Article citation info:
Zhao S, Pattanadech N. Construction of intelligent transformer anomaly detection model based on LSTM combined with optical flow features. Diagnostyka. 2025;26(1):2025115. https://doi.org/10.29354/diag/202094.

DIAGNOSTYKA, 2025, Vol. 26, No. 1

e-ISSN 2449-5220
DOI: 10.29354/diag/202094

CONSTRUCTION OF INTELLIGENT TRANSFORMER ANOMALY DETECTION MODEL BASED ON LSTM COMBINED WITH OPTICAL FLOW FEATURES

Shuzong ZHAO [*], **Norasage PATTANADECH**

Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
[*] Corresponding author, e-mail: lionogj@sina.com

Abstract

Transformers are essential for the transmission and distribution of electricity, but due to changes in load and the influence of the working environment, various faults may occur in transformers. To accurately and quickly detect faults in transformers and conduct effective fault diagnosis and equipment maintenance, this study solves the problems of data imbalance and temporal data in transformers by introducing a long short-term memory network with fatigue factors. In addition, a fusion model is ultimately constructed by combining the recursive all-pair field transformation streamer method to achieve more accurate and robust optical flow estimation in the model. The experiment indicated that the maximum accuracy of the predicted values combined with the model was around 95% and the minimum was around 35%. Compared to other models, the maximum accuracy of actual values was around 80% and the minimum accuracy was better at 10%. In the application experiment, the frequency of insulation faults was the least obvious, with only 10 faults. The resistance fault was evident, with a total of 100 faults. The combined model could well reflect the fluctuation of fault current and the collection of fault numbers by different sensors. Therefore, the proposed model has high accuracy, good precision, and outstanding application effects, which can provide new ideas for constructing intelligent transformer anomaly detection models.

Keywords: LSTM, optical flow characteristics, transformer abnormality, invisible factor, RAFT streamer method

1. INTRODUCTION

As important equipment for energy distribution and transmission, the stable operation and reliability of power transformers are crucial for the normal operation of the power grid. However, due to long-term work and the complexity of the operating environment, power transformers are prone to faults, and the types of faults are diverse [1-2]. Common types include winding short circuits, partial discharge, insulation aging, etc. Partial discharge is one of the most common transformer faults, and it has a serious impact on the insulation system of transformers. The specific phenomenon is the frequency, characteristics, and short-circuit faults caused by abnormal discharge, so it is necessary to adopt corresponding diagnoses for different types of power transformer faults [3-4]. Traditional fault detection techniques are mostly based on image processing methods, which can accurately diagnose faults such as insulation aging and short circuits inside transformers through infrared and X-ray images of the transformer surface. Online monitoring of power transformers can also be achieved through sensor technology, enabling prediction, warning, and forecasting of transformer faults. However, traditional technology has problems such as low diagnostic accuracy and measurement efficiency, making it difficult to meet the requirements of high-precision fault diagnosis for transformers [5]. Based on this background, to improve the accuracy and efficiency of fault diagnosis of power transformers and ensure the stability of power system operation, this study is based on a Long Short-Term Memory (LSTM) network and introduces fatigue factor to detect subtle abnormal signals of transformers. An LSTM-RAFT model is constructed by combining LSTM with Recursive All-pair Field Transformation (RAFT) optical flow to address the issue of temporal correlation in transformer fault data. The RAFT optical flow estimation algorithm also makes the model more innovative in the feature extraction of transformer fault information. By analyzing the performance and application effects of the combined model through certain model parameter values and software, it is hoped that the proposed LSTM-RAFT model can improve the accuracy of Transformer Fault Detection (TFD).

This study is divided into four parts to analyze the construction of an intelligent transformer anomaly detection model. Part 1 is a discussion by

Received 2024-08-24; Accepted 2025-02-19; Available online 2025-02-24

industry researchers on TFD. Part 2 is the construction process of the LSTM-RAFT model and the methods and processes for TFD using the model. Part 3 tests the performance and application effectiveness of the research model. Part 4 summarizes the test results and application effects, as well as analyzes the shortcomings and improvement directions of the model.

2. RELATED WORKS

Transformers are the core equipment in the power system. For a long time, to detect transformer faults more efficiently, researchers have proposed many effective methods. He and Pang proposed a dynamic Bayesian network-based model to predict the state of transformers to reduce the cost of equipment maintenance and operation in aging power systems. With this model, the risk probability of the transformer subsystem could be obtained, and the overall system risk coefficient could also be obtained through individuals. This method could effectively control the operation and maintenance costs of the power system [6]. The content of furan and the CO2/CO ratio could solve the two most dangerous faults in transformers, and the key factors were insulation. However, when a small amount of paper was involved, they were often unreliable or not sensitive enough. Duval and Buchacz developed a new inexpensive method to detect arcs on paper and distinguish them from arcs in petroleum. Experimental results have shown that the reliability and sensitivity of this method have been significantly improved [7]. Kia et al. proposed to develop models for three common randomized concept attacks and their associated detection methods, aiming to create models suitable for real-time situations through software. This approach was designed to address the issue of network attacks that have the potential to damage digital protection relays and alter their output commands to circuit breakers. Finally, using wavelet analysis of substation current and energy, a new detection metric called network attack detection metric was defined. This indicator had acceptable performance, with an accuracy rate of 100% and error free triggering [8]. Doorwar et al. proposed a new scheme for synchronous generator stator fault detection and classification based on a novel Differential Component (DC) and Amplifying Discrete Teager-Kaiser Energy Operators (ADTKEO) for better classification of transformer fault types. Detection was based on threshold to check the second harmonic content in the secondary current of current transformers. This scheme could detect other low resistance faults within a quarter cycle [9]. In practice, frequency response analysis may not always be available, and subsequent analysis measurements themselves cannot detect faults. In view of this, Pramanik et al. proposed an innovative method for measuring the frequency response of transformer windings under sinusoidal excitation with equal and opposite polarities at both ends. The results obtained from the verification of this method on actual single-phase auto-transformers demonstrated that the response frequency of this method was superior to that of traditional methods [10].

