Multi-source Traveling Wave Signal Fusion and Fault Identification Supported by Cloud Computing

Xingyuan Xu *, Quanhui Liu, ZeMing Chen, Lianchen Li, Na You

Guangdong Power Grid Co., Ltd. Guangzhou Zengcheng Power Supply Bureau, Zengcheng 511300, Guangzhou, China
**Corresponding Author:Xingyuan Xu*
Email: xingyuanxu11@outlook.com

Received 20 March 2026; Accepted 23 June 2026

Abstract: The accuracy and real-time performance of fault identification of multi-source traveling wave signals in current power systems are difficult to guarantee due to problems such as asynchronous acquisition, data heterogeneity, and processing delay. In view of the above problems, this paper constructs a multi-source traveling wave signal fusion and fault identification method based on cloud computing platform, relying on the advantages of cloud computing in resource elastic scheduling and distributed computing to achieve efficient concurrent processing of large-scale heterogeneous signals and global information collaborative analysis. The signal time alignment is completed through the global clock synchronization mechanism, and the unified data format is used for standardized processing. The transient characteristics of each source signal are extracted based on wavelet transform, and the multi-source characteristics are fused in the cloud before inputting into the deep neural network model to complete fault identification and location. The adopted fault identification model is constructed based on long short-term memory network and spatio-temporal attention mechanism, which can realize dynamic weighted fusion of multi-source signal characteristics in time dimension and space dimension, and output fault type discrimination results and location estimation values in an end-to-end manner. Experimental results show that the proposed method achieves a fault identification accuracy of 98.6% on the IEEE 39-node power system simulation platform, and the average system response time is 1.3 seconds, which meets the real-time identification requirements. The research results verify the effectiveness of this method in improving the efficiency and reliability of power system fault handling, and provide technical support for smart grid fault monitoring in the Internet of Things environment.

Keywords: Power System Fault; Traveling Wave Signal; Data Fusion; Cloud Computing; Time Synchronization

1. Introduction

With the deep integration of smart grids and the Internet of Things (IoT), power systems are required to achieve higher levels of real-time performance, accuracy, and intelligence in fault detection and identification [1-2]. Traveling wave signals, as high-frequency transient signals with fast response characteristics, have become an important basis for fault identification and location in modern power systems. However, as grid structures become increasingly complex and the number of measurement points continues to grow, traveling wave signals exhibit multi-source distribution, data heterogeneity, and massive data volume characteristics [3-4]. These features pose significant challenges to traditional signal processing and fault identification methods [5-6].

In practical power system monitoring scenarios, traveling wave signals collected from different measurement points often suffer from time asynchrony, inconsistent data formats, and strong noise interference [7-8]. These factors reduce the timeliness, consistency, and availability of the acquired signals, making it difficult to effectively align and fuse multi-source traveling wave data within a unified time domain. Consequently, the efficiency and reliability of fault judgment are adversely affected [9-10]. In addition, most existing signal fusion methods rely on edge devices or local computing architectures. Owing to limited computing resources, fixed resource allocation, and processing delays, such methods are often unable to meet the real-time response requirements of large-scale power grid fault detection [11-12]. In IoT-based power system environments, the multidimensionality, dynamic variation, and complex processing flow of monitoring data further intensify these challenges [13-14]. Therefore, efficient collaborative fusion and processing of multi-source traveling wave signals on a unified platform, as well as fast and accurate fault identification, have become key issues in power system monitoring and protection [15].

To address the above problems, this paper proposes a cloud-computing-based multi-source traveling wave signal fusion and fault identification method. Compared with traditional local centralized architectures, cloud computing platforms provide advantages in elastic resource scheduling, distributed computing, and collaborative processing of heterogeneous data. Specifically, containerized deployment and distributed resource scheduling mechanisms are adopted to support dynamic task allocation and elastic system scaling through the Kubernetes cluster management framework. When the number of measurement points increases significantly, the system can automatically expand feature extraction nodes to ensure timely signal processing. Meanwhile, distributed message middleware is introduced to realize orderly distribution and unified caching of multi-channel traveling wave data, thereby reducing time dislocation caused by transmission jitter. In addition, the heterogeneous resource pool equipped with graphics processing units (GPUs) provides computational support for deep neural network models, enabling efficient feature fusion and fault identification.

Based on the above design, this paper constructs an integrated system architecture covering data acquisition, preprocessing, feature extraction, fusion analysis, and fault identification. A high-precision clock synchronization mechanism is deployed, and the Global Positioning System (GPS) is used to achieve time alignment among measurement points. Anti-noise filtering and standardized data conversion are then performed to construct a unified feature vector set. In the cloud-based fusion stage, the time-frequency characteristics of signals from different measurement points are integrated through a multi-channel feature fusion network [16]. Furthermore, a deep learning model based on the Long Short-Term Memory (LSTM) network is employed to identify fault types and estimate fault locations. The proposed method enables efficient processing and rapid inference of large-scale multi-source traveling wave data, thereby promoting the development of power system fault diagnosis toward cloud-based collaborative intelligence.

2. Related Works

The research on multi-source traveling wave signal fusion technology focuses on improving fault identification accuracy, improving ranging methods, and optimizing fusion mechanisms. LI H et al. [17] proposed a fault ranging method based on a multi-terminal traveling wave frequency matrix. By establishing a mathematical relationship between the main frequency component and the transmission distance, accurate location can be achieved under complex distribution network conditions, reducing the dependence on line parameter frequency changes and fault location types. Zhang H [18]'s team constructed an identification model based on U-Net and multimodal fusion mechanism. By extracting features from time-frequency graphs and wave frequency graphs, the identification rate of four-signal aliasing exceeds 97.3% under low signal-to-noise ratio conditions, verifying the effectiveness of multimodal fusion. Sun H [19]'s team applied multi-source information fusion technology to the fault diagnosis of key devices in railway signal systems and promoted the development of fault analysis towards data-driven through sensor data fusion. Although existing research has made progress in some local links, there are still problems such as lack of unified standards for fusion strategies, insufficient algorithm generalization ability, and low engineering matching. Traditional local centralized or edge computing modes are limited by resource constraints and are difficult to support large-scale heterogeneous data collaborative processing. The multimodal fusion mechanism of Zhang H's team relied on single device or local network modeling, which lacked flexibility in task scheduling in high-concurrency multi-point scenarios and failed to take advantage of cloud-based distributed computing. In addition, the static weight allocation strategy is difficult to adapt to the dynamic evolution of complex fault modes in power systems.

Although the cloud computing architecture of the parallel programming model proposed by Al-Jumaili A H A et al. [20] emphasized resource elastic scheduling and large-scale concurrent task management capabilities, it was not optimized for the scenario of burst collection of massive traveling wave signals in power grids. The idea of Maurya M's [21] team on integrating artificial intelligence algorithms in the cloud to improve the level of intelligent device state perception provided a theoretical basis for the end-to-end signal feature extraction and fault identification process in this paper, but did not involve the real-time synchronization problem of streaming data caused by the high-frequency transient characteristics of traveling wave signals. Although the lightweight deployment strategy on the edge side of Lu S's [22] team provided a reference direction for model compression and inference acceleration, it did not solve the dynamic scheduling mechanism of signal feature fusion in heterogeneous resource environments. The above studies reveal the general advantages of cloud computing and edge computing in intelligent fault diagnosis, but do not map their technical characteristics to the specific needs of traveling wave signal processing in power systems. The traditional local centralized method is limited by the fixed resource allocation and communication bottlenecks of the measuring points, and it is difficult to support the global collaborative analysis of large-scale signals; the edge computing fusion strategy lacks a unified scheduling mechanism, resulting in insufficient ability to model the temporal correlation across measuring points, and is difficult to adapt to high concurrent dynamic loads. In contrast, the cloud computing platform relies on distributed resource scheduling to support unified access and elastic computing of multi-source data, realizes dynamic scaling of tasks through containerized deployment, and combines message middleware to ensure orderly reception and caching of data streams. In view of the adaptability defects of the static weight allocation strategy under complex power grid fault modes, the method in this paper integrates heterogeneous resource pools and parallel computing frameworks to achieve spatio-temporal joint optimization of multi-measurement point signals, significantly enhancing feature expression capabilities and identification robustness.

3. Methods

3.1 Data Acquisition and Clock Synchronization of Multi-source Traveling Wave Signals

The acquisition system of multi-source traveling wave signals is based on distributed intelligent terminal devices deployed at key nodes of the power system [23-24]. Each terminal device is equipped with a high-precision atomic clock or GPS synchronization module as a global time reference. During the acquisition process, all signals are timestamped with the UTC time standard to achieve cross-device timing unification. Clock synchronization adopts a synchronization mechanism based on the Precision Time Protocol (PTP) and dynamically adjusts the local clock deviation through the master-slave clock architecture to ensure that the time error of each terminal is controlled at the sub-microsecond level.

In the process of multi-source signal acquisition, time synchronization is an important prerequisite for ensuring signal consistency. To achieve high-precision time alignment across measurement points, this method constructs a global clock synchronization mechanism based on the cloud computing platform and uses the Precision Time Protocol (PTP) with a master-slave structure to complete the time calibration of each acquisition terminal. This mechanism dynamically manages the local clock deviations of multiple edge nodes through a centralized cloud scheduler and uses a containerized microservice module to execute a recursive filtering algorithm to improve the stability and response speed of clock synchronization. Compared with the traditional local centralized architecture, the synchronization mechanism in the cloud computing environment has higher scalability and fault tolerance and can maintain sub-microsecond time error control when a large number of nodes are connected.

