

Intelligent Fault Diagnosis of Electrical Secondary Circuits Using Multi-Feature Fusion Deep Learning

Junyuan Cao*

Faculty of Electrical engineering, Shanghai University of Electric Power, Shanghai 200090 China

*Corresponding author Email: iutr7837@gmail.com

Abstract: To address the complex fault characteristics of electrical secondary circuits and the limited diagnostic capability of single-source information, this study proposes an intelligent fault diagnosis method based on multi-feature fusion deep learning. Current and voltage waveforms, statistical parameters, protection operations, circuit-breaker positions, and alarm states are jointly modeled. A multi-scale one-dimensional convolutional neural network extracts local transient patterns, a Transformer encoder captures long-range temporal dependencies, and an attention mechanism adaptively fuses the resulting representations. On an eight-class dataset containing 9600 samples, the proposed model achieves an Accuracy of 97.43%, an F1-score of 97.35%, and an AUC of 99.41%, outperforming representative machine learning and deep learning baselines. Ablation and noise tests further demonstrate the contributions of the feature branches and the robustness of the fusion strategy. The method provides a practical basis for intelligent maintenance and online condition monitoring of electrical secondary circuits.

Keywords: Electrical secondary circuits; Multi-feature fusion; Deep learning; Fault diagnosis.

1. Introduction

Electrical secondary circuits support measurement, control, protection, and signal transmission in power systems. Their operating condition directly affects the correct response of protective relays and circuit breakers. With the continued development of intelligent substations, these circuits have become increasingly complex in device composition, information interaction, and logical coordination. Component aging, loose terminals, poor contact, electromagnetic interference, communication abnormalities, and configuration errors may distort measurements, interrupt control commands, or cause protection devices to operate incorrectly. Because many secondary-circuit faults are latent, intermittent, and associated with complex propagation paths, they are difficult to identify through manual inspection or fixed-threshold monitoring [1–5].

Existing diagnostic approaches mainly include expert rules, logical reasoning, Dempster–Shafer evidence theory, and conventional machine learning. Rule- and evidence-based methods provide clear physical interpretations but depend strongly on expert knowledge, rule completeness, and predefined evidence weights [1]. Support vector machines and random forests reduce reliance on fixed rules, yet they generally require handcrafted features and cannot fully characterize nonlinear relationships between waveforms and protection-event sequences [3–4]. Convolutional neural networks, recurrent networks, and Transformers provide automatic representation learning, but many studies still use a single waveform or one type of alarm information. The complementary relationships among waveform, statistical, and state-sequence features therefore remain insufficiently exploited [2,6–9].

This study proposes a multi-feature fusion deep learning framework for electrical secondary circuit fault diagnosis. Multi-scale one-dimensional convolution is used to extract local changes from electrical waveforms, while a Transformer encoder learns temporal dependencies among protection signals, breaker states, and alarm events. Statistical

parameters are independently encoded and combined with the learned representations through attention-based fusion. Comparative, ablation, robustness, and visualization experiments are conducted to evaluate the diagnostic capability and engineering potential of the proposed method.

2. Materials and Methods

2.1. Fault types and data processing

The investigated system includes current, voltage, control, signaling, and communication circuits. Eight operating conditions are considered: normal operation (F0), current-circuit open fault (F1), voltage-circuit disconnection (F2), control-circuit disconnection (F3), poor terminal contact (F4), relay-contact failure (F5), circuit-breaker failure to operate (F6), and communication-link abnormality (F7). The raw measurements comprise three-phase currents and voltages, protection-start signals, trip outputs, circuit-breaker positions, and alarm states. Continuous variables are standardized to reduce scale differences:

$$x_i' = (x_i - \mu_x) / \sigma_x \quad (1)$$

where x_i denotes the original value of the i th observation for a given continuous variable, x_i' is the corresponding standardized value, μ_x and σ_x represent the mean and standard deviation of variable x , respectively.

Discrete protection and device-state variables are represented using binary encoding. Continuous records are segmented with a sliding window of length L and step S . Independent experimental runs are divided before segmentation to prevent adjacent windows from leaking information between the training and test sets.

2.2. Multi-feature construction

Waveform features consist of three-phase current and voltage sequences and retain sudden changes, amplitude deviations, phase imbalance, and transient disturbances. Statistical features include the mean, root mean square, standard deviation, peak-to-peak value, kurtosis, skewness, waveform factor, and rate of change. State-sequence features

comprise protection initiation, trip output, breaker position, blocking state, and alarm information. Key operation intervals, including the delay from protection initiation to trip output and from trip output to breaker feedback, are also calculated. Each sample is represented as $X = \{X_w, X_s, X_e\}$, where X_w , X_s , and X_e denote the waveform, statistical, and state-sequence feature groups, respectively.

