Membrane Technology
ISSN (online): 1873-4049

Research on the Construction of Fault Prediction Model for Flexible DC Converter Valve Submodule Network Based on Deep Learning

Peng Li[1,2], Taihu Ma[3,*], Haiwang Jin[1,2], Wenya Li[3], Jingzheng Xu[1,2], Hui Zhou[3]

[1] *State Grid Jibei Electric Power Co., Ltd. BEIJING EHV Power Transmission Company.Beijing,102488, China*
[2] *Hebei Key Laboratory of Equipment and Technology Demonstration of Flexible DC Transmission, Tianjin 300100, China*
[3] *Xuji Flexible Power Transmission Co., Ltd. Xuchang, 461000, China.*
` *Corresponding author

Abstract:
Along with the wide use of DC power network in the modern electric power system, the fault of the sub-module network of converter valve has become a key factor which influences the stability of the power system. For the purpose of forecasting these failures and enhancing the stability of the power system, a new model of the deep learning fault forecast is presented in this paper, which is based on the Transformer and GCN. This model combines the global time series modeling capability of Transformer with the topological relationship modeling capability of GCN, and can simultaneously process the time series data of converter valve submodule and the spatial topological structure information between modules, thereby enhancing the accuracy of fault prediction. Experimental results show that this model has significant advantages over traditional machine learning methods (SVM, RF) and other deep learning models (LSTM, CNN + LSTM). The Transformer + GCN model achieved 96.2%, 94.7% and 95.4% in accuracy, recall and F1 score respectively, which greatly surpassed the performance of traditional models. It is proved that the combination of time series and topology characteristics of the deep learning model can increase the precision and robustness of the fault forecast.

Keywords: Flexible DC Grid; Converter Valve Submodule; Fault Prediction; Transformer; Graph Convolutional Network; Deep Learning; Time Series Data; Topological Data.

INTRODUCTION

With the advancement of global energy transformation, flexible DC grid (FDC) as a new type of power transmission has gradually played an important role in modern power systems. With its unique advantages, flexible DC grid provides a more flexible and economical solution in efficient, long-distance, and cross-regional power transmission. However, with the gradual deepening of its application, the stability problems of flexible DC grid systems have gradually been exposed, especially the networking failure of converter valve submodules, which has become an important factor restricting the reliability and safety of the system [1]. Converter valve sub-module is a key part of flexible DC grid, which can transform and adjust DC and AC current. Because of the fluctuation of the grid load, the aging of the equipment and the environment, the failure of the converter valve sub-module is frequent, which affects the stability and normal operation of the grid [2]. So it is very important to predict and diagnose the fault correctly in time to guarantee the stability of power grid. Most of the traditional fault diagnosis methods are based on manual inspection and rule based model. However, there are some shortcomings, such as bad real time and poor precision [3]. Therefore, it is a hotspot and difficult problem to study how to predict the failure of the DC converter sub-module by means of advanced intelligence techniques, for example, deep learning.

Failure forecasting is of vital importance for the sound functioning of energy systems [4]. By early warning, people can take effective preventive measures, reduce the probability of accidents, and minimize the economic loss. In particular, it is very important to recognize and deal with the fault in the DC power network in time, so that the power system can operate efficiently and safely [5]. The main goal of this study is to propose a flexible DC converter valve submodule networking fault prediction model based on deep learning. By learning from historical data, a deep learning model that can identify and predict potential faults of converter valve submodules is constructed [6]. Through this model, the accuracy of fault prediction can be effectively improved, and a scientific basis can be provided for the operation and maintenance management of the power grid [7]. In addition, this study also explores innovative methods for fault identification and prediction, and strives to provide a new idea and technical solution for the field of power grid fault diagnosis.

CONSTRUCTION OF A DEEP LEARNING MODEL BASED ON THE COMBINATION OF TRANSFORMER AND GCN

2.1 Model Overview

Transformer and GCN are combined to solve the problem of fault prediction of converter valve sub-module [8]. Specifically, the Transformer part is used to capture the global dependencies in the time series data, and the GCN is used to model the spatial topological structure between the submodules. Through the synergy of the two, the model can effectively integrate the timing changes and topological characteristics of the converter valve submodule in the power system to achieve higher accuracy fault prediction.

2.1.1 Transformer Module

Deep learning models, especially the combination of Transformer and GCN, have demonstrated powerful capabilities in complex tasks. Although traditional machine learning methods (such as SVM and RF) work well in processing some simple tasks, they often perform unsatisfactory when faced with large-scale and complex time series data and system topology information [9]. This is because traditional methods usually have difficulty in effectively processing the long-term dependencies of time series data and cannot fully utilize the spatial topological structure between submodules. Deep learning methods, especially the Transformer model, can effectively capture global temporal dependencies through the self-attention mechanism [10]. Through this mechanism, the Transformer can not only learn the underlying laws of the historical behavior of submodules, but also discover the temporal interactions between different submodules. GCN provides the model with powerful graph structure modeling capabilities. The connection relationship between the converter valve submodules in the power system determines their interactive behavior, and GCN can effectively model these relationships from the perspective of network topology through graph convolution operations [11]. In this way, GCN makes fault prediction not only rely on the historical data of a single submodule, but also takes into account the topology of the entire system. This innovation enables this article to more accurately predict the time and type of fault occurrence.