The clock deviation Δt_i is defined as the difference between the local time of the i-th acquisition device and the standard time. Its adjustment process is shown in Formula 1:

$$t_i^{adjusted} = t_i^{local} - \Delta t_i, \Delta t_i = f(\Delta t_{i-1}, \Delta t_{ref}, \delta t) \tag{1}$$

Among them, t_i^{local} is the local acquisition time of the i-th device; Δt_{ref} is the reference master clock deviation; δt is the communication delay compensation function; function f achieves stable convergence based on the recursive filtering algorithm of the clock deviation. This method effectively solves the time asynchrony problem caused by clock drift and network delay in distributed acquisition nodes, ensures the high consistency of multi-source signal timestamps, and meets the strict requirements of subsequent fusion algorithms for time alignment.

The data acquisition module designs a unified data transmission protocol, uses edge computing nodes to pre-process the collected data, and completes data integrity verification and preliminary format conversion in real-time. The signal sampling rate is uniformly set to 1MHz to ensure sufficient time resolution to cover the transient characteristics of the traveling wave signal. Figure 1 illustrates the clock synchronization and data transmission process framework in the multi-source signal acquisition system and shows the collaborative working structure of the synchronization module, edge pre-processing, and cloud reception.

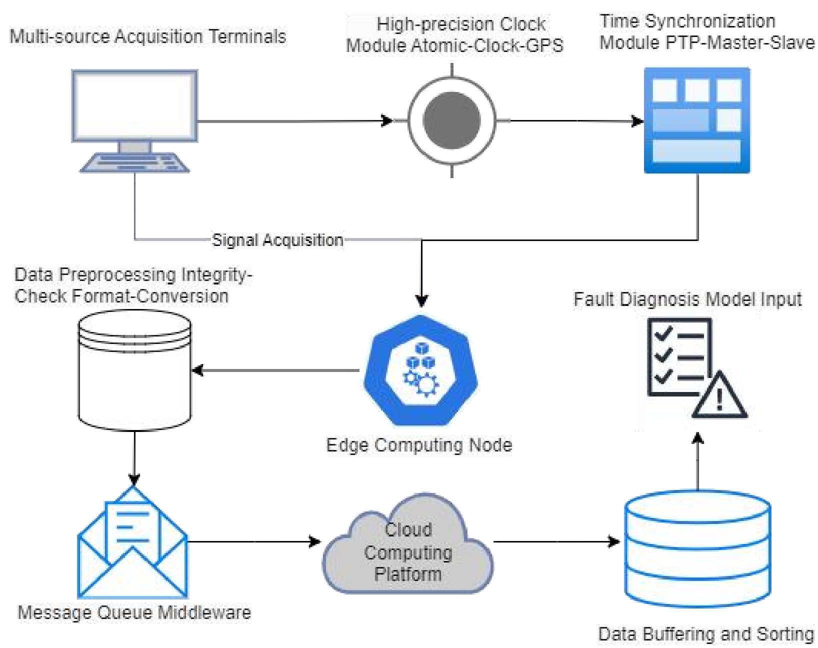

Figure 1. Multi-source traveling wave signal acquisition and clock synchronization system architecture

Figure 1 shows the system architecture of the entire process of multi-source traveling wave signal acquisition to cloud fault identification. The multi-source acquisition terminal is responsible for accessing the primary device signal. The high-precision clock module composed of the atomic clock and GPS provides a unified time base to the time synchronization module and completes the timestamp consistency through the PTP master-slave structure. The synchronized and calibrated signal is pre-processed in the edge computing node and is transmitted to the local database after integrity verification and format conversion operations. The data is uploaded to the cloud computing platform through the message queue middleware and is uniformly cached and reordered in the cloud to construct a standardized input vector. The system finally inputs the fused signal features into the fault diagnosis model to achieve precise identification and location tasks based on time-frequency features.

To quantify the acquisition time difference of multiple measurement points, the signal time alignment error index ϵ_{time} is defined, and its expression is:

$$\epsilon_{time} = \max_{i,j} |t_i^{adjusted} - t_j^{adjusted}| \tag{2}$$

This index is used as the core parameter for synchronization performance evaluation in the system design stage. By adjusting the PTP synchronization period and filtering parameters, ϵ_{time} can be controlled below 1 microsecond, meeting the technical requirements of strict timing alignment of traveling wave signals. Table 1 analyzes the synchronization accuracy differences of each terminal.

Table 1. Statistics of time synchronization errors of multi-source acquisition terminals

Collection terminal number	Mean time error (μs)	Time error standard deviation (μs)	Maximum time error (μs)
1	0.45	0.07	0.62
2	0.38	0.06	0.55
3	0.52	0.09	0.74
4	0.47	0.08	0.68
5	0.41	0.05	0.59

Table 1 lists the time error statistical indicators of each acquisition terminal in the system under the unified time synchronization mechanism. Each terminal number corresponds to a set of synchronization error data; the mean time error reflects the average offset of the terminal relative to the master clock; the standard deviation indicates the degree of fluctuation of the time error; the maximum time error is used to measure the maximum deviation amplitude during the synchronization process. All data are in microseconds, aiming to quantify the stability and consistency of terminal synchronization accuracy and provide basic support for subsequent timing alignment and feature fusion.

During the signal acquisition process, there are differences in the output formats of heterogeneous devices. The data standardization conversion mechanism is used to unify the signal format. To ensure the integrity and real-time performance of the data acquisition process, an asynchronous multi-threaded data acquisition and time synchronization management module is designed. The acquisition delay D_c and processing delay D_p are defined, and the sum of the two does not exceed the system real-time response time threshold. The delay model formula is as follows:

$$D_{total} = D_c + D_p, D_{total} \leq T_{max} \tag{3}$$

where T_{max} is the maximum allowable delay set by the system, which is controlled at the millisecond level to ensure the timeliness of fault identification. The priority scheduling strategy is used to allocate transmission and processing thread resources to achieve high-speed and stable flow of multi-source data.

3.2 Preprocessing and Standardization of Traveling Wave Signals

In the preprocessing and standardization of multi-source traveling wave signals, the key goal is to remove the noise interference in the original signal, unify the data format of different acquisition sources, and provide structured input for subsequent feature extraction and fusion. The original traveling wave signal is affected by factors such as sensor accuracy, sampling noise, and electromagnetic interference in the field environment. The time domain waveform of the signal has interference components such as high-frequency spikes and low-frequency drift [25-26]. To suppress high-frequency random noise, the signal is first subjected to wavelet threshold denoising. The db4 wavelet function is selected, and the maximum number of decomposition layers is set to 5. The high-frequency component is denoised according to the soft threshold function to filter out high-frequency disturbances not related to faults. In this process, the approximate coefficient and detail coefficient of the signal in the j-th layer wavelet domain are recorded as A_j and D_j, respectively, and the detail coefficient is denoised using the following soft threshold function:

$$\widehat{D}_j = \text{sign}\ (D_j) \cdot max(|D_j| - \lambda_1, 0) \tag{4}$$

where $\lambda_1 = \sigma_1\sqrt{2\log n}$ is the denoising threshold; σ_1 is the noise standard deviation; n is the sample length. The denoised signal is reconstructed for subsequent processing. The approximate coefficient A_j contains the main features of the signal and is directly retained without participating in the threshold processing.

Considering the differences in sampling rate, channel structure and data encapsulation methods of different devices, a unified standard data reconstruction process is used to normalize multi-source signals into a unified format. All signals are resampled to 1 MHz and reconstructed using the bilinear interpolation method. The resampled data is stored in a three-dimensional tensor structure. It is assumed that $X_i(t) \in \mathbb{R}^{C \times T}$ represents the signal sample of the i-th measurement point; C is the number of channels; T is the length of the time series.

To improve the efficiency of signal preprocessing and standardization conversion, a parallel data processing process is designed on the cloud computing platform, and the format parsing, field mapping, and resampling operations of the original signal are distributed to multiple computing nodes for parallel execution. After the signal is structured, it is further normalized to eliminate the dimension effect. Using the z-score standardization method, the time series of each channel is transformed to make its mean 0 and standard deviation 1, and the expression is:

$$\widetilde{X}_i(t) = \frac{X_i(t) - \mu_i}{\sigma_i} \tag{5}$$

where μ_i and σ_i are the mean and standard deviation of the signal at the i -th measurement point, respectively. This processing step improves the comparability of each source data in subsequent feature fusion and reduces the deviation caused by scale differences.

Figure 2 shows the waveform comparison of the signal before and after filtering. To quantify the filtering effect, an ideal traveling wave signal model is constructed as a theoretical reference, which is defined as follows:

$$s(t) = \begin{cases} 0.5t, & t < t_0 \\ 1 - \alpha_1(t - t_0), & t_0 \le t < t_1 \\ \beta_1 + \gamma_1 t, & t \ge t_1 \end{cases} \tag{6}$$

where $t = 0.003$ s represents the fault occurrence time, and α_1, β_1, and γ_1 are the set response coefficients, which reflect the slope change and amplitude offset in the transient process and are used to simulate the traveling wave attenuation and reflection effects. The ideal model refers to a theoretically constructed noise-free traveling wave signal model, which is used to simulate the ideal voltage or current waveform when a transient disturbance propagates in a transmission line.

