ISSN 2616-5775 Vol. 6, Issue 6: 100-107, DOI: 10.25236/AJCIS.2023.060616

GCA-CNN based transformer digital twin model construction and fault diagnosis and condition evaluation analysis

Ziyi Wang[1], Haoran Xu[2]

[1]*Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan, 650000, China*
[2]*School of Computer Science, Beijing Institute of Technology, Beijing, 102400, China*

Abstract: *Oil-immersed power transformer is the most important piece of equipment in the transmission system, and the stable operation of this equipment is of great significance to the normal work of the power system. At present, deep learning has been widely used in transformer condition evaluation and fault detection. About the shortage of deep learning algorithm models, this paper proposes a transformer digital twin model construction and fault diagnosis and condition evaluation analysis based on gray clustering algorithm (GCA) and conventional neural network (CNN). The data are first collected and filtered by combining the operation data, condition information amount and faulty features of power transformers, and then the condition features and fault features are combined to evaluate and detect the condition and faults of power transformers using GCA-CNN respectively. The whitening weight function is determined by expert scoring, the state evaluation matrix is established, the evaluation coefficients are obtained to calculate the evaluation weights, and the transformer state is obtained according to the clustering coefficients; 2000 pieces of raw data are input into the model to obtain the output fault types. Finally, the results are derived and compared with the real results. This paper uses real data from a power plant in Yunnan, and according to the results, the model established in this paper has higher accuracy and better evaluation and detection effects.*

Keywords: *Status Assessment; Fault Detection; Gray clustering algorithm; Conventional neural network*

1. Introduction

Oil-immersed power transformers are an important part of the power system, and it is important to accurately assess their operational status and discover potential faults to ensure the safe and stable operation of the power system [1]. In recent years, with the rapid development of artificial intelligence, plenty of studies on the condition assessment and fault detection of power transformers using deep learning have emerged, and the main networks include DNN, CNN and RNN [2]. With the development of the new generation of information technology, fault detection of the system through the digital world established using digital twin technology is important to understand the dynamic laws within the system [3]. Therefore, a real-time monitoring system with digital twin is established to verify the feasibility of the artificial intelligence algorithm model by accurately and quantitatively assessing the status and faults of power transformers.

At present, many scholars at home and abroad have made a lot of research results for the condition evaluation and fault diagnosis strategy research of large power transformers, mainly around two aspects of multi-algorithm fusion model and single algorithm optimization model. Firstly, in terms of multi-algorithm fusion model research, the literature [4] by constructing a condition assessment system for power transformers based on AHP and association rules. AHP was used to divide the system into three levels, association rules were selected to determine the weight of the scheme level, and hierarchical analysis was selected to determine the weight of the criterion level. The literature [5] combines the advantages of fuzzy theory, expert systems, and artificial neural network techniques for the effective analysis of power transformers. The literature [6] used Bayesian-optimized CNN-bi-RNN to predict short-term electric loads. Secondly, in terms of single algorithm optimization model research. The literature [7] constructs an FCM-based condition assessment model. The unsupervised learning algorithm is used to represent different transformer condition classes by clustering results. In the literature [8], FMADM is constructed to evaluate power transformers. According to the weight of each attribute, the transformer status is classified into four stages: "Very good, good, poor, and very poor". A continuous

ISSN 2616-5775 Vol. 6, Issue 6: 100-107, DOI: 10.25236/AJCIS.2023.060616

hidden semi-Markov model to improve the accuracy of state assessment was constructed in the literature [9]. The coupled hidden semi-Markov model of different states is established to improve the accuracy of transformer condition assessment. The literature [10] uses an AHP-Grey fixed-weight clustering model to make a quantitative assessment of the transformer, which is judged to be in an abnormal state based on the clustering coefficients, and a preliminary estimation of the cause. The literature [11] evaluates the predicted transformer load based on the DBN model with reverse correction by BP algorithm to optimize the model parameters. The BO was introduced in the literature [12] to tune the hyperparameters to optimize the performance and improve the prediction accuracy. However, the current research algorithms are still lacking to a certain extent: the lack of accuracy of the detection method and the insufficient amount of state information is an urgent problem to be solved.

