energies

MDPI

Article

Transient Fault Signal Identification of AT Traction Network Based on Improved HHT and LSTM Neural Network Algorithm

Huan Zhou, Jianyun Chen *, Manyuan Ye, Qincui Fu and Song Li

State Key Laboratory of Performance and Guarantee of Rail Transportation Infrastructure, East China Jiaotong University, Nanchang 330013, China

* Correspondence: tiger984@163.com; Tel.: +86-0791-83973506

Abstract: This paper aims to address the difficult to pinpoint fault cause of the full parallel AT traction power supply system with special structure. The fault characteristics are easily covered up, and high transition impedance only affects the singularity of the wavehead, making the traveling waves hard to identify. Moreover, the classification accuracy of the traditional time-frequency analysis method is not sufficiently high to distinguish precisely. In this paper, a fault classification method of traction network based on single-channel improved Hilbert–Huang transform and deep learning is proposed. This method extracts effective fault features directly from the original fault signals and classifies the fault types at the same time. The accuracy of data categorization is increased by directly applying the Hilbert–Huang transform to fault signals to extract transient fault features and produce one-dimensional feature data, which are analyzed by the time-frequency energy spectrum. Using the similarity recognition method of long-short-term memory neural network, the extracted high-frequency one-dimensional feature data are trained and tested to classify fault signals more accurately. In order to verify the effectiveness of this method, several kinds of short-circuit and lightning strike faults are continuously simulated and verified in this paper. Considering various fault conditions and factors, the proposed improved HHT+LSTM method is compared with the LSTM method for direct processing of the original signals. The improved HHT + LSTM classification algorithm achieves an accuracy of 99.99%.

Keywords: traction power supply system; transient fault signal; Hilbert–Huang transform; long-short-term memory

Citation: Zhou, H.; Chen, J.; Ye, M.; Fu, Q.; Li, S. Transient Fault Signal Identification of AT Traction Network Based on Improved HHT and LSTM Neural Network Algorithm. *Energies* **2023**, *16*, 1163. https://doi.org/10.3390/en16031163

Academic Editors: José Matas, Jose Luis Calvo-Rolle and Silvio Simani

Received: 27 October 2022
Revised: 25 November 2022
Accepted: 29 December 2022
Published: 20 January 2023

1. Introduction

Faults of traction network mainly include short circuits and lightning strikes, which can disrupt train operations and even result in more serious accidents. Although the current traction power supply protection monitoring apparatus can quickly turn off and isolate the power supply fault section, it lacks the ability to identify and categorize faults. How to decrease maintenance time and restore the power supply in a timely manner by promptly and precisely determining the causes of faults and identifying the sorts of problems represent the trickiest issues concerning the running and upkeep of the traction power supply system. At present, China's high-speed railway primarily adopts the AT autotransformer power supply special mode, and the uplink and downlink power supply lines operate in parallel. The network structure is complex, leading to the monitored fault electromagnetic transient signals containing multiple refraction and reflection aliased signals. Using conventional signal identification techniques, it is challenging to achieve satisfactory results [1]. Given the particularity of the traction power supply system and the difficulty of signal recognition, it is necessary to develop an adaptive and high-accuracy recognition algorithm.

A full parallel AT traction network features complex operation structure and asymmetric power supply mode. Low leakage reactance AT autotransformers are connected

in parallel lines every 10~15 km. Up and down traction network lines are connected in parallel. Traveling waves not only propagate on fault lines, but also through parallel structures. Simultaneously, there is electromagnetic coupling between phases and lines. They influence each other through coupling. Traveling waves propagate with dispersion, making frequency-varying rail parameters obvious. In addition, high transition impedance only affects the singularity of the wavehead. As a result, the above traveling wave's propagation in the traction network is distinct and intricate. The method of the power system cannot be directly transplanted and used. Instead, a fault identification method better suited should be tailored according to the structural characteristics of the high-speed railway traction network. The challenge of the research is how to maximize the accuracy and speed of identification while taking full advantage of signal propagation characteristics.

Transient fault classification methods can be divided into three categories based on physical measurement, signal processing, and artificial intelligence. The physical measurement method is mainly based on the phase angle difference [2], impedance measurement [3], waveform peak value [4], and other parameters. In [4], based on kurtosis distribution characteristics, the single-phase-to-ground fault can be identified. However, the change in phase angle brought on by a lightning fault results in low detection accuracy and subpar generalization performance. Signal processing techniques which can be roughly divided into S-Transform [5], mathematical morphology [6,7], wavelet transform [8–12], and Hilbert–Huang transform (HHT) [13,14], are more widely used than physical measurement techniques. Wavelet transform and S transform have good time-frequency localization functions, but the accuracy of wavelet transform depends on the decomposition level and the selected mother wave, and there is wavelet aliasing. In recent years, Wavelet transforms have been treated by relevant researchers. In [15], wavelet entropy is used to reduce the impact of wavelet aliasing on the precision of feature extraction accuracy. Due to the randomness and irregularity of the fault signals, the unique entropy result cannot be guaranteed. The disadvantage of the S-transform is the large computation and inflexibility of the Gaussian window function. HHT transform, on the other hand, is an adaptive time-frequency analysis method, which can adaptively decompose the signal based on the characteristics of the signal and extract features from the transient signal.

Given the potential feature extraction and classification of data, machine learning provides improved performance, primarily including neural network [16–19] methods and support vector machines (SVM) [20,21] as fault classifiers. Deep learning has been investigated in recent years in transient fault classification [22–25]. This shows that deep learning classification accuracy is superior to various shallow learning approaches [26]. The combination of deep learning, artificial neural network, and SVM combined with multiple algorithms results in full play for swarm intelligence. For example, wavelet transform and neural network [27], wavelet transform and SVM [28], wavelet transform and fuzzy control [29], HHT and convolutional neural network [30], HHT and SVM. Major drawbacks of all of all these methods are related to the following factors: computational burden for real-time implementation, and complex generalization. In [20,21], the combination of HHT and SVM presented was applied to fault classification of the power distribution system. In [31], in the case of a single-phase grounding fault, the energy difference is small, which is easily translated into a misjudgment. In [30], HHT transform is used to construct a time-frequency energy spectrum, and CNN is used to classify image data, which causes a huge computational burden. Since the time-frequency energy images of the traction network are not well distinguished, the accuracy cannot be guaranteed. Long-short term memory (LSTM) neural network is frequently used for processing time-series data. As a special structure of deep learning neural network, it can analyze signal characteristics in the spatio-temporal domain to perform fault classification.

This paper proposes a full parallel transient fault identification method combining HHT-LSTM for rapid anomaly detection, which integrates time-frequency domain learning into a neural network architecture to solve the problem of fault classification of the full parallel traction network: 1. learn the traveling wave transient fault characteristics

effectively of all-parallel traction network through HHT analysis. (1) Impedance only affects the singularity of the wavehead. The improved HHT in this paper can extract the high-frequency characteristic one-dimensional signal of the fault signal, and it is unaffected by wavehead singularity. (2) Increase the classification accuracy. 2. Through the LSTM similarity recognition method, fault information is captured and classified adaptively into seven types of short-circuit and lightning strike faults. 3. It overcomes the problem that the traditional time-frequency of energy distribution is not obvious, and that the faults are hard to distinguish from one another. Finally, the effectiveness of the proposed fault classification method is verified under a variety of fault conditions and factors.

2. Full-Parallel AT Traction Power Supply System

To address the issue that faults are difficult to identify due to the parallel structure, autotransformer, and electromagnetic coupling of full parallel AT traction network (FPATTN), resulting in the continuous superposition of traveling waves, it is necessary to analyze the unique network structure of the full-parallel AT traction network and the propagation characteristics of multi-conductor traveling waves. The full-parallel AT traction network has unique complex structural features and power supply mode. A parallel AT autotransformer located in the middle of a multi-lead structure can be used to solve the backflow and electromagnetic interference problem. Short-circuit and lightning traveling wave refraction and reflection laws on the traction array are complex. Thus, in order to reveal the mechanism of influence of transient features of the complex multi-lead structure of the traction network, it is first necessary to establish a network model.

2.1. Traction Network Structure and Modelling

The conductor structure of the AT traction network consists of feeder wire (FW), messenger wire (MW), trolley wire (TW), protective wire (PW), rail (R), and ground wire (GW), as shown in Figure 1.

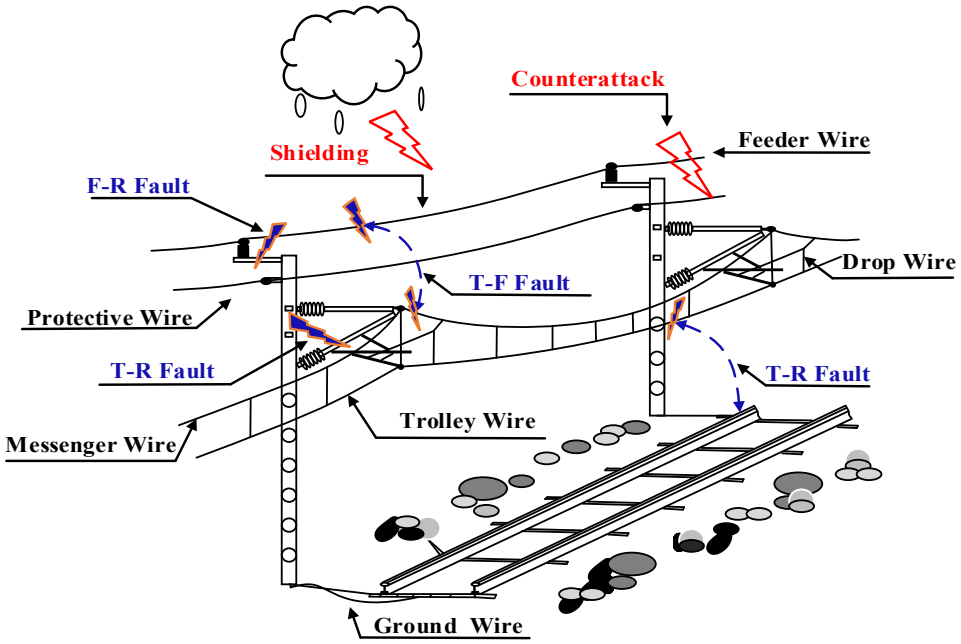

Figure 1. Structure configuration of the AT traction Network.