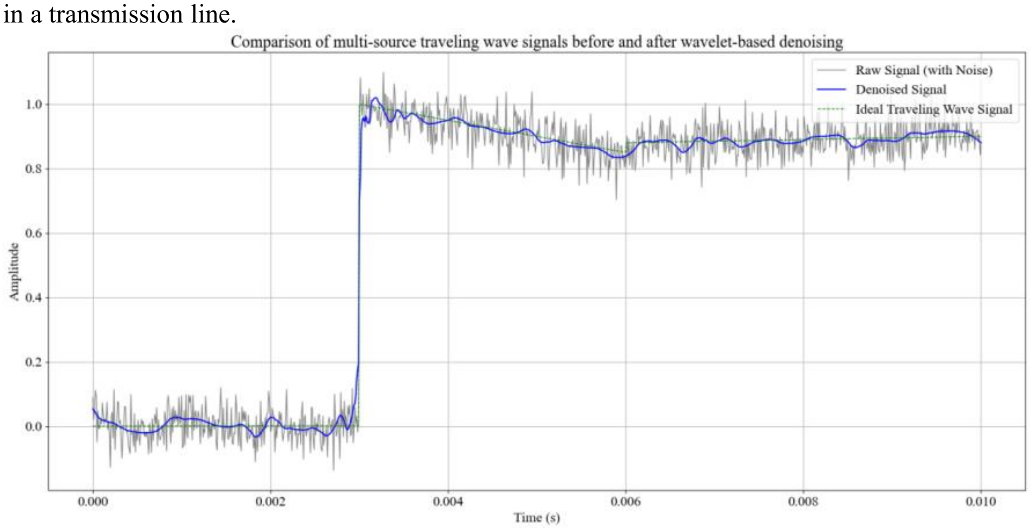

Figure 2. Comparison of waveforms of multi-source traveling wave signals before and after filtering

As shown in Figure 2, the waveform comparison of multi-source traveling wave signals before and after wavelet filtering is presented. The gray curve in the figure is the noisy original signal, which has a large fluctuation amplitude and presents significant high-frequency disturbances, interfering with the identification of transient characteristics of the signal. The blue curve is the signal after wavelet threshold filtering. On the basis of retaining the main mutation structure, high-frequency noise is effectively suppressed; the edge contour remains clear; the position of the mutation point is highly consistent with the ideal model. The green dotted line represents the ideal traveling wave signal, which is used to test the ability of the filtering algorithm to restore the mutation structure. The filtered signal is highly consistent with the ideal model at 0.003 seconds, indicating that the filtering process has achieved a good balance between fidelity and denoising performance. The waveform comparison reflects the basic supporting role of the preprocessing stage in improving feature stability and location accuracy.

In the process of data structure standardization, a unified data frame format field is set, as shown in Table 2, which defines the timestamp, channel number, voltage, current key fields, and their byte lengths in the data frame. This structure ensures the consistency of data transmission and decoding on the cloud platform.

Table 2. Definition table of unified data frame format for multi-source signals

Field Name	Data Type	Byte Length	Description
Timestamp	uint64	8	Unified timestamp
ChannelID	uint8	1	Signal channel identifier
VoltageValue	float32	4	Voltage sampling value
CurrentValue	float32	4	Current sampling value

As shown in Table 2, the listed fields constitute the data frame structure of the unified encapsulation of multi-source traveling wave signals during the preprocessing process, in which each field has a clear functional positioning in the data transmission and analysis process. The Timestamp field is in the form of a 64-bit unsigned integer, which is used to identify the global unified time of each frame of data to ensure the alignment consistency of data across measurement points in the time dimension. The ChannelID field is represented by an 8-bit unsigned integer to identify the number of the data source channel, which helps to distinguish and index the channel dimension during the fusion process. The VoltageValue and CurrentValue fields are represented by 32-bit floating point numbers, respectively, carrying the real-time sampling values of voltage and current, which are the basic inputs for subsequent signal analysis and feature extraction. The length of each field is designed based on data accuracy and storage efficiency. The overall structure supports the standardized encapsulation of high-frequency, high-density, and multi-channel signals, and maintains the consistency and stability of the format in the cloud data decoding and model input interface.

3.3 Multi-scale Time-frequency Feature Extraction

In the multi-scale time-frequency feature extraction process, the wavelet packet decomposition (WPD) method is used to perform multi-resolution analysis on each source signal in view of the significant transient characteristics and complex frequency band distribution of the traveling wave signal [27-28]. In the multi-scale time-frequency feature extraction process, to cope with the computing power requirements brought by the concurrent processing of multiple source signals, this paper relies on the heterogeneous resource pool of the cloud computing platform to allocate the wavelet packet decomposition task to multiple central processing unit (CPU) nodes for parallel execution. The resource scheduling strategy dynamically adjusts the weight parameters according to the task priority and historical consumption to ensure that the feature extraction tasks on the critical path obtain sufficient computing resources.

This method achieves a finer frequency domain division by recursively filtering and downsampling the low-frequency and high-frequency parts of the signal at the same time, thereby effectively capturing the energy mutation and frequency component changes of the signal at the moment of fault occurrence. In the specific implementation process, the db4 wavelet is selected as the mother wavelet function, and the number of decomposition layers is set to 5 layers to balance the time resolution and frequency resolution. Each layer of decomposition generates a set of subband coefficients, forming a time-frequency distribution diagram of a binary tree structure. Among them, the subband signal corresponding to the k-th node after the j-th layer decomposition is represented as $c_{j,k}(n_1)$, and its mathematical expression is:

$$c_{j,k}(n_1) = \sum_m x(m) h_k(2^j m - n_1) \tag{7}$$

where $x(m)$ is the original input signal; $h_k(\cdot)$ is the impulse response of the k-th filter; j represents the number of decomposition layers; n_1 is the discrete time index. By calculating the energy entropy value of each subband coefficient, a feature vector representing the local singularity of the signal is constructed. The definition of energy entropy is as follows:

$$E_k = -\sum_{i=1}^{N} p_i^{(k)} \log p_i^{(k)} \tag{8}$$

where $p_i^{(k)}$ represents the energy proportion of the i-th time period in the k-th subband, and N is the total number of time periods. This indicator can reflect the complexity of the signal in different frequency bands and help distinguish the dynamic differences between the signal in normal operation and fault state.

To further enhance the feature expression capability, after completing the wavelet packet decomposition, the Teager Energy Operator (TEO) is applied to perform nonlinear energy estimation on each subband signal to highlight the transient mutation characteristics of the signal. The calculation formula of TEO is as follows:

$$\Psi(c_{j,k}(n_1)) = c_{j,k}^2(n_1) - c_{j,k}(n_1 - 1) c_{j,k}(n_1 + 1) \tag{9}$$

This operator can quickly detect the transient impact component in the signal without adding additional computational burden and is suitable for the task of extracting the features of traveling wave signals caused by sudden faults such as lightning strikes and short circuits in power systems.

On this basis, to improve the robustness and discriminability of the features, the principal component analysis (PCA) dimensionality reduction processing is performed on the extracted multi-scale time-frequency feature vectors to remove redundant information and retain the most representative low-dimensional feature space. It is assumed that the original feature matrix is: $X \in \mathbb{R}^{d \times N_s}$, where d is the feature dimension and N_s is the number of samples. PCA constructs the projection matrix $W \in \mathbb{R}^{d \times r}$ by solving the maximum eigenvector of the covariance matrix and maps the original data to the low-dimensional space $Y = W^T X$, where $r < d$ is the target dimension.

3.4 Design of Multi-source Feature Fusion Algorithm

In the process of multi-source traveling wave signal feature fusion, a feature layer fusion strategy [29] based on dynamic weight allocation and spatial attention mechanism [30] is designed for the multi-channel feature vector [31] after time alignment. This method captures the topological dependency between measurement points by constructing a feature correlation matrix and uses a nonlinear mapping function to achieve dimensional alignment and information complementarity in the feature space. Specifically, assuming that the wavelet packet decomposition feature vector of the i-th measurement point is $\mathbf{F}_i \in \mathbb{R}^{d_i}$, where d_i is the feature dimension of the i-th measurement point, the features of each measurement point are projected to the common feature space $\mathbb{R}^{d_{\text{com}}}$ through the shared weight matrix $\mathbf{W}_p \in \mathbb{R}^{d_{\text{com}} \times d_i}$, and the mapping relationship is expressed as:

$$\mathbf{F}_i^{\text{com}} = \mathbf{W}_p^T \mathbf{F}_i + \mathbf{b}_p \tag{10}$$

In the formula, $\mathbf{b}_p \in \mathbb{R}^{d_{\text{com}}}$ is the bias term, and d_{com} is the common space dimension. This step eliminates the inconsistency of feature scale caused by the difference in measurement point position, while retaining the discriminative ability of local features.