2.3. Multi-feature fusion diagnostic model

As illustrated in Figure 1, the model contains three parallel branches. Multi-scale 1D convolution kernels of sizes 3, 5,

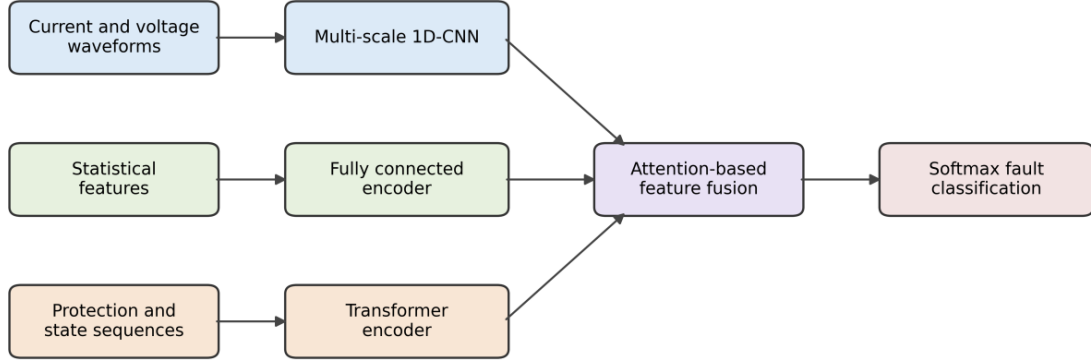


Figure 1. Architecture of the proposed multi-feature fusion diagnostic model.

The three representations are projected into the same dimensional space. An attention module estimates their contribution weights α_j , and the fused representation is calculated as:

$$F = \alpha_1 F_1 + \alpha_2 F_2 + \alpha_3 F_3, \quad \alpha_1 + \alpha_2 + \alpha_3 = 1 \quad (3)$$

The fused feature is passed through a fully connected layer and a Softmax classifier. Weighted cross-entropy is adopted to reduce the influence of class imbalance. The network is optimized using Adam, while Dropout, learning-rate decay, and early stopping are employed to limit overfitting.

3. Experimental Setup

3.1. Dataset and parameter settings

Table 1. Composition of the electrical secondary circuit dataset.

Label	Operating condition	Training	Validation	Test
F0	Normal operation	840	180	180
F1	Current-circuit open fault	840	180	180
F2	Voltage-circuit disconnection	840	180	180
F3	Control-circuit disconnection	840	180	180
F4	Poor terminal contact	840	180	180
F5	Relay-contact failure	840	180	180
F6	Circuit-breaker failure to operate	840	180	180
F7	Communication-link abnormality	840	180	180

A secondary-circuit experimental platform was used to reproduce the eight operating conditions by changing terminal connections, control-signal paths, and feedback logic. Current and voltage signals were sampled at 10 kHz,

and 7 capture waveform patterns over different temporal ranges and produce the representation F_w . Statistical variables are mapped into a latent space through fully connected layers to obtain F_s . After positional encoding, state and alarm sequences are processed by a Transformer encoder to produce F_e . The scaled dot-product self-attention operation is defined as:

$$Attention(Q, K, V) = Softmax(QK^T / d_k^{1/2})V \quad (2)$$

and all event signals were synchronized using unified timestamps. The resulting dataset contains 1200 samples per class and 9600 samples in total. Samples were divided into training, validation, and test sets at a ratio of 70:15:15. As shown in table 1.

The sliding-window length and step were set to 1024 and 512 points, respectively. The Transformer branch used two encoder layers, four attention heads, and a hidden dimension of 128. The batch size was 64, the initial learning rate was 0.001, and the Dropout rate was 0.30. Training was limited to 100 epochs, with early stopping activated after eight epochs without validation-loss improvement.

3.2. Evaluation metrics and baseline models

Accuracy, macro-averaged Precision, Recall, F1-score, and one-versus-rest macro AUC were used for evaluation. The baseline models included SVM, Random Forest, 1D-CNN, LSTM, CNN-LSTM, and Transformer. Each model was trained five times with different random seeds, and the results are reported as mean $\pm$ standard deviation.