The aforementioned conductors in the full parallel AT traction network line are divided into uplink and downlink. The wire model, equivalent radius, and DC resistance parameters are given in Table 1. The equivalent radius of the rail conductor is set based on the results of the finite element analysis in [32]. The V/X transformer of the traction substation contains two single-phase three-winding transformers which supply the left and

right AT traction network, respectively. The neutral point of the secondary side winding is pulled out and ground so that the voltage of the two windings is 27.5 kV. The two windings are connected to the T bus and F bus, respectively, to form AT power supplymod. The ideal transformer model (1:1) is used to form a self coupling connection in order to simulate the neutral point extraction from the secondary side winding of the transformer. The rated capacity of the autotransformer is 10 MVA, the no-load loss is 5.0 kW, the load loss is 23.0 kW, the no-load current is 0.45%, and the short-circuit voltage is 0.59%. According to [33], the short-circuit reactance at the 27.5 kV side is 0.45 Ω, and the resistance at 27.5 kV side is 0.17 Ω, which may well satisfy the requirements of a full parallel AT traction power supply system.

Table 1. Simulation model related parameters.

Wire	Model	Equivalent Radius (cm)	DC Resistance (Ω/km)
TW	TCG-100	0.460	0.179
MW	TJ-95	0.650	0.315
PF	LGJ185	0.903	0.163
PW	LGJ120	0.722	0.255
R	P60	1.279	0.135
GW	—	0.406	0.280

According to the principle of wire merging, the full parallel AT traction network is equivalent to the parallel connection of three uplink and downlink TRF wires, as shown in Figure 2.

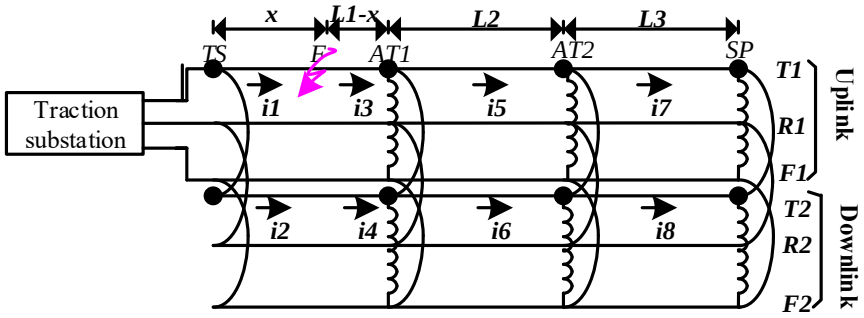

Figure 2. Structure configuration of AT traction Network.

Low leakage reactance AT autotransformers are connected in parallel between lines every 10~15km, and 2~3 autotransformers are connected in parallel for each power supply arm [5]. The autotransformer divides the power supply arm into three sections. Set the distance between TS~AT1 as L1, AT1~AT2 as L2, and AT2~SP as L3.

Short circuit faults are divided into the short circuit between trolley wire and rail (TR), short circuit between positive feeder line and rail (FR), and short circuit between trolley wire and positive feeder line (TF). Considering that the nonlinear resistance model under the action of the lightning wave is not accurate enough, the lightning stroke model adopts the simplified version of the IEEE model proposed by P. Pinceti.

Judging whether a lightning strike flashover or not depends on the current value of the flashover analog switch. If the current value is zero, the flashover does not occur. If the current value is not zero, the flashover occurs. The double exponential lightning wave is used to simulate the lightning current waveform. There are two types of direct lightning strike on the traction network line, namely counter strikes and shielding strikes. The counterattack is a lightning strike on the tower or protective line, which causes the tower potential to rise and discharges the line. The shielding failure is caused by lightning bypassing the protection of the lightning conductor and directly hitting the line.

2.2. Influence of AT Autotransformer on Traveling Wave Propagation Characteristics and Selection of Current Traveling Wave

AT autotransformers can not only solve the electromagnetic interference problem, but also reduce the voltage loss of the traction network and the cost of the line. Leakage impedance has a great influence on the performance in terms of preventing electromagnetic interference and reducing voltage loss. An autotransformer cannot be regarded as an open circuit [33]. The influence of the full parallel structure of the traction network and the AT autotransformer structure on the traveling wave propagation characteristics is shown in Figure 2. The autotransformer constitutes an impedance discontinuity point. The influence of a full parallel AT structure, no AT autotransformer, and no full parallel AT on a traveling wave is shown in Figure 3.

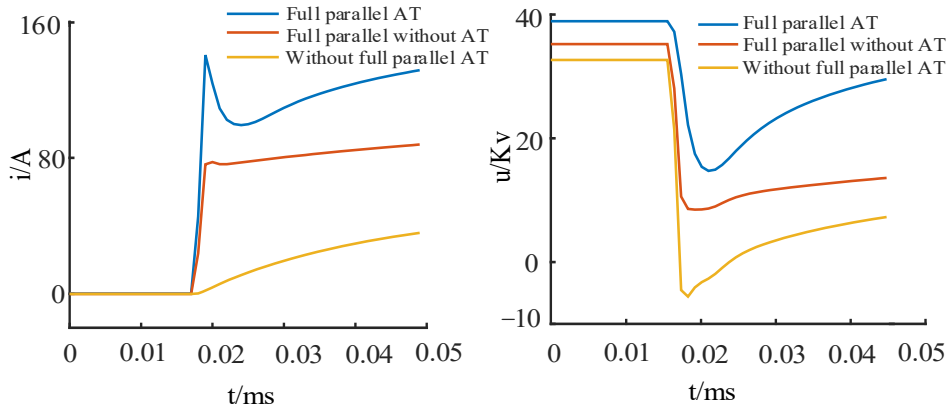

Figure 3. Influence of full parallel and AT structure on traveling wave.

The initial traveling wave amplitude of current without full-parallel connection and without AT autotransformer is reduced to half of the former respectively, and the initial voltage amplitude is increased by half. In the case of high-frequency transient fault, it is necessary to consider the influence of the capacitance between the windings of the autotransformer and the capacitance of the winding to the ground on the traveling wave propagation characteristics, as shown in Figure 4.

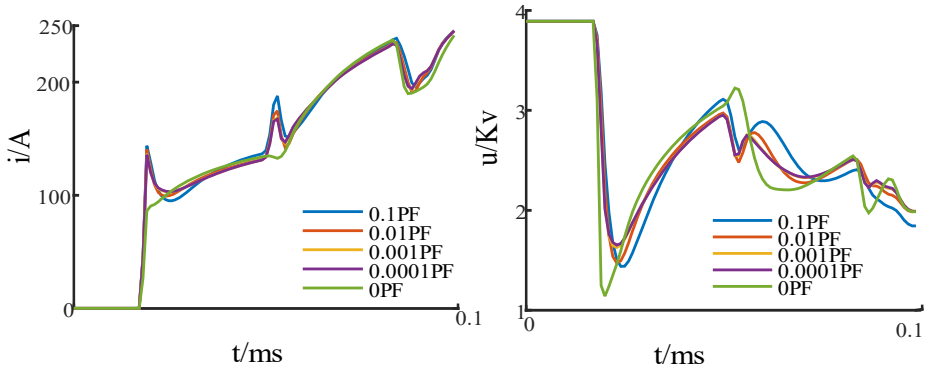

Figure 4. Influence of winding capacitance on traveling wave characteristics.

With the increase of the winding capacitance, the peak value of the initial electric current wave head gradually increases, and the winding capacitance has less influence on the current traveling wave, so the current traveling wave is selected as the transient waveform for signal classification.

2.3. Problems Arising from Traveling Wave Propagation Law of Full-Parallel Traction Network

Due to the different mechanisms of lightning current shielding or counterattack, when lightning strikes the lightning wire or the top of the tower, it is a counterattack.

The phenomenon of bypassing the lightning wire and directly striking the conductor is called shielding. A part of the thundercloud charge is injected into the earth through the lightning wire and protective wire before the flashback is injected into the conductor through the insulator after the flashover, which makes the amplitude and steepness of the current traveling wave after the flashover jump. The rapid decline soon became flat, and the waveform is similar to the double exponential waveform. The distance between AT posts is short, and the catadioptric propagation of traveling waves between AT posts, traction transformers, and fault points is complex, which makes the traveling waves continuously superimposed in the propagation process. Due to the existence of positive and negative polarity of the propagation coefficient, the amplitude of current traveling waves increases in the subsequent attenuation propagation process. The current traveling wave on the counterattack non-flashover conductor is mainly induced by electromagnetism. The current rises suddenly, then drops rapidly, and tends to be flat. The current attenuates rapidly compared with the shielding current, and the current oscillation is not obvious. The shielding oscillation frequency is smaller than the counterattack, but the oscillation amplitude is much larger than the counterattack. The amplitude of the lightning current of shielding non-flashover is larger than that of counterattack, and the oscillation does not take the form of the chopped wave. The amplitude range of the traveling wave of a short circuit fault, counterattack fault, and shielding fault cannot be distinguished by amplitude characteristics, and the frequency range cannot be distinguished.

When the traveling wave passes through the fault point, the AT autotransformers and traction transformer reach a discontinuous point of impedance, the catadioptric traveling wave continuously produces superposition, and the energy is almost generated, irregularly attenuated after superposition, and then superimposed and attenuated again.

This paper has verified the method in [30,31], and it is impossible to distinguish the transient fault by energy ratio [31] or generate a time-frequency image by constructing a time-frequency energy matrix to distinguish images [30]. It is expected to differentiate fault signals adaptively by analyzing one-dimensional signals from time-frequency amplitude characteristics.

3. Fault Classification Method

In order to solve the problem concerning which method is used to extract the effective fault features and to solve the problem that the fault traveling wave is difficult to identify, the solution can be found based on the search for Hilbert–Huang transform and depth learning method. Due to the ability of the Hilbert–Huang transform to process non-stationary data, it has found many different applications in the real world.

Due to the attenuation of signals of different fault types in the transmission process, the attenuation of transient signals at different fault distances is different, the traveling wave superposition is different, the frequency energy distribution range cannot be distinguished, and the time-frequency energy matrix of equal frequency bands cannot be constructed for differentiation. In order to extract the characteristics of fault signals, HHT is used to process the transient fault signal, the high-frequency characteristic mode components are separated, and the effective high-frequency mode features are extracted. The extracted high-frequency mode one-dimensional data are taken as the fault signal feature, and LSTM is used to identify the similarity of one-dimensional data to classify the fault types. However, the mode components decomposed by HHT are prone to spectral aliasing, so this paper improves the traditional HHT.