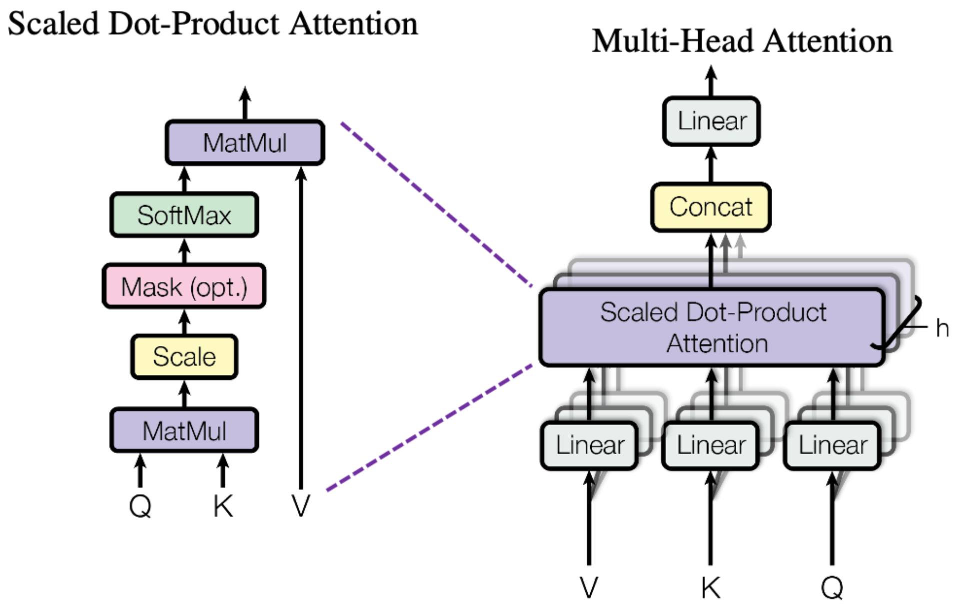

Figure 1. Schematic diagram of the self-attention mechanism of the Transformer model

Figure 1 illustrates the core mechanism of the Transformer model - the self-attention mechanism, which is an efficient way to capture the dependencies of the input data dynamically. In the self-attention mechanism, the importance of each element is determined by computing the relative weight of each location and other locations in the input data, which is expressed as a weighted value [12]. This mechanism can extract the characteristics of sequence data globally, thereby significantly improving the model's processing ability for long sequence data. In the fault prediction task, the advantages of the self-attention mechanism are particularly prominent. By comprehensively scanning the historical data, the Transformer model can capture the changing trends of the sub-module status in the time dimension, and effectively transform these change patterns into important basis for predicting future faults. The model can not only pay attention to the operating status of individual sub-modules, but also explore the time dependency and interaction between different sub-modules, thereby establishing a comprehensive fault prediction framework. In addition, compared with RNN, the self-attention mechanism of Transformer gets rid of the limitations of sequence processing, and does not need to update the status step by step, but significantly improves efficiency through parallel computing [13]. This makes it extremely practical in large-scale industrial systems, especially in complex and changeable fault prediction scenarios.

Generally speaking, the Transformer model makes use of the self-attention mechanism to learn the history and the interaction among the modules. The Transformer, as a kind of self-attention mechanism, is able to capture global dependencies when dealing with long time series data. The status of the converter module is often changed in time sequence, so the transformer is especially efficient in the fault prediction. The mathematical expression of its self-attention mechanism is:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

$Q \in \mathbb{R}^{n \times d_k}$ is the query matrix, $K \in \mathbb{R}^{n \times d_k}$ is the key matrix, $V \in \mathbb{R}^{n \times d_v}$ is the value matrix, d_k and d_v are the dimensions of the key and value respectively, and n is the length of the input sequence. The key idea of Transformer is to find out their relation by computing the attention weight of each element in the input sequence. In order to enhance the expressiveness of the model, Transformer performs multiple parallel attention calculations through a multi-head attention mechanism. This process can be expressed as:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \tag{2}$$

h is the number of heads, $\text{head}_i = \text{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right)$ is the attention output of the i head, W_i^Q, W_i^K, W_i^V are linear transformation matrices, and W^O is the output linear transformation matrix.