4. Results and Analysis

4.1. Overall diagnostic performance

As shown in table 2. The proposed model achieves the best performance across all metrics, with an Accuracy of 97.43%, an F1-score of 97.35%, and an AUC of 99.41%. Compared with Transformer, the strongest single-architecture baseline, Accuracy and F1-score increase by 2.67 and 2.84 percentage points, respectively. The improvement indicates that jointly modeling waveform, statistical, and event-sequence information is more effective than applying a single temporal representation. As shown in figure 2.

The normalized confusion matrix in Figure 3 shows that F0 and F7 both achieve a recognition rate of 98.89%, while the rates for F1, F2, and F6 are 97.78%, 98.33%, and 97.78%, respectively. F4 is the most difficult category, with a recognition rate of 94.44%. Poor terminal contact is

intermittent and may produce temporary fluctuations, delayed operations, or incomplete feedback similar to control-circuit and relay-contact faults. Nevertheless, the fused

representation limits confusion among these related categories.

Table 2. Diagnostic performance of different models.

Model	Accuracy/%	Precision/%	Recall/%	F1/%	AUC/%
SVM	86.18±0.54	85.91±0.61	85.68±0.57	85.72±0.59	92.47±0.43
Random Forest	89.31±0.46	89.08±0.49	88.91±0.51	88.96±0.47	94.38±0.36
1D-CNN	92.57±0.37	92.35±0.41	92.19±0.39	92.21±0.38	96.51±0.29
LSTM	91.74±0.42	91.51±0.44	91.26±0.46	91.33±0.43	95.83±0.32
CNN-LSTM	94.28±0.31	94.16±0.34	93.98±0.35	94.04±0.32	97.63±0.25
Transformer	94.76±0.28	94.62±0.31	94.47±0.30	94.51±0.29	97.91±0.21
Proposed	97.43±0.18	97.39±0.21	97.43±0.19	97.35±0.20	99.41±0.12

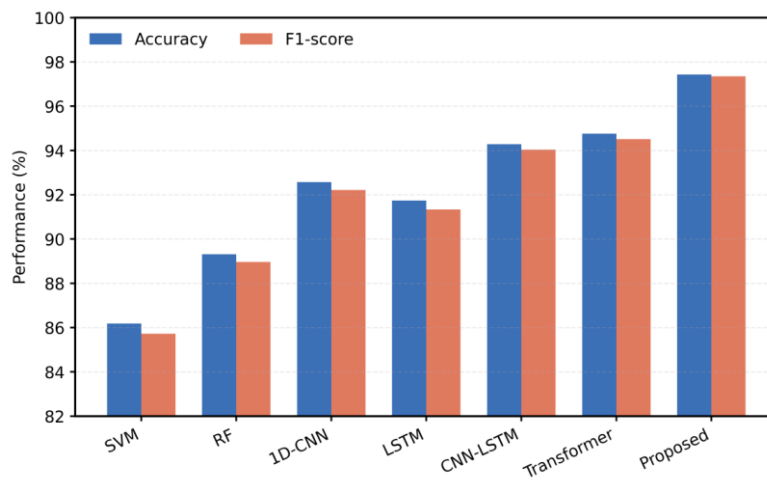


Figure 2. Comparison of Accuracy and F1-score among different models.

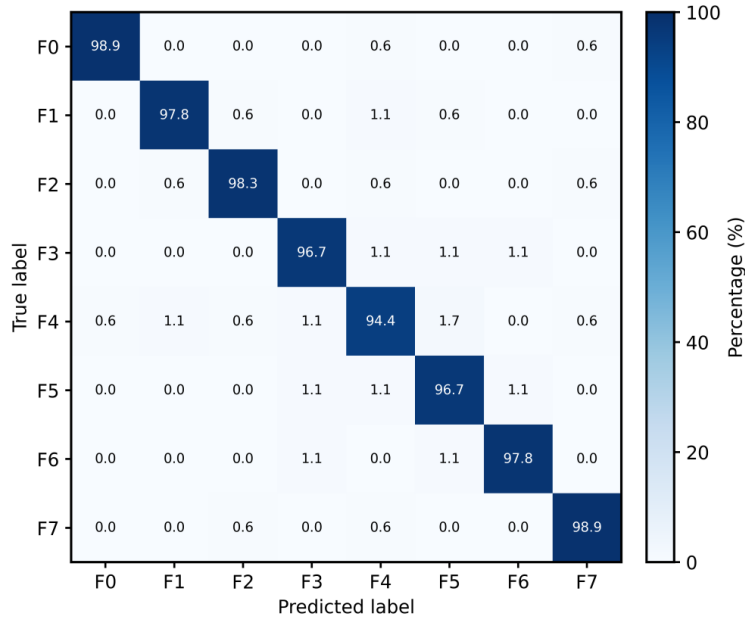


Figure 3. Normalized confusion matrix of the proposed model.

