

Article

Signal Identification of Wire Breaking in Bridge Cables Based on Machine Learning

Guangming Li ¹, Heming Ding ¹, Yaohan Li ², Chun-Yin Li ^{2,*} and Chi-Chung Lee ²¹ Department of Mechanical, Electrical and Information Engineering, Shandong University, Weihai 264209, China² School of Science and Technology, Hong Kong Metropolitan University, Hong Kong, China

* Correspondence: cyli@hkmu.edu.hk

Abstract: With the booming development of bridge construction, bridge operation and maintenance have always been major issues to ensure the safety of the community. Affected by the long-term service of bridges and natural factors, the safety and durability of cables can be threatened. Cables are critical stress-bearing elements of large bridges such as cable-stayed bridges. Realizing the health monitoring of bridge cables is the key to ensuring the normal operation of bridges. Acoustic emission (AE) is a dynamic nondestructive testing method that is increasingly used in the local monitoring of bridge cables. In this paper, a testbed is described for generating the acoustic emission signals for signal identification testing with machine learning (ML) models. Owing to the limited number of measured signals being available, an algorithm is proposed to simulate acoustic emission signals for model training. A multi-angle feature extraction method is proposed to extract the acoustic emission signals and construct a comprehensive feature vector to characterize the acoustic emission signals. Seven ML models are trained with the simulated acoustic emission signals. Long short-term memory (LSTM) has been specially applied for deep learning demonstration which requires a large amount of training data. As all machine learning models (including LSTM) provide desired performance, it shows that the proposed approach of simulating acoustic emission signals can be effective.

Keywords: acoustic emission; bridge cable; deep learning; health monitoring; synthetic data**MSC:** 68T05

Citation: Li, G.; Ding, H.; Li, Y.; Li, C.-Y.; Lee, C.-C. Signal Identification of Wire Breaking in Bridge Cables Based on Machine Learning. *Mathematics* **2022**, *10*, 3690. <https://doi.org/10.3390/math10193690>

Academic Editor: Catalin Stoean

Received: 21 July 2022

Accepted: 6 October 2022

Published: 9 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the rapid development of bridge construction, the management and maintenance of bridges have become a key issue. In the past, the health monitoring of bridges was mostly performed by manual inspection. With the increasing number of bridges being built and the increasing complexity of bridge structures, the traditional manual inspection exposes the shortcomings, such as a small monitoring range, large workload, and low inspection efficiency, which may also endanger the safety of inspectors under harsh environments. Therefore, a complete bridge structural health monitoring and early warning system [1,2] is needed. Such a system can realize real-time monitoring of field data, mitigate risk, and provide timely warning by promptly detecting early defects during the operation of bridges.

The current health monitoring of bridges is mostly based on overall monitoring. This method is mainly based on the monitoring of bridge service environment, operational loads, and structural response data to achieve bridge structural conditions and safety performance [3]. However, compared with the overall condition, monitoring local structures such as cables and suspenders is also important for bridge operation and maintenance. Therefore, the local damage monitoring methods should be introduced in conjunction with overall bridge monitoring techniques. Acoustic emission (AE), as a non-destructive testing technique [4], can be applied. This technique can record the signals generated

by the target acoustic emission process in real time with deploying sensors to the target for data acquisition. By applying it to bridge cable monitoring, only a small number of sensors are required to achieve efficient detection of the broken wires, thus effectively assessing the damage level of bridge cables. Researchers in the field of acoustic emission monitoring have conducted numerous studies based on different types of cable damage. Such damage analyses include corrosion of the natural environment in which the bridge is located [5–8], fatigue loading from long-term vehicle passage [9–11], and external tensile breaking action [12–18].

Today, the wave of artificial intelligence is sweeping the world. Machine learning and deep learning models have become common research hotspots in the fields of artificial intelligence. They have been widely used to solve complex problems in engineering applications and scientific fields with theories and methods. For instance, Son et al., 2021 proposed a deep learning model to locate the damaged cables and conduct the severity assessment of cable-stayed bridges [19]. A machine learning-based approach was developed to detect bridge cable damage subjected to stochastic effects caused by corrosion and fire [20]. Han et al., 2019 applied a deep learning algorithm to distinguish the AE signal from damage signals [21]. Wu and Li, 2022 implemented a method considering both qualitative and quantitative analysis [22]. They employed AE rate process theorem and a machine learning algorithm to evaluate the damage of masonry. However, incorporating machine learning or deep learning algorithms in the signal recognition of broken wires of bridge cables has not been thoroughly investigated. It should be noted that it is not simple to apply machine learning and even deep learning models to analyze acoustic emission signals for bridge monitoring. Training is the first fundamental requirement for machine learning to build useful models to analyze data. However, in acoustic emission signal monitoring and detection of wire breaking, the available samples are often far from sufficient for proper training in machine learning models. This is because such samples are normally generated and measured from testbeds in a controlled laboratory environment. Considering the testbed setup and data measurement, the data generation process will be slow, expensive, and with limited supply. Due to this data issue, the transfer learning technique has been proposed to reduce the required samples for training [23]. However, transfer learning can complicate the modelling process during the application. Therefore, building useful models with limited training data becomes a significant challenge for applying machine learning (deep learning) to the acoustic emission tests of bridges.