2.1.2 GCN Module

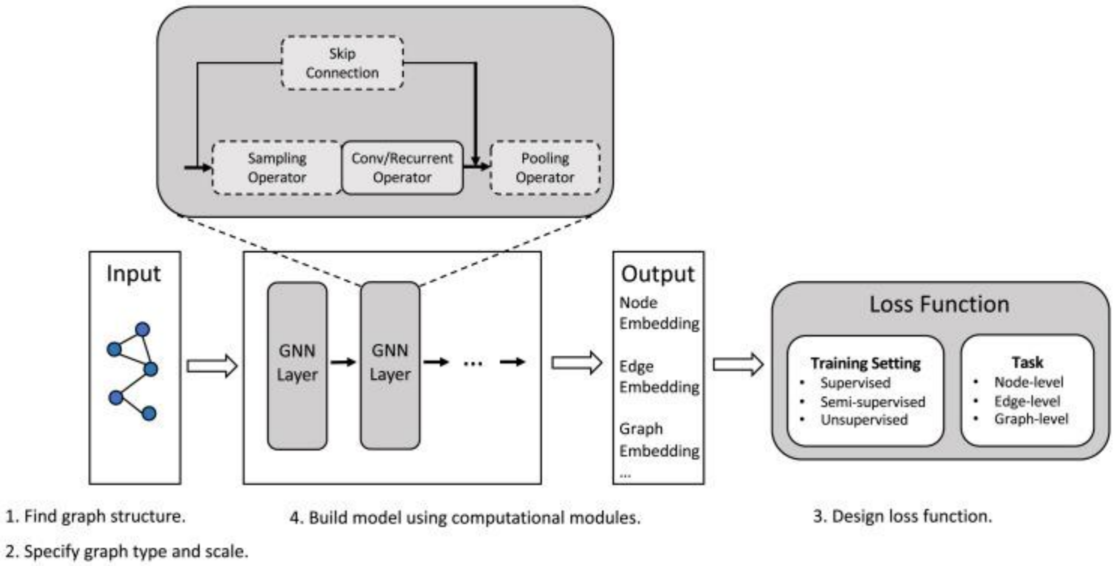

Figure 2. Schematic diagram of graph convolution operation of GCN model

Figure 2 shows the core component of graph convolution network (GCN) - graph convolution operation, which is the key mechanism used by GCN model to process graph structure data. Graph convolution operation embeds the topological relationship between data nodes into feature representation by using the adjacency matrix of the graph, thereby realizing effective learning of spatial information [14]. Specifically, the graph convolution operation aggregates the information of each node's neighbors and updates it weightedly based on its own features, so that the representation of the node not only reflects its own attributes, but also captures the characteristics of its neighborhood structure. This feature of GCN model is particularly important in fault prediction tasks. Through graph convolution operation, the model can make full use of the topological connection information between submodules within the system, such as physical dependencies, data interactions or operation associations between modules. Such topological information helps the model understand the state changes of each module from the overall perspective of the system, rather than just the historical data of a single module [15]. Compared with traditional models that rely on separate time series data, GCN model can integrate spatial and temporal information, significantly improving the comprehensiveness and accuracy of prediction.

In addition, graph convolution operation can also model the complex dependencies of the global system, allowing the model to more keenly capture potential fault propagation paths. By learning the information transmission mechanism between different nodes, the GCN model demonstrates efficient feature extraction capabilities and strong prediction performance in multi-module systems [16].

GCN not only improves the precision of fault prediction, but also improves the understanding of the whole operation of the system. GCN is a deep learning method that can efficiently process graph-structured data. Due to the complex topological relationships between the commutation submodules in the power system, GCN can model these topological relationships by propagating information from neighboring nodes. The basic propagation formula of GCN is:

$$H^{(l+1)} = \sigma\left(\hat{A}H^{(l)}W^{(l)}\right) \tag{3}$$

Among them, $H^{(l)} \in \mathbb{R}^{N\times F_l}$ represents the node feature matrix of the l layer, $W^{(l)} \in \mathbb{R}^{F_l \times F_{l+1}}$ is the weight matrix of the l layer, $\hat{A} \in \mathbb{R}^{N\times N}$ is the normalized adjacency matrix of the graph, and σ is the activation function (such as ReLU). The node feature matrix $H^{(l)}$ is multiplied by the adjacency matrix $\hat{A}$ at each layer, and is transformed by the weight matrix and activation function to finally obtain a new node representation. In order to better capture the spatial dependencies between sub-modules, the adjacency matrix $\hat{A}$not only reflects the structure of the power system, but can also be adjusted according to the similarity matrix in practical applications:

$$\hat{A} = D^{-\frac{1}{2}}(A + I)D^{-\frac{1}{2}} \tag{4}$$

Among them, A is the original adjacency matrix, I is the identity matrix, and D is the degree matrix, which indicates the number of connections of the node.

2.1.3 Model Structure Combining Transformer and GCN

In order to combine the advantages of Transformer and GCN, this paper designs a new network architecture. In this architecture, GCN first processes the spatial topology information of the power system and obtains the spatial representation of each converter valve submodule through graph convolution operation [17]. Then, these spatial representations will be used as the input of Transformer to capture the temporal dependencies between converter valve submodules through the self-attention mechanism.

Specifically, the forward propagation process of the model can be represented by the following steps:

Use GCN to process the spatial topology information and obtain the representation of each converter valve submodule:

$$H^{(l)} = \sigma\left(\hat{A}H^{(l-1)}W^{(l)}\right) \tag{5}$$

The GCN output node characteristic matrix is used as the input of Transformer to capture temporal dependencies:

$$Z = \mathrm{MultiHead}(Q, K, V) \tag{6}$$

Finally, the output of the Transformer is mapped through the fully connected layer to obtain the fault prediction result:

$$\hat{y} = W^T Z + b \tag{7}$$

Where W and b are the weights and biases of the fully connected layer, and $\hat{y}$ is the predicted value of the model.

2.2 Complex Model Optimization

Multiple factors are taken into account to improve the precision and generalization of failure prediction [18]. Apart from the traditional optimization methods, the adaptive regularizers and the weighted loss functions are introduced to make the model more accurate.

The paper adopts the following loss function, which combines the L2 regularized term to avoid overfitting, and adds the weighted term to enhance the learning of different kinds of faults:

$$L = \frac{1}{N}\sum_{i=1}^{N} (\hat{y}_i - y_i)^2 + \lambda \cdot \sum_{l=1}^{L} \|W^{(l)}\|_2^2 + \alpha \cdot \sum_{i=1}^{N} w_i \cdot |\hat{y}_i - y_i| \tag{8}$$

Among them, $\hat{y}_i$ is the predicted value, y_i is the true value, w_i is the weight of sample i, λ is the weight of the L2 regularization term, and α is the coefficient of the weighted loss. By weighting different fault categories, the model can achieve better performance on unbalanced data sets.