Eruvai and Chilaka proposed a new fuzzy logic model for evaluating the health status of power transformers, which can detect early faults in multiple power grid systems. 31 fuzzy rules were designed built on the severity of these gases to decide the Health Index (HI) of oil products. The results showed that the model overcomes the lacks of traditional ratio encoding methods in identifying early faults in typical situations (i.e. multiple early fault situations) [11]. Su et al. proposed a weak pulse signal detection scheme grounded on chaotic time series phase space reconstruction and LSTM in deep learning models to better distinguish between signal and non-signal points of weak faults. Reconstructing the its phase space could extract chaotic message from sequences, thereby obtaining better detection results. In the case of low signal-to-noise ratio, this model could detect weak pulse-signals in chaotic backgrounds [12]. Windmann and Westerhold designed a regularized LSTM and auto-encoder hybrid model for fault detection in automated production systems to better observe the differences between sensor data and predicted probability density distributions. The model combined the advantages of both, with auto-encoders utilizing the correlation between individual sensor signals, while LSTM utilized neurons to capture temporal dependencies. The effectiveness of the model for device anomaly detection has been demonstrated [13]. Aljemely A H et al. proposed a method combining LSTM and Large Margin Nearest Neighbor (LSTM-LMNN) for multi-fault identification of mechanical rotating equipment to address the problem of rolling bearings being prone to failure due to excessive working stress. This method utilized powerful orthogortical weight initialization techniques to memorize critical information about failures during parameter updating and effectively organize samples for each condition during pattern classification. The experimental results showed that the proposed LSTM-LMNN model outperformed other existing methods in terms of diagnostic efficiency, stability, and reliability[14] . HaiBin S et al. proposed a hybrid model of single Layer-Wide Convolutional Neural Network and LSTM (LWCNN-LSTM) for the problem of complex and changeable data and noise interference in bearing fault diagnosis. The model first extracted features using a wide convolution kernel to reduce the effect of noise and then learned new sequence features via LSTM. Experimental results showed that the model had higher diagnostic accuracy and robustness under mixed load and noise conditions [15]. Papin et al. employed distinctive identification techniques to ascertain that the Climate Potential Eddy (PVS) index in the subtropical North Atlantic Basin surpassed the

preceding PVS climatology. This investigation sought to elucidate the influence of humidity and wind anomalies on the associated PVS activity. The indices collected by this method provided a comprehensive measure of PVS activity [16].

In summary, scholars mostly use fuzzy models to analyze TFD methods and also optimize traditional image techniques to enhance the sensitivity and precision of fault recognition. The detection and analysis of weak fault signals adopt the LSTM model based on chaotic time series. However, streamer features are mostly used for identifying and predicting data indices. There is a lack of analysis on transformers using LSTM combined with streamer characteristics in various studies. Therefore, the LSTM-RAFT model proposed in this study has innovative value for TFD.

3. CONSTRUCTION OF TRANSFORMER ANOMALY DETECTION MODEL BASED ON LSTM COMBINED WITH OPTICAL FLOW FEATURES

3.1 Construction of LSTM-RAFT detection model

In transformer anomaly detection, methods combining LSTM networks and optical flow features have been extensively studied. LSTM networks are good at processing time-series data and are able to learn long-term dependencies, while optical flow features help to capture dynamic changes in transformer operation [17]. However, LSTM lacks accuracy in detecting subtle changes in equipment and diagnosing faults. Therefore, this study introduces fatigue factors to improve the sensitivity of the algorithm to signals with subtle changes in faults. Meanwhile, an LSTM-RAFT model is constructed using the RAFT algorithm, which allows for more significant results in terms of accuracy and robustness. The network data training of LSTM with fatigue factor introduced is shown in Fig. 1.

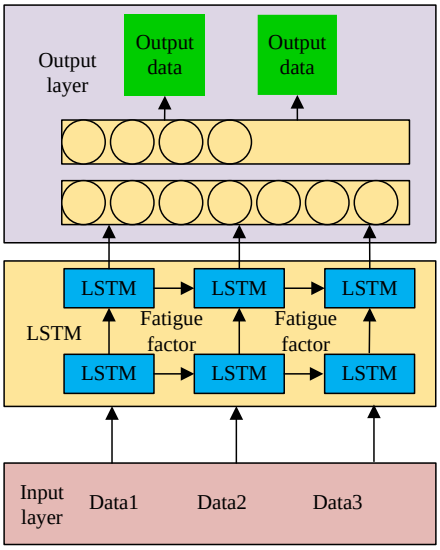


Fig. 1. LSTM network data training with fatigue factor

In Fig. 1, first, the data collected by the transformer sensor need to be cleaned and standardized to improve the efficiency and effectiveness of model training. The features of time series data can be extracted through wavelet transform and empirical mode decomposition to enhance the ability of the model to identify anomalous modes. After the partitioning is completed, the data are transmitted to LSTM with fatigue factor introduced. If the received data are the fault dataset of the entire transformer, then the data are divided into sequences. The operation data of the transformer including the denoising and normalization are required before applying the LSTM network. Moreover, dynamic features are extracted by optical flow techniques that can reflect the behavior changes of the transformer under normal and abnormal conditions. A sequence is composed of vectors, and the amount of sequences corresponds to the quantity of time steps required to complete the analysis of the entire fault sequence data [18]. The data processed by the network layer can be transmitted to the monitoring system as fault prediction data after being coordinated by the output layer. However, due to insufficient collection of small signals by algorithms, there may be prediction errors, resulting in sequence anomalies. The generated error prediction vector needs to be restored to a normal sequence by calculating the fatigue factor of the LSTM network layer. The prediction error can be obtained as equation (1).

$$e^{(t)} = \left| y'^{(t)} - y^{(t)} \right| \tag{1}$$

In equation (1), t is the time point and $y'^{(t)}$ is the predicted value. $y^{(t)}$ can be expressed as equation (2).

$$y^{(t)} = x_i = x_i^{(t+1)} \tag{2}$$

In equation (2), x_i represents the observations above the time series. i is the dimension of the actual detection data value. Adding $e^{(t)}$ to the vector of prediction error, the error vector can be expressed as equation (3).