In summary, to address the shortcomings of the above artificial intelligence algorithm model, this paper constructs a digital twin system model, provides detailed statistics on power transformer attributes, and proposes an evaluation system and a fault detection system based on the GCA-CNN. Firstly, the status of power transformers is classified into five categories: "normal, warning, abnormal, serious, and damaged", secondly, power transformers are evaluated using a gray clustering algorithm, and each category is classified again into "A, B, and C", and finally, power transformers are classified into three categories using CNN. A neural network algorithm is used for fault detection of power transformers. The model established in this paper is highly accurate and has a guiding significance for the safe and stable operation of transformers.

2. Fundamentals of the algorithm model

2.1 Basic structure and operating principle of power transformers

This paper studies a three-phase, three-winding oil-immersed power transformer, as showed in Figure 1. It is composed of three single-phase transformers, and the main components are the core, windings, case, transformer oil, etc. The core is the core, which is made of silicon steel sheets alternately laminated according to E and I. The core columns on the primary and secondary sides are wound with coils; the transformer oil plays the role of heat dissipation and insulation protection.

Power transformers operate on the principle of electromagnetic induction, where a changing magnetic flux in the coil generates an induced electric potential, a changing current, and thus a magnetic field. The change in magnetic flux is transferred to the secondary coil. When the number of turns of the secondary coil is less than that of the primary coil, the resulting electric potential is smaller than the primary electric potential, and the opposite is greater, thus changing the voltage.

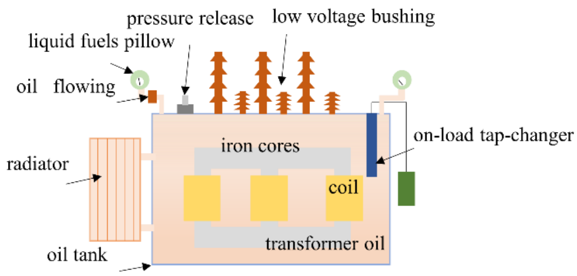

Figure 1: Physical entity model of oil-immersed power transformer

2.2 The physical architecture of the digital twin concept

With the proposal of "double carbon" and building a digital and intelligent power system, the digital twin grid technology was born. Designing the same digital space as the physical world, simulating the physical world to the virtual world. As shown in Figure 2, this paper will build a digital power transformer system to evaluate the operating conditions and faults of transformers. The twin digital power transformer creates a digital copy of the life cycle of the power transformer from design, operation, maintenance, for product updates.

ISSN 2616-5775 Vol. 6, Issue 6: 100-107, DOI: 10.25236/AJCIS.2023.060616

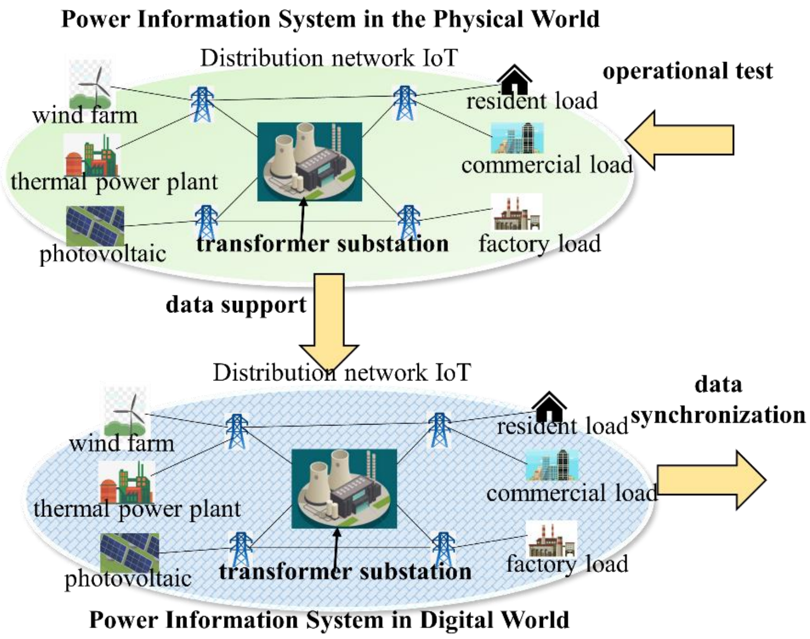

Figure 2: Relationship between the physical and digital worlds of powerful information systems