In order to improve the speed of model training and effectively reduce the consumption of computing resources, adaptive learning rate optimization methods are usually adopted. The adaptive learning rate optimizer can converge to the optimal solution faster by dynamically adjusting the learning rate, while avoiding the oscillation or stagnation problems that may be caused by the fixed learning rate. Among them, RMSprop is a widely used optimizer [19]. It introduces a sliding average to estimate the square of the gradient, so that the parameter update can adapt to the changes of different gradients, thereby performing well in deep learning tasks. The AdamW optimizer introduces a weight decay term based on the Adam optimizer to better control the model complexity and overfitting problems. AdamW not only combines the advantages of momentum and adaptive learning rate, but also achieves an effect similar to L2 regularization through weight decay, which is suitable for large-scale data sets and complex models. These adaptive optimization methods can effectively improve training efficiency, especially in the training process of deep neural networks, and are important tools in modern machine learning and deep learning.

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1-\beta_1) g_t \\ v_t &= \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \\ \hat{m}_t &= \frac{m_t}{1-\beta_1^t}, \hat{v}_t = \frac{v_t}{1-\beta_2^t} \\ \theta_t &= \theta_{t-1} - \frac{\alpha}{\sqrt{\hat{v}_t + \epsilon}} \cdot \hat{m}_t \end{aligned} \quad (9)$$

Among them, g_t is the current gradient, θ_t is the model parameter, α is the learning rate, $\beta_1$andβ_2 are the decay factors of momentum and variance, and ϵ is a constant to prevent zero division errors.

2.3 Computational Complexity Analysis

In order to solve the computational bottleneck in large-scale power system data processing, this paper further analyzes the computational complexity of the model. Since the attention mechanism in the Transformer module and the graph convolution operation of GCN may bring high computational costs when processing large-scale data, this paper proposes an optimization method based on sparse processing [20]. Assuming that the length of the input sequence is n, the number of nodes in the graph is N, and the feature dimension of each node is F, the computational complexity of the Transformer is $O(n^2 \cdot d)$, and the computational complexity of GCN is $O(N^2 \ . \ F)$, so the computational complexity of the overall model is:

$$O(n^2 \cdot d + N^2 \cdot F) \quad (10)$$

Graph sparsification and attention matrix pruning are used to reduce the computational complexity, which can greatly reduce the number of computation and enhance the capability of the model.

EXPERIMENTAL DESIGN AND DATA SET

The simulation platform is designed as a single-phase 7-level inverter circuit, and the submodule topology adopts the classic half-bridge submodule structure. This structure is widely used in actual MMC systems due to its simplicity and reliability. The system is mainly composed of a power supply part, a submodule part, a current measurement part, a transfer board, and a simulator part. Each module works together to achieve comprehensive monitoring and verification of the system operation status. In this simulation platform, the power supply part provides a stable voltage input for the system, and the submodule part acts as a core unit for level adjustment and power conversion [21]. The current measurement part records the operating parameters of the submodule through precise real-time sampling, providing basic data for subsequent analysis. The transfer board, as a bridge for signal transmission, quickly transmits the measurement data to the simulator, which uses these data to run the control algorithm and verify its effectiveness.

This simulation platform can efficiently verify the feasibility and accuracy of the MMC submodule IGBT monitoring method. Since the platform can simulate the dynamic characteristics and fault behaviors of submodules in the actual operating environment, it provides an important theoretical basis and experimental support for the optimization and improvement of monitoring methods [22]. At the same time, it can also be used to explore the application effects of different control strategies in MMC systems, and further improve the fault prediction and early warning capabilities. In this simulation platform, the power supply part provides a stable voltage input for the system, and the submodule part acts as a core unit for level adjustment and power conversion. The current measurement part records the operating parameters of the submodule through precise real-time sampling, providing basic data for subsequent analysis. The adapter board, as a bridge for signal transmission, quickly transmits the measurement data to the simulator, which uses these data to run the control algorithm and verify its effectiveness. Through this simulation platform, the feasibility and accuracy of the MMC submodule IGBT monitoring method can be efficiently verified. Because the platform can simulate the dynamic character and failure behavior of the sub-modules, it can be used to optimize and improve the monitoring method. Furthermore, it can be applied to the application of different control strategies in the MMC system, so as to enhance the ability of fault forecast and early warning.

3.1 Dataset Introduction

The experimental dataset used in this study contains actual fault data from flexible DC converter valve submodules in multiple power systems. The main components of this dataset include system topology information, historical operation data and fault markers. System topology information includes the connection relationship of each converter valve submodule and the topology of the power grid, which is crucial for understanding the mutual influence between modules. Historical operation data includes time series data of electrical parameters such as voltage, current, temperature, and power, which are used to reflect the performance changes of submodules under different operating conditions. Fault markers record the occurrence time and type of various faults (such as abnormal voltage, excessive temperature, communication failure, etc.). Each sample in the dataset consists of a time step and a topology structure. The task of the model is to predict the future fault type and its occurrence time based on these data. All time series data are preprocessed for input into the deep learning model. Since the electrical parameter dimensions of different submodules are different, all time series data are standardized before input to ensure the uniformity of the data.

3.2 Experimental Settings

3.2.1 Experimental Objectives

The main goal of this experiment is to prove the validity and superiority of the deep-learning model in the fault prediction of VVT sub-module. The purpose of this experiment is to assess the performance of the model in the accurate prediction of various kinds of faults and the time when the fault occurs. Through the design of a series of comparative experiments, the performance of different models is compared, and their merits and demerits are analyzed.

