquisition function.

Step 2: obtain the objective function value y_t based on the selected evaluation point x_t.

Step 3: add the newly obtained point (x_t, y_t) to the historical observation set and update the probabilistic regression model in preparation for the next iteration.

In this study, the Bayesian hyperparameter optimization is implemented using the scikit-optimize library in Python, which is easy to operate. Random forest is chosen as the probabilistic regression model.

Table III lists the structural features of the DPCNN before optimization, where FC represents the fully-connected layer, and GAP represents the global average pooling.

TABLE III
STRUCTURAL FEATURES OF DPCNN BEFORE OPTIMIZATION

Type	Size of convolution kernel	Number of convolution kernels	Stride	Activation function
Standard convolution	3×3×3	32	(2, 2)	ReLU6
DP convolution 1	3×3×1 (DW) 1×1×32 (PW)	32 (DW) 64 (PW)	(1, 1)	ReLU6
DP convolution 2	3×3×1 (DW) 1×1×64 (PW)	64 (DW) 128 (PW)	(2, 2)	ReLU6
DP convolution 3	3×3×1 (DW) 1×1×128 (PW)	128 (DW) 128 (PW)	(1, 1)	ReLU6
DP convolution 4	3×3×1 (DW) 1×1×128 (PW)	128 (DW) 256 (PW)	(2, 2)	ReLU6
DP convolution 5	3×3×1 (DW) 1×1×128 (PW)	256 (DW) 256 (PW)	(1, 1)	ReLU6
DP convolution 6	3×3×1 (DW) 1×1×256 (PW)	256 (DW) 512 (PW)	(2, 2)	ReLU6
DP convolution (×l)	3×3×1 (DW) 1×1×512 (PW)	512 (DW) 512 (PW)	(1, 1)	ReLU6
GAP				
FC		2		Softmax

The number of convolution kernels is adjusted by multiplying the existing number of convolution kernels by a factor k, where k is less than 1. In addition, the number of network layers is varied by adding l DP convolution layers. The numbers of convolution kernels and network layers as well as the learning rate and batch size of the DPCNN model are optimized to obtain the optimal parameter values.

B. Implementation of Proposed Protection Method

Figure 11 shows the flowchart of the proposed protection method. The specific implementation process is summarized as follows.

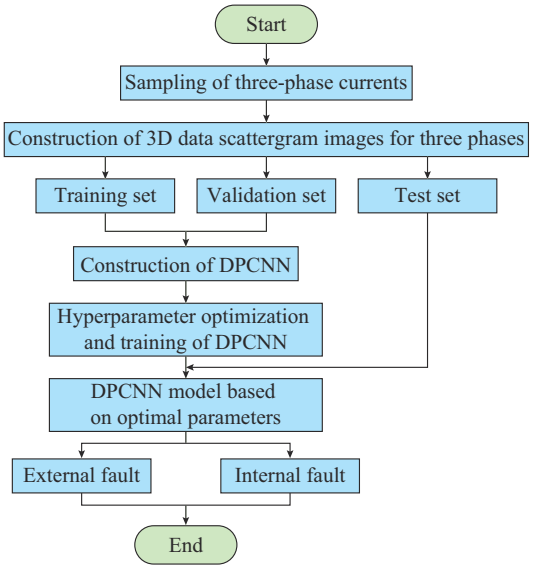

Fig. 11. Flowchart of proposed protection method.

Step 1: construct the dataset. One cycle of the three-phase current data on both ends of the transmission line is collected in real time. 3D data scattergram images are generated based on current data [i_b, i_g, i_{sum}], and these scattergram images are included in the dataset. The dataset is constructed considering multiple fault conditions, different operating states of BESS, CT saturations, and CT measurement errors. The dataset is segregated into training, validation, and test sets with a ratio of 8:1:1.

Step 2: construct and train the DPCNN. The DPCNN is constructed according to the structural features outlined in Table III. The DPCNN is trained using the training set, and the hyperparameters are continuously tuned using Bayesian hyperparameter optimization according to the accuracy of the validation set. The obtained hyperparameters are set, and the DPCNN is trained to obtain the final DPCNN.