2.3 Principle of condition assessment of powerful transformers

2.3.1 Feature extraction of power transformer operation status

Oil-immersed power transformers use oil as the insulating substance. As the working time increases the insulating substance is affected by heat and electricity, the insulation performance decreases and the safety of equipment operation decreases. As showed in Table 1 this paper evaluates the condition of the transformer from the electrical aspect.

(1) Temperature: consider the top oil temperature and winding hot spot temperature. When the oil temperature is very high, the oil quality deteriorates and the insulation material ages.

(2) Operating voltage and current value: transformer operation process exceeds the normal voltage and current value, in an abnormal state.

(3) Electric field strength: when the electric field strength exceeds the transformer insulation critical value, the internal discharge affects the transformed state.

(4) Gas components: the gas in the oil under normal condition consist of oxygen and nitrogen. Due to high temperature or discharge decomposition, there may be H_2, CO, CO_2 and other gases.

(5) The dielectric loss factor: the running time increases, the oil is mixed with impurities, the charged matter increases, and the dielectric loss factor increases.

(6) Acid value: the acid contained in the oil makes the conductivity of the oil increase, the insulation ability decreases, and reduces the transformer's ability.

Table 1: Transformer properties (condition assessment indicators)

Attribute	**Condition Assessment Index**				
	Normal	Early warning	Aberrant	Seriously	Spoilage
Acid value (mgKOH/g)	0-0.02	0.021-0.1	0.11-0.3	0.31-0.4	0.41-0.5
Dielectric loss factor	0-0.02	0.021-0.04	0.041-0.05	0.051-0.06	0.061-0.07
Winding end electric field strength(kV/mm)	≤50	50-60	60-70	70-80	≥80
Total hydrocarbon content of gas in oil	0-40	41-80	81-100	101-120	≥121
Oil temperature(℃)	≤85	86-89	90-92	93-95	≥96

ISSN 2616-5775 Vol. 6, Issue 6: 100-107, DOI: 10.25236/AJCIS.2023.060616

2.3.2 Principle of grey clustering algorithm for condition assessment of transformers

Power transformers are typically gray systems. For the evaluation index Y_i (i=1, 2, 3, 4, 5), the gray evaluation coefficient of the Nth gray class is $X_{i,N}$, the total gray evaluation coefficient is Xi, and the gray evaluation weight belonging to the Nth evaluation gray class is $r_{i,N}$. Then:

$$X_{i,N} = \sum_{k=1}^{p} f_N(d_i^k) \tag{1}$$

$$X_i = \sum_{N=1}^{n} X_{i,N} \tag{2}$$

$$r_{i,N} = X_{i,N} / X_i \tag{3}$$

The combined clustering coefficients of the evaluated objects concerning the Nth gray class are:

$$\xi_N = \sum_{i=1}^{m} (r_{i,N}\omega_i) \tag{4}$$

Because $\max\limits_{1\le N\le n}\{\xi_N\} = \xi_N^*$,it is possible to determine which gray category the evaluation index belongs to. The schematic diagram of transformer condition evaluation is shown in Figure 3.