$$e = \left| e^{(t-h)}, .., e^{(t-l_s)}, e^{(t-1)}, e^{(t)} \right| \tag{3}$$

In equation (3), h is the historical prediction error and l_s is the input length. The error threshold in the error set is shown in equation (4).

$$\varepsilon = \mu(e_s) + z\sigma(e_s) \tag{4}$$

In equation (4), μ is the mean, σ is the standard deviation, z is a positive integer, and e_s is a non-continuous sequence. The threshold can be determined according to equation (5).

$$X' = \frac{X-\mu}{\sigma} \tag{5}$$

In equation (5), X' is the random sequence value of the data and X is the continuous sequence value. The threshold is shown in equation (6).

$$\varepsilon = \arg\max(\varepsilon) = \frac{\frac{\Delta\mu(e_s)}{\mu(e_s)} + \frac{\Delta\sigma(e_s)}{\sigma(e_s)}}{|e_a| + |E_{seq}|^2} \tag{6}$$

In equation (6), e_a is a continuous sequence and E_{seq} is an abnormal sequence. The calculation formula for fatigue factor is shown in equation (7).

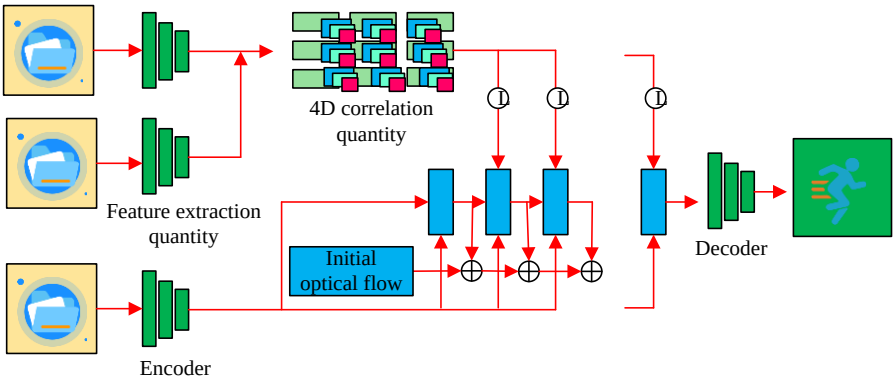


Fig. 2. Structure of RAFT model

$$s^{(i)} = \frac{\max\left(e_{seq}^{(i)}\right) - \arg\max(\varepsilon)}{\mu(e_s) + \sigma(e_s)} \tag{7}$$

In equation (7), i is the sequence coefficient and s is the severity of the abnormal sequence. The fatigue factor is optimized for LSTM by correcting the vector changes caused by the collection of weak fault signals. To further improve the accuracy of fault recognition in the network, the combination of RAFT algorithm is also needed. The structure of the RAFT is exhibited in Fig. 2.

In Fig. 2, RAFT includes feature extraction, 4D correlation, and upsampling of optical flow updates. Feature extraction uses a pixel by pixel approach to extract data features, and the upper and lower encoders focus on the second data extraction to obtain the content information of the data [19]. In the transformer anomaly detection, the RAFT algorithm can be employed to analyze the visual image of the transformer equipment and to extract the dynamic characteristics of the operating state of the equipment, including vibration frequency and displacement change, etc. This enables the provision of comprehensive spatial and temporal information, which is crucial for effective anomaly detection. The construction form of the encoder in the model is consistent, the only difference is the difference in the normalization method and operation method. The feature extraction operation is a way of normalizing actual examples. The operation method of the encoder is collective normalization. The operation of the 4D tensor in the RAFT model is given by equation (8).

$$C_{ijkl} = \sum_h F(I_1)_{ijh} \cdot F(I_2)_{klh} \tag{8}$$

In equation (8), j and k/l are the rows and columns of feature point one and feature point two. h is the number of channels, and it is one of the first and second feature points of $F(I_1)$ and $F(I_2)$. The updated gating representation through the streamer algorithm is shown in equation (9).

$$z_t = \sigma\left(Conv_{4\times4}([h_{t-1}, x_t], W_z)\right) \tag{9}$$

In equation (9), $Conv_{4\times4}$ is the convolutional neural network, x_t is the feature data from the previous iteration, and W_z is the gate feature dimension. The gate control and data after reset are given by equation (10).

$$\begin{aligned} r_t &= \sigma\left(Conv_{4\times4}([h_{t-1}, x_t], W_r)\right) \\ h_t &= \sigma\left(Conv_{4\times4}([h_{t-1}, x_t], W_h)\right) \end{aligned} \tag{10}$$

In equation (10), W_r and W_h are the dimensions of resetting gating and resetting data. By introducing fatigue factors and combining LSTM with RAFT models, an LSTM-RAFT model is constructed, as shown in Fig. 3.

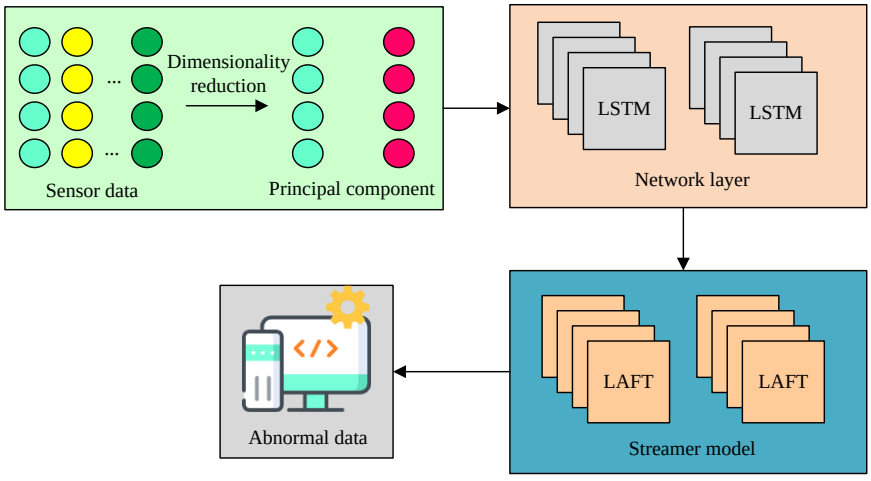


Fig. 3. LSTM-RAFT model and model structure

An end-to-end anomaly detection model is constructed by combining LSTM network with RAFT algorithm. In the model structure, the dynamic features extracted by the RAFT algorithm are used as inputs to the LSTM network, and the LSTM network's temporal analysis capability is utilized to monitor and predict the real-time operation status of the transformer. The structure of Fig. 3 consists of four parts. Part 1 is the dimensionality reduction of data transmitted by sensors. Part 2 is based on LSTM feature extraction, which extracts temporal features from the data collected by sensors in transformers through its long-term and short-term memory. In Part 3, the time feature data extracted by LSTM are transmitted to the RAFT model for information decoding to obtain the content information in the data and locate the location and type of transformer faults [20]. The fully connected layer in the fourth stage receives the decoded information content from LSTM-RAFT, and then transmits it to the output layer for judgment through weighted summation, thereby obtaining abnormal detection results. The formula for variance dimensionality reduction data is shown in equation (11).