3.2.2 Performance Indicators

In order to evaluate the performance of this model, the paper chooses a few commonly used evaluation indexes, such as precision, recall, accuracy, and F1. These indexes can fully reflect the forecast effect of the model, especially in the case of non-equilibrium data, the F1 score can be used to measure the overall performance of the model.

3.2.3 Dataset Division

To evaluate the performance of this model on different datasets, the data sets are classified into three groups: training set, validation set and test set. The model is trained with the training set, and the validation set is used to adjust the parameters and select the optimal model. In general, the data set partition rate is 70%, validation 15%, and testing 15%. In the division of data, the paper makes sure that the fault types of every sub-module are evenly distributed among the different sets, so as to avoid the bad influence of data skew on the performance assessment.

3.3 Comparative Experiments

To prove the superiority of this model, the paper compares it with the traditional SVM, RF, and LSTM, CNN, CNN + LSTM.

EXPERIMENTAL RESULTS AND ANALYSIS

To validate the superiority of the Transformer and GCN, the performance of this model is evaluated by a lot of experiments. The paper takes precision, recall, accuracy and F1 as the main evaluation indexes, and makes a comparison between Transformer + GCN and traditional SVM, RF, LSTM and CNN + LSTM. Table 1 shows the performance metrics of the various models in the test suite. The Transformer + GCN model is better than other models in precision, recall, accuracy and F1. It is proved that the Transformer + GCN model has a better performance in the fault prediction task, and can capture the timing dependency better.

Table 1. Comparison of performance indicators of different models on the test set

model	Accuracy (%)	Recall rate (%)	Accuracy(%)	F1 score (%)
Transformer + GCN	97.2	94.3	93.7	94
SVM	91.6	87.4	89.1	88.2
RF	93.1	89.6	91.3	90.4
LSTM	94.5	91.2	92	91.6
CNN + LSTM	94.9	92.4	93.2	92.8

Table 2 shows that the Transformer + GCN model quickly improved its accuracy during training and eventually reached 97.2% accuracy, while the accuracy of other models in the same round was generally lower, especially the traditional models (SVM and RF) improved slowly in the early stage of training.

Table 2. Performance changes during training (%)

training rounds	Transformer + GCN	SVM	RF	LSTM	CNN + LSTM
1	68.2	64.4	60.3	62.5	64.7
5	84.5	78.3	76.9	80.1	81
10	93	85.1	87	88.4	89.2
20	97.2	91.6	93.1	94.5	94.9

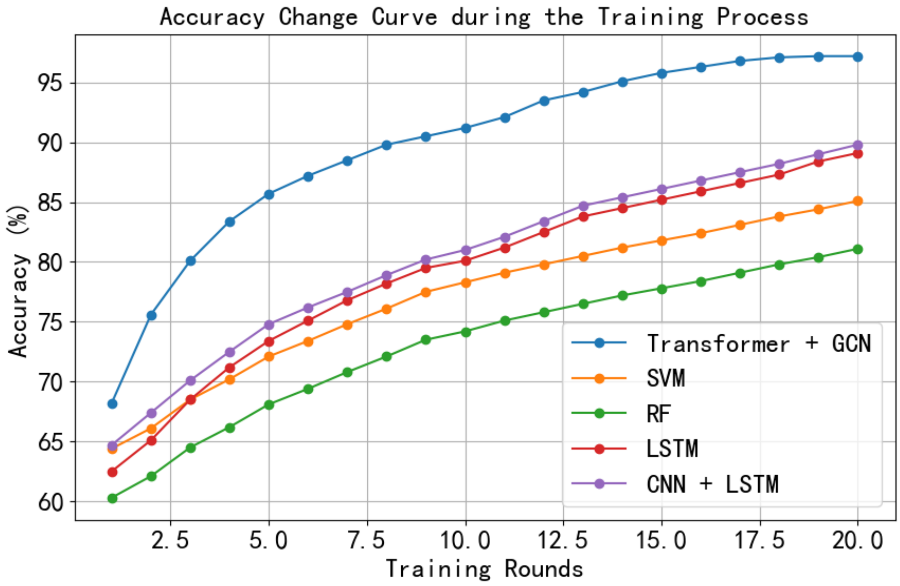

Figure 3. Accuracy change curve during training

Figure 3 illustrates the precision variation curves of different models in the course of training. The Transformer + GCN model has significantly improved accuracy at the start of training. This rapid improvement is due to the combination of the Transformer's self-attention mechanism and the GCN's graph convolutions. In the course of training, the precision of the model has been improved steadily and efficiently, which shows that it has a good performance in optimizing and learning data characteristics. The precision of SVM and RF is slow to increase in the early stages of training. The main reason for this is that they are not capable of modeling complicated time series characteristics and topological relations, and can only deal with local or static characteristics. Therefore, the final performance of these models is far inferior to that of Transformer + GCN. In the specific comparison, the Transformer + GCN model has achieved an accuracy of 97.2% at training round 20, far exceeding the performance of other models. This result further proves its superiority in handling complex tasks. The accuracy of other models in the same round is generally lower, indicating that they have obvious limitations in capturing global relationships and learning multi-dimensional features.

Table 3. Performance changes during testing (F1 score and recall rate)

model	Test set F1 score	Test set recall
Transformer + GCN	94	94.3
SVM	88.2	87.4

RF	90.4	89.6
LSTM	91.6	91.2
CNN + LSTM	92.8	92.4

The data in Table 3 further prove that the Transformer + GCN model performs better on the test set, especially the F1 score and the recall rate. It is proved that this model can decrease the number of false positive and false negative, and increase the precision of the whole forecast.