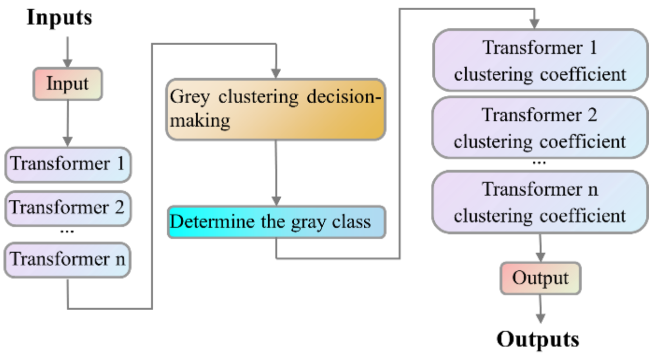

Figure 3: Power transformer condition assessment schematic

2.4 Fault detection principles of power transformer

2.4.1 Feature extraction for power transformer fault detection

This paper analyzes internal faults. Thermal faults produce gas that is CH_4 and C_2H_4. The temperature rises, the percentage of C_2H_4 rises, and severe overheating produces C_2H_2. Discharge faults produce the gas C_2H_2 and H_2, followed by C_2H_4 and CH_4. Detection is performed by comparing the special gas content with the oil note value. The modified three-ratio method fault types are shown in Table 2.

Table 2: Modified three-ratio method of fault type determination

C_2H_2/ C_2H_4	CH_4/ H_2	C_2H_4/ C_2H_6	Fault type
0	0	0	Low temperature overheating(<150℃)
0	2	0	Low temperature overheating(150-300℃)
0	2	1	Middle temperature overheating(300-700℃)
0	0,1,2	2	High temperature overheating(>700℃)
0	1	0	Partial discharge
2	0,1	0,1,2	Low energy discharge
2	2	0,1,2	Low energy discharge and overheating
1	0,1	0,1,2	Arc discharge
1	2	0,1,2	Arc discharge and overheating

2.4.2 Principle of CNN model for power transformer fault detection

Applying CNN to transformer fault detection can further improve the accuracy rate. As shown in Figure 4, CNN consists of an input layer, a convolutional layer, a pooling layer, a fully connected layer

ISSN 2616-5775 Vol. 6, Issue 6: 100-107, DOI: 10.25236/AJCIS.2023.060616

and an output layer. Among them, the convolutional layer extracts the input features; the pooling layer reduces the amount of data and suppresses overfitting; the fully connected layer enhances the model's ability to recognize features; and the activation function provides nonlinear features.

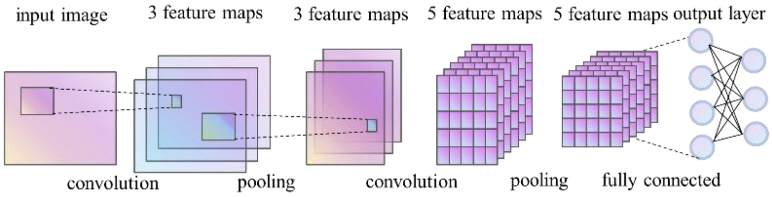

Figure 4: Basic structure of a convolutional neural network

The transformer produces CH_4, C_2H_6, C_2H_4, C_2H_2 and H_2. Middle and low-temperature overheating, high-temperature overheating, low-energy discharge, high-energy discharge, and partial discharge are selected as the model output fault types. Suitable data are selected, and the CNN model is used to test the data and validate it. The model structure is shown in Figure 5.

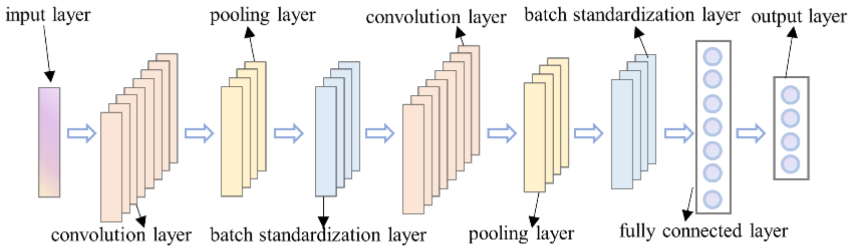

Figure 5: Convolutional neural network model diagram

The input layer is 15×1 one-dimensional arrays. Convolutional layer 1 has 15 convolutional kernels of size 5×1 with a step size of 1. Pooling layer 1 has 10 filters of size 3×1 with maximum pooling. Convolutional layer 2 is 25 convolutional kernels of size 5 × 1 and step size 1. Pooling layer 2 are 15 filters of size 3 × 1 with maximum pooling.

