



OPEN ACCESS

EDITED BY

Lipeng Zhu,
Hunan University, China

REVIEWED BY

Linfeng Yang,
Guangxi University, China
Sumei Liu,
Beijing Forestry University, China

*CORRESPONDENCE

Yuanyuan Wang,
yuanyuan.wang.1980@ieee.org

SPECIALTY SECTION

This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 20 July 2022

ACCEPTED 05 August 2022

PUBLISHED 06 September 2022

CITATION

Liu X, Wang Y, Luo X, Cao C, Li W, Wang B, Wang J and Wang Y (2022), Stator single-line-to-ground fault protection for powerformers based on HSGC and CNN. *Front. Energy Res.* 10:998797. doi: 10.3389/fenrg.2022.998797

COPYRIGHT

© 2022 Liu, Wang, Luo, Cao, Li, Wang, Wang and Wang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Stator single-line-to-ground fault protection for powerformers based on HSGC and CNN

Xiaohan Liu¹, Yuanyuan Wang^{1*}, Xiaomin Luo¹, Chengjun Cao¹, Wei Li², Buming Wang³, Jiabao Wang³ and Yin Wang¹

¹School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha, Hunan, China, ²China Electric Power Research Institute (CEPRI), Beijing, China, ³Xiangtan Electric Co., Ltd. Hualian, Xiangtan, Hunan, China

It is necessary for powerformers running in parallel to identify which powerformer occurs at the stator single-line-to-ground (SLG) fault. Some state-of-the-art fusion discriminations are used to identify stator SLG fault, but these methods extract fault features artificially, and application conditions are limited. Convolutional neural network (CNN) has shown superior automatic feature extraction ability in various fields, but it cannot directly extract features from one-dimensional time series vectors collected by powerformers. Therefore, this article proposed a novel SLG fault protection scheme based on the hub-and-spoke grid data converting algorithm (HSGC) and CNN. First, Pearson product-moment correlation coefficients (PCCs) are used to calculate the correlations of one-dimensional time series vectors, establish a correspondence between them and the distance of two-dimensional grid cells, and then convert one-dimensional time series vectors to two-dimensional grid-structured data by HSGC. Second, the trained CNN automatically extracts the features of two-dimensional grid-structured data. Finally, the faulty powerformer can be identified based on the output of CNN. The proposed protection scheme is verified through the simulation of ATP-EMTP and Python. The results show that the scheme can accurately detect a faulty Powerformer under different conditions where neutral point is high-resistance or reactance grounding, even if fault resistance is 8,000 Ω .

KEYWORDS

powerformer, single-phase-to-ground fault protection, HSGC, CNN, PCCs

Abbreviations: XLPE, cross-linked polyethylene; SLG, stator single-line-to-ground; CNN, convolutional neural network; HSGC, hub-and-spoke grid data converting algorithm; PCC, Pearson product-moment correlation coefficient.

1 Introduction

High-voltage equipment used in power generation is continuously being developed in order to reduce production costs, improve efficiency, and, more recently, minimize environmental impact (Tzelepis et al., 2020; Zhou et al., 2022). Over the past decades, a series of changes have taken place in high-voltage apparatus. One of them is the powerformer (Leijon, 1998), a new type of generator developed by ABB with a stator winding made up of cross-linked polyethylene (XLPE) cable (Shi et al., 2021). The cross section of the XLPE cable is a circular winding, which can avoid uneven electric field distribution along the conductor surface caused by the shape of rectangular stator bars of the conventional generator (Wang et al., 2020a), allowing the output voltage at the terminal of powerformer to exceed about 30–35 kV and reach theoretical levels of 400 kV. Thus, it can be directly connected to the power transmission systems, which leads to lower power losses.

The stator SLG fault is the most potential harm, and it is the most frequent fault that generators will experience (Xue et al., 2022; Huang and Jia, 2017). Especially for powerformers, because of the XLPE cable with greater capacitance to ground, the fault current is greater and the damage to the powerformer is greater after the stator SLG fault occurs. Meanwhile, the stator SLG fault is very likely to develop into a more severe inter-turn fault and phase-to-phase fault, leading to the damage of stator core and stator winding of powerformers and seriously affecting the safe and stable operation of the power grid. To avoid such consequences, reliable detection and isolation of stator SLG fault have become essential for powerformers.

For a group of generators running in parallel, if an SLG fault occurs on a generator, it is necessary to selectively detect the faulty generator. Several selective protection methods for detecting stator SLG fault of powerformers have been proposed. Wang et al. (2013a) proposed using the direction and the magnitude of leakage current as identification criteria to detect the SLG fault and combining the fundamental component and the third-harmonic component of leakage current to realize 100% coverage of fault detection. However, the protection criteria were analyzed separately, and the reliability was low. To address this problem, some protection methods based on fault feature fusion have been proposed in the study by Wang et al. (2013b). Wang et al. (2013c) proposed a stator SLG fault protection scheme based on a fuzzy clustering algorithm and hierarchical clustering algorithm, which synthesizes several fault features to calculate cluster centers and then compares the distance between the detected pattern and two cluster centers to identify the faulty powerformer. Also, Fang et al. (2020) further proposed a stator SLG fault protection scheme for powerformers based on multi-dimensional information fusion. They used the discriminant analysis method to synthesize four fault features and discriminated the faulty powerformer by comparing the Manhattan distance

between fault features. Even though these methods extracted fault features artificially, their application conditions were limited.

CNN is a successful algorithm of deep learning that has been widely used in multiple fields, especially in digital image processing, face recognition, audio retrieval, and other fields (Abdalla et al., 2019; Liu et al., 2019; Rajeev et al., 2019; Klompenburg et al., 2020; Roneel et al., 2021). CNN solves the dilemma that other algorithms need to artificially extract features. Also, it finds the optimal weight parameters matrix by error back propagation along with local connections and weight sharing based on correlation between data (Zhang et al., 2019), which can automatically extract abstract and valuable features from the data to complete specific tasks (Hao et al., 2022; Huang et al., 2021). Applying CNN to powerformer protection has the potential to achieve good results.