3.1. Improved HHT for Extracting One-Dimensional Fault Characteristic Signal

Time domain information can be obtained by collecting transient current signals. The HHT signal processing method can convert the time-domain fault signal into a time-frequency signal and separate the transient high-frequency mode component from the low-frequency mode component by decomposing it into multiple intrinsic mode functions

(IMF). The transient signal with multiple high-frequency mode components can reflect the characteristics of specific fault types.

Empirical mode decomposition (EMD) decomposes the fault current signal into a series of intrinsic IMF components. Therefore, the original signal $x(t)$ is decomposed into n IMF components and a residual signal. The formula is shown in (1).

$$s(t) = \sum_{j=1}^{n} c_j(t) + r_n(t) \tag{1}$$

where $c_j(t)$ is the order j IMFcomponent and $r_n(t)$ is the remainder. Taking the short-circuit waveform as an example, the original current signal is shown in Figure 5a, and the decomposed mode components are shown in Figure 5b–d.

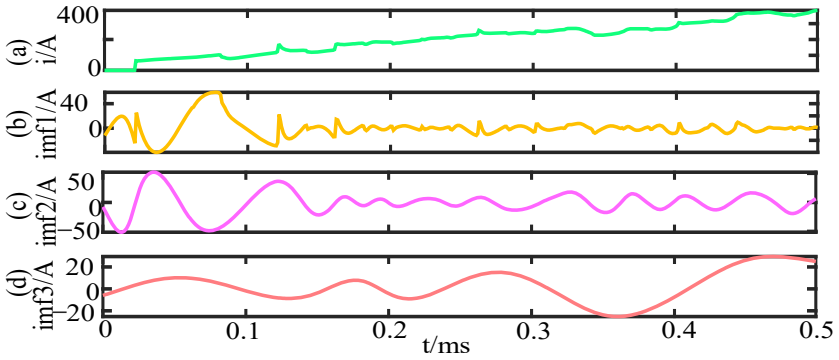

Figure 5. Original fault signal and IMF component.

The traditional HT of the continuous signal $x(t)$is defined as (2).

$$Y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{X(\tau)}{t-\tau} d\tau = x(t) \otimes \frac{1}{\pi t} \tag{2}$$

where $Y(t)$ is the converted signal, and $h(t) = 1/\pi$t is the transfer function. The transformed signal is shown in the following formula (3).

$$Y(\omega) = X(\omega) \cdot H(\omega) = -j \cdot \mathrm{sgn}(\omega) \cdot X(\omega) \tag{3}$$

When $X(t)$ and $Y(t)$ form a complex conjugate, an analytical signal $Z(t)$ can be obtained:

$$z(\mathrm{t}) = x(t) + i \cdot Y(t) = a(t) \cdot \updownarrow^{i\theta(t)} \tag{4}$$

where $a(t)$ and $\theta(t)$ are the instantaneous amplitude and phase, which can be calculated as:

$$a(t) = \sqrt{X^2(t) + Y^2(t)}, \theta(t) = \arctan\left(\frac{Y(t)}{X(t)}\right) \tag{5}$$

instantaneous frequency:

$$f(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt} \tag{6}$$

The IMF components of the corresponding order of the fault signal are obtained from EMD, as shown in Figure 5. The first two order IMF components are high-frequency components, which contain the transient characteristics of the fault signal. However, the frequency band ranges of different original signals may be different.

Therefore, this paper proposes that the improved HHT can effectively extract high-frequency transient fault features and identify different fault signals. Applying the Hilbert transform to each component of IMFs, one can obtain the corresponding instantaneous

amplitude and frequency of high-order mode components. The Hilbert spectrum of fault signals can be expressed as shown in (7).

$$H(w,t) = \mathrm{Re}\left(\sum_{j=1}^{n} \alpha_j(t) e^{i\int 2\pi f_j(t)dt}\right) \tag{7}$$

The data points of each mode component are s_{ij} (i = 1, 2,..., M and j = 1, 2,..., N). The original signal components within the frequency range of M-order mode components can be constructed by constructing reserved points. The instantaneous frequency $f > 0.1$ MHz amplitude point of all IMF components within the M-order frequency range remains unchanged, while the point beyond this segment is set to zero. The original high-frequency mode components are processed as shown in Figure 6a,b.

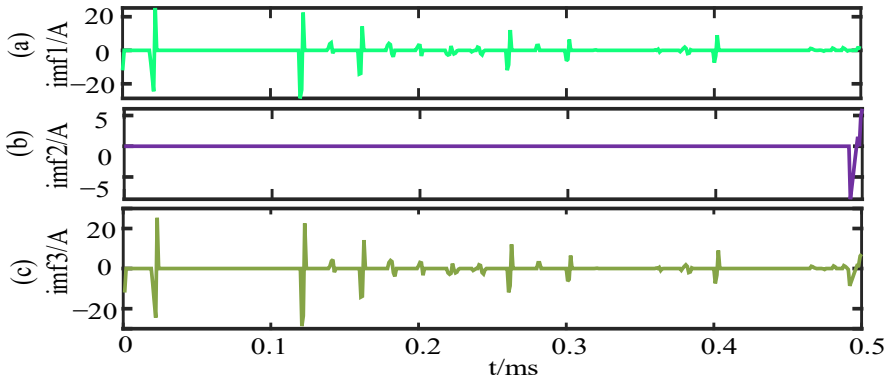

Figure 6. One Dimensional Fault Traveling Wave after Improved HHT.

The first two order IMF after improving HHT are shown in Figure 6a,b. The second and third-order IMFs do not have transient fault characteristics. The improved IMF is mixed to generate a one-dimensional signal waveform, as shown in Figure 6c.

3.2. Fault Classification Based on Multiscale LSTM

The traditional method uses the signal analysis method to construct the time-frequency spectrum and time-frequency energy matrix [30] and calculate the energy ratio of each frequency band [31]. Due to the existence of AT autotransformer, traction substation, and other special structures in the lines of the full parallel traction power supply system, back-flow occurs in the uplink and downlink at the same time. The traveling wave propagation mechanism is complex, the traditional method cannot extract effective signal features for identification, and the results of fault classification are not guaranteed.

In recent years, LSTM neural network is frequently used to process one-dimensional time series data. As a special structure of the neural network, it can adaptively extract features and classify one-dimensional data. The advantage of LSTM is that it can decide whether to forget or write new information to store long-term correlation by setting a "gate". The most significant feature is to automatically extract signal features through the training process.

In the LSTM network, each neuron is a storage unit, including an input gate, a forgetting gate, and an output gate. The output of the forgetting gate is f_t, the sigmoid activation function is used to map the output to the [0,1] interval. When the last time state C_{t-1} passes through the forgetting gate, it is multiplied by f_t. It determines how much of the previous state C_{t-1} enters the current state. The hidden layer state of the previous time is h_{t-1}, and x_t is the input of the current time. The input gate determines what is reserved. The information reserved at the previous time, together with the input reserved information, constitutes the current time state C_t. The output gate O_t determines which

information of h_{t-1} and x_t will be output. Mathematically, each neuron can be represented as follows:

$$f_t = \sigma\left(h_{t-1} \cdot W_f + x_t \cdot U_f + b_f\right) \tag{8}$$

$$i_t = \sigma\left(h_{t-1} \cdot W_i + x_t \cdot U_i + b_f\right) \tag{9}$$

$$c_t = f_t \times c_{t-1} + i_t \times \tanh(h_{t-1} \cdot W_c + x_t \cdot U_c + b_c) \tag{10}$$

$$o_t = \sigma(h_{t-1} \cdot W_o + x_t \cdot U_o + b_o) \tag{11}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{12}$$

where W_f, W_i, W_c, W_o, U_f, U_i, U_c, and U_o represent weight, and b_f, b_c, and b_o represent offset. The data used for training and testing in the neural network are extracted by improved HHT, and the structure of the multiscale LSTM network is shown in Figure 7.

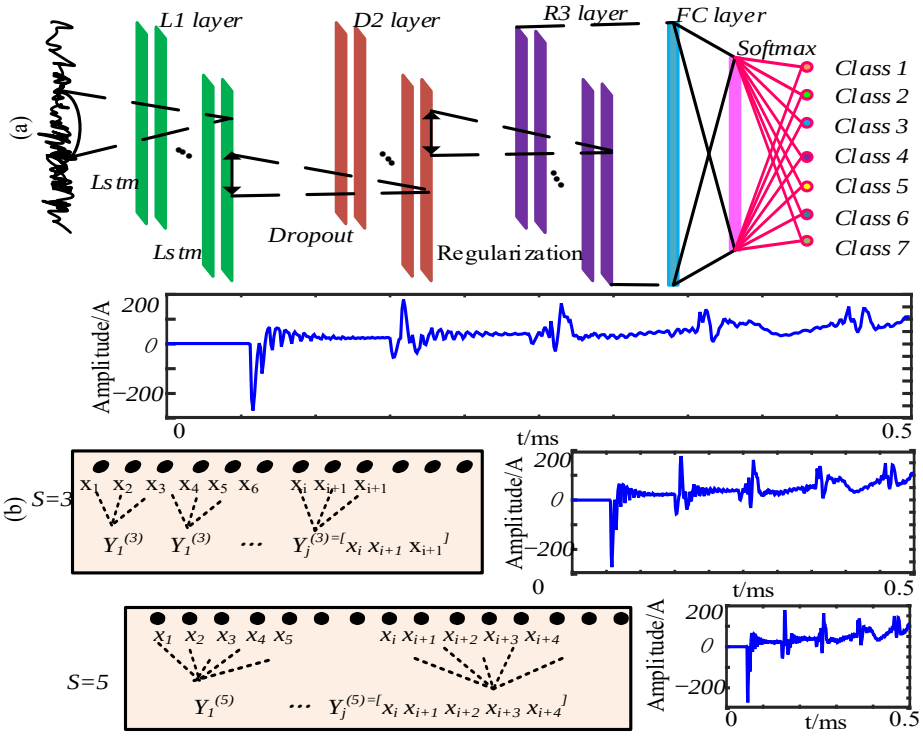

Figure 7. LSTM structure diagram (**a**) the LSTM network structure and (**b**)the multi-scale 1D data transformation structure.