Step 3: the 3D data scattergram images of the three phases are then fed into the trained DPCNN, and the image classification results are output to determine whether the images belong to the category of internal faults.

The proposed protection method has the natural function of identifying fault phases.

IV. Performance Evaluation

A. Dataset

The studied system with a BESS having a capacity of 150 MW, as shown in Fig. 1(a), is built in PSCAD/EMTDC.

The currents on the two ends of transmission line 33-19 are measured to construct 3D data scattergram images to generate a large dataset. Table IV lists the specific components of the dataset. Taking x=10% as an example, it represents a location at 10% of the length of transmission line 33-19 from bus 33. Different fault conditions, two operating states of the BESS, and the effects of CT saturation and errors are all considered in the dataset. Different conditions are considered to ensure dataset quality.

TABLE IV
Specific Components of Dataset

Operating state of BESS	Fault location	Fault type	Fault resistance R (Ω)	Condition	Number of 3D data scattergram images
Charging state	Internal fault: x=10%, 50%, and 90% External fault: bus 33	AG, ABG	0, 5, 10, …, 100	Normal condition	256×3
		AB, ABC	0, 5, 10, …, 50		
	Internal fault: x=50% External fault: bus 33	AG, ABG	0, 5, 10, …, 100	CT saturation	128×3
		AB, ABC	0, 5, 10, …, 50		
	Internal fault: x=50% External fault: bus 33	AG, ABG	0, 5, 10, …, 100	10% CT measurement error	128×3
		AB, ABC	0, 5, 10, …, 50		
Discharging state	Internal fault: x=10%, 50%, and 90% External fault: bus 33	AG, ABG	0, 5, 10, …, 100	Normal condition	256×3
		AB, ABC	0, 5, 10, …, 50		
	Internal fault: x=10%, 50%, and 90% External fault: bus 33	AG, ABG	0, 5, 10, …, 100	CT saturation	128×3
		AB, ABC	0,5, 10, …, 50		

Under different conditions, a total of 896 simulation experiments are conducted using PSCAD, and 2688 3D data scattergram images are obtained to generate the dataset. The dataset is segregated into training, validation, and test sets with a split ratio of 8:1:1.

B. Training and Test Results of DPCNN

The DPCNN is built according to the structural features listed in Table III, and the DPCNN is trained using the training set. The number of iterations is 80. The hyperparameters are optimized using a Bayesian optimization algorithm with the objective of maximizing the accuracy of the validation set. The hyperparameter search space and results are shown in Table V.

TABLE V
Hyperparameter Search Space and Results

Hyperparameter	k	l	Learning rate	Dropout rate
Search space	{0.25, 0.5, 0.75, 1}	[1, 5]	(0.001, 0.1)	{0.1, 0.2, 0.3, 0.4, 0.5}
Result	1	1	0.0078	0.1

In the search space, the learning rate is defined as a continuous variable with values in the range of (0.001, 0.1), and the remaining hyperparameters are defined as discrete variables with values in the ranges listed in Table V.

A comparison of the DPCNN and CNN is conducted, with the specific structural features of the CNN presented in Table VI, where C1-C6 denote six standard convolution layers.

TABLE VI
STRUCTURAL FEATURES OF CNN

Layer	Convolution layer		Activation function	Padding	Stride	Pooling type
	Number	Size				
C1	32	3×3	ReLU	Same	1	
C2	64	3×3	ReLU	Same	1	Max
C3	128	3×3	ReLU	Same	1	
C4	128	3×3	ReLU	Same	1	Max
C5	256	3×3	ReLU	Same	1	
C6	256	3×3	ReLU	Same	1	Max
FC1		512	ReLU			
FC2		2	Softmax			

Table VII shows the training and test results of the DPCNN and CNN. For DPCNN, the accuracy of the test set is 100%, and all 269 3D data scattergram images in the test set are correctly identified. Compared with the CNN, the DPCNN achieves higher recognition accuracy, and the number of parameters is significantly reduced, which is consistent with the analysis presented in Section III-A.