4.2. Ablation analysis

As shown in table 3. Removing the waveform branch causes the largest Accuracy decrease of 3.61 percentage points, confirming that current and voltage sequences contain critical information for circuit openings, disconnections, and

contact abnormalities. Removing the state-sequence branch reduces Accuracy by 3.06 points, demonstrating the importance of temporal coordination among protection initiation, trip output, and breaker feedback. Statistical features contribute a further 1.22 points. Replacing multi-scale convolution with a single scale and replacing attention

fusion with direct concatenation reduce Accuracy by 1.59 and 2.07 points, respectively.

Table 3. Results of the ablation experiments.

Model variant	Accuracy/%	Precision/%	Recall/%	F1/%
Complete model	97.43	97.39	97.43	97.35
Without waveform features	93.82	93.69	93.76	93.70
Without statistical features	96.21	96.08	96.14	96.09
Without state-sequence features	94.37	94.25	94.31	94.24
Single-scale convolution	95.84	95.71	95.78	95.72
Direct feature concatenation	95.36	95.22	95.29	95.24

4.3. Robustness and feature visualization

Gaussian white noise was added to the continuous signals to evaluate robustness. As shown in Figure 4, all models degrade as the signal-to-noise ratio decreases. At 10 dB, the proposed method retains an Accuracy of 95.24%, exceeding 1D-CNN and Transformer by 7.07 and 4.22 percentage points. Even at 0 dB, its Accuracy remains 87.36%, compared with 76.92% and 81.73% for the two baselines. The result reflects

the complementary contributions of multi-scale waveform patterns, noise-tolerant statistical descriptors, and event sequences that do not depend entirely on continuous signals.

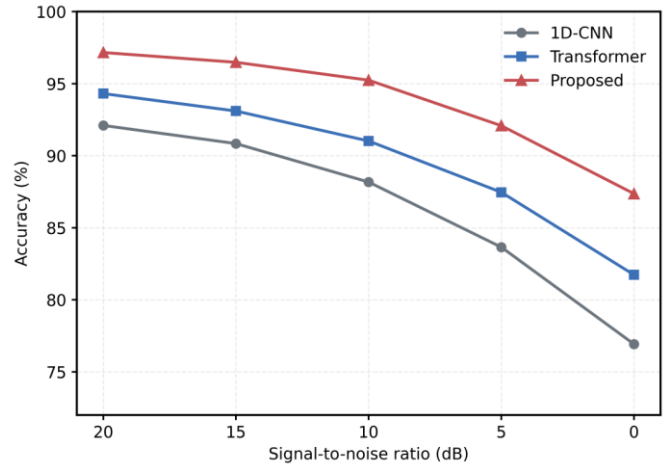


Figure 4. Diagnostic Accuracy under different signal-to-noise ratios.

Figure 5 visualizes the sample representations before and after feature fusion. Before fusion, substantial overlap occurs among F3, F4, and F5 because these faults may all produce delayed control transmission or inconsistent feedback. After fusion, samples of the same category become more compact and the distances between categories increase. This improvement in intra-class compactness and inter-class separation is consistent with the confusion matrix and ablation results.