The input of multi-scale LSTM hidden layer: $x1$ = {$x1$, $x2$,... xN}. It converts one-dimensional data into multi-scale s, which can not only improve the calculation efficiency, but also extract fault features from the multi-scale dimensions, increasing the calculation accuracy. For example, when s = 3, it is converted into two-dimensional data N/s × 3, and transmitted to the training network LSTM for current classification. The size of input layer is (N/s, s), the unit of LSTM layer is N/s. Each LSTM unit processes s one-dimensional arrays every time. Therefore, the LSTM network has an input layer containing N/s neurons, each neuron corresponding to a specific channel, the size of Dropout layer is 0.2, it discards 0.2 randomly, and then performs regularization to prevent overfitting. The size of Regularization is N/s, and the output layer is composed of 7 neurons.

3.3. Fault Classification Algorithm Flow Chart

The flow chart of the algorithm for the classification of lightning and short circuit faults in the traction network is shown in Figure 8.

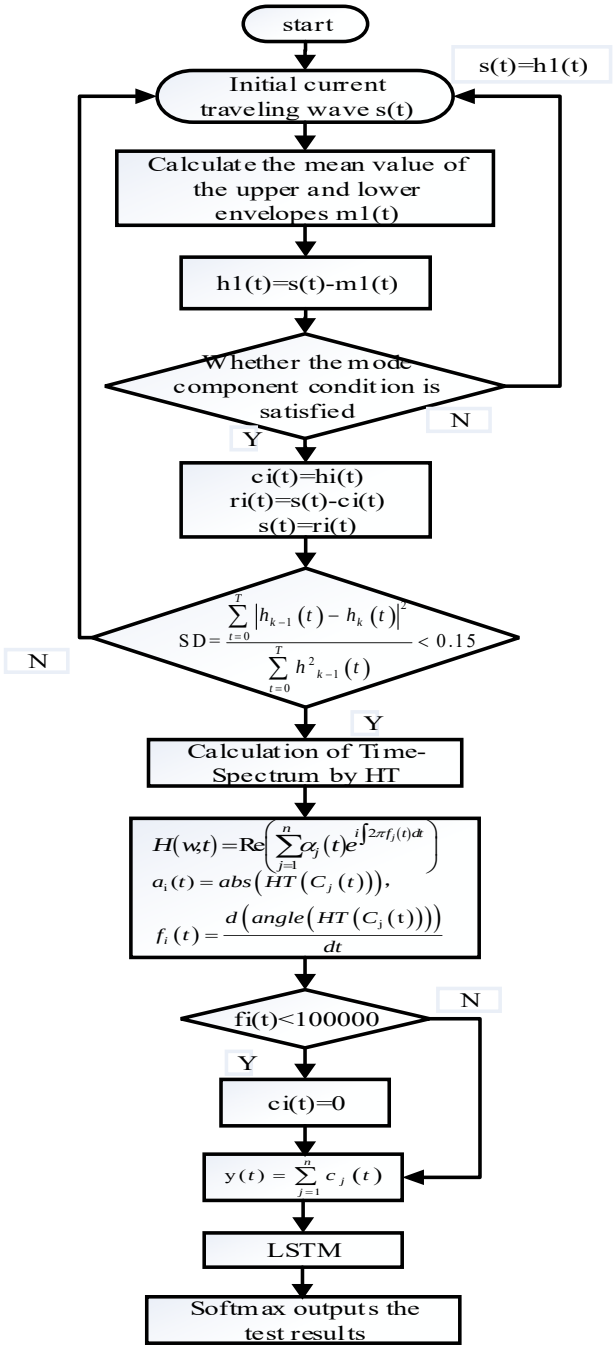

Figure 8. Flow Chart of Traction Network Transient Fault Classification.

4. Analysis and Discussion of Simulation Results

The electromagnetic transient simulation model of the traction network is built based on ATP-EMTP. The total length of the line is 45 km. Two ATs are set in the middle. The power supply arm is divided into three sections, each of which is 15 km. Seven models are simulated, including TR short circuit, TF short circuit, and FR short circuit fault models; flashover and non-flashover of the counter lightning strike, flashover, and non-flashover of surrounding lightning strike. The transient impedance of short-circuit fault simulation model is 1 Ω and 10 Ω respectively. The flashover lightning current of counter-strike lightning is set as 80 kA, 90 kA, 100 kA, and the non-flashover lightning current of counter-strike lightning is set as 50 kA, 60 kA, and 70 kA. The lightning current amplitude of shielding lightning flashover is 3 kA and 4 kA, and the lightning current amplitude of shielding lightning non-flashover is 1.5 kA and 2 kA. The fault angles of short circuits and

lightning strokes are 6, 17, 30, 49, and 90 degrees respectively. The fault sections are the first, the second, and the third sections. The TS end at the substation outlet is set as the signal monitoring point. HHT is implemented on MATLAB, and LSTM neural network algorithm is implemented on Tensorflow. Seven transient waveforms and their processing are analyzed as follows.

4.1. Counterattack Flashover

As shown in Figure 9, (a) represents the lightning current amplitude of 80 kA, S2 fault section, fault distance of 16 km, and fault angle of 6 degrees.

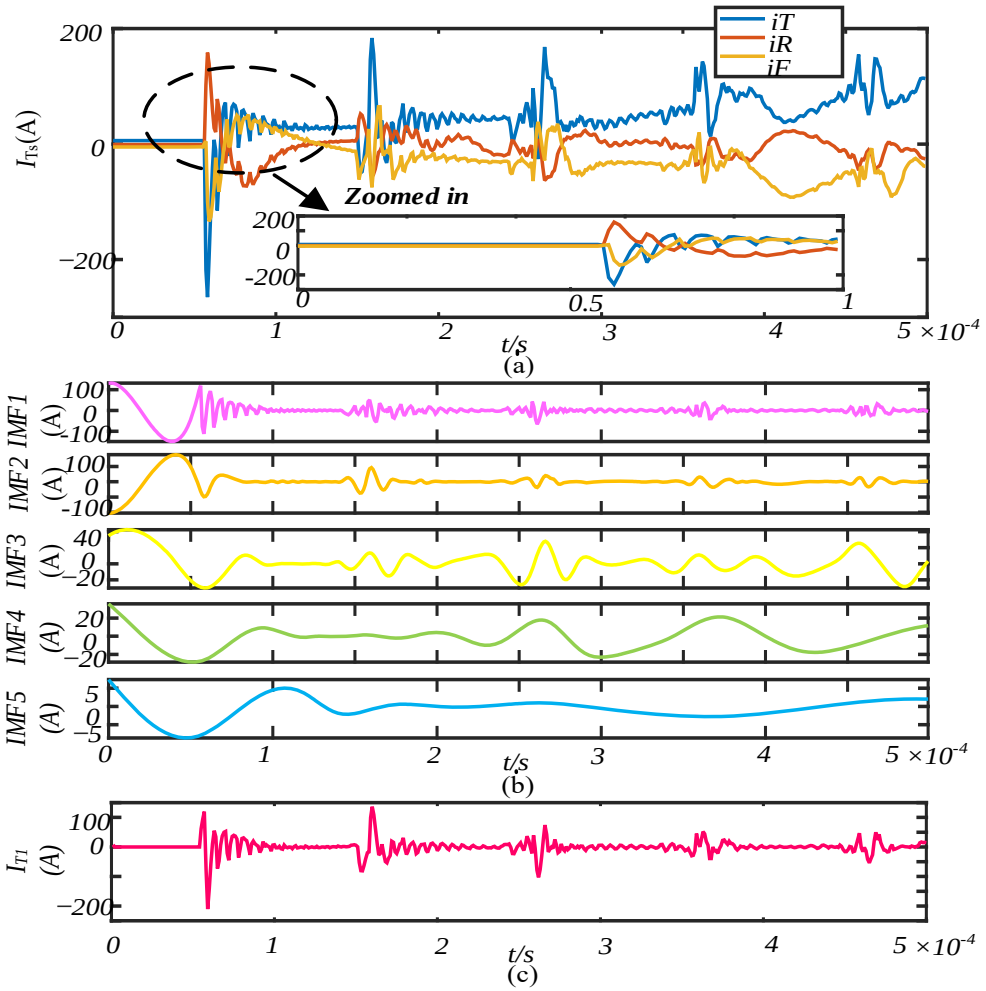

Figure 9. Analysis diagram of counterattack flashover waveform and improved HHT signal.

When the lightning current reaches TS at 0.053 ms, the TRF three-wire current suddenly rises rapidly. The amplitudes of T and F lines increase from 0.053 ms to 0.5 ms, but the amplitude of up and down oscillation decreases, and the amplitude energy is between −250 A and 200 A.

The EMD decomposition of the current traveling wave at the monitoring point is carried out to obtain each mode component as shown in Figure 9b. The Hilbert spectrum of each IMF component is calculated, and a three-dimensional time-frequency energy diagram is obtained, as shown in Figure 10a. The low-frequency energy of the transient fault signal is high, and the characteristics of the high-frequency transient signal are masked. The current traveling wave IT1 after being modified to HHT processing is shown in Figure 9c. At the initial moment, the current rapidly rises to the peak, then gradually decays, and then rises rapidly after a period of time. Before it gradually attenuates, the rising amplitude is

weakened, and the attenuation oscillation frequency is large. The oscillation amplitude ranges from −200 A to 150 A.