TABLE VII
TRAINING AND TEST RESULTS OF DPCNN AND CNN

Model	Accuracy (%)			Number of parameters
	Train set	Validation set	Test set	
DPCNN	100.00	100.00	100.00	545474
CNN	99.49	99.25	98.88	3224770

Figure 12 shows the dichotomous confusion matrices of the DPCNN and CNN, where labels 0 and 1 represent external and internal faults, respectively. All the 3D data scattergram images of the external and internal faults in the test set are correctly identified by the DPCNN. However, three images of internal faults are incorrectly identified by CNN.

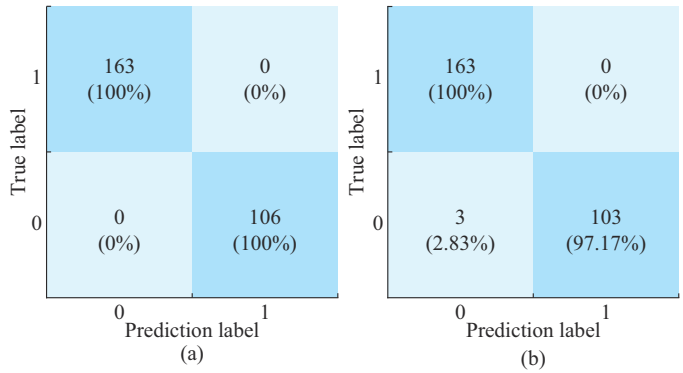

Fig. 12. Dichotomous confusion matrix. (a) DPCNN. (b) CNN.

C. Comparative Study of Different Algorithms

To highlight the notable functionality of the DPCNN, it is compared with both a CNN and an artificial neural network (ANN), and the results are shown in Table VIII. The DPCNN exhibits the highest accuracy, showing excellent feature extraction, learning, and generalization capabilities. The fault conditions involved in the test set are tested using conventional current differential protection (CDP).

TABLE VIII
COMPARISONS OF DIFFERENT ALGORITHMS

Protection method	Algorithm	Classification accuracy (%)
3D data scattergram image classification based protection	ANN	97.77
	CNN	98.88
	DPCNN	100.00
Conventional protection	CDP	88.10

Figure 13 shows the performance comparisons of different algorithms in Table VIII.

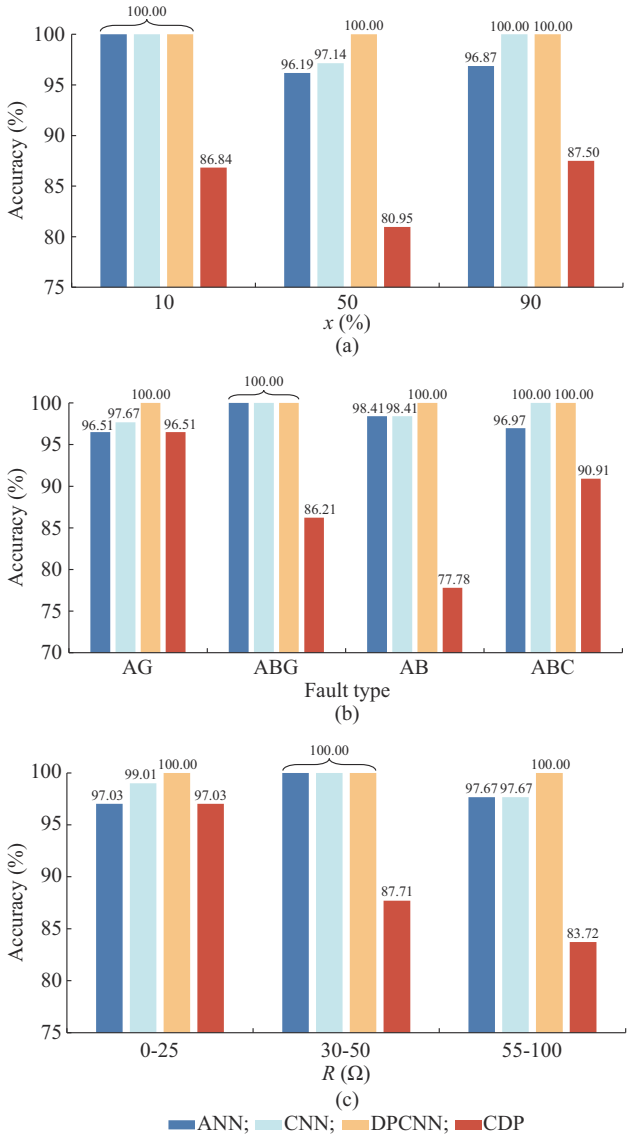

Fig. 13. Comparison of different algorithms. (a) Different fault resistances. (b) Different fault types. (c) Different fault locations.