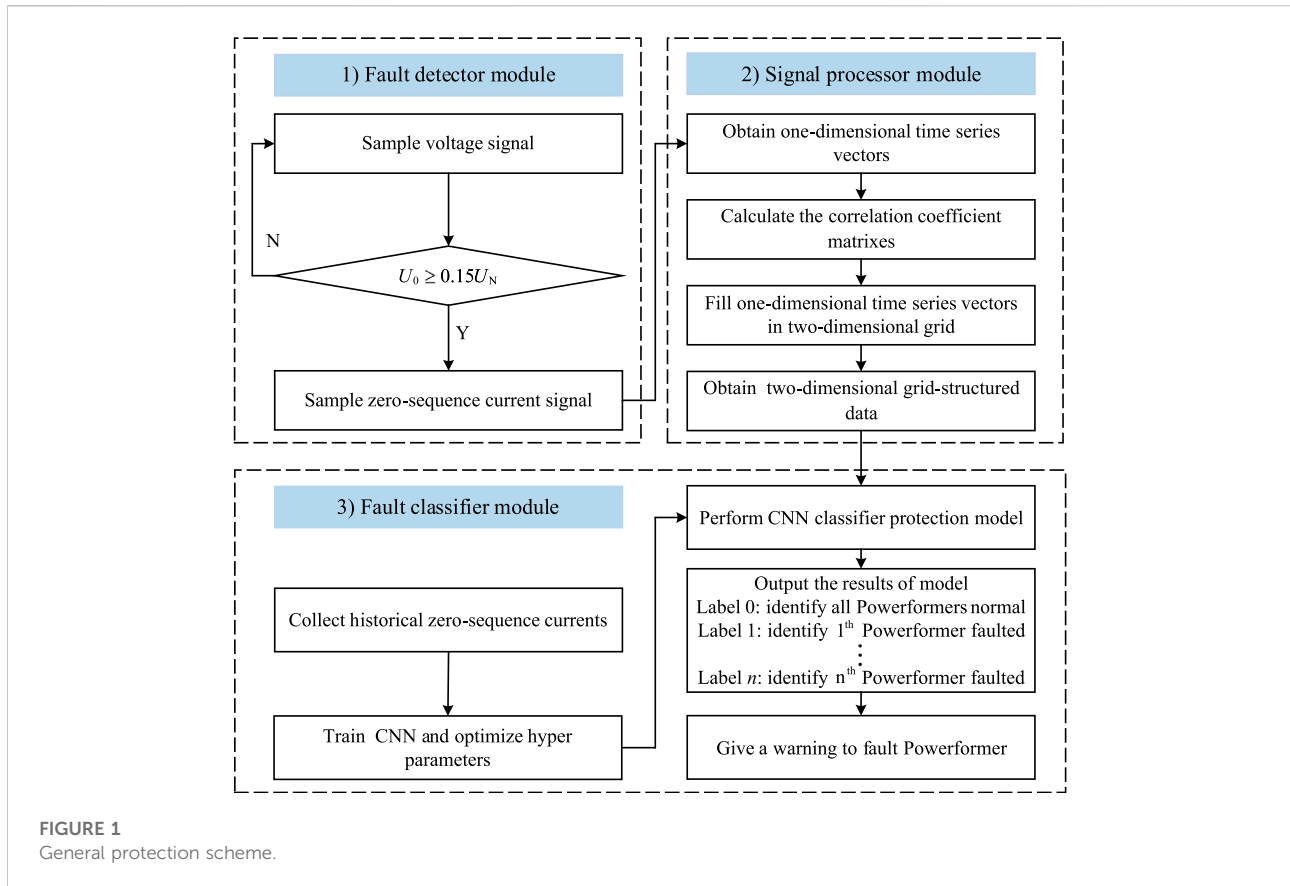
Considering that “the simpler the information, the more reliable the relay protection result”, the simplest case of the data collected by powerformer’s protection method is a one-dimensional time series vector, and there is a strong correlation between the data. CNN can only process multi-dimensional data (Liu et al., 2020) and cannot process one-dimensional time series data collected by powerformer. In order to accurately identify, which Powerformer has stator SLG fault in parallel operation, and use CNN with powerful automatic feature extraction capability to process one-dimensional time series vectors, this article proposes a data dimension transformation method based on HSGC and then uses CNN to implement powerformer stator SLG fault protection. Section 2 introduces the main modules and the framework of the proposed protection scheme. Section 3 explains the physical model of powerformer, sample set acquisition, fundamental theories of HSGC, and the protection principle. Section 4 shows the results of the ATP-EMTP and Python simulation and compares between the proposed protection scheme and other protection schemes.

2 General protection scheme

Basically, the protection scheme proposed in this article is divided into the following three modules: the fault detector module, the signal processor module, and the fault classifier module. A general protection scheme is shown in Figure 1.

2.1 Fault detector module

The zero-sequence voltage can be measured by the voltage transformer connected to the terminal of Powerformer. The effects of zero-sequence fundamental and third-harmonic voltages are eliminated by setting the reliable pickup of the overvoltage relay, which is the sum of zero-sequence fundamental voltage and zero-sequence harmonic voltage



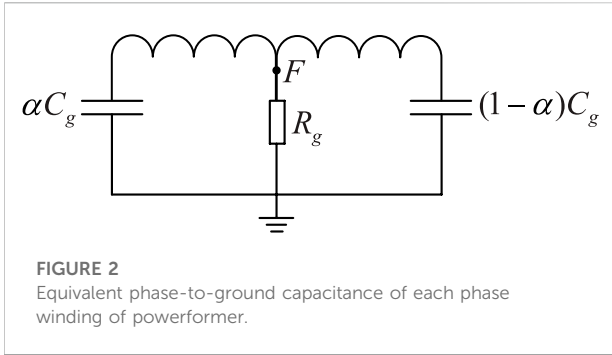
(15% phase voltage). When the zero-sequence voltage is detected to be greater than or equal to the pickup, the protection is activated, and then the zero-sequence currents can be measured by the current transformer connected to the terminal of Powerformer; conversely, it returns.

2.2 Signal processor module

First, if the fault detector protection is activated, the one-dimensional time series vectors, which consist of zero-sequence currents, are input to the signal processor module. Second, one-dimensional time series vectors are converted into two-dimensional grid-structured data by HSGC. Concretely, in order to reflect the correlation information of the one-dimensional time series vectors, the correlation coefficient matrixes of the one-dimensional time series vectors are calculated using PCCs. According to the correlation coefficient matrixes and the spatial neighborhood correlation of grid cells, the one-dimensional time series vectors are converted into two-dimensional grid-structured data. Finally, the two-dimensional grid-structured data was used as input to the fault classifier module.