3. Example analysis

3.1 Sample data set generation

The experiments were developed using PyTorch framework, based on python 3.6 environments, on a computer with the operating system for Windows 11 and CPU Intel Core i5 CPU.

The data in this paper were obtained from a power plant in Yunnan. The 2000 data were obtained by data cleaning, adding missing data, removing unwanted data, modifying the format, and performing correlation verification. The data set was divided into training set and test set according to 7:3, and the number distribution is shown in Table 3.

Table 3: Sample distribution of the training and test sets

Transformer ' s condition	Number of training sets	Number of test sets
Normal	196	84
Low-middle temperature overheating	280	120
High temperature overheating	280	120
Low energy discharge	287	123
High energy discharge	287	123
Partial discharge	70	30

3.2 Analysis of experimental results

3.2.1 Analysis of power transformer condition assessment model results

0~20 is "damage", 21~40 is "serious", 41~60 is "abnormal", 61~80 is "warning ", 81 ~ 90 is "normal". The five states of the transformer are divided into five gray categories, and the whitening power function is defined as follows:

ISSN 2616-5775 Vol. 6, Issue 6: 100-107, DOI: 10.25236/AJCIS.2023.060616

$$\begin{cases} f_j^1[-,10,20] \\ f_j^2[21,30,40] \\ f_j^3[41,50,60] \\ f_j^4[61,70,80] \\ f_j^5[81,90,100] \end{cases} \tag{5}$$

Where f^1_j means "damage", f^2j means "serious", f^3j means "abnormal", f^4j means "warning ", f^5j means "normal". Where j is the decision indicator.

Based on tracking and monitoring of the transformer, it was determined that it had not reached its life span and had no adverse operating records. First, experts were hired to score the transformer's condition and establish the transformer's condition assessment matrix D:

$$D=\begin{vmatrix} 50 & 35 & 70 & 82 \\ 18 & 85 & 57 & 65 \\ 37 & 19 & 84 & 70 \\ 53 & 30 & 15 & 92 \\ 43 & 65 & 29 & 18 \\ 69 & 38 & 88 & 19 \end{vmatrix} \tag{6}$$

Five assessment coefficients were obtained under the whitening weight function. According to equation (1) N=1, the gray evaluation coefficients are $X_{1,1}=f^{(50)}{}_1+f^{(35)}{}_1+f^{(70)}{}_1+f^{(82)}{}_1=0$; N=2, $X_{1,2}$=0.5; N=3, $X_{1,3}$=1.9; N=4,$X_{1,4}$=2.9;N=5,$X_{1,5}$=0.1;According to equation (2) the total evaluation factor $X_1=X_{1,1}+X_{1,2}+X_{1,3}+X_{1,4}+X_{1,5}$=5.4.According to equation (3), the five gray evaluation weight of indicator Y_1 is $r_{1,1}$=0,$r_{1,2}$=0.0926,$r_{1,3}$=0.3519,$r_{1,4}$=0.5370,$r_{1,5}$=0.0185. Similarly, evaluation weight of Y_2,Y_3,Y_4,Y_5 is obtained. Its weight index matrix is:

$$r=\begin{vmatrix} 0 & 0.0926 & 0.3519 & 0.5370 & 0.0185 \\ 0.0377 & 0.1887 & 0.2453 & 0.4528 & 0.0755 \\ 0.0152 & 0.1970 & 0.3030 & 0.4394 & 0.0455 \\ 0.0549 & 0.2088 & 0.2967 & 0.3297 & 0.1099 \\ 0.0263 & 0.2368 & 0.2895 & 0.4474 & 0 \\ 0.0172 & 0.0345 & 0.3448 & 0.4828 & 0.1207 \end{vmatrix} \tag{7}$$