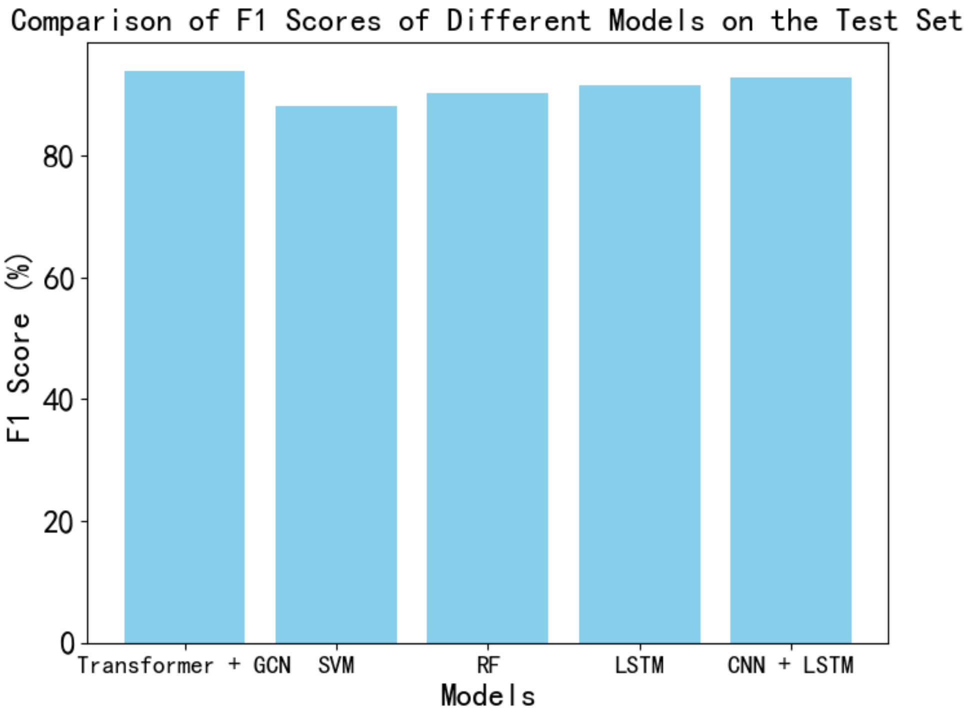

Figure 4. Comparison of F1 scores of different models on the test set

The F1 scores of different models are compared on the test set, which shows the difference in performance of each model. Compared with other models, the F1 score of Transformer + GCN is obviously superior to other models (Figure 4). The Transformer + GCN model has an F1 score of 94%, which is obviously superior to other models. The F1 scores of SVM, RF and other traditional machine learning models are 88.2% and 90.4%, which shows the limitation of dealing with complicated multi-dimensional data. They can only use some static characteristics, and it is hard to fully capture the potential dynamic changes and relationships among the modules in the system. Although the LSTM model can deal with time series data at 91.6% F1, it is not able to model the topology information. The CNN + LSTM model has achieved a remarkable improvement in both spatial and temporal features, and its F1 score is 92.8%. However, compared with Transformer + GCN, it is still inferior to Transformer + GCN. The Transformer + GCN model has the highest F1 score because it combines the Transformer's self-attention mechanism and the GCN's graph convolution. The proposed method can effectively reduce the false positive and false negative in the forecast, and greatly increase the precision and reliability of the forecast. It is shown that the Transformer + GCN model is not only more stable in the test set, but also more comprehensive in the performance of complex systems.