2.3 Fault classifier module

In the fault classifier module, the historical zero-sequence currents of each powerformer are collected under normal or fault conditions when n powerformers are running in parallel. The collected zero-sequence currents are divided into $n+1$ classifications. The zero-sequence currents are classified into one class when all powerformers are in a normal state, and the zero-sequence currents are classified into class n when the stator SLG fault occurs on the n^{th} powerformer. After converting these zero-sequence currents into two-dimensional grid-structured data separately, they are fed into the CNN classifier protection model to train it and optimize hyper-parameters. After training, the CNN classifier protection model can be used to identify the classification of two-dimensional grid-structured data and then accurately identify the faulty powerformer. If the stator SLG fault occurs on one of the powerformers running in parallel, this module starts and gives an alarm according to the output of the CNN classifier, so that maintenance personnel can take actions to reduce the loss caused by the fault. If the powerformers are all running normally, the results can be recorded and used as a reference for subsequent research.



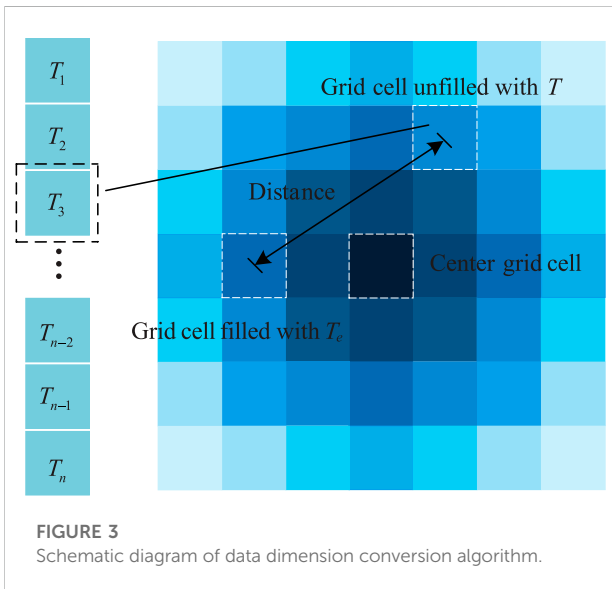
at fault. The equivalent phase-to-ground capacitance of each phase winding of the powerformer is shown in Figure 2.

After the stator SLG fault occurs on the Powerformer, the zero-sequence current of powerformers operating in parallel is measured. Also, the samples of zero-sequence currents measured from each Powerformer compose a one-dimensional time series vector.

The original data matrix T consists of one-dimensional time series vectors, which can be defined as follows:

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mn} \end{bmatrix}, \tag{1}$$

where n represents the number of powerformers operating in parallel, and m represents the number of zero-sequence current samples from each powerformer. Define each row of the original data matrix T as a one-dimensional vector T_i with n quantities.



3.2 Hub-and-spoke grid data converting algorithm

In order to take advantage of CNN, it is necessary to convert one-dimensional time series vectors into two-dimensional grid-structured data that can be directly processed by CNN. This section introduces the HSGC proposed in this article, which can increase the dimensionality of one-dimensional time series vectors.

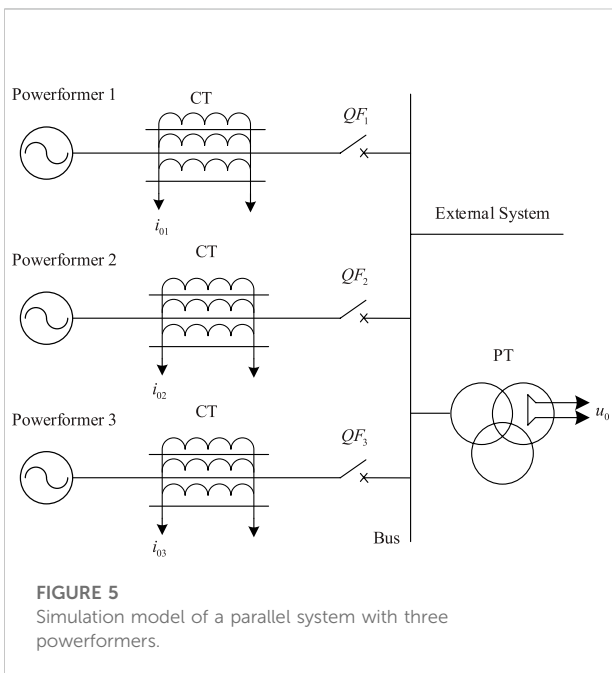
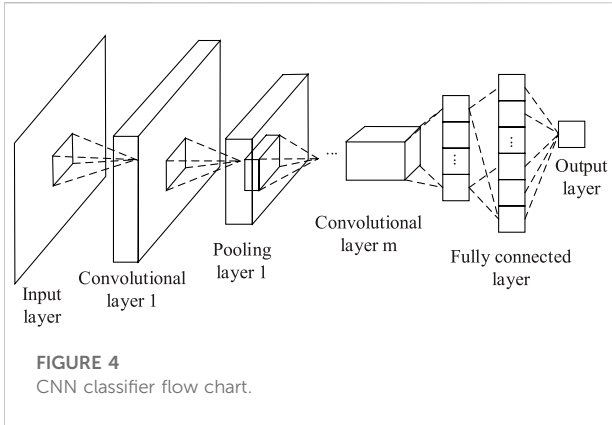
The relative positions and distances between two-dimensional grid cells can be used to describe the spatial neighborhood relations between grid cells, that is, the correlation between two-dimensional grid cells. The correlation is strongest for neighborhood grid cells and weakest for grid cells at diagonal vertices. Correspondingly, one-dimensional time series vectors also have strong correlations. Therefore, the correlation between two-dimensional grid cells can be used to reflect the correlation between one-dimensional vectors.

Based on the above mentioned correlation, in order to increase the dimension of one-dimensional time series vectors, this article first determines a one-dimensional vector to fill the center cell of the two-dimensional grid and then radially fills the remaining one-dimensional vectors into the unfilled two-dimensional grid cells one by one. Specifically, in order to fill the one-dimensional vector with the strongest correlation into the geometrical center cell of the two-dimensional grid, the two-dimensional grid should be a square with $b \times b$ cells (b is odd, $b > 1$, and $b \times b = m$) because the square has one and only one center grid cell, which meets the requirements of this article. The number of zero-sequence current samples should be consistent with the number of cells of the two-dimensional grid so that each one-dimensional vector T_i can be mapped to the cell of the two-dimensional grid one by one, and the spatial

3 Protection principle

3.1 Powerformer equivalent model and sample set acquisition

Establish an equivalent model for Powerformer. Since the Powerformer winding uses graded insulation (there are no linear relations between the ground capacitance of stator winding and turns) (Lin et al., 2018), the Powerformer cannot simply apply the equivalent model of conventional generators. When the stator SLG fault occurs on the Powerformer, this article equates the total ground capacitance C_g of one phase of the stator winding into two parts: one portion αC_g of the total ground capacitance from the neutral point to the fault location of the winding, while the rest $(1-\alpha)C_g$ of the total ground capacitance from the fault location to the powerformer terminal winding (Wang et al., 2020b; Tian et al., 2007). Then, the terminal voltage and the established equivalent capacitance are used to calculate the zero-sequence capacitance current when the powerformer is



neighborhood relations of the grid cells can well reflect the implicit correlation information of the one-dimensional vectors.