To further strengthen the cross-measurement point correlation between features, the spatial attention weight vector $\alpha \in \mathbb{R}^N$ is applied in the fusion process, where N is the total number of measurement points involved in the fusion. The attention weight is dynamically generated by calculating the cosine similarity between the features of each measurement point and the global feature mean. The calculation process is as follows:

$$\alpha_j = \frac{\exp\left(\mathbf{F}_j^{\text{com}T} \mathbf{F}_{\text{avg}}\right)}{\sum_{k=1}^{N} \exp\left(\mathbf{F}_k^{\text{com}T} \mathbf{F}_{\text{avg}}\right)} \tag{11}$$

where $\mathbf{F}_{\text{avg}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{F}_i^{\text{com}}$ is the global feature mean vector. This weight reflects the contribution of each measurement point feature to the overall fusion result, so that the system can still focus on the high-confidence feature area when there is local noise interference. The fused global feature vector $\mathbf{F}_{\text{fusion}} \in \mathbb{R}^{d_{\text{com}}}$ is obtained by weighted summation:

In the feature interaction modeling link, the gated recurrent unit (GRU) is used to construct a temporal association network to capture the dynamic evolution of multi-source features in the time dimension. Assuming that the state of the fused feature sequence at time step t is $\mathbf{h}_t \in \mathbb{R}^{d_h}$, its update process is coordinated by the reset gate $\mathbf{r}_t \in \mathbb{R}^{d_h}$ and the update gate $\mathbf{z}_t \in \mathbb{R}^{d_h}$:

$$\begin{aligned} \mathbf{r}_t &= \sigma(\mathbf{W}_r[\mathbf{F}_{\text{fusion}}^{(t)}, \mathbf{h}_{t-1}]) \\ \mathbf{z}_t &= \sigma(\mathbf{W}_z[\mathbf{F}_{\text{fusion}}^{(t)}, \mathbf{h}_{t-1}]) \\ \widetilde{\mathbf{h}}_t &= \tanh\left(\mathbf{W}_h\left[\mathbf{F}_{\text{fusion}}^{(t)}, \mathbf{r}_t \odot \mathbf{h}_{t-1}\right]\right) \\ \mathbf{h}_t &= (1-\mathbf{z}_t) \odot \widetilde{\mathbf{h}}_t + \mathbf{z}_t \odot \mathbf{h}_{t-1} \end{aligned} \tag{12}$$

In the formula, σ is the Sigmoid activation function; $\odot$ represents the Hadamard product; $\mathbf{W}_r$,$\mathbf{W}_z$, and $\mathbf{W}_h$ are trainable parameter matrices. This process realizes the deep coupling of multi-source features in time series and enhances the representation consistency of fault features in continuous time windows.

To quantify the changing characteristics of feature dimensions during the fusion process, a feature contribution evaluation index η_k is designed, which is defined as:

$$\eta_k = \frac{\|\mathbf{W}_k \mathbf{F}_k^{\text{com}}\|_2}{\sum_{i=1}^{N} \|\mathbf{W}_i \mathbf{F}_i^{\text{com}}\|_2} \tag{13}$$

where $\mathbf{W}_k \in \mathbb{R}^{d_{out} \times d_{com}}$ is the output layer weight matrix, and d_{out} is the target output dimension. This index is used to guide the optimization direction of the fusion network parameters to ensure that high-contribution features play a dominant role in subsequent classification tasks.

3.5 Construction of Fault Identification Model Based on Long Short-Term Memory Network

After the fusion of multi-source traveling wave signal features is completed, LSTM is used to build a fault identification model, and fault type discrimination and location estimation are realized through time series feature modeling [32]. The model input is a global feature vector sequence fused by the spatial attention mechanism, with a time resolution of $\Delta \mathrm{t} = 1\mathrm{ms}$, and the feature dimension is reduced by principal component analysis to retain the first $\mathrm{k} = 128$ principal components.

The spatial attention weight output by the feature fusion process is applied into the time series modeling link of the LSTM model [33] as a dynamic adjustment factor of the input feature to enhance the dominant role of high-confidence measurement point features in time series evolution. Specifically, the measurement point contribution vector generated in the fusion stage is multiplied element by element with the input feature of LSTM, so that the model can capture the trend of time dimension changes while retaining the selective enhancement ability of key spatial features. By combining the spatial attention mechanism with the time modeling capability of LSTM, an end-to-end identification framework with spatio-temporal joint optimization characteristics [34] is constructed, which enhances the consistency of feature expression and classification separability of multi-source traveling wave signals in complex fault scenarios.

The LSTM unit has a triple gate structure, including input gate i_t, forget gate f_t, and output gate o_t. The gate parameter matrix W_i, W_f, W_o and bias terms b_i, b_f, b_o are used to control the information flow. Specifically, the input gate calculates the concatenation vector of the feature vector x_t at the current moment and the hidden state h_{t-1} at the previous moment and generates the update weight through the Sigmoid activation function:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{14}$$

The forget gate maintains the long-term dependency through the exponential decay function, and its weight update follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{15}$$

The cell state C_t update process combines the input gate and forget gate outputs and generates the candidate state $\widetilde{C}_t$ through the hyperbolic tangent function:

$$\widetilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{16}$$

where W_c is the candidate state weight matrix, and b_c is the bias term. After the candidate state is weighted by the input gate activation value, it constitutes the cell state update together with the long-term memory controlled by the forget gate:

$$C_t = f_t \circ C_{t-1} + i_t \circ \widetilde{C}_t \tag{17}$$

where the operator $\circ$ represents the Hadamard product. The output gate calculates the hidden layer output based on the updated cell state:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), h_t = o_t \circ \tanh(C_t) \tag{18}$$

During the training process, the cross entropy loss function is used to measure the classification error, and the Adam optimizer is combined for parameter iteration. To enhance the sensitivity of the model to the fault location, a dual-branch structure is applied in the output layer. One branch outputs the probability distribution of the fault type through the Softmax activation function, and the other branch uses the linear regression layer to predict the fault distance. The loss function is defined as:

$$L = -\sum_{i=1}^{N_c} y_i \log(\hat{y}_i) + \lambda \sum_{j=1}^{N_d} (d_j - \hat{d}_j)^2 \tag{19}$$

The first term is the classification loss term: y_i and $\hat{y}_i$ represent the true label and the predicted probability respectively, and N_c is the number of fault type categories; the second term is the position regression loss term: d_j and $\hat{d}_j$ correspond to the actual and predicted fault distances, and $\lambda = 0.3$ is the balance coefficient.

After the feature sequence is processed by the LSTM network, the hidden layer output vector h_t is sent to the fully connected layer for nonlinear mapping. The weight matrix and bias term of the fully connected layer are optimized by the back propagation algorithm, and the output vector is normalized and input into the classification and regression branches. The classification branch weight matrix and the regression branch weight matrix are trained independently. The former generates the probability distribution through the Softmax function:

$$P(y = k|h_t) = \frac{\exp(w_k^T h_t)}{\sum_{j=1}^{N_c} \exp(w_j^T h_t)} \tag{20}$$

The latter directly outputs the fault distance prediction value:

$$\widehat{d} = W_{reg}^T h_t + b_{reg} \tag{21}$$

where $w_k \in \mathbb{R}^{64}$ and $w_j \in \mathbb{R}^{64}$ are the k-th and j-th row transposed vectors of the classification branch weight matrix, respectively, and $b_{reg} \in \mathbb{R}$ is the regression branch bias term.

3.6 Cloud Computing Platform Deployment and System Architecture Design

In the cloud computing platform deployment and system architecture design phase, a distributed computing environment is built based on containerization technology, and the Kubernetes (K8s) cluster management framework is used to realize dynamic resource scheduling [35] [36]. The system architecture is divided into data access layer, computing resource layer, and service application layer. Among them, the data access layer realizes real-time caching and orderly distribution of multi-source traveling wave signals through Kafka message queues; the computing resource layer is composed of GPU acceleration nodes and CPU computing nodes to form a heterogeneous resource pool; the service application layer encapsulates core algorithm modules such as feature extraction, fusion analysis, and fault identification through microservice architecture.

At the system deployment level, the cloud computing platform decouples and encapsulates algorithm modules such as feature extraction, fusion analysis, and model inference through microservice architecture, and uses Kubernetes to realize independent scheduling and version control of each component. The data access layer uses Kafka message queues to complete batch caching and streaming push of multi-source traveling wave signals to ensure data loss and sequential consistency in high-throughput scenarios. The computing resource layer is composed of GPU acceleration nodes and CPU computing nodes to form a heterogeneous resource pool. The former undertakes the matrix operation task of the LSTM model, and the latter is used to perform wavelet packet decomposition and timestamp alignment operations. The resource allocation model dynamically adjusts the weight parameters according to the task priority and historical consumption and gives priority to guaranteeing the computing resource supply of the critical path when multiple source signals arrive concurrently, thereby improving the overall system's response efficiency and operation stability.

To solve the contradiction between large-scale data concurrent processing and low delay response, a priority-based resource allocation model is designed:

$$R_i = \frac{\sum_{j=1}^{N} w_j \cdot Q_{ij}}{\sum_{k=1}^{M} \lambda_k \cdot T_{ik}} \tag{22}$$

In the formula, R_i represents the resource allocation weight of the i-th task; w_j is the task priority coefficient; Q_{ij} corresponds to the demand intensity of the task for the j-th resource; λ_k is the resource constraint factor; T_{ik} reflects the historical consumption time of the i-th task on the k-th resource. The model dynamically adjusts the weight parameter w_j and the constraint factor λ_k to achieve the optimal allocation of computing resources between feature extraction intensive tasks and real-time sensitive tasks.

In the model deployment stage, model sharding and pipeline parallel technology are used to optimize the inference efficiency of the LSTM network. The LSTM weight matrix $\mathbf{W} \in \mathbb{R}^{d\times h}$ is divided into three sub-modules: input gate $\mathbf{W}_i$, forget gate $\mathbf{W}_f$, and output gate $\mathbf{W}_o$ according to the gate structure, where d is the input feature dimension, and h is the number of hidden layer nodes. Each submodule is deployed on different GPU instances through the gRPC remote call interface, and the CUDA stream technology is used to implement asynchronous execution of gated computing. To reduce the cross-node communication overhead, a feature vector compression transmission protocol is designed to compress the transmission accuracy of the intermediate feature $\mathbf{h}_t \in \mathbb{R}^h$ from 32-bit floating point to 16-bit half-precision format, and the differential encoding technology is applied to eliminate redundant information of temporal features.