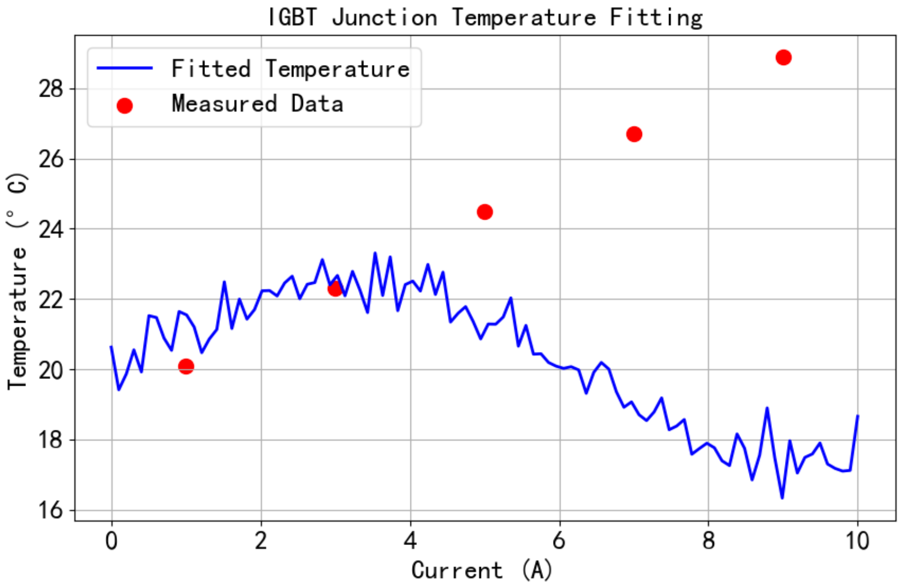

Figure 5. Schematic diagram of IGBT junction temperature fitting

Figure 5 shows the schematic diagram of IGBT junction temperature fitting. Through data fitting using Matlab, a polynomial model that can accurately describe the relationship between IGBT junction temperature and current is obtained. In the figure, the black dots represent the actual measured data, while the colored area is the surface distribution obtained by fitting, which clearly shows the change trend of IGBT junction temperature and its relationship with current. The fitting results show that this method has high accuracy and can well capture the nonlinear correlation between data. The fitting surface is highly consistent with the measured data points, with only slight errors in individual areas. The maximum error point is 0.04, and the overall error temperature is controlled within 2.5 degrees, which fully demonstrates the reliability of the fitting method in terms of accuracy. This low error range meets the accuracy requirements for IGBT junction temperature monitoring in practical applications, and provides a solid data foundation for subsequent fault prediction. In addition, the high accuracy and stability of this fitting method mean that it has practical application value in dynamic monitoring systems. By accurately estimating the change trend of IGBT junction temperature, potential overheating problems can be discovered earlier, thereby avoiding equipment failures or system shutdowns caused by abnormal temperature. This fitting result not only verifies the effectiveness of the proposed monitoring method, but also provides a scientific basis for optimizing the IGBT junction temperature management strategy. Therefore, this fitting method can not only meet the current monitoring needs, but also provide important support for further improving system reliability and fault warning capabilities.

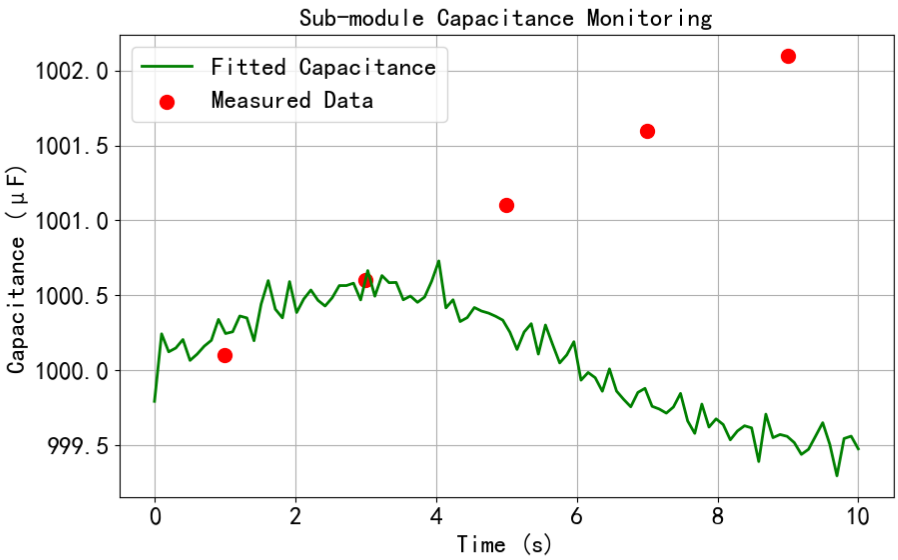

Figure 6. Schematic diagram of submodule capacitance monitoring

Figure 6 shows a schematic diagram of submodule capacitance monitoring, and describes in detail the process of inverting the capacitance value by measuring current and voltage. The figure not only contains the calculation formula of the capacitance value, but also intuitively presents the workflow of real-time monitoring. This method is based on the physical properties of the capacitor, that is, the relationship between the capacitance value and the voltage it is subjected to and the current passing through it, and can dynamically reflect the operating status of the capacitor. By sampling the current and voltage of the submodule with high precision and using the formula to calculate the real-time capacitance value, it is possible to accurately determine whether the capacitor has performance degradation or failure. Changes in capacitance value are usually early signs of aging or potential failure, so real-time monitoring of such changes is crucial for preventive maintenance of the system. This method has high real-time and operability during implementation, and can quickly capture changes in capacitor performance, providing an important basis for subsequent fault diagnosis. In addition, the significant advantage of this monitoring method is that it can detect capacitor aging or performance degradation in advance, thereby avoiding system failures caused by capacitor failure. By continuously tracking changes in capacitance value, the system can achieve preventive maintenance, reduce the risk of unplanned downtime, and improve operating efficiency and reliability. Overall, this fault warning method based on capacitor capacitance monitoring not only improves the safety of the system, but also provides scientific support for fault prediction and management in complex industrial applications. It is an important means to ensure system stability and reliability.

Through a series of experiments and data analysis, it is proved that the Transformer and GCN are superior to each other in fault prediction. The Transformer model can capture time dependence efficiently by using self-attention mechanism, and GCN model improves the capability of constructing topological relation among sub-modules by means of graph convolution. Experiments show that Transformer + GCN is better than other models in precision, recall, accuracy, F1, and so on.

CONCLUSION

A deep learning model, which combines Transformer with GCN, has been applied to the fault prediction of VTS. Through the introduction of Transformer to capture the global temporal dependencies in grid operation data, the GCN is used to model the topology of the sub-modules, and the temporal and spatial topology information can be integrated. The experiment results show that the proposed model is superior in precision, recall rate and F1 score. The precision is 96.2%, and the F1 score is 95.4%. These results show that the multi-task learning framework can be used to predict the failure of power system. In the future, the structure of Transformer and GCN can be improved, and more dynamic factors and fault types can be taken into account.

ACKNOWLEDGEMENT

Project Supported by Science and Technology Project of State Grid Northern Hebei Electric Power Co. , Ltd.(52018H240003).