3.2.1 Center grid cell selection strategy

The grid cell at the geometrical center of the square two-dimensional grid has the strongest overall correlation. Since the one-dimensional vector T_i is mapped to the grid cell one by one, the one-dimensional vector with the strongest overall correlation is selected and then filled into the center grid cell. The PCCs of the one-dimensional vectors are applied to describe the correlation between them. The equation is as follows:

$$Corr(T_i, T_j) = \frac{Cov(T_i, T_j)}{\sigma_i \sigma_j}, \quad (2)$$

where $Corr(T_i, T_j)$ represents the variances of the one-dimensional vectors T_i and T_j , σ_i and σ_j represent the

standard deviation of the one-dimensional vectors T_i and T_j , respectively. The PCCs between m one-dimensional vectors are calculated by (2) to build an $m \times m$ correlation coefficient matrix. The sum of each row of the matrix is defined to represent the overall correlation of the corresponding one-dimensional vector. The greater the sum of PCCs of a one-dimensional vector, the stronger its overall correlation.

3.2.2 Rest cells' selection strategy

The rest of the grid cells are filled radially outward from the center grid cell. The filling order is based on the correlation of one-dimensional data, that is, the spatial neighborhood relation of the two-dimensional grid. Based on the relative distance between the unfilled grid cell and all filled grid cells, we select the one-dimensional vector with the largest score and fill it into the corresponding grid cell each time. The specific score rule is shown in Eq. 3.

3.2.3 Score rule

When filling the one-dimensional vectors into the rest of the grid cells, the average of the PCCs weighted by the relative distance is defined as the score rule for the rest cells' selection strategy.

The set of one-dimensional vectors filled into the grid is \mathbf{Q} , the set of the rest of one-dimensional vectors unfilled into the grid is \mathbf{P} , and the unfilled grid cell is defined as A ; let A_e be the grid cell filled with the one-dimensional vector T_e . For any one-dimensional vectors $T \in \mathbf{P}$, $|\mathbf{Q}| = n$, the scoring equation is as follows:

$$score(T) = \frac{1}{n} \sum_{e=1}^n \frac{Corr(T, T_e)}{Dist(A, A_e)}, \quad (3)$$

where $Dist(A, A_e)$ is the length of the connection of the geometrical center between grid cells A and A_e (the relative linear distance between them).

The $score(T)$ of the rest of the one-dimensional vectors is calculated by Eq. 3, and the one-dimensional vector with the largest $score(T)$ is selected to fill the grid cell A .

Figure 3 shows the schematic diagram of the HSGC, in which the depth of the grid cell filling color represents the priority of the filling order. The darker color of the grid cell is filled first.

3.3 Powerformer protection principle

The basic idea of identifying faulty Powerformer with the help of CNN is to perform hierarchical processing on image data, which is two-dimensional grid-structured data in this article, to extract features autonomously, to fuse features at the high level of the network, and to complete the classification task of stator SLG fault in Powerformer. Specifically, in order to identify the faulty Powerformer, the two-dimensional data converted from one-dimensional vectors are used as the input of CNN, and then the

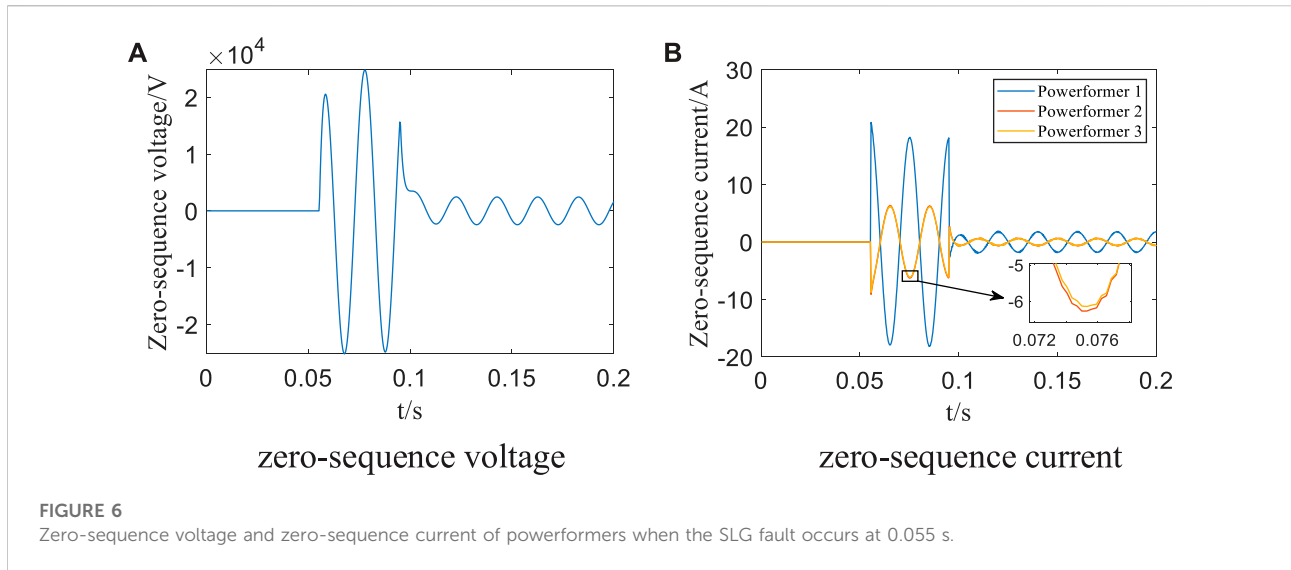


FIGURE 6
Zero-sequence voltage and zero-sequence current of powerformers when the SLG fault occurs at 0.055 s.

TABLE 1 Zero-sequence currents of powerformer 1 under different fault conditions.