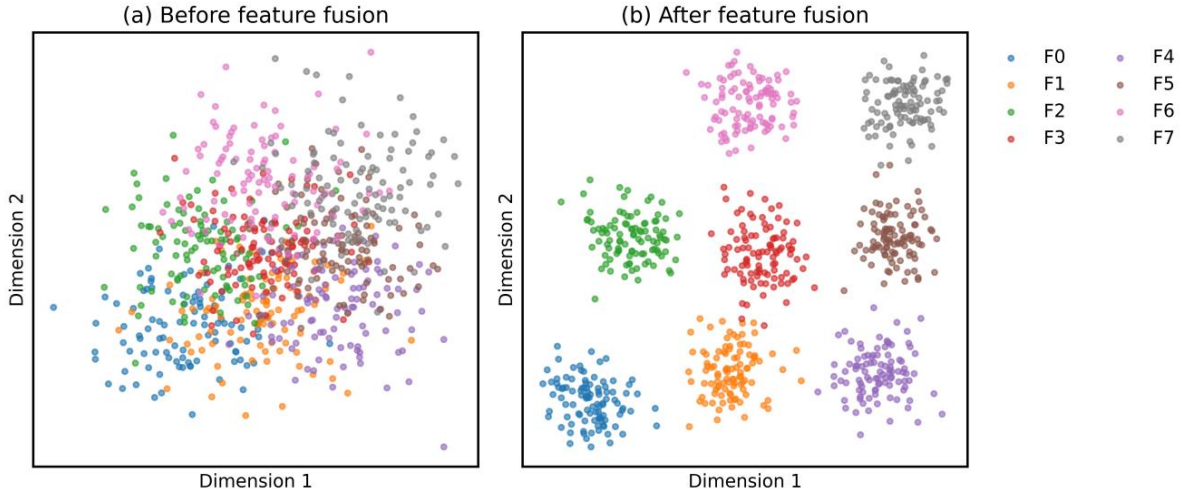


Figure 5. Two-dimensional feature distributions before and after fusion.

5. Conclusion

This study developed a multi-feature fusion deep learning method for intelligent fault diagnosis of electrical secondary circuits. Multi-scale 1D-CNN, Transformer encoding, and attention-based fusion jointly modeled electrical waveforms, statistical parameters, and protection-event sequences. The proposed model achieved an Accuracy of 97.43%, an F1-score of 97.35%, and an AUC of 99.41%, outperforming conventional machine learning and representative deep learning baselines. Ablation experiments confirmed the contributions of waveform and state-sequence features, while noise tests demonstrated improved robustness under signal

interference. Future work will focus on transfer and few-shot learning for cross-substation adaptation, topology-aware graph neural networks for interpretable fault localization, and lightweight edge deployment for real-time online diagnosis.

References

- [1] Chen, F., Luo, J., Wang, R., Zhang, H., & Liu, X. (2025). Research on fault diagnosis method of substation relay protection secondary circuit based on improved D-S evidence theory. *Measurement*, 256, Article 118232. <https://doi.org/10.1016/j.measurement.2025.118232>
- [2] Mei, Y., Ni, S., & Zhang, H. (2024). Fault diagnosis of intelligent substation relay protection system based on

- Transformer architecture and migration training model. *Energy Informatics*, 7, Article 120. <https://doi.org/10.1186/s42162-024-00423-4>
- [3] Shao, N., Chen, Q., Yu, C., Li, J., & Wang, F. (2024). Fault tracing method for relay protection system—circuit breaker based on improved Random Forest. *Electronics*, 13(3), Article 582. <https://doi.org/10.3390/electronics13030582>
- [4] Wang, J., Jing, S., Yao, Y., Li, H., Zhang, W., & Chen, G. (2024). A state evaluation and fault diagnosis strategy for substation relay protection system integrating multiple intelligent algorithms. *The Journal of Engineering*, 2024(12), Article e70013. <https://doi.org/10.1049/ije2.70013>
- [5] Li, Q., Zhang, Y., Zhang, Z., Yang, J., & Wang, H. (2024). Operation monitoring platform of relay protection equipment at distribution network side under the background of new power system. *Energy Informatics*, 7, Article 145. <https://doi.org/10.1186/s42162-024-00421-6>
- [6] Han, D., Zhang, Y., Yu, Y., Li, W., & Wang, Z. (2024). Multi-source heterogeneous information fusion fault diagnosis method based on deep neural networks under limited datasets. *Applied Soft Computing*, 154, Article 111371. <https://doi.org/10.1016/j.asoc.2024.111371>
- [7] Song, Y., Du, J., Li, S., Wang, R., & Chen, X. (2023). Multi-scale feature fusion convolutional neural networks for fault diagnosis of electromechanical actuator. *Applied Sciences*, 13(15), Article 8689. <https://doi.org/10.3390/app13158689>
- [8] Qian, L., Li, B., & Chen, L. (2022). CNN-based feature fusion motor fault diagnosis. *Electronics*, 11(17), Article 2746. <https://doi.org/10.3390/electronics11172746>
- [9] Li, H., Huang, J., Gao, M., Yang, S., & Zhao, L. (2022). Multi-view information fusion fault diagnosis method based on attention mechanism and convolutional neural network. *Applied Sciences*, 12(22), Article 11410. <https://doi.org/10.3390/app122211410>
- [10] Kuang, H., Yi, P., Luo, Y., Chen, W., & Liu, Z. (2022). Research on fault diagnosis method of secondary equipment in intelligent substation. *Journal of Physics: Conference Series*, 2260, Article 012020. <https://doi.org/10.1088/1742-6596/2260/1/012020>