As observed in Fig. 13, the proposed 3D data scattergram image classification based protection method exhibits better performance across different fault resistances, types, and locations compared with CDP, with a higher accuracy in the classification of internal and external faults. A comprehensive comparison of the recognition accuracies shows that the DPCNN exhibits the highest accuracy.

Five conditions from the test set are considered to demonstrate the simulation results of the CDP and the proposed protection method based on 3D data scattergram images, as shown in Fig. 14. The five conditions are detailed as follows.

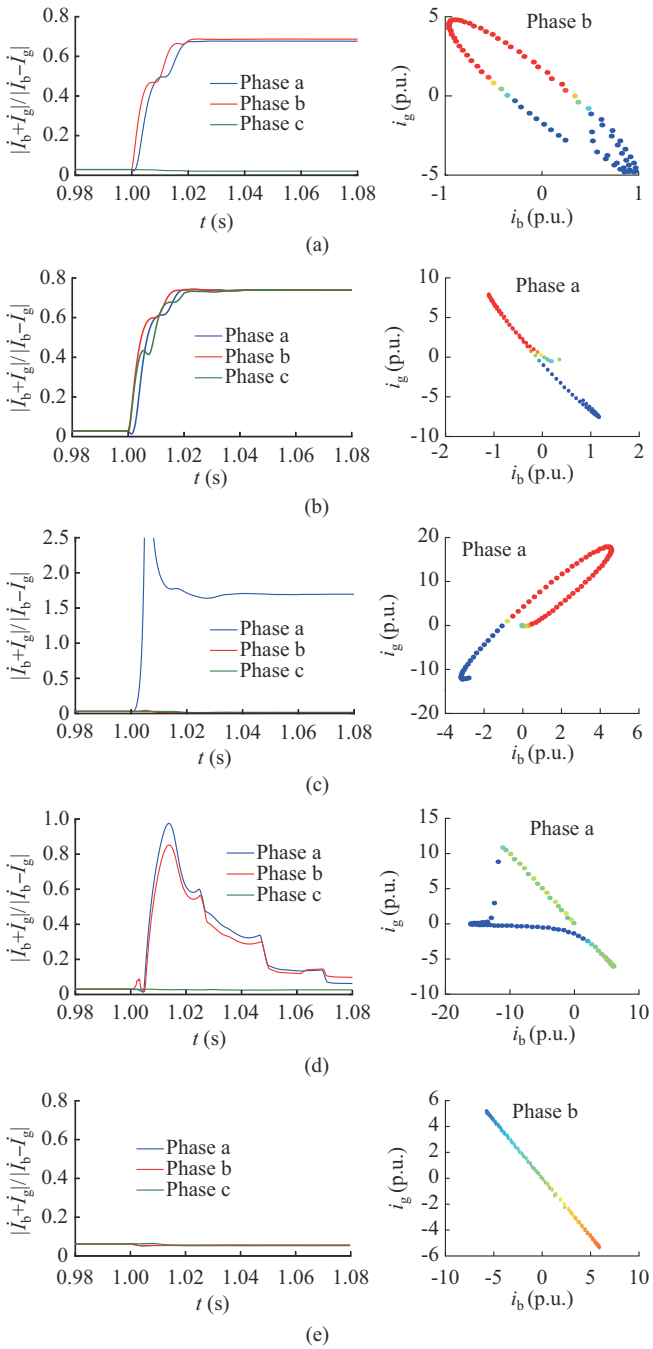

Fig. 14. Simulation results of CDP and 3D data scattergram images for fault phase under five conditions. (a) Condition 1. (b) Condition 2. (c) Condition 3. (d) Condition 4. (e) Condition 5.