Neutral grounding method	R_g	α	Zero-sequence current samples of powerformer 1
High-resistance grounding	5	0%	[-0.60788917; 0.30851999; -0.27880605; 0.10633977; -0.12194124; 0.02240053; -0.04838275; -0.00232426; -0.01602729]
	100	25%	[55.93843142; 35.60664558; 23.96540705; 18.69411469; 13.92200724; 11.99219259; 8.835983276; 7.47810173; 4.878698071]
	500	50%	[27.02622859; 25.18531545; 22.91681035; 21.36199188; 19.14410273; 17.6465613; 15.43405787; 13.83491516; 11.644967]

	7,750	75%	[2.749214808; 2.802613576; 2.765174866; 2.755957286; 2.648429235; 2.581607183; 2.420020103; 2.295909882; 2.094366391]
8,000	100%	[3.552270253; 3.621972402; 3.574234009; 3.562864939; 3.424406687; 3.338462194; 3.130012194; 2.969912052; 2.709702412]	
Reactance grounding	5	0%	[-0.59140396; 0.33029556; -0.27327219; 0.10557429; -0.13298098; 0.00188382; -0.07752037; -0.04104487; -0.06283362]
	100	25%	[55.63127136; 34.33812141; 21.76983643; 16.41232427; 12.39474996; 11.85453606; 10.50195313; 11.17452272; 10.70542415]
	500	50%	[27.03229459; 25.12030411; 22.68122482; 20.95497386; 18.6288414; 17.1318175; 15.05434736; 13.73387051; 11.96792761]

	7,750	75%	[3.16090711; 3.228677114; 3.187924703; 3.17470932; 3.043077469; 2.952977498; 2.749420484; 2.584888776; 2.328799963]
8,000	100%	[3.559144338; 3.639770508; 3.597915649; 3.586538951; 3.44129626; 3.342091242; 3.114433606; 2.929886659; 2.641470591]	

CNN autonomously extracts key information and features from it through multiple convolutions and pooling operations (Ren et al., 2020). These key information and features are fed into a fully connected layer. Also, the overall information can be obtained after fusing key information and features at the fully connected layer. The output layer outputs the classification labels

of powerformers based on the overall information. Finally, the faulty powerformer is determined corresponding to the classification label, and the classification task of CNN is completed.

Define $Y \in R^{(b \times b \times n)}$ and $Y_0 \in Y, Y_1 \in Y, \dots, Y_n \in Y$, where Y is the sample set of two-dimensional grid-structured

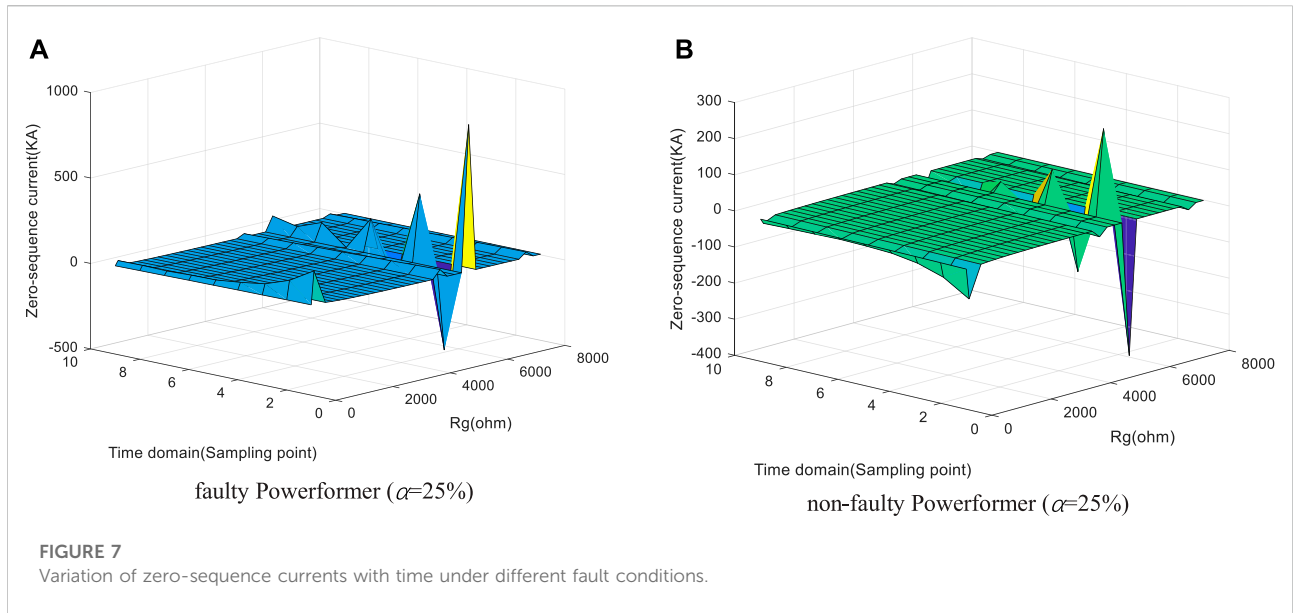


FIGURE 7
Variation of zero-sequence currents with time under different fault conditions.

TABLE 2 Correlation coefficient matrix of samples.

	T_1	T_2	T_3	...	T_8	T_9
T_1	1	0.99999	0.99999	0.99995	0.99992
T_2	0.99999	1	0.99999	...	0.99996	0.99993
T_3	0.99999	0.99999	1	...	0.99997	0.99994
T_4	0.99999	0.99999	0.99999	...	0.99997	0.99995
T_5	0.99999	0.99999	0.99999	...	0.99998	0.99996
T_6	0.99998	0.99998	0.99999	...	0.99999	0.99998
T_7	0.99997	0.99997	0.99999	...	0.99998	0.99998
T_8	0.99995	0.99996	0.99997	...	1	0.99999
T_9	0.99992	0.99993	0.99994	...	0.99999	1

data, Y_0 is the sample set when all powerformers are in a normal situation, Y_1 is the sample set when powerformer 1 occurs at the stator SLG fault, and Y_n is the sample set when Powerformer n occurs at the stator SLG fault. The key information and features of the sample y are extracted by convolution operation of convolution kernel k and sample y . The convolution operation equation is as follows:

$$Cov(x, k) = \sum_{t=1}^p x(t)k(t), \tag{4}$$

where x is the receptive field of sample y (Zhang et al., 2022), and the size of both the convolution kernel k and the receptive field x is p .