The index weight is:w_1=0.4, w_2=0.01, w_3=0.15, w_4=0.25, w_5=0.3, w_6=0.5. According to equation (4), the integrated clustering coefficients about the five gray classes are x^*_1=0.0329, x^*_2=0.2240, x^*_3=0.5220, x^*_4=0.7433, x^*_5=0.1028; thus, the evaluation index belongs to the gray 4th class, the early warning status. Repeating the above steps, it is found to be in class " A" of the early warning status, indicating that it has some kind of fault. As shown in Figure 6, the results of five transformers are listed, which are in the abnormal, warning, normal, abnormal and severe states.

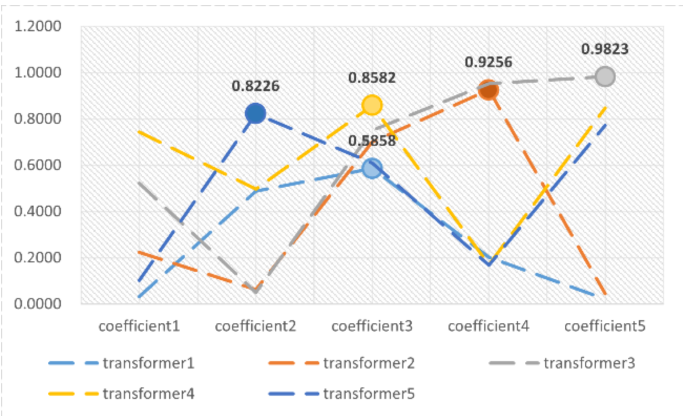

Figure 6: Transformer condition evaluation diagram

3.2.2 Power transformer fault detection model results analysis

The training and test sets were separated in the ratio of 7:3. After 500 simulations of training, as shown in Figure 7, the fault diagnosis accuracy of the model was finally maintained at about 91%.

ISSN 2616-5775 Vol. 6, Issue 6: 100-107, DOI: 10.25236/AJCIS.2023.060616

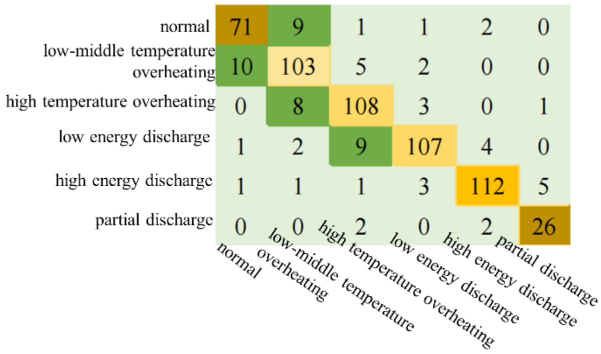

Figure 7: Transformer fault detection confusion diagram

For transformer fault detection, the original samples were selected as input data, and the first five transformer raw data are shown in Table 4 to judge the actual accuracy of the model to identify the fault type, and the model fault detection results are shown in Table 5.

Table 4: Gas data in power transformer oil

H_2	CH_4	C_2H_6	C_2H_4	C_2H_2	Transformer ' s condition
18.25	80.70	4.45	3.35	0	Low-middle temperature overheating
27.80	70.20	4.90	130.80	0	High temperature overheating
177.00	28.11	5.12	17.69	53.95	Low energy discharge
25.57	90.45	47.80	94.67	0.45	Low-middle temperature overheating
198.06	47.89	14.00	118.05	130.00	High energy discharge

Table 5: CNN model diagnostic results

Normal	Low-middle temperature overheating	High temperature overheating	Partial discharge	Low energy discharge	High energy discharge
0.00	0.99	0.01	0.00	0.00	0.00
0.00	0.05	0.94	0.00	0.00	0.01
0.00	0.00	0.00	0.00	0.97	0.03
0.00	0.99	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	1.00

As shown in Table 5, the CNN model was able to accurately identify the fault type and further detect the transformer with the results of "A, B, B, C, C". Based on the partial fault detection results obtained from Table 4 and Table 5, a scatter plot of the actual transformer faults and the model detected faults are made as showed in Figure 8. It can be seen that the CNN model detection results can correctly detect the transformer fault type with 100% recognition rate.