$$v_k = \frac{\lambda_i}{\sum_{i=0}^{k} \lambda_i} \tag{11}$$

In equation (11), k is the eigenvector and λ_i is the eigenvalue. The accumulated dimensionality reduction data are defined as equation (12).

$$T_V = \sum_{j=1}^{i} V_k = \frac{\sum_{j=1}^{i} \lambda_j}{\sum_{j=1}^{k} \lambda_j} \tag{12}$$

In equation (12), V_k is the variance dimensionality reduction value. k is characteristic coefficient.

3.2 Detection method and process of LSTM-RAFT model

The advantage of using LSTM to extract temporal features of data is to achieve dimensionality reduction and simplification of transformer fault data. This study combines the RAFT model to decode complex content in data information, locate fault locations in transformers, and distinguish fault types. The method of integrating models for fault diagnosis is shown in Fig. 4.

In Fig. 4, due to the lack of fault data in the production environment, a signal prediction model is constructed using RSFT to decode the information content of the data through feature processing. Due to the lack of information on equipment aging and wear, as well as loopholes in collecting small signals in the environment, fatigue factors are added to detect hidden anomalies. The fatigue factor can be used for transformer fault diagnosis by resampling after detecting abnormal data. Due to the complex temporal correlation, the classification model of faults must rely on computational analysis combined with the model to screen effective samples. Afterwards, the dataset is optimized and updated to obtain the incremental error of the pre-trained model for transformer fault diagnosis. The functional modules of the anomaly detection model are shown in Fig. 5.

Fig. 5 includes four modules: data management, user management, anomaly detection, and anomaly warning. Data management is carried out through data storage and classification, data decoding, and data transmission. User management is carried out through two methods: managing user operation logs and managing user information. The anomaly detection module is managed through two methods: manual detection and automatic detection. The part of anomaly warning is divided into two methods: device anomaly warning and historical information tracing. The anomaly detection model is shown in Fig. 6.