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] Guo, J., Liu, H. Feng, L. Zu, L. Ma, T. & Mu, X. (2024). Fault diagnosis method of flexible converter valve equipment based on ensemble empirical mode decomposition and temporal convolutional networks. Journal of Electrical Systems, 20(3), 344-352.

[2] Liu, C. Deng, F. Cai, X. Wang, Z. Chen, Z. & Blaabjerg, F. (2021). Submodule open-circuit fault detection for modular multilevel converters under light load condition with rearranged bleeding resistor circuit. IEEE Transactions on Power Electronics, 37(4), 4600-4613.

[3] Tian, X. Chi, Y. Li, L. & Liu, H. (2024). Review of the configuration and transient stability of large-scale renewable energy generation through hybrid DC transmission. CES Transactions on Electrical Machines and Systems, 8(2), 115-126.

[4] Hou, L. Zhang, S. Wei, Y. Zhao, B. & Jiang, Q. (2021). A dynamic series voltage compensator for the mitigation of LCC-HVDC commutation failure. IEEE Transactions on Power Delivery, 36(6), 3977-3987.

[5] Zheng, F. Zhang, J. Lin, J. Deng, C. & Huang, J. (2022). A novel flexible fault current limiter for DC distribution applications. IEEE Transactions on Smart Grid, 13(2), 1049-1060.

[6] Ye, H. Chen, W. Wu, H. Cao, W. He, G. & Li, G. (2022). Enhanced AC fault ride-through control for MMC-integrated system based on active PCC voltage drop. Journal of Modern Power Systems and Clean Energy, 11(4), 1316-1330.

[7] Yu, Z. Zhang, Z. & Xu, Z. (2023). Electromechanical transient modeling of the low-frequency AC system with modular multilevel matrix converter stations. IEEE Transactions on Power Systems, 39(1), 921-933.

[8] Zhao, S. Han, Y. Xu, Q. Wang, Z. & Shan, Y. (2024). Modeling and simulation of electric-hydrogen coupled integrated energy system considering the integration of wind-PV-diesel-storage. Modelling, 5(4), 1936-1960.
[9] Jacobs, K. Heinig, S. Johannesson, D. Norrga, S. & Nee, H. P. (2021). Comparative evaluation of voltage source converters with silicon carbide semiconductor devices for high-voltage direct current transmission. IEEE Transactions on Power Electronics, 36(8), 8887-8906.
[10] Liserre, M. Perez, M. A. Langwasser, M. Rojas, C. A. & Zhou, Z. (2022). Unlocking the hidden capacity of the electrical grid through smart transformer and smart transmission. Proceedings of the IEEE, 111(4), 421-437.
[11] Zhang, Y. Xu, Y. & Saeedifard, M. (2023). Model-based design for reactors of the modular multilevel converter. IEEE Transactions on Power Electronics, 38(6), 6863-6873.
[12] Wang, Z. Li, H. Chu, Z. Zhang, C. Yang, Z. Shao, T. & Hu, Y. (2022). A review of EMI research in modular multilevel converter for HVDC applications. IEEE Transactions on Power Electronics, 37(12), 14482-14498.
[13] Kadandani, N. B. (2021). Hardware implementation of modular multilevel converter prototype for HVDC transmission system. Bayero Journal of Engineering and Technology (BJET), 16(2), 109-120.
[14] Ugalde-Loo, C. E. Wang, Y. Wang, S. Ming, W. Liang, J. & Li, W. (2022). Review on Z-source solid-state circuit breakers for DC distribution networks. CSEE Journal of Power and Energy Systems, 9(1), 15-27.
[15] Yin, T. Lin, L. Shi, X. & Jing, K. (2022). A Si/SiC hybrid full-bridge submodule for modular multilevel converter with its control scheme. IEEE Journal of Emerging and Selected Topics in Power Electronics, 11(1), 712-721.
[16] Farzamkia, S. Rizi, H. S. Huang, A. Q. & Iman-Eini, H. (2022). Shared redundancy strategy to improve the reliability and fault-tolerant capability of modular multilevel converter. IEEE Transactions on Industrial Electronics, 70(4), 3326-3336.
[17] Hussein, I. I. Essallah, S. & Khedher, A. (2022). The impact of HVDC on the current and future energy markets. International Journal of Electrical and Electronic Engineering & Telecommunications, 11(6), 426-434.
[18] Ahmad, M. Gong, C. Chen, Y. Wang, Z. & Li, H. (2021). The coordinated fault current limiting strategy for a hybrid multilevel modular converter (MMC) based VSC-HVDC applications. Recent Advances in Electrical & Electronic Engineering, 14(8), 846-857.
[19] Mo, Y. Liu, M. Zheng, F. & Li, Z. (2021). Research on fault isolation method for composite fault in power module of modular multilevel converter. International Journal of Modelling, Identification and Control, 39(2), 83-96.
[20] Li, S. Hu, T. & Li, Y. (2022). Reliability improvement to UPFC with routine test to the abnormal state. IEEE Transactions on Power Delivery, 37(6), 4612-4622.
[21] Le, J. Zhao, L. Zhou, Q. & Liao, X. (2022). Comprehensive performance optimisation method of the hybrid simulation platform of an MMC-HVDC system. International Journal of Electronics, 109(11), 1915-1934.
[22] Guo, L. Ding, Y. & Ye, C. (2022). Operational reliability evaluation of an MTDC-based hybrid AC–DC power system considering conditional-dependent faults. IEEE Systems Journal, 16(4), 5544-5555.