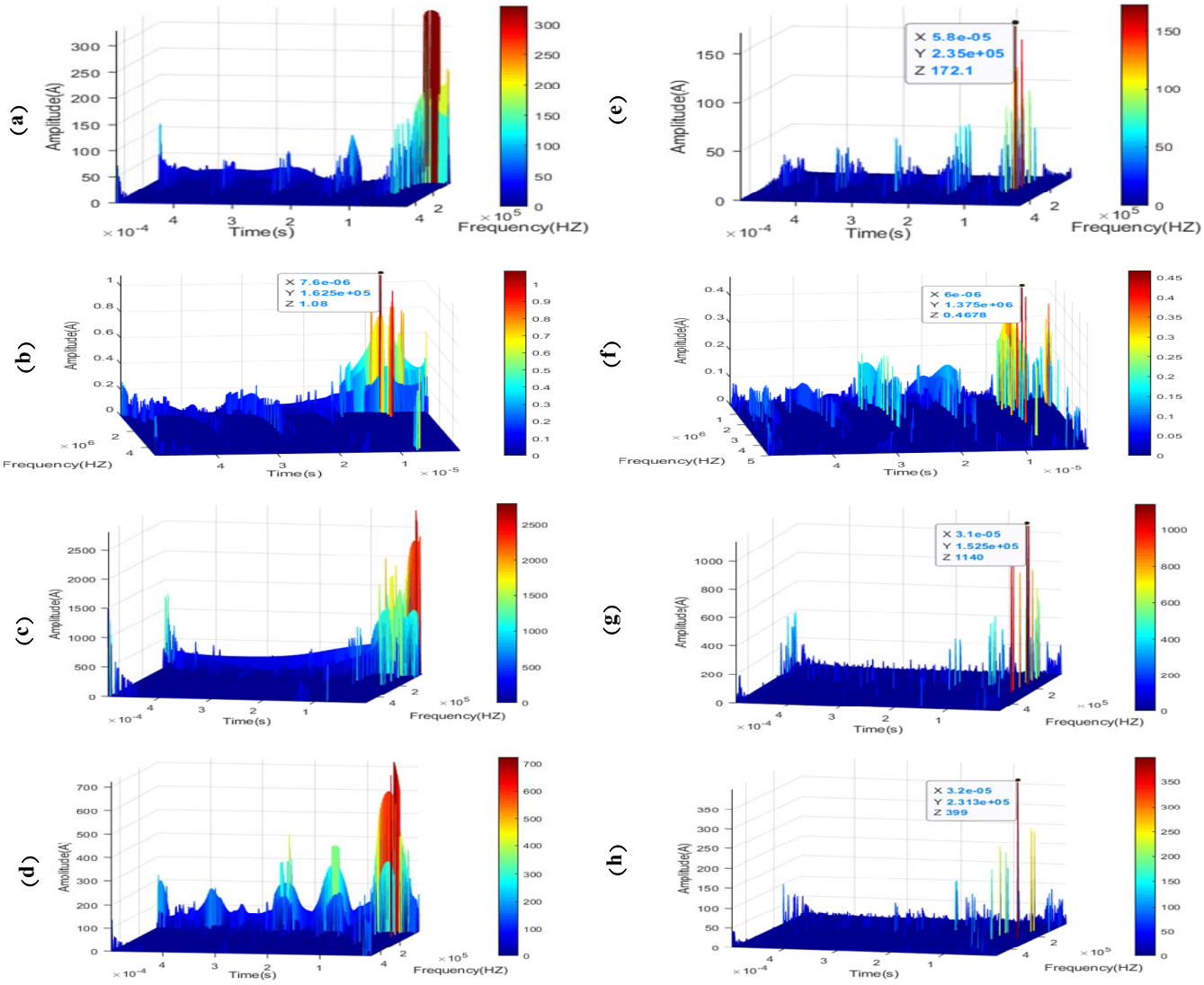

Figure 10. Lightning flashover and non flashover time frequency energy diagram before and after improvement HHT.

The time-frequency energy diagram after the improved HHT is analyzed and calculated as shown in Figure 10e. The high-frequency features are extracted. The maximum energy point amplitude in the diagram is 172.1 A, the frequency is 0.235 MHz, and the time is 0.58 ms when the fault traveling wave is transmitted from the fault point to the TS.

4.2. Counterattack Non-Flashover

As shown in Figure 11, (a) is not a flashover current traveling wave with a lightning current amplitude of 50 Ka, S2 section fault, fault distance of 18 km, fault angle of 30 degrees, and lightning strike PW line.

When the lightning current reaches the initial traveling wave head of TS at 0.6 ms, the TRF three wires current suddenly rises, and the amplitude energy of the Trolley Wire current is between 4.4 A and 5.4 A. The TW current is decomposed by EMD to obtain its mode components, as shown in Figure 10b. The current waveform attenuation oscillation is not obvious, and the amplitude is as low as −0.4 A to 0.6 A.

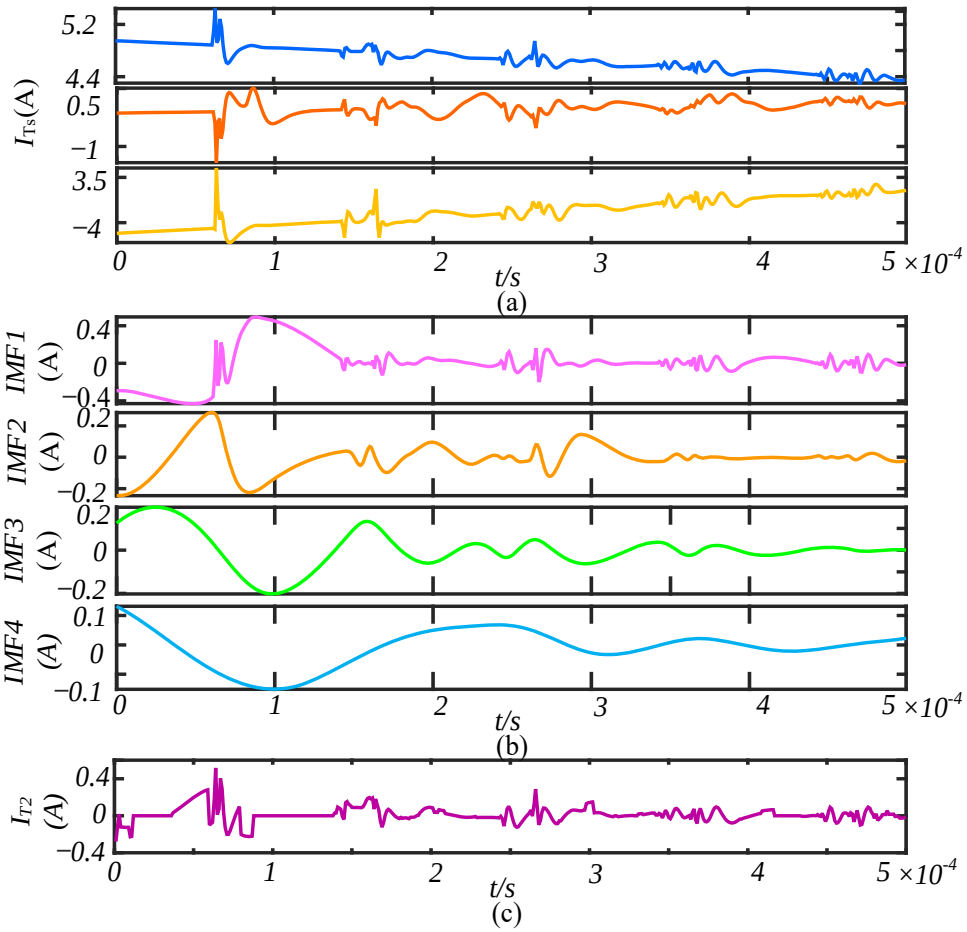

Figure 11. Analysis diagram of counterattack non-flashover waveform and improved HHT signal.

The three-dimensional time-frequency energy diagram obtained by calculating the Hilbert spectrum is shown in Figure 10b. The characteristics of high-frequency transient signals are masked and cannot be identified. The corresponding time of the highest energy point is not the initial traveling wave head, which leads to misjudgment. The current traveling wave IT2 after improved HHT is shown in Figure 10c.

The TW traveling wave generates a three-dimensional time-frequency diagram after improved HHT transformation, as shown in Figure 10f. The highest energy point is shown in Figure 10f, with an amplitude of 0.4678 A, frequency of 1.375 MHz, and time of 0.6 ms, namely the initial traveling wave head of lightning flashover fault waveform.

4.3. Shielding Flashover

As shown in Figure 12, (a) represents the TS three-line lightning current traveling wave with an amplitude of 3 Ka, S1 section fault, fault distance of 9 km, fault angle of 6 degrees, lightning strike T line flashover, and the TW current is decomposed by EMD to obtain its mode components, as shown in Figure 12b.

It can be obtained by enlarging the image. At 0.03 ms, the flashover occurs, and the current rises rapidly. The current traveling wave IT3 after the HHT is shown in Figure 12c. The TW current generates a time-frequency energy diagram before and after HHT transformation as shown in Figures 10c and 10g, respectively.

The TW current waveform starts to rise rapidly to the peak value, and then gradually attenuates. After a while, it rises rapidly and then decays rapidly. The rising amplitude is weakened, and the oscillation frequency is smaller than that of the lightning strike PW line, and the amplitude is larger, and the oscillation amplitude range is −800–1200 A.

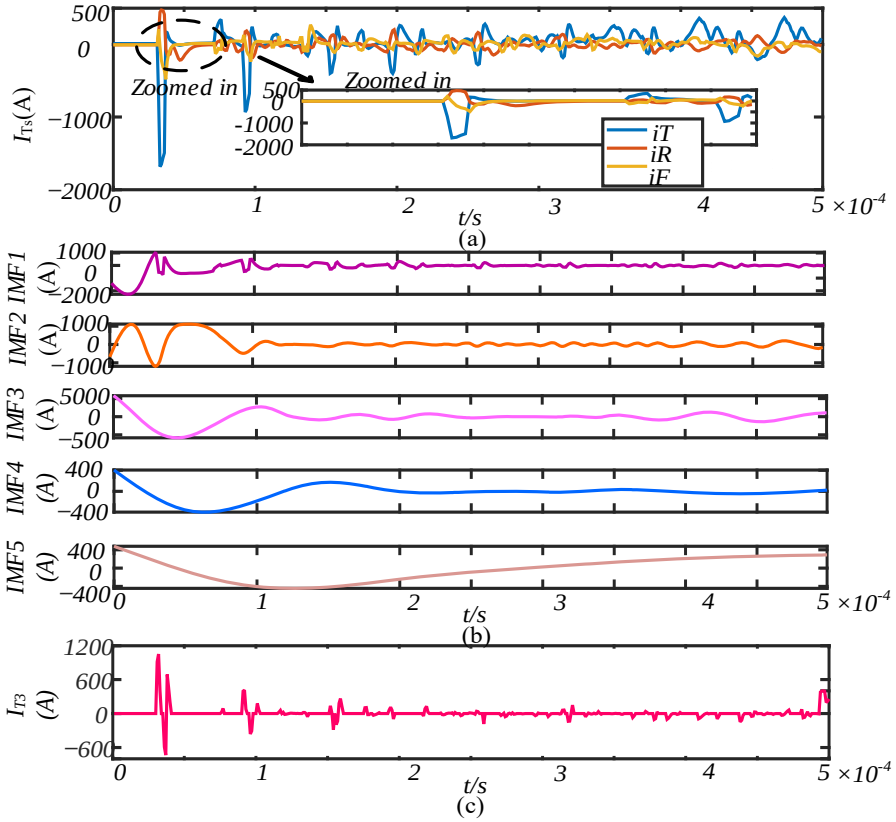

Figure 12. Analysis diagram of shielding flashover waveform and HHT signal.