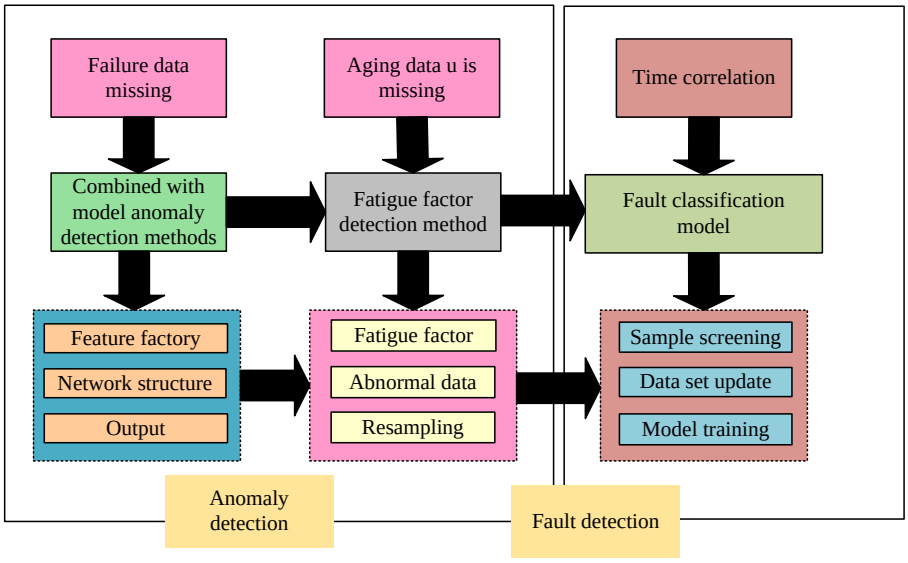


Fig. 4. Fault diagnosis method combined with model

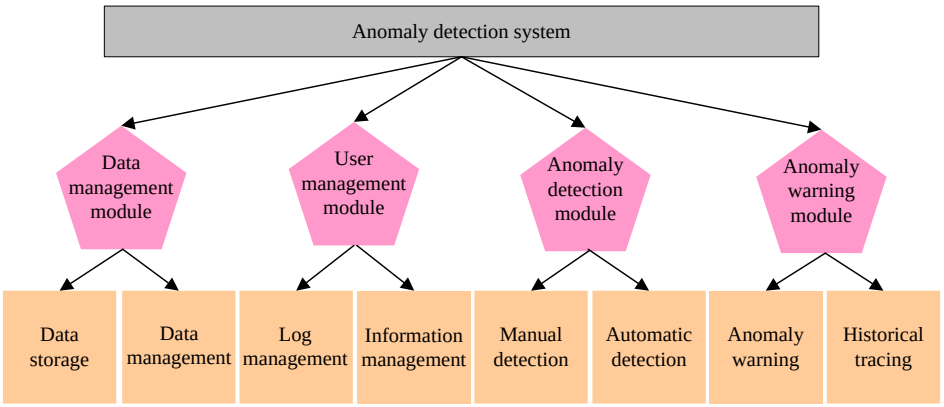


Fig. 5. Function module of anomaly detection model

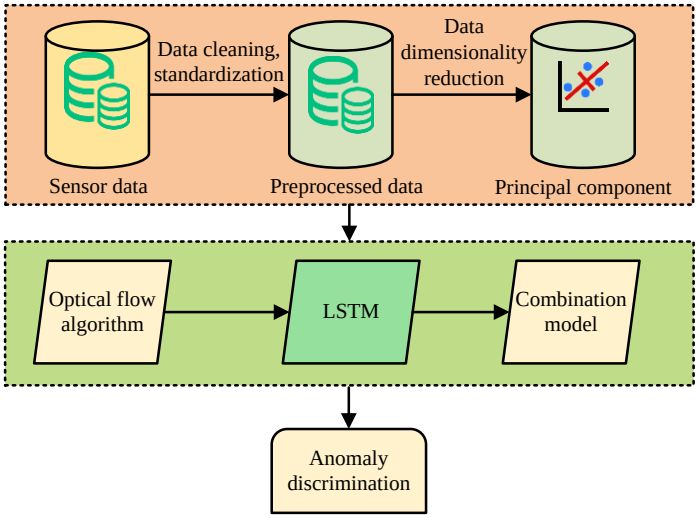


Fig. 6. The process of anomaly detection in the model

In the detection process of Fig. 6, the data collected by the sensors are cleaned and standardized before being transmitted to the data preprocessing stage. The data are initialized through the input layer of LSTM. After initialization, the principal components of the fault data are obtained by simplifying the operation of the data through dimensionality reduction calculation. This is the data processing module in model anomaly detection. After data processing, the data flow to the training module of the model. The RSFT algorithm first performs simple filtering on the data, and then identifies anomalies in the training error transformer through LSTM and fusion models. The evaluation indicators for anomaly detection are the accuracy of fault identification and detection rate. The accuracy is calculated using equation (13).

$$Hcc = \frac{BP+CN}{BP+DN+BP+EN} \tag{13}$$

In equation (13), Hcc and BP are the actual number of predicted and faulty samples. EP and DN are the sample numbers for normal detection and accurate detection as faults. CN is the number of samples with normal fault detection. The accuracy of the detection is shown in equation (14).

$$FDR = \frac{BP}{BP+EN} \tag{14}$$

To better evaluate the training indicators, this study chooses to use the average training time as the treatment for efficiency, as shown in equation (15).

$$Att = \frac{\sum_{i=1}^{l} t_i}{L} \tag{15}$$

In equation (15), Att is the average training time. L is the number of training sessions. t is the training time. i is the line coefficient. Only by maintaining a high fault detection rate and high accuracy in fault diagnosis can the fusion model demonstrate good performance in fault recognition. The trained LSTM-RAFT model is applied to the actual transformer abnormality detection task, monitoring the running status of the transformer in real time to find the potential abnormalities and faults in time. At the same time, the model can be further adjusted and optimized according to the feedback in the practical application, thereby enabling its adaptation to disparate detection scenarios and requirements.

4. PERFORMANCE TESTING AND APPLICATION EFFECT ANALYSIS OF LSTM-RAFT MODEL

4.1 Performance testing of LSTM-RAFT

The fusion model can extract temporal features of transformer faults and process small signals

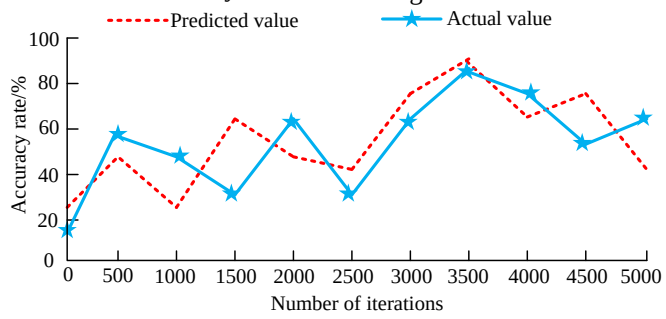
through fatigue factors. Select LSTM [21] and RAFT [22] models from other research designs for comparison. The RAFT model achieves fault diagnosis by decoding data information content. To further confirm the performance of the research model, performance tests are conducted on the fusion model. Table 1 shows the training parameters of the model.

Table 1. Model training parameter setting

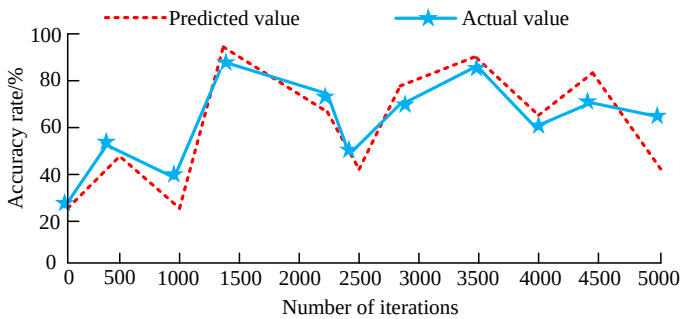
Network parameter	LSTM	Network parameter	RAFT
Input vector dimension	1	Input vector dimension	4
Hidden ganglion number	128,64,32,1	Hidden ganglion number	128,64,32,1
Network layer number	1	Network layer number	2
Learning rate	0.02	Learning rate	0.005
Maximum iterations	100	Maximum iterations	100
Output vector dimension	1	Output vector dimension	5

In Table 1, when the input vector dimension of LSTM is 1, the output vector dimensions of RAFT model are 4 and 5. The number of hidden neurons for both is 128, 64, 32, and 1. The network layer of LSTM is 1, and RSFT is 2. The learning rate of LSTM is 0.02, RAFT is 0.005, and the maximum iteration for both is 100. The LSTM-RAFT model is designed to diagnose faults by analyzing current changes in transformers. Therefore, the sensors used must be capable of detecting small current changes, particularly those with weak signals resulting from early failures, which require a wide dynamic range. The hall effect sensor is suitable for the measurement of DC and low-frequency AC with high sensitivity and a wide dynamic range. The sensor model is LEM LA25-PNP-N, with a sensitivity of 10 mV/mA, a frequency range of 0 Hz to 10 kHz, a resolution of 1μA, and a temperature range of -40 ℃ to +125 ℃. These characteristics enable the sensor to play an important role in transformer fault diagnosis, ensuring that the model can accurately and stably extract and analyze fault signals. The accuracy error between the predicted and actual fault values of different models is shown in Fig. 7.