The pooling layer applies subsampling to extract key information and features and ensure the invariance of them (Chang and Shen, 2021). After multiple convolutions and

pooling operations, the key information and features are used as input to the fully connected layer to obtain overall information, which is used to obtain the probability distribution $[g_0, g_1, \dots, g_{n-1}, g_n]$ that the sample belongs to each class (g_0 is the probability of all powerformers being in a normal state, and g_n is the probability of n^{th} powerformer being in a fault state). The protection criterion is as follows:

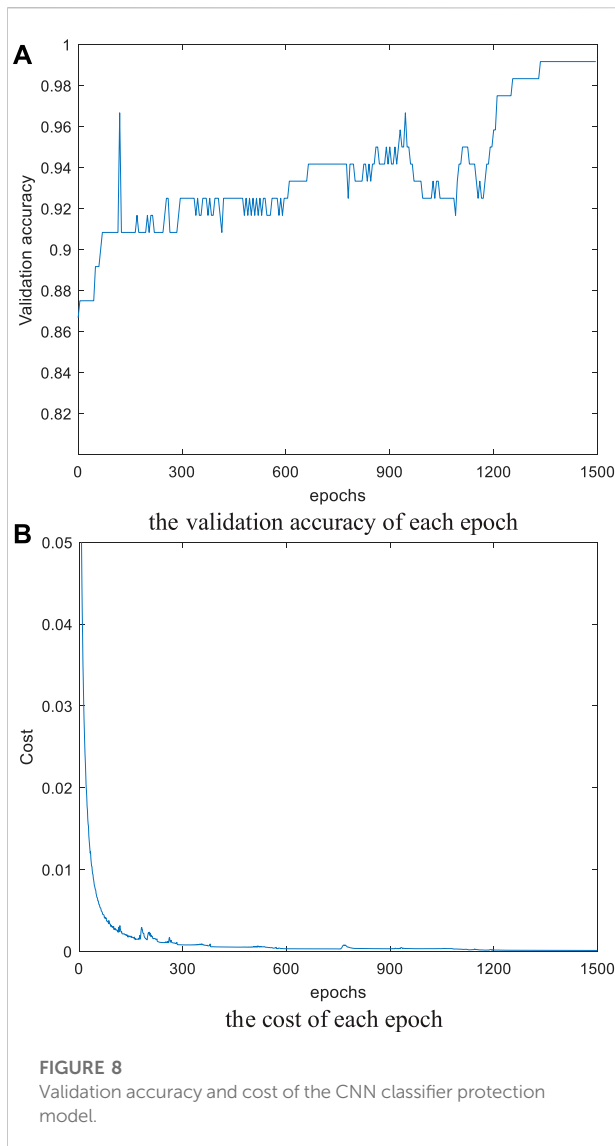
$$\begin{cases} g_0 = \max \{g_0, g_1, \dots, g_n\}, y \in Y_0 \\ g_1 = \max \{g_0, g_1, \dots, g_n\}, y \in Y_1 \\ \vdots \\ g_n = \max \{g_0, g_1, \dots, g_n\}, y \in Y_n \end{cases} \tag{5}$$

If $g_0 = \max\{g_0, g_1, \dots, g_n\}$, all powerformers are in normal situation, and the output label is 0; if $g_1 = \max\{g_0, g_1, \dots, g_n\}$, it means the stator SLG fault occurs on the Powerformer 1, and the output label is 1; \dots ; if $g_n = \max\{g_0, g_1, \dots, g_n\}$, it means the stator SLG fault occurs on the Powerformer n , and the output label is n . The CNN classifier flow chart is shown in Figure 4.

4 Simulation verification and analysis

4.1 Simulation model

To test the reliability of the powerformer protection method, a simulation model of the stator SLG fault of powerformer was established by ATP-EMTP. A parallel system with three powerformers is taken as an example (fault started at 0.055 s). The model is shown in Figure 5. Specifically, the related parameters of the model are as follows:



the rated voltage U_N of powerformers 1–3 is 150 kV; the rated frequency f_N is 50Hz; the winding ground capacitances C_{g1} , C_{g2} , and C_{g3} per phase is 0.555, 0.5574, and 0.555 μF , respectively; the ground capacitance C_t per phase of the external system is 0.15 μF ; the inception angle is 0° ; the inductance L_n is 1,838 mH, the equivalent resistance R_n is 57 Ω ; the neutral grounding resistance R_l is 1,900 Ω ; the set of fault resistance R_g is 5, 100, 500, 750, 1,000 Ω ,, 7,500, 7750, 8,000 Ω ; the stator SLG fault occurs when 0, 25, 50, 75, and 100% (α) of the stator winding are grounded.

Figure 6 shows the zero-sequence voltage and zero-sequence current of powerformers when the SLG fault occurs at 0.055 s. When the Powerformer operating in parallel occurs at a stator SLG fault, zero-sequence voltage and zero-sequence current components will be measured by the fault detector module. If $U_0 \geq 0.15U_N$, the protection is activated.

4.2 Simulation analysis

This section presents the experimental results which are performed on the parallel system with three powerformers. To increase the diversity of samples and to evaluate the efficiency of fault detection, the fault conditions, including the transition resistance, fault location, and neutral grounding method of the powerformers, were tested. The obtained data are then fed into a Python code which executes the protection scheme proposed by this article according to the block diagram in Figure 1.

Table 1 shows the one-dimensional time series vectors under different fault conditions (only partial data of Powerformer 1 is listed). In this simulation, a two-dimensional grid has nine cells ($b = 3$ and $m = b \times b = 9$). So, the one-dimensional time series vectors are composed of the first nine current sampling samples after the stator SLG fault occurs. A total of 3,813 one-dimensional time series vectors, which are the zero-sequence current samples, are obtained as the sample set. Specifically, 960 one-dimensional time series vectors are obtained from cases where all powerformers are in a normal situation, and 2,853 one-dimensional time series vectors are obtained from cases where a stator SLG fault occurs on one of the powerformers running in parallel under different fault conditions.

Figure 7 shows the variation of zero-sequence currents with time under different fault conditions. It has been found from the figure that the amplitude of zero-sequence currents of the powerformer changes smoothly under most fault conditions, while it changes drastically only in few cases. Also, the zero-sequence currents of faulty and non-faulty powerformers have similar trends over time in most cases. CNN is used to avoid the dilemma of artificial feature extraction and automatically extract features from zero-sequence currents.

Table 2 shows some data from the correlation coefficient matrix of one sample. Since the zero-sequence currents of the parallel system with three powerformers are sampled nine times at equal time intervals, nine one-dimensional vectors with three components (each representing a powerformer) can be obtained. The PCCs between each of the nine one-dimensional vectors are calculated to obtain 1,271 correlation coefficient matrixes of the sample set. According to the correlation coefficient matrixes of the sample set and the selection strategy of HSGC, the one-dimensional vectors of the sample set are converted to obtain $1,271 \times 3 \times 3$ two-dimensional grids. Each cell in the two-dimensional grid is mapped to the one-dimensional vector one to one, and the spatial adjacent relationship between the grid cells can reflect the correlation information between the one-dimensional vectors well. At the same time, two-dimensional grid-structured data can be used as input to the CNN classifier protection model.

Figure 8 shows the validation accuracy and cost of the CNN model; 90% of the two-dimensional grid-structured data is used as the training set, and the remaining 10% is used as the test set.