4.4. Shielding Non-Flashover

As shown in Figure 13a, the amplitude of the lightning current is 1.5 Ka, the fault is in section S1, the fault distance is 9 km, the fault angle is 90 degrees, and the TRF current waveform at the TS end is not flashed by lightning strike T line. After waveform amplification, the jitter at approximately 0.0071 ms and 0.0091 ms is formed by AT1 reflected wave and secondary reflected traveling wave at the fault point, respectively.

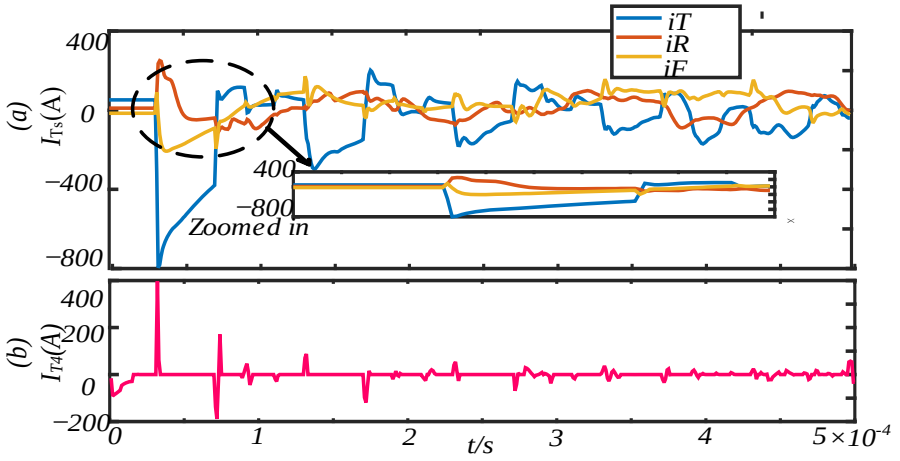

Figure 13. Analysis diagram of shielding non-flashover waveform and HHT signal.

The time-frequency energy diagram of the signal generated by the HHT transformation of TW current is shown in (d) of Figure 10. Compared with the non-flashover of the PW line by lightning, the non-flashover signal energy of the T line by lightning is larger, and the signal characteristics cannot be analyzed. The current waveform after HHT is shown in Figure 13b.

The current starts to rise rapidly to the peak value and then decays rapidly. After a while, it rises rapidly and then decays rapidly. Then, the rising amplitude is weakened, and the attenuation oscillation frequency is very small. The oscillation amplitude range is −200–400 A. The three-dimensional energy diagram of the improved HHT time-frequency diagram is shown in Figure 10h. The amplitude is 399 A, the frequency is 0.2313 MHz, and the time is 0.32 us, which is the wave head of the fault initial traveling wave.

4.5. Short Circuit (TR, TF, FR Short Circuit)

Figure 14a–c show the TR, TF, FR short circuit current traveling wave of TS terminal with fault section S1, fault distance of 6 km, fault angle of 30 degrees, transition resistance of 1 Ω. When a short circuit fault occurs, the line current rises rapidly, and the TR and TF short circuit current amplitude difference is small. The RF short circuit amplitude is the smallest, and is indistinguishable. The results calculated by HHT are shown in Figure 14d–f.

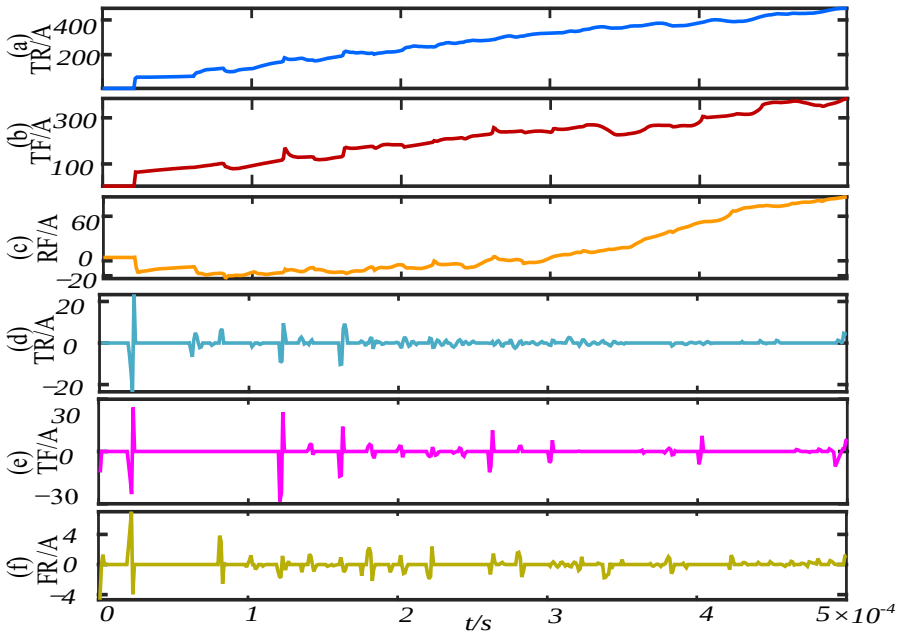

Figure 14. Short circuit fault signal before and after improved HHT.

The current amplitude of three different short circuits is much smaller than that of a lightning stroke. The amplitude difference between TR and TF is small, ranging from −25 A to 25 A. The minimum RF amplitude is between −4 A and 7 A. The oscillation frequency of TR is larger than that of TF short-circuit traveling waves. The current traveling wave analysis results before and after improved HHT are shown in Figure 15. Figure 15a,d shows the TR short-circuit time-frequency energy diagram before and after the improved HHT, Figure 15b,e shows the TF short circuit time-frequency energy diagram before and after the improved HHT, and Figure 15c,f shows the RF short circuit time-frequency energy diagram before and after the improved HHT.

The difference in the time-frequency distribution diagram of different short circuit faults can be seen. For several different lightning strikes and short-circuit faults, the frequency of transient signals is widely distributed. With the different fault angle transit impedance, the signal energy changes obviously, and it is difficult to distinguish. To prevent the impact of related physical quantities, feature extraction is carried out through HHT to decompose high-frequency traveling wave signals with characteristic information. The LSTM network is used to process time series to find out the characteristics of the transient signal catadioptric waveform at different times, so as to classify the signals.

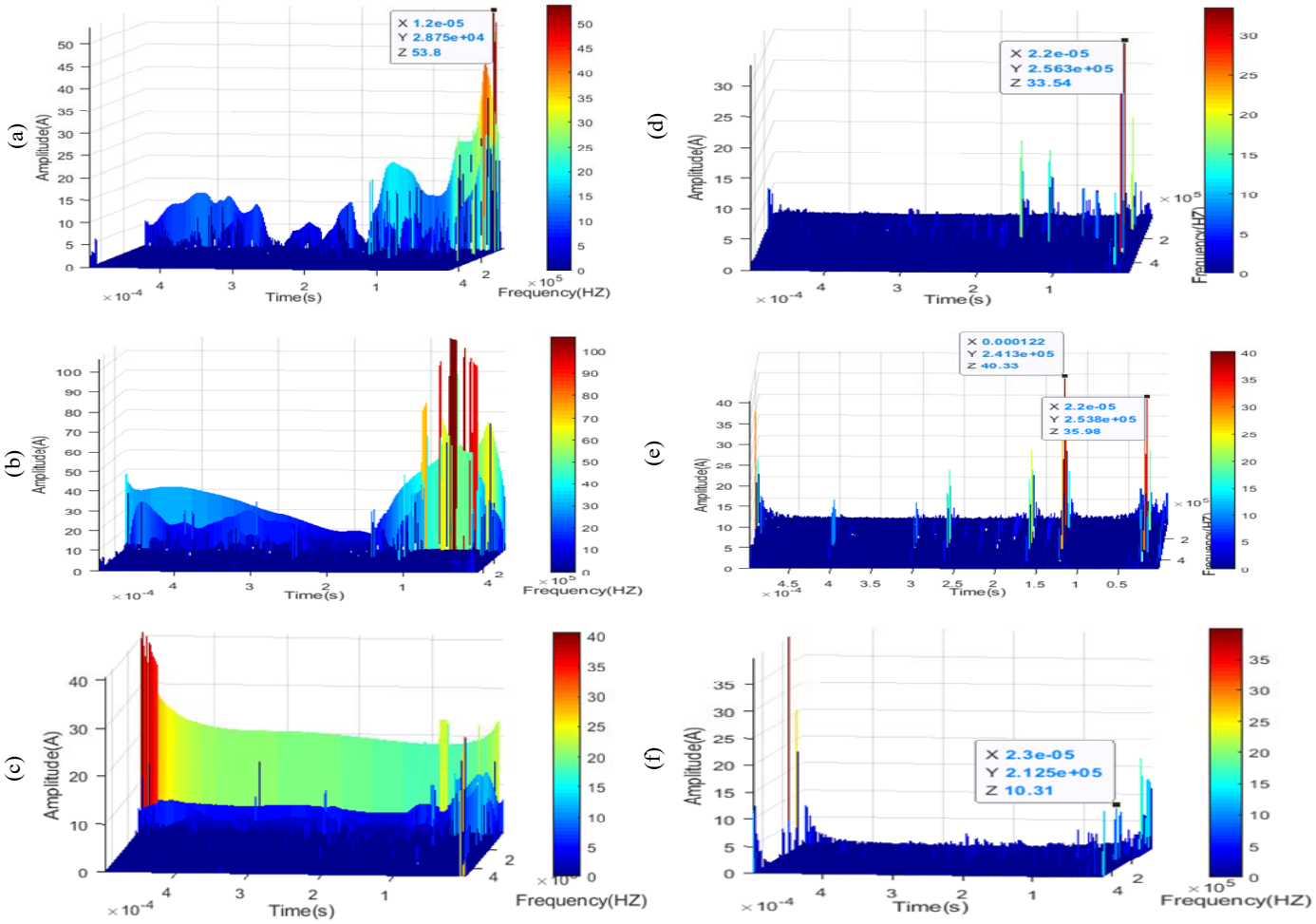

Figure 15. Time frequency energy diagram of short circuit fault before and after improved HHT.