1) Condition 1: an ABG-type internal fault with $R = 90\,\Omega$ and $x = 90\%$ when BESS operates in the charging state.

2) Condition 2: an ABC-type internal fault with $R = 40\,\Omega$ and $x = 10\%$ when BESS operates in the charging state.

3) Condition 3: an AG-type internal fault with $R = 5\,\Omega$ and $x = 50\%$ when BESS operates in the discharging state.

4) Condition 4: an AB-type external fault with $R = 0\,\Omega$ considering CT saturation when BESS operates in the discharging state.

5) Condition 5: an ABG-type external fault with $R = 50\,\Omega$ considering CT measurement error when BESS operates in the charging state.

As shown in Fig. 14(a), the CDP may fail to function under an internal fault when BESS operates in the charging state. Figure 14(d) shows that the CDP may misoperate under external faults with CT saturation. In Fig. 14, the 3D data scattergram images for the fault phase have the same conditions as CDP simulation. These images are identified correctly by DPCNN.

D. Comparison of Different Data Processing Methods

The 3D data scattergram highlights and magnifies the differences among the three sets of current data under internal and external faults. As described in Section II-C, the 3D data scattergram integrates information from [i_b, i_g, i_{sum}]. The 3D data scattergrams for internal and external faults differ significantly in the shape of their scatter trajectories and colors. This makes it easier for a DPCNN to identify internal and external faults. To verify the necessity and superiority of constructing 3D data scattergrams, a new dataset is constructed directly using unprocessed vector of [i_b, i_g, i_{sum}].

The machine learning algorithms including *K*-nearest neighbor (KNN), decision tree, and support vector machine (SVM) are used to train the new dataset, and the comparison results are shown in Table IX. The results clearly show that the preprocessing of 3D data to construct the 3D data scattergram is meaningful for improving the classification accuracy.

TABLE IX
COMPARISONS OF RESULTS OF DIFFERENT DATA PROCESSING ALGORITHMS

Input	Algorithm	Classification accuracy (%)
Unprocessed vector of [i_b, i_g, i_{sum}]	KNN	81.78
	SVM	84.39
	Decision tree	84.01
3D data scattergram	DPCNN	100.00

E. Performance Evaluation Based on Fault Recording Data from RT-Sim Based Experimental Platform

The system, which includes a BESS with a capacity of 150 MW, is built on the RT-Sim based experimental platform to further test the performance of the proposed protection method. Figure 15 shows the overall structure of the experimental platform. The BESS is integrated into a power system via a 50 km 220 kV transmission line. The positive- and zero-sequence impedances of transmission line are $(0.076 + j0.338)\,\Omega$/km and $(0.284 + j0.824)\,\Omega$/km, respectively,

and the positive- and zero-sequence capacitances are 0.0086 μF/km and 0.0061 μF/km, respectively. The fault recording data can be acquired using the experimental platform shown in Fig. 15, which is used to produce 3D data scattergram images.

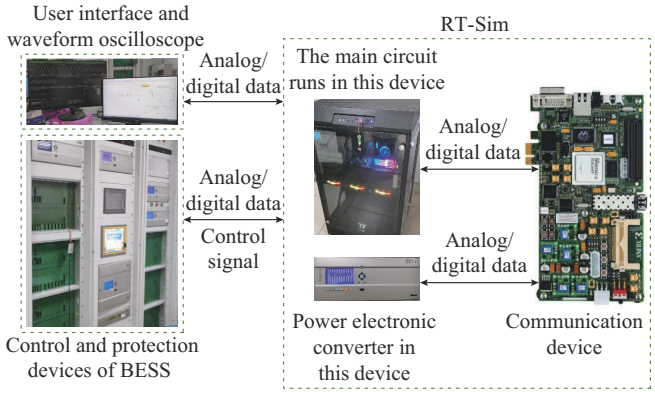

Fig. 15. Experimental platform based on RT-Sim.