TABLE 3 Simulation results under different fault conditions.

Sample number	Probability distribution	Label	Result of classification
13	[1, 0, 0, 0]	0	True
72	[0, 1, 0, 0]	1	True
117	[0, 0, 0, 1]	3	True
183	[1, 0, 0, 0]	0	True
277	[0, 0, 1, 0]	2	True
324	[0, 0, 0, 1]	3	True
372	[0, 0, 1, 0]	2	True
429	[0, 1, 0, 0]	1	True

TABLE 4 Comparison results between the proposed schemes.

Fault condition	Protection scheme	Measurement data	Result
Reactance grounding powerformer 2 (50%–10%)	Fuzzy clustering algorithm	Distance from fault cluster left $d_{1g} = 2.576$	Distance from non-fault cluster left $d_{2g} = 2.9471$ $d_{1g} < d_{2g}$ (true)
	Multi-dimensional fusion	None	Unable to identify
	Scheme of this article	Probability distribution [0, 0, 1, 0]	Label = 2 (true)
High-resistance grounding powerformer 1 (5–50%)	Fuzzy clustering algorithm	Distance from fault cluster left $d_{1g} = 4.0314$	Distance from non-fault cluster left $d_{2g} = 3.8326$ $d_{1g} > d_{2g}$ (fault)
	Multi-dimensional fusion	Faulty Manhattan distance $D_{1f} = 0.442$	Sound Manhattan distance $D_{1s} = 4.624$ $D_{1f} < D_{1s}$ (true)
	Scheme of this article	Probability distribution [0, 1, 0, 0]	Label = 1 (true)

By inputting the training set into the CNN classifier protection model, it can be found that the model has high validation accuracy, low cost, and stable results.

Table 3 shows the results of feeding the test set into the trained CNN classifier protection model (due to the one-hot algorithm, the probability is only zero and one). Label 0 represents all powerformers that are in a normal situation; Label 1 represents the 1st Powerformer that occurs at the stator SLG fault; Label 2 represents the 2nd Powerformer that occurs at the stator SLG fault; Label 3 represents the 3rd Powerformer that occurs at the stator SLG fault. The simulation results show that the faulty powerformer can be accurately identified by the protection scheme based on the HSGC and CNN proposed in this article.

4.3 Comparative analysis of algorithms

A comparison between the proposed protection scheme and the other existing protection schemes of multi-dimensional

fusion (Fang et al., 2020) and a protection scheme based on a fuzzy clustering algorithm (Wang et al., 2016) is conducted. The comparison results are shown in Table 4. The protection scheme of multi-dimensional fusion is not applicable when the neutral point is grounded by an arc suppression coil due to the compensation effect of the arc suppression coil on capacitance current; the fuzzy clustering algorithm experiences misclassification under the fault condition of powerformer 1 (5–50%) due to the influence of the initial clustering center selection caused by artificial feature extraction.

In contrast, the protection proposed in this article converts the one-dimensional time series vectors into two-dimensional grid-structured data, which can be processed directly by CNN. The protection scheme avoids the artificial selection of features. Meanwhile, since local connectivity and weight sharing of CNN are based on the correlation between input data, the protection scheme uses the spatial adjacent correlation between two-dimensional grid cells to retain the correlation information between one-dimensional vectors. CNN is used to automatically extract the features of one-dimensional time

series vectors to discriminate whether the powerformer occurs at the stator SLG fault.

The simulation results show that the protection scheme proposed in this article is not affected by the system operation mode, and it can accurately identify the faulty Powerformer when the abovementioned two comparison methods cannot correctly identify the fault. The protection scheme based on HSGC and CNN can better realize the protection of parallel operation of powerformers with high accuracy, low cost, and stable results.

5 Conclusion

Some state-of-the-art fusion discriminations are used to identify which powerformer occurs at a stator SLG fault when they are operating in parallel, but these methods extract fault features artificially, and the application conditions are limited. A novel protection scheme based on HSGC and CNN has been developed in order to overcome these problems. Its specific advantages are as follows:

- 1) The protection model has superior automatic feature extraction capability and takes correlation information between one-dimensional data into account. The protection scheme has high validation accuracy, low cost, and stable results, which improves the reliability of stator SLG fault protection for powerformer.
- 2) Using the deep learning capabilities, strong generalization ability, and robustness of CNN, abstract and valuable features are automatically extracted from the zero-sequence current data to accurately identify which powerformer occurs at the stator SLG fault.
- 3) The protection model has a strong adaptive ability. There is no need to set the threshold, which can realize protection without a threshold. Also, the scheme is not easily affected by fault resistance, fault location, and neutral grounding mode. The protection scheme can be used to identify the faulty powerformer under different fault conditions.

References

- Abdalla, Y., Tariq Iqbal, M., and Shehata, M. (2019). Convolutional neural network for copy-move forgery detection[J]. *Symmetry* 11 (10). doi:10.3390/sym11101280
- Chang, M., and Shen, Y. X. (2021). Fault diagnosis strategy of a wind power bearing based on an improved convolutional neural network[J]. *Power Syst. Prot. Control* 49 (06), 131–137. doi:10.19783/j.cnki.pspc.200585
- Fang, T., Zhou, Y., and Liao, W. P. (2020). A new discrimination of stator single-phase grounding fault based on multi-dimensional fusion for Powerformer[J]. *J. Phys. Conf. Ser.* 1633 (1), 1–7. doi:10.1088/1742-6596/1633/1/012113
- Hao, H., Zhenhua, Z. H., Ziyi, Z. Y., Xu, X. X. Y., and Xu, Y. C. (2022). Insulator fouling monitoring based on acoustic signal and one-dimensional convolutional neural network. *Front. Energy Res.* 10, 906107. doi:10.3389/FENRG.2022.906107

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

XL carried out the simulation and drafted the manuscript. YW designed the project and revised the manuscript. XL participated in the design of the study and performed the statistical analysis. YW and WL participated in its design and coordination. CC, BW, and JW revised the manuscript. All authors have read and approved the final manuscript.

Funding

This work was supported by National Natural Science Foundation of China (No. 52177069), Hunan Student Innovation Training Project (No. S202210536087) and Hunan Graduate Research and Innovation Project (No. CX20210795).