5. Analysis and Verification of Fault Classification Results

5.1. Improve HHT + LSTM Training Results and Verification

Each layer structure and activation function of the model is critical to the classification performance. The LSTM network consists of seven cascaded layers, namely an input layer, an LSTM hidden layer, a drop-out layer, a regularization layer, a full connection layer, and a softmax output layer. The structure of the proposed neural network is shown in Figure 7. The optimizers that are widely used are SGD, Adagrad, Adam, Adamelta, RMSprop, Nadam, and Adamas. Adam is selected as the optimizer, the weight and deviation are updated according to the gradient of the loss function, and the mean square error is selected as the loss function. LSTM is trained offline, and the initial learning rate is set to 0.001. Adam is used to update all weights of the network. Through backpropagation, the weight and deviation of the trained neural network are realized through gradient descent and derivative chain rules.

Due to a large amount of data and the offline training algorithm, 3510 samples of short-circuit and lightning-stroke fault data are taken from the input data. In this paper, seven kinds of fault data are extracted from the original data. Each group of data samples contains 500 points. Counterattack flashover and non-flashover represent 270 groups of data respectively, shielding flashover and non-flashover represent 180 groups of data, respectively, TR, TF, and FR short circuits represent 870 groups of data, respectively, and training set and test set to account for 70% and 30%, respectively. Adding 20 db white Gaussian noise, 2457 groups of samples for training and 1053 groups of samples for testing are obtained, as shown in Table 2.

Table 2. Fault type training and test conditions.

Fault Type	Fault Impedance (Ω)	Fault Angle	Fault Location (km)	Fault Section	Numbers
Counterattack flashover	-	6, 17, 30, 49, 90	1, 3, 6, 9, 12, 14	s1, s2, s3	270
Counterattack non flashover	-	6, 17, 30, 49, 90	1, 3, 6, 9, 12, 14	s1, s2, s3	270
TR, TF, RF	1, 10	6, 17, 30, 49, 90	0.5, 1, 1.5, ..., 14.5	s1, s2, s3	870
Shielding flashover	-	6, 17, 30, 49, 90	1, 3, 6, 9, 12, 14	s1, s2, s3	180
Shielding non flashover	-	6, 17, 30, 49, 90	1, 3, 6, 9, 12, 14	s1, s2, s3	180

The final output layer selects seven sample tags for training and testing by the softmax activation function, and the counterattack flashover tag is defined as 0, counterattack non-flashover is 1, shielding flashover is 2, shielding non-flashover is 3, TR short circuit is 4, TF short circuit is 5; FR short circuit is 6. The whole neural network is composed of an input layer, a hidden layer (LSTM layer, a drop-out layer, a regularization layer, a full connection layer), and an output layer. The data flow in from the input layer, and the input data can be converted into multi-scale data. Through experimental comparison and analysis, S is taken as 10, the training speed is the fastest, and the accuracy is high. The one-dimensional data of the input signal is converted into 50 × 10, and the LSTM of 50 units is set. Through the LSTM neural network, each unit processes 10 data points at the same time, discards 0.2 randomly, and then performs regularization to prevent overfitting. Then, the full connection layer uses the softmax function to finally achieve output. Specific parameters are shown in Table 3.

Table 3. Related parameters of neural network.

Network Structure	Unit	Coefficients
LSTM	50	-
Dropout	-	0.2
Regularization	-	-
Softmax	7	-

For the current fault signal processed by the improved HHT algorithm, after 500 iterations of LSTM depth neural network training and testing, the accuracy and loss function size results are as shown in Figures 16 and 17, respectively.

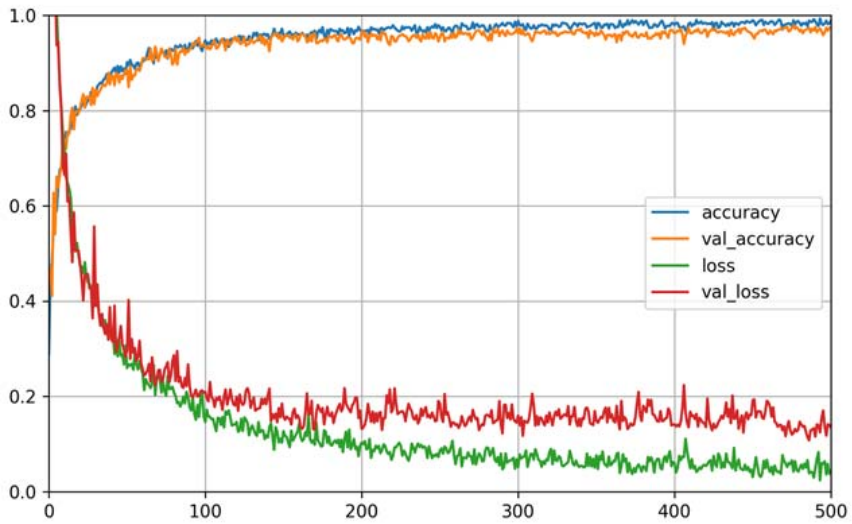

Figure 16. Accuracy and loss function after improved HHT post-processing training and testing.

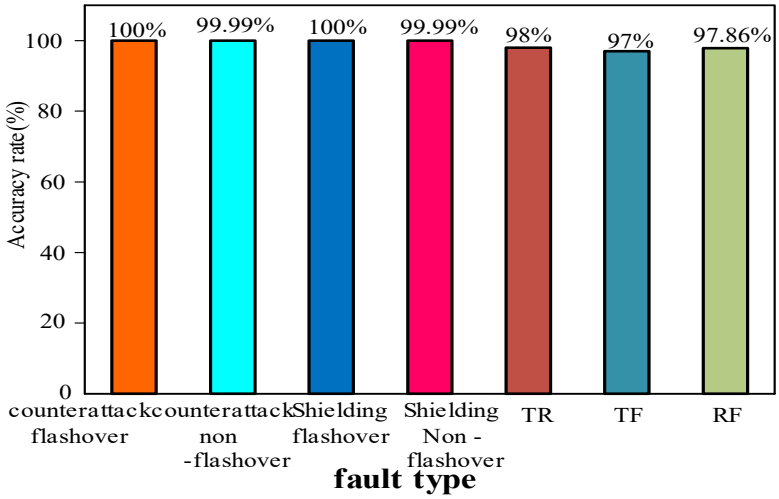

Figure 17. Test accuracy of different fault types.

As shown in Figure 16, the classification accuracy of the data set processed by improved HHT continues to improve with the training. After 500 times of training, the accuracy of the training set reaches 99.99%, and the accuracy of the test set reaches 97.67%. The loss function decreases rapidly in the first 100 periods and then converges to remain stable.

For the proposed improved HHT + LSTM method, 100 data sets are selected for testing for different fault types, and the statistical test results are shown in Figure 17. The minimum accuracy rate of TF short circuit fault is 97%, which means that three out of 100 faults are wrongly classified. The accuracy rate of lightning failure is 99.99% at the lowest and 100% at the highest.

5.2. Train and Test Results Only Using LSTM

In order to compare and verify the accuracy of the proposed features and improve the accuracy of fault classification, LSTM is used alone to classify the original data. As shown in Figure 18, the highest accuracy rates of the training set and test set are only 80% and 70%, respectively. The loss function value is very large and does not converge.

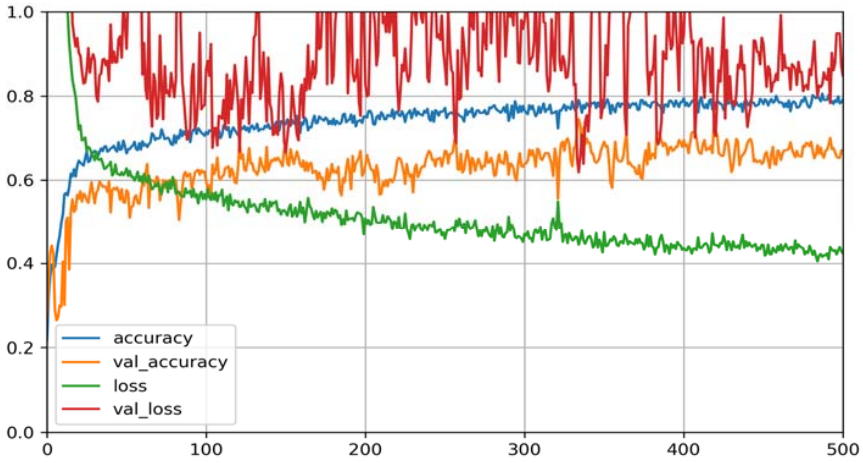

Figure 18. Raw data training and testing accuracy and loss functions.

In this study, the processor, memory, and operating system of the computer are Intel(R) Xeon(R) W-2123 CPU @ 3.60 GHz 3.60 GHz, 32.00 GB, and Win64 respectively. The execution time of the MATLAB fault classification algorithm program training and testing is about 1 min, and the execution time of the fault classification program after training is approximately 0.1 s. These results meet the requirements of practical engineering applications.

5.3. Comparison and Verification of Different Algorithms

The method proposed in this paper is compared with SVM, DWT + LSTM, HHT + CNN, DWT + CNN, and DWT + SVM. The results are shown in Table 4. It can be concluded that the method has higher accuracy and better performance than other classifiers.

Table 4. Classification accuracy of different algorithms.

Algorithm Model	TR	TF	FR	Counterattack Flashover	Counterattack Non Flashover	Shielding Flashover	Shielding Non Flashover
SVM	71.5%	71.5%	71.5%	72.5%	72%	71%	71.5%
LSTM	80%	82%	80.5%	83%	82%	80%	80%
HHT + LSTM	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%
DWT + LSTM	82%	83%	82%	80%	79%	80%	79%
HHT + CNN	80%	81%	80%	81%	82%	80%	81%
DWT + CNN	60%	65%	62%	60%	61%	65%	66%
DWT + SVM	70%	75%	75%	70%	70%	69%	70%
HHT + SVM	85%	83%	86%	80%	80%	85%	85%

5.4. Experimental Verification of Field-Measured Data

The short circuit data of the traction substation of the high-speed railway Wuhan Guangzhou line are measured by the power quality analyzer in the traction substation. The PQ-Box 200 power quality analyzer with a sampling frequency of 1 MHz is taken to measure the short circuit data from the traction substation, PQ-Box 200 Power quality analyzer in field experiments is shown in Figure 19. The total length of the full parallel AT traction network is 44.5 km, and the length of L1, L2, and L3 is 14.833 km. The measured results are shown in Table 5.