In Fig. 7, the accuracy error of the LSTM-RAFT is smaller than that of RAFT. The maximum accuracy of the predicted values in the RAFT model in Fig.7 (a) is around 90%, the minimum is around 25%, and the fluctuation range is 25%-90%. The maximum accuracy of the actual value is around 80%, and the minimum is 10%. The fit between the actual and the predicted values is not good. The accuracy of the predicted values of the fusion model in Fig. 7 (b) ranges from a maximum of 95% to a minimum of around 35%, with a fluctuation range of 35%-95%. The maximum accuracy of the actual value is around 90%, and the minimum is 35%, indicating a good fit between the two. The fault diagnosis accuracy of different models is shown in Fig. 8.

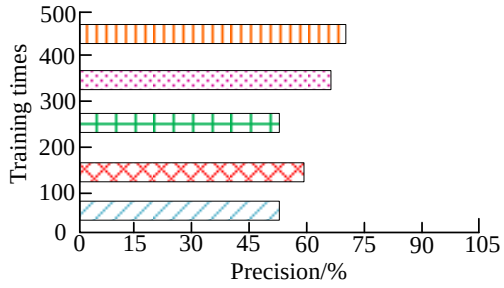


(a) Accuracy of predicted values in the RAFT model

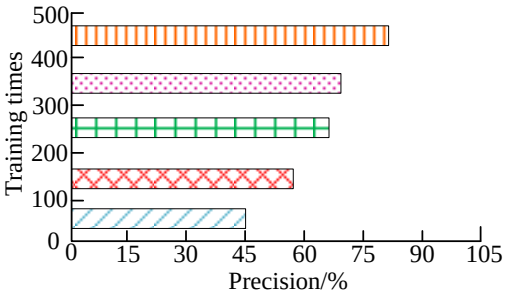


(b) Accuracy of predicted values in the LSTM-RAFT model

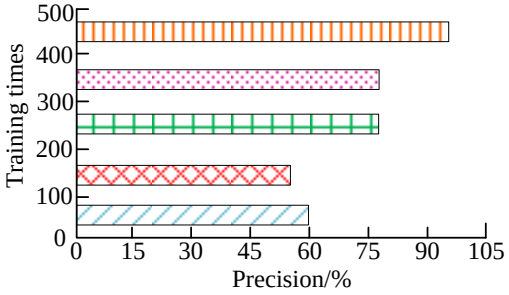
Fig. 7. Accuracy errors of predicted and actual fault values of different models



(a) The variation of fault diagnosis accuracy of LSTM model with training times



(b) The variation of fault diagnosis accuracy of LAFT model with training times



(c) The variation of fault diagnosis accuracy of LSTM-LAFT model with training times

Fig. 8. The precision of different models varies with the training times

In Fig. 8, among the comparison of fault diagnosis precision among all models, the fusion model has the highest precision. When trained 100 times, the minimum fault diagnosis precision of the LSTM model is 50%, RAFT is 45%, and LSTM-RAFT is 60%. As the training iterations increase, the maximum fault diagnosis precision of LSTM, RAFT, and LSTM-RAFT becomes 75%, 85%, and 95%.

4.2 Analysis of the application effect of LSTM-RAFT model

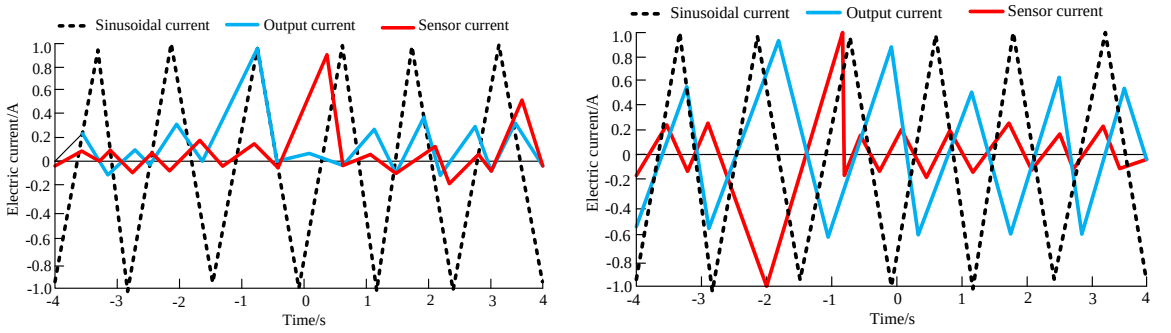
The accuracy of the LSTM-RAFT and the precision of fault diagnosis have shown stable performance and maintained at a high level in performance testing. Compared to other models, the fusion model has higher fault diagnosis performance. To further investigate the effectiveness of the model, it is applied to a fixed software setting for analysis. Table 2 shows the software parameter configuration of the model.

Table 2. Software dependency list

Designation	Versions	Feature
NFS	3.0	Through the network to achieve system, machine read and write
Java	1.9	Cross-platform design language, multi-linear program development
Python	2.8.4	Cross-platform design language, machine learning, deep learning
Mpi4py	2.7.0	Pass the interface, write the program
Keras	2.3.5	Neural network library for deep learning

Table 2 shows the names, versions, and corresponding functions of each software in the model application analysis. The matching version of software NFS is 3.0. The version of Java is 1.9, which is mainly designed through cross platform programming language and widely used in web applications. Python is compatible with version 2.8.4, and its main function is also to develop cross platform programming languages. The version requirement for Mpi4py is version 2.7.0. Keras uses version 2.3.5, mainly written in Python, as an open-source artificial neural network library for implementing deep learning. The normal variation of transformer current through the fusion model and the variation during fault occurrence are shown in Fig.9.