To reinforce the superiority of the proposed protection method, it is compared with the CS-based method proposed in [13] and [14]. Table X lists the comparison results of the two methods. A total of 57 experiments are conducted by varying the operating states and fault conditions. All 57 experiments are correctly identified by the proposed protection method, whereas the CS-based protection method with thresholds of −0.9 and −0.2 is incorrect in 6 and 9 instances, respectively. The threshold affects the dependability and security of the CS-based protection method. When the threshold is set to be −0.9, the CS-based protection method exhibits poor security and is vulnerable to the negative effects of CT saturation. When the threshold is set to be −0.2, the security of the CS-based protection method is significantly improved. However, the dependability is markedly reduced. In comparison, the proposed protection method does not require setting the threshold. In addition, the proposed protection method is insensitive to fault resistance and can correctly identify internal faults under different fault resistances. In the case of external faults, the proposed protection method is unaffected by CT saturation and CT measurement errors and offers a higher level of security.

TABLE X
COMPARISONS OF RESULTS OF CS-BASED METHOD AND PROPOSED PROTECTION METHOD

Operating state of BESS	Fault type	Conditions	Number of experiments	Number of experiments correctly identified		
				CS-based method		Proposed protection method
				Threshold is −0.9	Threshold is −0.2	
Charging state	Internal fault	$R \in (0, 50)\,\Omega$	5	5	3	5
		$R \in [50, 100]\,\Omega$	5	5	0	5
	External fault	Without nonideal conditions	4	4	4	4
		CT saturation	3	0	2	3
		CT measurement error	4	4	4	4
Discharging state	Internal fault	$R \in (0, 50)\,\Omega$	12	12	12	12
		$R \in [50, 100]\,\Omega$	10	10	10	10
	External fault	Without nonideal conditions	6	6	6	6
		CT saturation	3	0	2	3
		CT measurement error	5	5	5	5
Total			57	51	48	57

V. CONCLUSION

Based on fault current characteristics of BESS, a 3D data scattergram image classification based protection method using DPCNN is proposed, with the following conclusions.

1) The 3D data scattergram contains almost all the information of fault currents, including the currents on both ends of transmission line and their sum. The 3D data scattergram exhibits markedly different features under internal and external faults. Therefore, it can be employed to ascertain whether faults are internal or external using the CNN-based image classification method.

2) Using DP convolution instead of standard convolution enables the DPCNN to significantly reduce the number of parameters and computational cost without sacrificing accuracy. Compared with ANN and CNN, the DPCNN achieves higher accuracy in classifying internal and external faults.

3) Compared with the CDP, the proposed protection method functions accurately when BESS operates in both the charging and discharging states, regardless of fault conditions such as fault type, resistance, and location. In addition, the proposed protection method can ensure excellent security, even with CT saturation and measurement errors.

REFERENCES

[1] M. Rouholamini, C. Wang, H. Nehrir *et al*., "A review of modeling, management, and applications of grid-connected Li-ion battery storage systems," *IEEE Transactions on Smart Grid*, vol. 13, no. 6, pp. 4505-4524, Nov. 2022.

[2] U. Datta, A. Kalam, and J. Shi, "The relevance of large-scale battery energy storage (BES) application in providing primary frequency control with increased wind energy penetration," *Journal of Energy Storage*, vol. 23, pp. 9-18, Jun. 2019.

[3] L. M. S. de Siqueira and W. Peng, "Control strategy to smooth wind power output using battery energy storage system: a review," *Journal of Energy Storage*, vol. 35, p. 102252, Mar. 2021.

[4] Z. Guo, W. Wei, L. Chen *et al*., "Impact of energy storage on renewable energy utilization: a geometric description," *IEEE Transactions*