To improve the system's scalability, a metadata management mechanism based on etcd distributed key-value storage is designed, and the elastic expansion and contraction of computing nodes is realized through heartbeat detection and service registration mechanism. When the cluster load rate L exceeds the threshold θ, the node expansion strategy is automatically triggered:

$$\Delta N = \left\lceil \frac{L-\theta}{\eta} \right\rceil \tag{23}$$

In the formula, ΔN is the number of new nodes, and η represents the single node load bearing coefficient. In the system architecture design, this paper combines the elastic expansion and contraction mechanism unique to cloud computing to realize the on-demand allocation of computing resources. When the cluster load rate exceeds the set threshold, the operation of adding computing nodes is automatically triggered to dynamically adapt to the burst traffic pressure in multi-source traveling wave signal processing.

4. Experiments

4.1 Experimental Environment and Hardware Configuration

The experimental environment of this paper is built on a distributed cloud computing architecture, and the containerization technology is used to achieve dynamic resource scheduling. The overall system framework consists of a data access layer, a computing resource layer, and a service application layer. The data access layer uses the Kafka message queue to achieve real-time caching and orderly distribution of multi-source traveling wave signals, and its throughput is designed to achieve a concurrent processing capacity of 100,000 data packets per second. The computing resource layer consists of a heterogeneous resource pool, including GPU acceleration nodes and CPU computing nodes. Among them, the GPU node uses the NVIDIA A100 accelerator card, equipped with a PCIe 4.0 interface and the third-generation Tensor Core technology, and the computing power of a single card can reach 19.5 TFLOPS (FP16). The CPU node uses the Intel Xeon Platinum 8380 processor, with a main frequency of 2.3GHz, supporting the 512-bit AVX-512 instruction set, 8 memory channels, and a bandwidth of 204.8 GB/s.

The service application layer encapsulates core algorithm modules such as feature extraction, fusion analysis, and fault identification through a microservice architecture. The cluster management uses the Kubernetes (K8s) framework, configures the etcd distributed key-value storage system to implement metadata management, and ensures real-time synchronization of node status through heartbeat detection and service registration mechanisms. The load balancing module deploys the HAProxy component, uses the minimum connection number algorithm to allocate task flows, and supports SSL offloading and HTTP/2 protocol transmission.

Table 3 shows the resource configuration of each functional node of the cloud computing platform. Among them, the feature extraction node is configured with a high-concurrency multi-threaded processing unit, and the model inference node uses the NVIDIA A100 GPU accelerator card to meet the computing power requirements of deep neural networks.

Table 3. Resource configuration table of functional nodes of the cloud computing platform

Node Type	CPU Cores	Memory(GB)	GPU Model	Storage(TB)
Feature Extraction	32	128	None	10
Model Inference	16	64	NVIDIA A100	5
Data Storage	8	32	None	100
Cluster Management	4	16	None	2
Load Balancer	8	32	None	1

As shown in Table 3, the resource configuration parameters of the five types of functional nodes directly correspond to the core functional module requirements of the cloud computing platform. The feature extraction node is configured with a 32-core CPU and 128GB memory to support the high concurrent computing requirements of parallel preprocessing and feature extraction of multi-source traveling wave signals. The design choice of its GPU-free configuration is due to the technical characteristics of this stage that mainly focuses on multi-threaded task scheduling and does not require floating-point operation acceleration. The model inference node uses a combination of a 16-core CPU with 64GB memory and an NVIDIA A100 GPU. The matrix parallel computing capability of the CUDA core accelerates the time series feature processing of LSTM. The memory capacity design meets the storage requirements of deep learning model parameter cache and intermediate feature tensors. The data storage node is configured with an 8-core CPU and 32GB memory, with a 100TB mechanical hard disk array. The RAID5 redundant architecture ensures the persistent storage reliability of the original data and feature library of multi-source traveling wave signals. Its hardware selection focuses on storage density and cost control rather than random read and write performance. The cluster management node implements metadata management and task scheduling with a streamlined configuration of 4-core CPU and 16GB memory. The control plane resource consumption is reduced through lightweight service design to ensure the efficient operation of etcd distributed key-value storage and Kubernetes scheduler. The load balancing node is configured with 8-core CPU and 32GB memory. The software load balancing solution is used to implement the traffic distribution of Kafka message queue and microservice call chain tracking. Its storage capacity design only meets the temporary storage requirements of log records and state snapshots.

4.2 Dataset Construction and Parameter Setting

This paper focuses on the construction of multi-source traveling wave signal dataset and the setting of key parameters. This dataset is generated based on the IEEE 39-node power system simulation platform and covers the characteristics of traveling wave signals under various fault types and operating conditions. To ensure the wide applicability and statistical validity of the experimental results, the design of the dataset fully considers the diverse distribution of fault locations, fault types, and environmental noise. The fault simulation includes typical power system fault forms such as single-phase grounding fault, two-phase short circuit fault, and three-phase short circuit fault. At the same time, by adjusting the fault resistance, fault occurrence time, and line load level, the complexity of the dataset and the representativeness of actual engineering are further enhanced.

To accurately reflect the time-space characteristics of multi-source traveling wave signals, the acquisition terminals are deployed at key monitoring points of the power system, covering different sections of the transmission lines. Each measuring point is equipped with a high-precision synchronization module, and the signal is marked with the UTC time standard to ensure the consistency of multi-source signals in the time domain. The signal sampling rate is uniformly set to 1 MHz to ensure sufficient time resolution for capturing the transient characteristics of traveling wave signals. Each record in the dataset contains complete voltage and current waveform information and is accompanied by a precise timestamp, fault type identifier, and fault distance parameter. These fields constitute the core input for subsequent feature extraction and model training.

Table 4 lists the main parameter configurations of the multi-source traveling wave signal dataset, including key indicators such as signal sampling rate, total number of dataset samples, number of fault types, fault distance range, and fault resistance value range. The basic scale and complexity of the dataset are defined, reflecting its adaptability and coverage in simulating real power grid fault scenarios.

Table 4 describes in detail the specific values or value ranges of each parameter in the dataset, covering the entire process from signal sampling to data format standardization. The signal sampling rate is 1 MHz to ensure the integrity of high-frequency transient signals. The total number of samples in the dataset is 50,000, covering a variety of scenarios with different fault types. There are three types of faults, corresponding to single-phase grounding faults, two-phase short circuit faults, and three-phase short circuit faults. The fault distance ranges from 0.5 km to 150 km, covering the typical conditions of short-distance and long-distance transmission lines. The fault resistance ranges from 0 Ω to 500 Ω, simulating the transition state of low-impedance and high-impedance faults. There are five monitoring points, distributed at multiple key nodes in the power system. The data format standardization part clarifies the data type and byte length of voltage, current, timestamp, and channel number, providing structured support for subsequent signal processing.

In the process of dataset construction, a multi-level noise interference mechanism is applied to simulate the electromagnetic interference and sensor measurement errors that may occur in the field environment. The noise injection adopts a combination of white noise and impulse noise, which is controlled between 10 dB and 30 dB to ensure that the dataset has a certain anti-interference ability while retaining the fault characteristics. This design not only improves the practicality of the dataset but also provides a basis for evaluating the robustness of the signal preprocessing algorithm.

Table 4. Multi-source traveling wave signal dataset parameter configuration table

Parameter Description	Value OR Range
Signal Sampling Rate	1 MHz
Total Number of Samples	50,000
Number of Fault Types	3 (Single-phase Grounding, Two-phase Short Circuit, Three-phase Short Circuit)
Fault Distance Range	0.5 km - 150 km
Fault Resistance Range	0 Ω - 500 Ω
Number of Monitoring Points	5
Data Format Standardization	Voltage: float32, Current: float32, Timestamp: uint64, ChannelID: uint8

4.3 Experimental Process and Performance Index Definition

This paper focuses on the experimental process and performance index definition and comprehensively verifies the multi-source traveling wave signal fusion and fault identification method based on the cloud computing platform through a systematic method.

The experimental process includes data acquisition, preprocessing, feature extraction, multi-source feature fusion, and fault identification model training and testing. In the data acquisition stage, all traveling wave signals are marked according to a unified time standard, and the time consistency of the data at each measuring point is ensured by a high-precision clock synchronization mechanism. Subsequently, the original signal is filtered, denoised, and standardized to eliminate noise interference and unify the data format. On this basis, the wavelet packet decomposition method is used to perform multi-scale time-frequency analysis on the signal to extract features such as transient energy distribution and frequency band energy ratio that can reflect the fault characteristics. After these features are integrated by the feature layer fusion strategy, they are sent to LSTM as input data for fault type discrimination and location estimation tasks.

To comprehensively evaluate the effectiveness of the method, multiple performance indicators are defined, covering aspects such as fault identification accuracy, fault location error, system response time, and multi-source data fusion effect. Among them, the fault identification accuracy is used to measure the model's ability to classify different fault types; the fault location error reflects the system's estimation accuracy of the fault location; the system response time counts the total time from data collection to the final output of the fault identification result, which is an important basis for evaluating the real-time performance of the system; the multi-source data fusion effect is reflected by comparing the changes in the discrimination of the feature vectors before and after fusion, which is used to verify the contribution of the feature fusion algorithm to improving the overall identification performance.

During the experiment, attention should also be paid to the resource utilization of the cloud computing platform and the scalability of the system to ensure that the method has good deployment capabilities and operating efficiency in actual engineering applications. To this end, parameters such as resource allocation weight, task priority coefficient, and load rate threshold are set to dynamically adjust the allocation ratio of computing resources between different tasks. The parameter configuration of the dataset is detailed in Table 4. Its sampling rate, fault type, and distance range design ensure the wide applicability of the experiment.