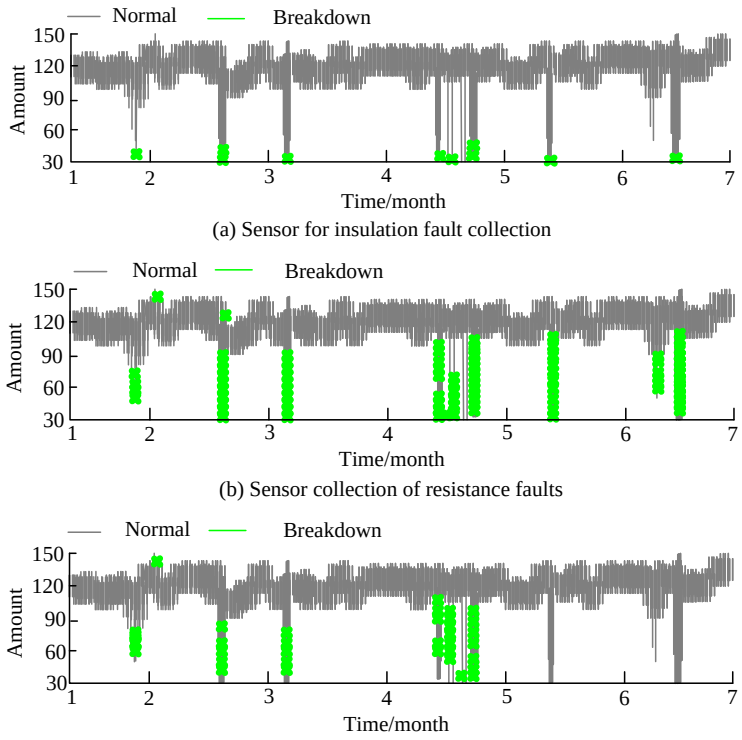
The current variation in Fig. 9 shows that the actual output value of the fault diagnosis current is significantly abnormal. The black curve represents the fluctuation of sine current, the blue curve denotes the fluctuation of input current, and the red curve means the change of current after passing through the sensor. Fig. 9 (a) shows the current output of the transformer through the sensor during normal operation. The fluctuation of output current is between -0.1-1, with a duration of 2s, while the variation of sensor current is between -0.1-1, with a shortened duration of 1s. Fig.9 (b) shows the current output of the transformer when a fault occurs. The fluctuation of output current is between -0.5-1, with a duration of 2s, while the variation of sensor current is between -1-1, with a shortened duration of 1s. To better analyze the application effect of the model, multiple sensors, including current transformer, Hall effect sensor, and Roche coil, are deployed at different positions of the transformer. The study is expressed as sensors 1, 2, and 3 to monitor the current change. It is imperative to record the current change data of each sensor during normal operation and in the event of a fault, including the amplitude, frequency, and time stamp of the current. The collected data are divided into two distinct subsets: a training set and a test set. The training set is used for



(a) The current output of the transformer through the sensor under normal operation

(b) The current output of the transformer through the sensor under fault operation

Fig. 9. Different variations of fusion models



(a) Sensor for insulation fault collection

(b) Sensor collection of resistance faults

(c) Sensor collection for electrical power faults

Fig. 10. Fault prediction capability of sensors at different locations

the training and validation of the model. The test set is employed for evaluating the fault detection ability of the sensors at different locations and the fault collection and prediction ability of each sensor. The prediction of sensors located at different positions of the transformer by the model is displayed in Fig. 10.

Figures 10 (a), (b), and (c) show the collection of insulation faults, resistance faults, and electrical power by sensors. As the model training iteration changes, the frequency of insulation faults is the least obvious in the collection of sensor faults, with only 10 faults collected. The maximum number of collected resistance fault sensors is 100. In the second month of fault collection, there is no further change in the situation, and sensor 2 still has the strongest ability to collect and predict faults.

5. CONCLUSION

With the widespread application of the power grid system, fault diagnosis of core equipment in the power system is becoming increasingly important. However, the complexity of the device environment and the current technological conditions have limitations in terms of precision for fault diagnosis. To improve the precision of fault diagnosis and reduce the cost of equipment maintenance, this study constructed an LSTM-RAFT model based on LSTM. The RAFT model has improved the collection ability of LSTM for weak fault signals and the precision and robustness for fault diagnosis. Performance experiments showed that the maximum accuracy of the predicted values in the fusion model was around 95%, and the minimum was around 35%. The maximum accuracy of the actual values was around 90%, and the minimum was around 35%. The actual values fitted well with the predicted values. The fault diagnosis accuracy was 60%, and as the training iterations increased, the maximum fault diagnosis accuracy became 95%, which is higher and more stable than the precision of other models. In the application experiment, it was found through the

fusion model that there were significant abnormal changes in the actual output value of the fault diagnosis current. Among the collected sensor faults, insulation faults had the fewest number, only 10. The maximum number of sensors collected for resistance faults was 100. The analysis of the fusion model found that the resistance fault of sensor 2 collected the highest number of faults. In summary, the research model has higher accuracy and precision in fault diagnosis compared to other extraction models, and can also respond well to abnormal current fluctuations and sensor fault collection prediction functions in application effects. However, this study did not conduct performance tests on the impact of different types of faults and noise issues, which is an area that can be improved in the future.

Source of funding: *This research received no external funding.*

Author contributions: *research concept and design, S.Z.; Collection and/or assembly of data, S.Z.; Data analysis and interpretation, N.P.; Writing the article, S.Z.; Critical revision of the article, N.P.; Final approval of the article, S.Z., N.P.*

Declaration of competing interest: *The author declares no conflict of interest.*