on Sustainable Energy, vol. 12, no. 2, pp. 874-885, Apr. 2021.
[5] Y. Liang, W. Li, and G. Xu, "Performance problem of current differential protection of lines emanating from photovoltaic power plants," *Sustainability*, vol. 12, no. 4, p. 1436, Feb. 2020.
[6] A. Hooshyar, M. A. Azzouz, and E. F. El-Saadany, "Distance protection of lines emanating from full-scale converter-interfaced renewable energy power plants – part I: problem statement," *IEEE Transactions on Power Delivery*, vol. 30, no. 4, pp. 1770-1780, Aug. 2015.
[7] M. M. Alam, H. Leite, J. Liang *et al*., "Effects of VSC based HVDC system on distance protection of transmission lines," *International Journal of Electrical Power & Energy Systems*, vol. 92, pp. 245-260, Nov. 2017.
[8] J. Jia, G. Yang, A. H. Nielsen *et al*., "Impact of VSC control strategies and incorporation of synchronous condensers on distance protection under unbalanced faults," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1108-1118, Feb. 2019.
[9] K. Jia, J. Chen, Z. Xuan *et al*., "Active protection for photovoltaic DC-boosting integration system during FRT," *IET Generation, Transmission & Distribution*, vol. 13, no. 18, pp. 4081-4088, Sept. 2019.
[10] C. Chao, X. Zheng, Y. Weng *et al*., "Adaptive distance protection based on the analytical model of additional impedance for inverter-interfaced renewable power plants during asymmetrical faults," *IEEE Transactions on Power Delivery*, vol. 37, no. 5, pp. 3823-3834, Oct. 2022.
[11] Z. Yang, Q. Zhang, W. Liao *et al*., "Harmonic injection based distance protection for line with converter-interfaced sources," *IEEE Transactions on Industrial Electronics*, vol. 70, no. 2, pp. 1553-1564, Feb. 2023.
[12] P. Mishra, A. K. Pradhan, and P. Bajpai, "Adaptive distance relaying for distribution lines connecting inverter-interfaced solar PV plant," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 3, pp. 2300-2309, Mar. 2021.
[13] L. Zheng, K. Jia, T. Bi *et al*., "Cosine similarity based line protection for large-scale wind farms," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 7, pp. 5990-5999, Jul. 2021.
[14] L. Zheng, K. Jia, W. Wu *et al*., "Cosine similarity based line protection for large scale wind farms part II – the industrial application," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 3, pp. 2599-2609, Mar. 2022.
[15] A. Saber, M. F. Shaaban, and H. H. Zeineldin, "A new differential protection algorithm for transmission lines connected to large-scale wind farms," *International Journal of Electrical Power & Energy Systems*, vol. 141, p. 108220, Oct. 2022.
[16] Z. Yang, W. Liao, H. Wang *et al*., "Improved euclidean distance based pilot protection for lines with renewable energy sources," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 12, pp. 8551-8562, Dec. 2022.
[17] Y. Liang and C. Pan, "Time-domain differential protection based on operating and restraining trajectory indices for lines connecting battery storage energy stations," *Journal of Modern Power Systems and Clean Energy*, vol. 12, no. 4, pp. 1074-1086, Jul. 2024.
[18] Y. Liang, Y. Ren, and Z. Zhang, "Pilot protection based on two-dimensional space projection of dual differential currents for lines connecting MMC-HVDC stations," *IEEE Transactions on Industrial Electronics*, vol. 70, no. 5, pp. 4356-4368, May 2023.
[19] Y. Liang, C. Pan, and J. Zhang, "Current trajectory coefficient based time domain line protection for battery storage energy stations," *Journal of Energy Storage*, vol. 51, p. 104468, Jul. 2022.
[20] Y. Liang, X. Yang, Y. Wang *et al*., "Internal fault probability-based time domain differential protection applied to transmission lines connecting battery energy storage stations," *Journal of Energy Storage*, vol. 55, p. 105707, Nov. 2022.
[21] A. Jamehbozorg and S. M. Shahrtash, "A decision-tree-based method for fault classification in single-circuit transmission lines," *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2190-2196, Oct. 2010.
[22] Y. Chen, O. Fink, and G. Sansavini, "Combined fault location and classification for power transmission lines fault diagnosis with integrated feature extraction," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 1, pp. 561-569, Jan. 2018.
[23] Y. Zhu and H. Peng, "Multiple random forests based intelligent location of single-phase grounding fault in power lines of DFIG-based wind farm," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 5, pp. 1152-1163, Sept. 2022.
[24] S. Afrasiabi, M. Afrasiabi, B. Parang *et al*., "Integration of accelerated deep neural network into power transformer differential protection," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 865-876, Feb. 2020.
[25] J. Mei, R. Ge, Z. Liu *et al*., "An auxiliary fault identification strategy of flexible HVDC grid based on convolutional neural network with branch structures," *IEEE Access*, vol. 8, pp. 115922-115931, Jun. 2020.
[26] S. K. Murugan, S. P. Simon, and R. R. Eapen, "A novel signal localized convolution neural network for power transformer differential protection," *IEEE Transactions on Power Delivery*, vol. 37, no. 2, pp. 1242-1251, Apr. 2022.
[27] S. Hou, W. Guo, Z. Wang *et al*., "Deep-learning-based fault type identification using modified CEEMDAN and image augmentation in distribution power grid," *IEEE Sensors Journal*, vol. 22, no. 2, pp. 1583-1596, Dec. 2022.
[28] X. Wang, C. Li, Z. Liang *et al*., "Fault feeder detection method of distribution network based on Grampian angular field and convolutional neural network," *CSEE Journal of Power and Energy Systems*, doi: 10.17775/CSEEJPES.2022.00080.
[29] Z. Xu, Z. Tang, T. Ji *et al*., "Detection and identification of underground cable incipient faults based on denoising autoencoder and optimized convolutional neural network," in *Proceedings of 2023 International Conference on Power System Technology* (*PowerCon*), Jinan, China, Sept. 2023, pp. 1-9.
[30] Y. Liang, Y. Ren, J. Yu *et al*., "Current trajectory image-based protection algorithm for transmission lines connected to MMC-HVDC stations using CA-CNN," *Protection and Control of Modern Power Systems*, vol. 8, no. 1, pp. 1-15, Jan. 2023.
[31] Manitoba Hydro International Ltd. (2018, May). IEEE 39 bus system. [Online]. Available: https://hvdc.ca/knowledge-base/read,article/28/ieee-39-bus-system/v
[32] B. Wu, "Fundamentals of Mechanical Vibration," in *Physics*, Xi'an: Xi'an Jiaotong University Press, 2009.
[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceedings of 3rd International Conference on Learning Representations*, *ICLR 2015 – Conference Track Proceedings*, San Diego, USA, May 2015, pp. 1-12.
[34] C. Szegedy, W. Liu, Y. Jia *et al*., "Going deeper with convolutions," in *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*, Boston, USA, Jun. 2015, pp. 1-9.
[35] A. G. Howard, M. Zhu, B. Chen *et al*. (2017, Apr.). "MobileNets: efficient convolutional neural networks for mobile vision applications. [Online]. Available: https://arxiv.org/abs/1704.04861
[36] B. Shahriari, K. Swersky, Z. Wang *et al*., "Taking the human out of the loop: a review of Bayesian optimization," *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148-175, Jan. 2016.
[37] B. Bischl, M. Binder, M. Lang *et al*., "Hyperparameter optimization: foundations, algorithms, best practices, and open challenges," *WIREs Data Mining and Knowledge Discovery*, vol. 13, no. 2, pp. 1-43. Mar. 2023.