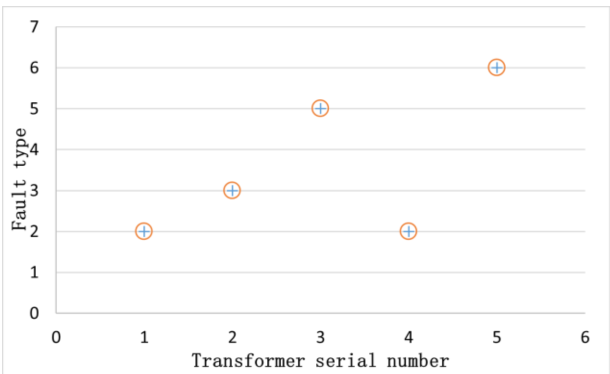

Figure 8: CNN model fault detection and diagnosis results

4. Conclusion

This paper applies digital twin technology to system condition assessment and fault detection, provides detailed statistics on power transformer operation data, amount of condition information, and

ISSN 2616-5775 Vol. 6, Issue 6: 100-107, DOI: 10.25236/AJCIS.2023.060616

proposes an assessment system and a fault detection system based on GCA-CNN.

Firstly, the status of power transformers is classified as "normal, warning, abnormal, serious, and damaged". Secondly, GCA is used to evaluate the power transformers, and each category is again classified into "A, B, and C". Finally, CNN is used for fault detection of power transformers. This model is highly accurate and has a guiding meaning for safe and stable transformer operation.

References

[1] Xian Richang, Fan Huifang, Li Fei, et al. Fault diagnosis of power transformers based on improved GSA-SVM model [J]. Intelligent Power, 2022, 50(6):50-56.

[2] Fan Lu, Lu Yuncai, Tao Fengbo, et al. Application of artificial intelligence in partial discharge detection (II): pattern recognition and state assessment [J]. Insulation Materials, 2021, 54(7):10-24.

[3] Chen Qipeng. Research on the key technology for monitoring the state of automated production line manufacturing process oriented to digital twin [D]. Guizhou: Guizhou University, 2021.

[4] K.Y. Tan. Condition assessment and life prediction of power transformers [D]. Hunan: Hunan University, 2018.

[5] Fu Yang, Qiao Fei. A comprehensive analysis method of power transformer condition based on intelligent technology [J]. Journal of Shanghai Electric Power Institute, 2004, 20(3):1-6.

[6] Wu H, Shen G Q, Lin X, et al. A transformer-based deep learning model for recognizing communication-oriented entities from patents of ICT in construction[J]. Automation in Construction, 2021, 125(May 2021):103608.

[7] Luo Weiming. Research on power transformer condition assessment and fault diagnosis method [D]. Guangdong: Guangdong University of Technology, 2019.

[8] Deng Y, Wang XH. Research on transformer condition evaluation method based on CHSMM [J]. Journal of Electric Power, 2021, 36(1):7-16.

[9] Xie P. Research on health management system of oil-immersed power transformers based on data and model [D]. Guangdong: South China University of Technology, 2020.

[10] Liu C. F., Luo R. C., Jiang Q. Q., et al. Transformer group health state assessment based on gray clustering and fuzzy comprehensive evaluation [C]. // Proceedings of the Twenty-seventh Annual Academic Conference of Chinese Higher Education Institutions on Power Systems and Their Automation. 2011: 1-5.

[11] Wang Xingxun, Wang Yuexiang, Ni Kangting, et al. Deep learning based on transformer load capacity assessment and application [J]. Electrical Engineering Technology, 2022(15):42-45, 55.

[12] Xuejiao Gong, Ruijin Zhu, Bo Tang. Short-term peak load prediction based on Bayesian optimization XGBoost [J]. Power Engineering Technology, 2020, 39(6): 76-81.