Table 5. Test and verification results of traction substation of Wuhan Guangzhou high-speed railway.

Fault Type	Time	Validation Results
TF(uplink)	10:18:23	TF short circuit
TR(uplink)	11:52:29	TR short circuit

Figure 19. PQ-Box 200 Power quality analyzer in field experiments.

6. Conclusions

Intended for the full parallel AT traction power supply system with a special structure, in this paper, the influence of full parallel structure, AT autotransformer, and winding capacitance on the traveling wave propagation characteristics is analyzed and the current traveling wave is selected as a transient fault classification signal. An improved HHT algorithm is used to deal with the transient fault signal of traction network, and one-dimensional feature data are obtained, which are analyzed in terms of time-frequency energy spectrum. In comparison to the results of the three-dimensional time-frequency energy analysis prior to improved processing of the HHT algorithm, it is concluded that the signal features after improved HHT processing are evident, and the high frequency transient signal features can be extracted effectively.

The one-dimensional feature data are converted to two-dimensional data as input to the LSTM, and each LSTM unit processes multivariate data at once. These data as fault feature signals are classified by LSTM. The proposed improved HHT + LSTM method is compared with the LSTM method to directly process the original signal. The accuracy of the improved HHT + LSTM classification algorithm is 99.99%, and the accuracy of using LSTM alone is 80%.

In comparison to many other classifiers, the classification algorithm proposed in this paper has better classification performance than the other classifiers. The method is reliable and robust to different fault points, loads, angles, flash current magnitudes, and other fault conditions.

Author Contributions: Conceptualization, methodology, validation and writing, H.Z.; methodology, conceptualization and guidance, J.C.; conceptualization, M.Y.; conceptualization, Q.F.; conceptualization, S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 51467004,and the double-class discipline of Traffic Information Engineering and Control of Jiangxi, grant number 1600218049.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Guillen, D.; Serna, J.A.D.L.O.; Zamora-Mendez, A.; Salinas, F. Taylor–Fourier Filter-Bank Implemented With O-Splines for the Detection and Classification of Faults. *IEEE Trans. Ind. Inform.* **2021**, *17*, 3079–3089. [CrossRef]
2. Doorwar, A.; Bhalja, B.R.; Malik, O.P. A New Internal Fault Detection and Classification Technique for Synchronous Generator. *IEEE Trans. Power Deliv.* **2019**, *34*, 739–749. [CrossRef]
3. Moyo, N.M.; Bansal, R.C.; Naidoo, R.; Sprong, W. Line Impedance Measurement to Improve Power Systems Protection of the Gautrain 25 kV Autotransformer Traction Power Supply System. *IEEE Access* **2019**, *7*, 136962–136974. [CrossRef]
4. Liu, P.; Huang, C. Detecting Single-Phase-to-Ground Fault Event and Identifying Faulty Feeder in Neutral Ineffectively Grounded Distribution System. *IEEE Trans. Power Deliv.* **2018**, *33*, 2265–2273. [CrossRef]
5. Chavez, J.J.; Popov, M.; López, D.; Azizi, S.; Terzija, V. S-Transform based fault detection algorithm for enhancing distance protection performance. *Int. J. Electr. Power Energy Syst.* **2021**, *130*, 952–962. [CrossRef]
6. Yu, X.; Xiao, L. A DC Fault Protection Scheme for MMC-HVDC Grids Using New Directional Criterion. *IEEE Trans. Power Deliv.* **2021**, *36*, 441–451. [CrossRef]
7. Salehi, M.; Namdari, F. Fault classification and faulted phase selection for transmission line using morphological edge detection filter. *IET Gener. Transm. Distrib.* **2017**, *12*, 1595–1605. [CrossRef]
8. Hu, K.; Liu, Z.; Lin, S. Wavelet Entropy-Based Traction Inverter Open Switch Fault Diagnosis in High-Speed Railways. *Entropy* **2016**, *18*, 78. [CrossRef]
9. He, Z.; Fu, L.; Lin, S.; Bo, Z. Fault Detection and Classification in EHV Transmission Line Based on Wavelet Singular Entropy. *IEEE Trans. Power Deliv.* **2010**, *25*, 2156–2163. [CrossRef]
10. Liu, Z.; Han, Z.; Zhang, Y.; Zhang, Q. Multiwavelet Packet Entropy and its Application in Transmission Line Fault Recognition and Classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2010**, *25*, 2043–2052. [CrossRef]
11. Sartori, C.; Sevegnani, F.X. Fault Classification and Detection by Wavelet-Based Magnetic Signature Recognition. *IEEE Trans. Magn.* **2010**, *46*, 2880–2883. [CrossRef]
12. Chen, J.; Dou, Y.; Yang, L.; Li, J. Application of Shannon Wavelet Entropy and Shannon Wavelet Packet Entropy in Analysis of Power System Transient Signals. *Entropy* **2016**, *18*, 437. [CrossRef]
13. Li, Y.; Lin, J.; Niu, G.; Wu, M.; Wei, X. A Hilbert-Huang Transform-Based Adaptive Fault Detection and Classification Method for Microgrids. *Energies* **2021**, *14*, 5040. [CrossRef]
14. Anand, A.; Affijulla, S. Hilbert-Huang transform based fault identification and classification technique for AC power transmission line protection. *Int. Trans. Electr. Energy Syst.* **2020**, *30*, 999–1012. [CrossRef]
15. Dehghani, M.; Khooban, M.H.; Niknam, T. Fast fault detection and classification based on a combination of wavelet singular entropy theory and fuzzy logic in distribution lines in the presence of distributed generations. *Int. J. Electr. Power Energy Syst.* **2016**, *78*, 455–462. [CrossRef]
16. Ghashghaei, S.; Akhbari, M. Fault detection and classification of an HVDC transmission line using a heterogenous multi-machine learning algorithm. *IET Gener. Transm. Distrib.* **2021**, *15*, 2319–2332. [CrossRef]
17. Fahim, S.R.; Sarker, S.K.; Muyeen, S.M.; Das, S.K.; Kamwa, I. A deep learning based intelligent approach in detection and classification of transmission line faults. *Int. J. Electr. Power Energy Syst.* **2021**, *133*, 1156–1170. [CrossRef]
18. Tong, H.; Qiu, R.C.; Zhang, D.; Yang, H.; Ding, Q.; Shi, X. Detection and Classification of Transmission Line Transient Faults Based on Graph Convolutional Neural Network. *CSEE J. Power Energy Syst.* **2021**, *007*, 456–471.

19. Zhang, C.; Wang, J.; Huang, J.; Cao, P. Detection and classification of short-circuit faults in distribution networks based on Fortescue approach and Softmax regression. *Int. J. Electr. Power Energy Syst.* **2020**, *118*, 105–117. [CrossRef]
20. Hemapriya, C.K.; Suganyadevi, M.V.; Krishnakumar, C. Detection and classification of multi-complex power quality events in a smart grid using Hilbert-Huang transform and support vector machine. *Electr. Eng.* **2020**, *102*, 1681–1706. [CrossRef]
21. Wang, J.; Sima, W.; Yang, Q.; Yuan, T. Classification and identifica-tion of over-voltage based on HHT and SVM. *Proc. International. Conference. High Volt. Eng.* **2010**, *8*, 85–88.
22. Wang, T.; Zhang, G.; Zhao, J.; He, Z.; Wang, J.; Perez-Jimenez, M.J. Fault Diagnosis of Electric Power Systems Based on Fuzzy Reasoning Spiking Neural P Systems. *IEEE Trans. Power Syst.* **2015**, *30*, 1182–1194. [CrossRef]
23. de Bruin, T.; Verbert, K.; Babuska, R. Railway Track Circuit Fault Diagnosis Using Recurrent Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 523–533. [CrossRef] [PubMed]
24. Rai, K.; Hojatpanah, F.; Ajaei, F.B.; Grolinger, K. Deep Learning for High-Impedance Fault Detection: Convolutional Autoencoders. *Energies* **2021**, *14*, 3623. [CrossRef]
25. Wang, Q.; Yu, Y.; Ahmed, H.; Darwish, M.; Nandi, A. Fault Detection and Classification in MMC-HVDC Systems Using Learning Methods. *Sensors* **2020**, *20*, 4438. [CrossRef]
26. Ince, T.; Kiranyaz, S.; Eren, L.; Askar, M.; Gabbouj, M. Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks. *IEEE Trans. Ind. Electron.* **2016**, *63*, 7067–7075. [CrossRef]
27. Silva, K.M.; Souza, B.A.; Brito NS, D. Fault Detection and Classification in Transmission Lines Based on Wavelet Transform and ANN. *IEEE Trans. Power Deliv.* **2006**, *21*, 2058–2063. [CrossRef]
28. Yu, J.; Hou, Y.; Lam, A.; Li, V. Intelligent Fault Detection Scheme for Microgrids With Wavelet-Based Deep Neural Networks. *IEEE Trans. Smart Grid* **2017**, *56*, 158–165. [CrossRef]
29. Jamil, M.; Singh, R.; Sharma, S.K. Fault identification in electrical power distribution system using combined discrete wavelet transform and fuzzy logic. *J. Electron. Syst. Inf. Technol.* **2015**, *2*, 257–267. [CrossRef]
30. Guo, M.F.; Yang, N.C.; Chen, W.F. Deep-Learning-based Fault Classification using Hilbert-Huang Transform and Convolutional Neural Network in Power Distribution Systems. *IEEE Sens. J.* **2019**, 1–9. [CrossRef]
31. Dai, Z.; Liu, N.; Zhang, C.; Pan, X.; Wang, J. A Pilot Protection for HVDC Transmission Lines Based on Transient Energy Ratio of DC Filter Link. *IEEE Trans. Power Deliv.* **2020**, *35*, 1695–1706. [CrossRef]
32. Peng, T.; Chen, J. Calculation of the High Frequency-dependent Impedance of Rail Using a Equivalent Tubular Conductor Model. *J. China Railw. Soc.* **2019**, *41*, 45–49.
33. Dai, P.; Liu, T.; Zhou, H. Characteristics of Travelling Wave Propagation in Catenary of High-speed Railway. *J. China R Ailway Soc.* **2014**, *36*, 25–30.