Yingyu Liang received the B.S. degree in electrical engineering and its automation from Northeast Electric Power University, Jilin, China, in 2011, and the Ph.D. degree in power system and its automation from the North China Electric Power University, Beijing, China, in 2016. He is currently an Associate Professor with the China University of Mining and Technology (Beijing), Beijing, China. His research interests include protection and control of new power system.

Yi Ren received the B.S. degree in electrical engineering and its automation from the North University of China, Taiyuan, China, in 2018, and the M.Sc. degree in electrical engineering with the China University of Mining and Technology (Beijing), Beijing, China, in 2023. Her research interests include power system protection and deep learning.

Xiaoyang Yang received the B.S. degree in electrical engineering and its automation from the China University of Mining and Technology (Beijing), Beijing, China, in 2021, and the M.S. degree in electrical engineering with the China University of Mining and Technology (Beijing), in 2024. He is currently pursuing the Ph.D. degree in electrical engineering in China University of Mining and Technology (Beijing). His research interests include protection and control of new power system.

Wenting Zha received the Ph.D. degree in control theory and control engineering from Southeast University, Nanjing, China, in 2016. She is currently an Associate Professor with the China University of Mining and Technology (Beijing), Beijing, China. Her research interests include control for complex systems and application of artifical intelligence (AI) in power system.