REFERENCES

1. Choudhuri S, Adeniye S, Sen A. Distribution alignment using complement entropy objective and adaptive consensus-based label refinement for partial domain adaptation. Artificial Intelligence and Applications. 2023;1(1):43-51. https://doi.org/10.47852/bonviewAIA2202524.
2. Long XM, Chen YJ, Zhou J. Development of AR experiment on electric-thermal effect by open framework with simulation-based asset and User-Defined Input. Artificial Intelligence and Applications. 2023;1(1):52-57. https://doi.org/10.47852/bonviewAIA2202359.
3. Amanul I, Othman F, Sakib N, Babu HMH. Prevention of Shoulder-Surfing attack using shifting condition with the digraph substitution rules. Artificial Intelligence and Applications. 2023;1(1):58-68. https://doi.org/10.48550/arXiv.2305.06549.
4. Zhang ZP, Wang GB. LSTM network optimization and task network construction based on heuristic algorithm. Journal of Computational Methods in Sciences and Engineering. 2024;24(2):697-714. https://doi.org/10.3233/JCM-237124.
5. Feng Q, Tu Y, Hou C, Cao B. TLN-LSTM: An automatic modulation recognition classifier based on a two-layer nested structure of LSTM network for extremely long signal sequences. International Journal of Web Information Systems. 2024;20(3):248-267. https://doi.org/10.1108/IJWIS-12-2023-0248.
6. He Y, Pang Y. Automatic detection of transformer health based on bayesian network model. Applied Mathematics and Nonlinear Sciences. 2023;8(2):2069-2076. https://doi.org/10.2478/amns.2023.1.00311.
7. Duval M, Buchacz J. Detection of carbonization of paper in transformers using duval pentagon 2 and triangle 5. IEEE Transactions on Dielectrics and Electrical Insulation. 2023;30(4):1534-1539. https://doi.org/10.1109/TDEI.2023.3278623.
8. Kia MY, Saniei MS, Seyyed G. A novel cyber-attack modelling and detection in overcurrent protection relays based on wavelet signature analysis. IET Generation, Transmission & Distribution. 2023;17(7): 1585-1600. https://doi.org/10.1049/gtd2.12766.
9. Doorwar A, Bhalja BR, Malik OP. Novel approach for synchronous generator protection using new differential component. Transactions on Energy Conversion 2023;38(1):180-191. https://doi.org/10.1109/TEC.2022.3196005.
10. Pramanik S, Ganesh A, Duvvury Chaitanya VSB. Double-end excitation of a single isolated transformer winding: An improved frequency response analysis for fault detection. IEEE Transactions on Power Delivery. 2022;37(1):619-626. https://doi.org/10.1109/TPWRD.2021.3067863.
11. Eruvai M, Chilaka R. Oil health index calculation and incipient fault diagnosis in power transformers using fuzzy logic. Insight: Non-Destructive Testing and Condition Monitoring. 2022; 64(1): 28-37. https://doi.org/10.1784/insi.2022.64.1.28.
12. Su L, Yin M, Zhao S. PSR-LSTM model for weak pulse signal detection. Multimedia Tools and Applications. 2023;82(23):35853-35877. https://doi.org/10.1007/s11042-023-14987-w.
13. Windmann S, Westerhold T. Fault detection in automated production systems based on a long short-term memory autoencoder. at – Automatisierung-stechnik. 2024;72(1):47-58. https://doi.org/10.1515/auto-2023-0031.
14. Aljemely AH, Xuan J, Al-Azzawi O, Jawad FKJ. Intelligent fault diagnosis of rolling bearings based on LSTM with large margin nearest neighbor algorithm. Neural Computing and Applications. 2022;34(22): 19401-19421. https://doi.org/10.1007/s00521-022-07353-8.
15. Sun HB, Fan YG. Fault diagnosis of rolling bearings based on CNN and LSTM networks under mixed load and noise. Multimedia Tools and Applications. 2023; 82(28): 43543-43567. https://doi.org/10.1007/s11042-023-15325-w.
16. Sampat C, Ramachandran R. Optimizing energy efficiency of a twin-screw granulation process in Real-Time Using a Long Short-Term Memory (LSTM) Network. ACS Engineering. Au 2024;4(2):278-289. https://doi.org/10.1021/acsengineeringau.3c00038.
17. Bouziane SE, Khadir MT. Towards an energy management system based on a multi-agent architecture and LSTM networks. Journal of Experimental & Theoretical Artificial Intelligence. 2024;3(4):469-487. https://doi.org/10.1080/0952813X.2022.2093407.
18. Li WT. Construction and analysis of QPSO-LSTM model in network security situation prediction. Journal of Cyber Security and Mobility. 2024;13(3):417-438. https://doi.org/10.13052/jcsm2245-1439.1334.
19. Bian C, Huang G. Predicting PM2.5 concentration with enhanced state-trend awareness and uncertainty analysis using bagging and LSTM neural networks. Journal of Environmental Quality. 2024;53(4):441-455. https://doi.org/10.1002/jeq2.20589.
20. Shen Z, Liu X, Li W, Li XY, Wang Q. Classification of visually induced motion sickness based on phase-locked value functional connectivity matrix and CNN-

LSTM. Sensors. 2024;24(12): 3936-3936. https://doi.org/10.3390/s24123936.
21. Ding C, Zhao J, Sun S. Concept drift adaptation for time series anomaly detection via transformer. Neural Processing Letters. 2023;55(3):2081-2101. https://doi.org/10.1007/s11063-022-11015-0.
22. Jin X, Ma C, Luo S, Zeng P, Wei Y. Distributed IIoT anomaly detection scheme based on blockchain and federated learning. Journal of Communications and Networks. 2024;26(2):252-262. https://doi.org/10.23919/JCN.2024.000016.

Shuzong ZHAO, born in 1982, Weifang, Shandong, China, graduated from School of Electrical Engineering and Automation in Wuhan University of Technology in 2007, Bachelor's degree, and in 2013, Master's degree. Focusing on the research of high voltage and insulation technology, he has been a doctoral candidate majoring in High Voltage Laborator, Department of Electrical Engineering, Faculty of Engineering King Mongkut's Institute of Technology Ladkrabang since August, 2020.
Working Experiences: From 2007 to 2009, technician, in Energy Experimental Research Center, Wuhan Iron and Steel (Group) Corp., China; From 2009 until now, professor, in Guangxi Electrical Polytechnic Institute, China; From 2020 until now, doctoral candidate, in Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang.
Academic Experiences: Published 2 academic papers; Led or participated in 5 scientific research projects; Applied for 5 utility model patents.
e-mail: lionogj@sina.com

Norasage PATTANADECH received B. Eng. and M. Eng. degree in electrical engineering from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 1998 and Chulalongkom University, Thailand, in 2002 respectively. He is also rewarded for Dr.techn. from the Institute of High Voltage Engineering and System Management, Graz University of Technology, Austria, in 2013.
Working Experience: Currently, he works as an associate professor at Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand.
Academic Experiences: Published 173 research document contributed on Mineral Oil; Dielectrics; Power Transformer; Dielectrics; including 2 academic textbooks.
E-mail" norasage.pa@kmitl.ac.th