Conflict of interest

Authors BW and JW were employed by Xiangtan Electric Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Huang, S. F., and Jia, W. C. (2017). A new fault location method for stator single-phase ground fault in large turbine generator[J]. *Power Syst. Prot. Control* 45 (9), 35–40. doi:10.7667/PSPC161504

Huang, X. H., Li, Y. Y., and Chai, Y. (2021). Intelligent fault diagnosis method of wind turbines planetary gearboxes based on a multi-scale dense fusion network[J]. *Front. Energy Res.* 9, 747622. doi:10.3389/FENRG.2021.747622

Klompenburg, T. V., Kassahun, A., and Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review[J]. *Comput. Electron. Agric.* 177. doi:10.1016/j.compag.2020.105709

Leijon, M. (1998). Powerformer TM-radically new rotating machine[J]. *ABB Rev.* 2 (2), 21–26.

- Lin, C., Wang, Y. Y., and Zeng, X. J. (2018). Stator single phase ground Fault protection of powerformers based on grading insulation[J]. *Proc. CSEE* 38 (23), 7052–7060+7139. doi:10.13334/j.0258-8013.psee.180189
- Liu, C. Y., Wang, Q., and Bi, X. J. (2020). Research on rain removal method for single image based on multi-channel and multi-scale CNN[J]. *J. Electronics&Information Technol.* 42 (09), 2285–2292. doi:10.11999/JEIT190755
- Liu, J. H., Chen, K. X., and Ma, J. (2019). Classification of three-phase voltage dips based on CNN and random forest[J]. *Power Syst. Prot. Control* 47 (20), 112–118. doi:10.19783/j.cnki.pspc.181337
- Rajeev, R., Vishal M, P., HyperFace, C. R., and [J] (2019). A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. *IEEE Trans. pattern analysis Mach. Intell.* 41 (1). doi:10.1109/TPAMI.2017.2781233
- Ren, H. Y., Hou Z, J., Bharat, V., Wang, H., and Pavel, E. (2020). Power system event classification and localization using a convolutional neural network[J]. *Front. Energy Res.* 8, 607826. doi:10.3389/FENRG.2020.607826
- Roneel, S. V., Hao, X., and Shlomo, B. (2021). Benchmarking audio signal representation techniques for classification with convolutional neural networks [J]. *Sensors* 21 (10). doi:10.3390/s21103434
- Shi, J. F., Ge, B. J., and Lv, Y. L. (2021). Analysis of excitation loss fault of cable winding asynchronous high-voltage generator[J]. *Electr. Mach. Control* 25 (03), 46–55+66. doi:10.15938/j.emc.2021.03.006
- Tian, Q., Lin, X. N., and Liu, L. (2007). A novel self-adaptive compensated differential protection design suitable for the generator with considerable winding distributed capacitance. *IEEE Trans. Power Deliv.* 22 (2), 836–842. doi:10.1109/tpwr.2007.893590
- Tzelepis, D., Psaras, V., Tsotsopoulou, E., Mirsaedi, S., Dysko, A., Hong, Q., et al. (2020). Voltage and current measuring technologies for high voltage direct current supergrids: A Technology review identifying the options for protection, fault location and automation applications. *IEEE Access* 8, 203398–203428. doi:10.1109/access.2020.3035905
- Wang, Y. Y., Guo, Y. S., Zeng, X. J., Chen, J., Kong, Y., and Sun, S. (2020). Stator single-line-to-ground Fault protection for bus-connected powerformers based on S-transform and bagging ensemble learning. *IEEE Access* 8, 88322–88332. doi:10.1109/access.2020.2993692
- Wang, Y. Y., Huang, X. C., and Zeng, X. J. (2020). Studies on stator single-line-to-ground faults protection for a Powerformer considering the winding electromotive force distribution[J]. *Int. J. Electr. Power Energy Syst.* 119, 321–328. doi:10.1016/j.ijepes.2020.105940
- Wang, Y. Y., Zeng, X. J., Dong, J. B. Z. Y., Li, Z. W., and Huang, Y. (2013). Studies on the stator single-phase-to-ground Fault protection for a high-voltage cable-wound generator. *IEEE Trans. Energy Convers.* 28 (2), 344–352. doi:10.1109/tec.2013.2240303
- Wang, Y. Y., Zeng, X. J., and Huang, Z. Y. (2013). Novel protection scheme of stator single-phase-to-ground fault for powerformers. *Int. J. Electr. Power & Energy Syst.* 53 (1), 321–328. doi:10.1016/j.ijepes.2013.05.010
- Wang, Y. Y., Zeng, X. J., Xu, Z. Y. Y., Yuan, J., and Huang, Y. (2013). Stator single-phase-to-ground Fault protection for bus-connected powerformers based on hierarchical clustering algorithm. *IEEE Trans. Energy Convers.* 28 (4), 991–998. doi:10.1109/tec.2013.2281491
- Wang, Y. Y., Zhou, J. M., Wei, G., Dong, Z., and Chen, H. (2016). Stator winding single-phase grounding faults protective scheme based on discriminant analysis for Powerformers with selectivity. *Int. J. Electr. Power & Energy Syst.* 77 (1), 145–150. doi:10.1016/j.ijepes.2015.11.037
- Xue, L., Sun, G. H., and Wang, X. H. (2022). Novel location method for stator single-phase grounding fault of large non-salient pole generator[J]. *Electr. Power Autom. Equip.* 42 (1), 221–224. doi:10.16081/j.epae.202109023
- Zhang, D. H., Zhang, X. W., Sun, H., and He, J. H. (2022). Fault diagnosis for AC/DC transmission system based on convolutional neural network[J]. *Automation Electr. Power Syst.* 46 (05), 132–145. doi:10.7500/AEPS20210201006
- Zhang, Q., Wang, J. P., and Li, W. T. (2019). Insulator state detection of CNN based on feedback mechanism[J]. *Trans. China Electrotech. Soc.* 34 (16), 3311–3321. doi:10.19595/j.cnki.1000-6753.tces.180962
- Zhou, H. Y., Yao, W., Ai, X. M., and Li, D. H. (2022). Comprehensive review of commutation failure in HVDC transmission systems[J]. *Electr. Power Syst. Res.* 205. doi:10.1016/j.epr.2021.107768

Nomenclature

Variables C_g total ground capacitance of one phase of the stator winding

α location of the stator winding grounded

T_i i th one-dimensional vectors of the original data matrix T

n number of powerformers operating in parallel

m number of zero-sequence current samples

σ_i standard deviation of the one-dimensional vector T_i

A_e grid cell filled in the one-dimensional vector T_e

p size of the convolution kernel and receptive field

g_0 probability of all powerformers being in a normal state

g_n probability of n^{th} powerformer being in a fault state

u result of the output of CNN

R_g ground fault resistance.

Sets

Q set of one-dimensional vectors filled into the grid

P set of one-dimensional vectors unfilled into the grid

Y set of the two-dimensional grid-structured data samples.