The parameter definition covers the entire process from signal acquisition to data format normalization, reflecting the adaptability and extensiveness of the dataset in simulating real power grid fault scenarios. By setting a variety of fault types, distances, and resistance parameters, the experimental results are ensured to be highly representative and practical, providing solid data support for subsequent model training and performance evaluation.

5. Results

5.1 Fault Identification Accuracy

In the task of fault identification in power systems, this paper constructs a multi-source signal fusion and fault identification method based on a cloud computing platform to address the problems of asynchronous multi-source traveling wave signal acquisition, data heterogeneity, and processing delay. To verify the identification performance of this method under different fault types, the experiment selects five representative models for comparative analysis, including traditional methods (wavelet transform combined with support vector machine (wavelet+SVM)), single LSTM, multi-source feature splicing multi-layer perceptron (MLP), fusion model without attention mechanism, and the spatio-temporal long short-term memory network (ST-LSTM) of this paper. These categories represent the development path from early shallow classifiers to modern deep neural networks, and their structural complexity and feature extraction capabilities are gradually improved. The experiment uses the identification accuracy of three typical fault types as evaluation indicators, covering single-phase grounding, two-phase short circuit, and three-phase short circuit, aiming to fully reflect the model's generalization ability and robustness under different fault modes. The standard deviation is used to measure the model's stability in multiple experiments; the error bar reflects the confidence interval of the result; the red dotted line marks the overall accuracy of the method in this paper, which is 98.6%.

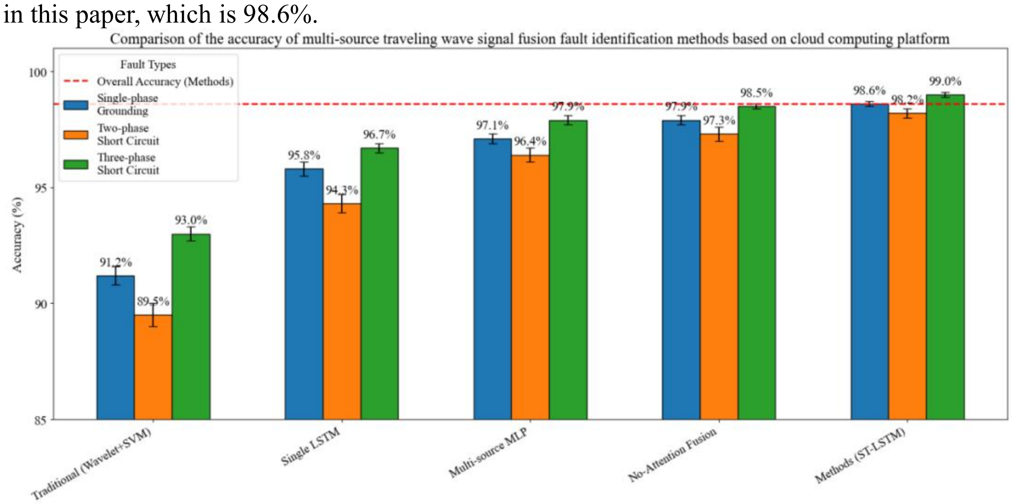

Figure 3. Comparison of the accuracy of multi-source traveling wave signal fusion fault identification methods based on cloud computing platform

Figure 3 shows the difference in identification accuracy of five types of models on three types of faults. Traditional methods are limited by manual feature extraction and the design of shallow classifiers, and the accuracy of various types of faults is generally lower than 93%. Although a single LSTM has the ability to model time series, its overall performance still lags behind the fusion method because it only relies on single measurement point information. After applying multi-source signal splicing, the MLP model improves the accuracy to more than 96%, but fails to achieve dynamic weight allocation, resulting in limited identification of some fault types. The fusion model without attention mechanism further improves the identification performance, but it is still insufficient in capturing detailed features. In contrast, the ST-LSTM method proposed in this paper applies the spatio-temporal attention mechanism to simultaneously optimize the feature expression in the time dimension and the space dimension, so that the identification accuracy of the three types of faults exceeds 98%, among which the three-phase short circuit fault reaches 99.0%. This result is due to the effective collaborative use of multi-source signals and the support of cloud-based distributed computing resources for model training efficiency. Experimental data shows that with the gradual optimization of the model structure, the identification performance shows a progressive improvement trend, which confirms the significant enhancement of the fusion strategy and deep learning architecture improvement on the power system fault diagnosis effect.

5.2 Fault Location Error Evaluation

In the power system fault identification task, a multi-source signal fusion and fault identification method is constructed based on the cloud computing platform to address the location error problem caused by asynchronous acquisition and data heterogeneity of multi-source traveling wave signals. The experiment uses fault location error (unit: meter) as the core evaluation indicator to reflect the accuracy of the model's estimation of the fault location; the standard deviation is used to quantify the degree of discreteness of multiple experimental results, and the error band (± standard deviation) further reflects the data fluctuation range. Figure 4 shows the changing trend of the location error of the five models in the range of 0.5 km to 150 km, and the semi-transparent shaded area indicates the error fluctuation range.

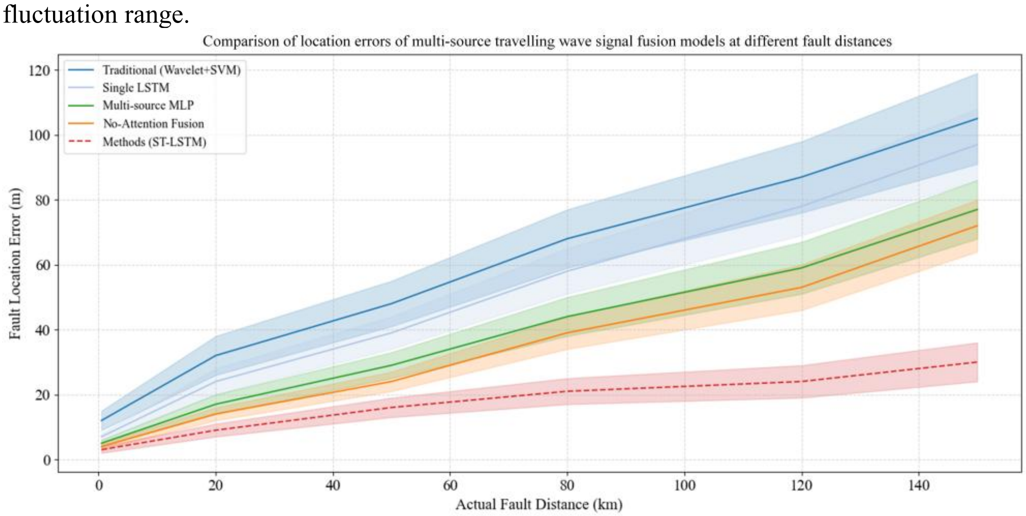

Figure 4. Comparison of location errors of multi-source traveling wave signal fusion models at different fault distances

As shown in Figure 4, the traditional method (Wavelet+SVM) has an error of 105 m at 150 km; the single LSTM has an error of 97 m; the ST-LSTM has a stable error of 30 m, and the error band width is significantly lower than other models. This result is due to the fact that ST-LSTM dynamically allocates multi-source signal weights through the spatio-temporal attention mechanism, suppressing the influence of noise interference on feature expression in long-distance scenarios, and the distributed computing resources provided by the cloud computing platform guarantee the real-time processing capability of complex models. In contrast, the traditional method is limited by the limitations of manual feature extraction; the single LSTM fails to fully integrate the information of multiple measurement points; the multi-source MLP and non-attention fusion models lack dynamic weight adjustment, resulting in local feature redundancy. As the fault distance increases, the errors of each model show an increasing trend, but the increase of ST-LSTM is the smallest, and its error band is always maintained within ±6 m, which verifies the stability of the location accuracy and engineering applicability of this method in long-distance transmission lines. This result is due to the effective weighted fusion of multi-point features by the spatio-temporal attention mechanism, and the support of cloud-based distributed computing resources for model training efficiency. It can be concluded that the coordinated optimization of the fusion strategy and the deep learning architecture has a significant effect on improving the accuracy of power system fault location.

5.3 System Response Time Measurement

In the research on power system fault monitoring and identification, the experiment focuses on the real-time bottleneck problem caused by asynchronous acquisition, data heterogeneity, and processing delay of multi-source traveling wave signals. The response characteristics of three typical faults (single-phase grounding, two-phase short circuit, and three-phase short circuit) are compared and analyzed around different model architectures. The experiment quantitatively evaluates the performance differences of various methods under the cloud computing platform through three core indicators: response time (total time from data acquisition to fault output), standard deviation (reflecting model stability), and real-time threshold (1.3 seconds). The horizontal bar chart in Figure 5 distinguishes the three types of faults with different colors; the error bar reflects the data fluctuation range; the red dotted line marks the real-time threshold.

As shown in Figure 5, the response time of traditional methods generally exceeds 1.9 seconds; the single LSTM is shortened to 1.71-1.83 seconds; the multi-source MLP is further reduced to 1.51-1.68 seconds; the non-attention fusion model achieves a breakthrough in the range of 1.33-1.45 seconds; the ST-LSTM achieves real-time requirements with a stable performance of 1.28-1.32 seconds. The difference in response time is due to the hierarchical evolution of the model architecture and feature fusion mechanism: the traditional method relies on manual feature extraction, resulting in redundant calculations; the single LSTM is limited to the single-point information input dimension; the multi-source MLP enhances information integrity through feature splicing but lacks dynamic optimization capabilities; the non-attention fusion model initially applies a weight distribution strategy; the ST-LSTM realizes the coordinated optimization of multi-source features in the time and space dimensions through the spatio-temporal attention mechanism, and significantly reduces the redundant calculation overhead in combination with the distributed resource scheduling of the cloud computing platform. The standard deviation analysis shows that as the model complexity increases, the response stability shows a gradient enhancement trend, and the standard deviation of ST-LSTM is controlled within 0.05-0.06 seconds, reflecting its robustness in high-concurrency scenarios. This result verifies that the multi-source signal fusion framework based on the cloud computing platform improves the engineering applicability of the fault identification system while ensuring real-time requirements through algorithm architecture innovation and dynamic resource allocation.

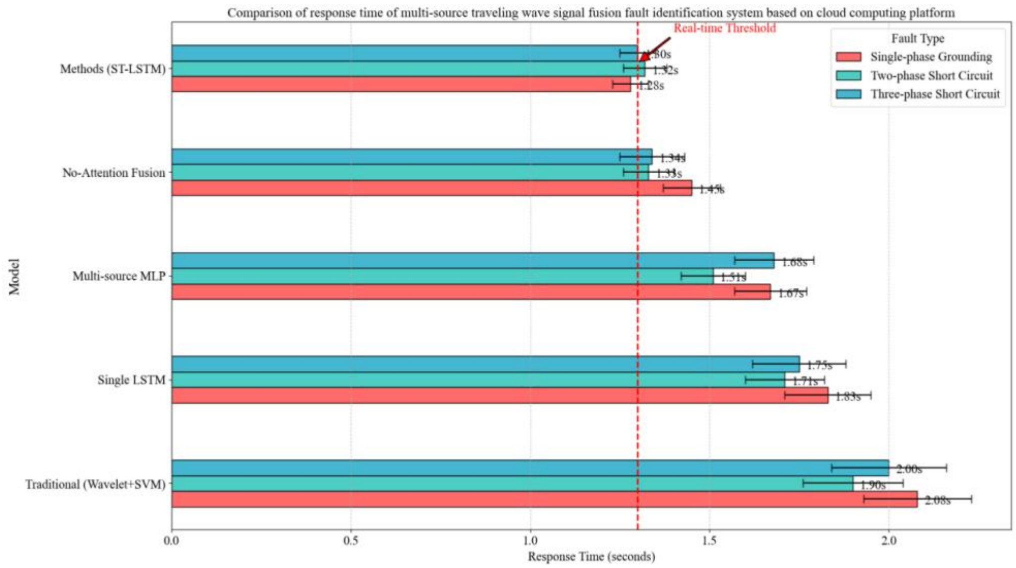

Figure 5. Comparison of response time of multi-source traveling wave signal fusion fault identification system based on cloud computing platform

5.4 Verification of Multi-source Data Fusion Effect

The experiment uses feature discrimination as the core evaluation indicator to quantify the feature separation ability of fault categories under different models. The higher the value, the clearer the feature expression. In Figure 6, the X-axis shows the names of five types of models, and the Y-axis represents the average Euclidean distance.

As shown in Figure 6, the feature discrimination of traditional methods under three types of faults is less than 0.75, which is limited by the limitations of manual feature extraction and the discrimination ability of shallow classifiers. Single LSTM improves the discrimination to above 0.71 through time series modeling, but the gain is limited due to the lack of sufficient fusion of multi-measurement point information. After applying multi-source feature concatenation, the discrimination of Multi-Source MLP reaches above 0.80, but the fixed weight allocation limits the improvement of some fault types (such as three-phase short circuit). The fusion model without attention mechanism further increases the discrimination to above 0.81 by weighted integration of multi-source features, but there is still redundant feature interference. In contrast, ST-LSTM dynamically optimizes the measurement point weights through the spatio-temporal attention mechanism, and the discrimination of the three types of faults reaches 1.02–1.15, which is an average improvement of more than 60% compared with traditional methods. This result stems from the ability of ST-LSTM to suppress noise interference and strengthen high-credibility feature areas in the spatial dimension, as well as the advantage of capturing dynamic evolution laws in the temporal dimension. At the same time, the distributed resource scheduling of the cloud computing platform ensures efficient training and inference of complex models. The data shows that the coordinated optimization of the multi-source signal fusion strategy and the deep learning architecture significantly improves the feature expression ability, providing key support for the robustness and real-time performance of power system fault identification.

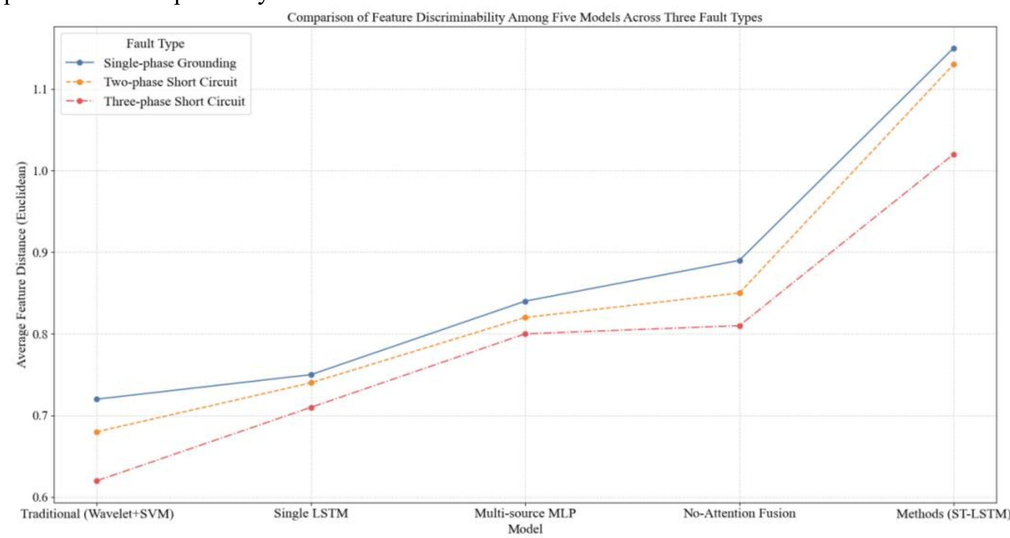

Figure 6. Comparison of feature discrimination of five models under three types of faults

5.5 Resource Utilization of Cloud Computing Platform

In the study of multi-source traveling wave signal fusion and fault identification in power systems, resource utilization aims to reveal the dynamic correlation characteristics between heterogeneous resource scheduling strategies and load changes. The experiment collects resource utilization data of CPU, memory, and model inference GPU by time period. The resource utilization index quantifies the computing power requirements of each component in data preprocessing, feature fusion, and deep learning inference by calculating the core occupancy rate of the calculation unit, memory bandwidth consumption, and GPU video memory peak ratio.

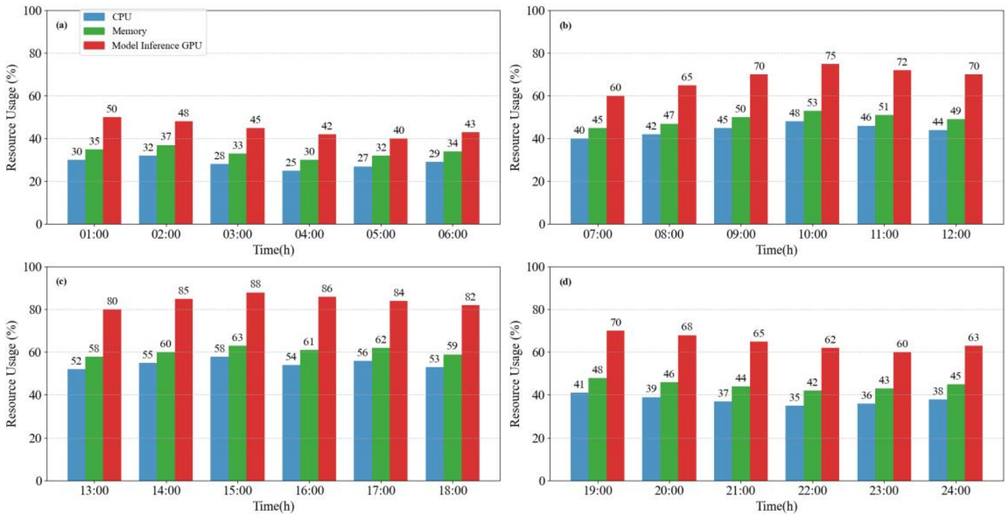

Figure 7. Comparative analysis of resource utilization of multi-source traveling wave signal processing on cloud computing platform by time period. (Figure 7(a). 01:00–06:00 (Low-Load Period), Figure 7(b). 07:00–12:00 (Working Period), Figure 7(c). 13:00–18:00 (Peak Load Period), Figure 7(d). 19:00–24:00 (Post-Peak Period).)

As shown in Figure 7, during the peak load period, the GPU utilization rate of model inference climbs to 80-88%, and the low load period increases by more than 30%. This increase is directly related to the gradient adjustment mechanism of the priority coefficient in the dynamic resource allocation model: the highly complex spatio-temporal attention calculation triggers the elastic expansion of GPU resources when the load rate exceeds the threshold, and the CPU and memory utilization rates of the feature extraction module simultaneously climbed to 52-58% and 58-63%, reflecting its continuous computing power demand as a preprocessing core. The GPU and CPU utilization fluctuate in a positive correlation during the working cycle, indicating that the feedback optimization of the historical consumption time parameters effectively balances the resource competition between feature extraction and deep learning inference. During the fallback period, the GPU utilization rate drops to 60-70%, and the CPU utilization rate drops to 35-41%, revealing the priority guarantee characteristics of the cloud resource recovery strategy for deep learning tasks. The load rate-driven node expansion strategy effectively balances the computing load requirements of the deep model through dynamic resource allocation, avoiding performance bottlenecks in heterogeneous resource scheduling. The study verifies the effectiveness of dynamic scheduling of containerized heterogeneous resource pools through Kubernetes cluster management, and combined with the feature weight allocation of the spatio-temporal attention mechanism, significantly improves the resource adaptation efficiency and system expansion capabilities of multi-source signal processing.

6. Conclusions

This paper proposed a cloud-computing-based method for multi-source traveling wave signal fusion and fault identification in power systems. The proposed framework uses global clock synchronization for signal time alignment, wavelet packet decomposition for multi-scale time-frequency feature extraction, and an LSTM model enhanced by a spatio-temporal attention mechanism for fault type identification and location estimation. At the system level, cloud-native technologies such as containerized deployment, automatic scaling, heterogeneous resource scheduling, and microservice-based module decoupling are introduced to improve the flexibility, robustness, and scalability of multi-source signal processing. These mechanisms enable dynamic resource allocation and efficient concurrent processing, thereby enhancing system response capability and maintainability. Experimental results on the IEEE 39-bus power system simulation platform show that the proposed method achieves a fault identification accuracy of 98.6% and an average response time of 1.3 s, meeting real-time fault identification requirements. The dynamic-weight-based feature fusion strategy improves the discrimination of three typical fault types to 1.02–1.15, which is over 60% higher than that of traditional methods. However, the peak GPU utilization reaches 88%, indicating a potential energy consumption bottleneck for computation-intensive deep learning tasks. Future work will focus on lightweight feature extraction, cloud-edge collaborative computing, federated learning, and 5G-based low-latency transmission to further improve the efficiency and scalability of smart grid fault monitoring in IoT environments.

Funding

This work was supported by the Science and Technology Project of China Southern Power Grid Company Limited under (GDKJXM20231090).

References

[1] Koohang A, Sargent C S, Nord J H, et al. Internet of Things (IoT): From awareness to continued use. International Journal of Information Management, 2022, (62), pp. 102442.
[2] Yu F, Liu J, Yu H, et al. Multimodal Wearable System With Dual-Frequency Enhancement Network for Risk Recognition. IEEE Internet of Things Journal, 2025, 12(11), pp.17634-17648.
[3] Cui H, Yang Q, Sun J, et al. Electromagnetic time reversal fault location method based on characteristic frequency of traveling waves. IEEE Transactions on Power Delivery, 2023, 38(5), pp. 3033-3044.
[4] Deng F, Zhong Y, Zeng Z, et al. A single-ended fault location method for transmission line based on full waveform features extractions of traveling waves. IEEE Transactions on Power Delivery, 2023, 38(4), pp.2585-2595.
[5] Yu F, Rao Z, Chen N, et al. ArmBCIsys: Robot Arm BCI System With Time–Frequency Network for Multiobject Grasping. IEEE Transactions on Neural Networks and Learning Systems, 2025, 36 (10), pp.18327-18341..

[6] Feng Y, Xiao Z, Liu L, et al. BiaCanDet: Bioelectrical impedance analysis for breast cancer detection with space–time attention neural network. Expert Systems with Applications, 2025, (269), pp. 126223.

[7] Zhao H, Gong Z, Yan J, et al. Unsynchronized underwater localization with isogradient sound speed profile and anchor location uncertainties. IEEE Transactions on Vehicular Technology, 2024, 73(6): 8864-8877.

[8] Shu H, Deng Y, Cao P, et al. Transient pilot protection based on directional traveling wave principle for wind power transmission line. International Journal of Electrical Power & Energy Systems, 2024, 157, pp. 109835-109846.

[9] Wu J, Li Q, Chen Q, et al. Evaluation, analysis and diagnosis for HVDC transmission system faults via knowledge graph under new energy systems construction: A critical review. Energies, 2022, 15(21), pp. 8031-8051.

[10] Liu R, Fu R, Xu K, et al. A review of knowledge graph-based reasoning technology in the operation of power systems. Applied Sciences, 2023, 13(7), pp. 4357-4381.

[11] Hassani H, Razavi-Far R, Saif M, et al. Intelligent decision support and fusion models for fault detection and location in power grids. IEEE Transactions on Emerging Topics in Computational Intelligence, 2021, 6(3), pp. 530-543.

[12] Ni Q, Zhan Z, Li X, et al. A real-time fault diagnosis method for grid-side overcurrent in train traction system using signal time series feature pattern recognition. IEEE Transactions on Industrial Electronics, 2023, 71(4), pp. 4210-4218.

[13] Sun R, Ren Y. A multi-source heterogeneous data fusion method for intelligent systems in the Internet of Things. Intelligent Systems with Applications, 2024, 23, pp. 200424-200433.

[14] Sahu A, Mao Z, Wlazlo P, et al. Multi-source multi-domain data fusion for cyberattack detection in power systems. IEEE Access, 2021, 9, pp. 119118-119138.

[15] De La Cruz J, Gómez-Luna E, Ali M, et al. Fault location for distribution smart grids: Literature overview, challenges, solutions, and future trends. Energies, 2023, 16(5), pp. 2280-2317.

[16] F. Yu, Y. Zhang, H. Li, C, et al. Phase Contour Enhancement Network for Clothing Parsing. IEEE Transactions on Consumer Electronics, 2024, 70(1), pp. 2784-2793.

[17] LI H, ZENG H, YU K, et al. Fault location method for complex distribution networks based on multi- terminal traveling wave frequency matrix. Journal of Electric Power Science and Technology, 2024, 39(3), pp. 19-30, 37.

[18] Zhang H, Li L, Pan H, et al. Time-Frequency Aliased Signal Identification Based on Multimodal Feature Fusion. Sensors, 2024, 24(8), pp. 2558-2581.

[19] Sun H, Qin B, Wei Z, et al. A review of fault diagnosis for train signal system based on multi-source information fusion. Engineering Research Express, 2025, 7(2), pp. 022503-022528.

[20] Al-Jumaili A H A, Muniyandi R C, Hasan M K, et al. Big data analytics using cloud computing based frameworks for power management systems: Status, constraints, and future recommendations. Sensors, 2023, 23(6), pp. 2952-2989.

[21] Maurya M, Panigrahi I, Dash D, et al. Intelligent fault diagnostic system for rotating machinery based on IoT with cloud computing and artificial intelligence techniques: a review. Soft Computing, 2024, 28(1), pp. 477-494.

[22] Lu S, Lu J, An K, et al. Edge computing on IoT for machine signal processing and fault diagnosis: A review. IEEE Internet of Things Journal, 2023, 10(13), pp. 11093-11116.
[23] Li B, Tang L, Gu Z, et al. Research on Distribution Network Fault Location Based on Electric Field Coupling Voltage Sensing and Multi-Source Information Fusion. Energies, 2025, 18(4), pp. 913-939.
[24] Yuan Q, Pi Y, Kou L, et al. Multi-source data processing and fusion method for power distribution internet of things based on edge intelligence. Frontiers in Energy Research, 2022, 10, pp. 891867-891876.
[25] Li T, Sun C, Fink O, et al. Filter-informed spectral graph wavelet networks for multiscale feature extraction and intelligent fault diagnosis. IEEE Transactions on Cybernetics, 2023, 54(1), pp. 506-518.
[26] Ahmed N, Hashmani A A, Khokhar S, et al. Fault detection through discrete wavelet transform in overhead power transmission lines. Energy Science & Engineering, 2023, 11(11), pp. 4181-4197.
[27] Shao H, Xia M, Wan J, et al. Modified stacked autoencoder using adaptive Morlet wavelet for intelligent fault diagnosis of rotating machinery. IEEE/ASME Transactions on Mechatronics, 2021, 27(1), pp. 24-33.
[28] Huang D, Zhang W A, Guo F, et al. Wavelet packet decomposition-based multiscale CNN for fault diagnosis of wind turbine gearbox. IEEE Transactions on Cybernetics, 2021, 53(1), pp. 443-453.
[29] Yu F, Cao J, Liu L, et al. SuperLightNet: Lightweight Parameter Aggregation Network for Multimodal Brain Tumor Segmentation. Proceedings of the Computer Vision and Pattern Recognition Conference. 2025, pp. 5197-5206.
[30] Liao C, Du C, Liu Y, et al. SterAF: A Scene Text Recognizer with Appearance-Flow Rectification. International Journal of Pattern Recognition and Artificial Intelligence, 2024, 38(06): 2450011.
[31] Yu F, Yu C, Tian Z, et al. Intelligent wearable system with motion and emotion recognition based on digital twin technology. IEEE internet of things journal, 2024, 11(15), pp. 26314-26328.
[32] Al-Selwi S M, Hassan M F, Abdulkadir S J, et al. RNN-LSTM: From applications to modeling techniques and beyond—Systematic review. Journal of King Saud University-Computer and Information Sciences, 2024, 36(5), pp. 102068.
[33] Yu F, Chen Z, Jiang M, et al. Smart clothing system with multiple sensors based on digital twin technology. IEEE Internet of Things Journal, 2022, 10(7), pp. 6377-6387.
[34] Cheng W, Zhou F, Liao C, et al. STHM: A skinning transformer approach for predicting clothing deformation from human motion. Journal of Computational Methods in Sciences and Engineering, 2025, pp. 14727978251385168.
[35] Soni D, Kumar N. Machine learning techniques in emerging cloud computing integrated paradigms: A survey and taxonomy. Journal of Network and Computer Applications, 2022, 205, pp. 103419.
[36] Parast F K, Sindhav C, Nikam S, et al. Cloud computing security: A survey of service-based models. Computers & Security, 2022, 114: 102580.