Another problem is the feature extraction for acoustic emission signals. We may analyze the acoustic emission signals from time, frequency, and time-frequency domains. Hence, there are many sets of features/parameters that can be extracted for analysis. So far, the feature selection for analysis is often by experience. Generally, only common acoustic emission parameters are analyzed. Some studies are only based on statistical analysis of parameters to distinguish damage. Surely, unsuitable features having been extracted for acoustic emission signals can have an impact on the accuracy of signal identification even when powerful machine learning (deep learning) models are used.

There are many machine learning models available for acoustic emission signal analysis. Some methods such as support vector machine (SVM) and decision trees are easy to be implemented but may only provide simple classification functions. Others such as long short-term memory (LSTM) are powerful with deep learning capability but the modelling process may be more complicated and needs a large number of training samples. Since the acoustic emission analysis may be required for the whole process (i.e., from cable breaking detection to health status estimation), there can be many factors affecting the analysis. Hence, the proposed model should strike a balance between multiple factors to ensure effectiveness and accuracy.

In this paper, we propose to use machine learning models for signal identification of wire breaking in bridge cables. According to the problems listed in the above paragraphs, we first describe our testbed for generating acoustic emission signals in Section 2. Owing to the limited number of signal samples available, the synthetic data approach is proposed to

solve the problem. An algorithm is developed to generate the simulated acoustic emission signals for training machine learning models. From the time, frequency, and time-frequency domains, 22 features being extractable from the acoustic emission signals are described in Section 3. Such a comprehensive features list is used as input to machine learning models. In Section 4, the structure of a deep learning LSTM model is described. LSTM is powerful for waveform analysis and needs a large number of training samples. It can also be used to demonstrate the usefulness of simulated acoustic emission signals being generated by the proposed algorithm. In Section 5, the performance of LSTM and other machine learning models are compared. All machine learning models are trained with the simulated acoustic emission signals. As all machine learning models (including LSTM) have the desired performance, it demonstrates the simulated acoustic emission signals to be effective. Finally, conclusions are given in Section 6.

2. Generating the Acoustic Emission Signals

In the acoustic emission monitoring of bridges, there are roughly three kinds of acoustic signals: environmental noise, acoustic emission signals (i.e., broken wire signals) at the moment of wire breaking, and acoustic emission signals of the wires under certain pulling forces when it is still non-broken (i.e., non-broken wire signals). The environmental noise can be easily filtered out because its spectrum is rather different from that of acoustic emission signals. Acoustic emission signals can be generated from the tension of bridge cables. It often occurs when materials such as metal, carbon fiber, and rock are under external stress. Due to the discontinuity of materials, some micro structures in materials are deformed because of the stress concentration. Within the elastic deformation range, energy is stored in the micro structures. As the stress exceeds the limit of micro structures, the stored deformation energy will be instantly released in the form of elastic waves (or stress waves), i.e., the acoustic emission signal. Using machine learning (deep learning) to monitor and detect the health of the bridge cables, sufficient samples will be needed for both kinds of acoustic signals. In order to obtain the broken and non-broken wire signals of bridge cable, tensile tests were conducted on the cable steel strand under laboratory conditions. Note that the bridge cables may be subjected to intermittent loadings and fatigue loading to the passing vehicles on the bridge when it is in real situations. Such loadings will surely have an impact on the detection of acoustic emission signals. To ensure the accuracy of signal detection, the unwanted signals/noises must be filtered. We have therefore measured four bridges to figure out the characteristic of such unwanted signals/noise for preprocessing the acoustic emission signals detection [24].

2.1. Experiment Setup for Capturing AE Signals

The testbed for the tensile test is shown in Figure 1. The strand selected for the test was the same as bridge cable with single filament epoxy coated pre-stressing strand, i.e., diameter of about 15.24 mm (1×7 strand) and an ultimate tensile strength of 1860 MPa. Three kinds of acoustic emission sensors (WG50, SR150M, and SR40, from www.en.ae-ndt.com (accessed on 15 May 2022)) were used to cover the whole frequency range (15 to 1000 KHz) measurement. The available parameters of these sensors are listed in Table 1. The amplification factor of the preamplifier is 20 dB. In real circumstances, bridge cables may break at random locations. In the experiment, we therefore prefabricated a notch on the steel strand to set the cable break position. We used an electric wheel saw to cut out a single wire with a depth of 2 mm prefabricated notch as shown in Figure 2. The strand was first arranged in the reaction frame, and the acoustic emission sensor was deployed on the surface of the strand as shown in Figure 3. One end of the reaction frame was fixed with pre-stressing clamps and anchorages. The other end was tensioned at a uniform rate by a jack. The tensioning process of the strand continues until wires are broken (see Figure 4). Starting from the pulling force reaching 60% of the ultimate tensile strength, the acoustic emission sensor captures the acoustic emission signals in real time and transmits it to the high-speed acquisition card of a computer. As the upper frequency bound of acoustic emission sensors

is 1 MHz, the sampling clock rate used in the high-speed acquisition card is 3 MHz to ensure all details of the acoustic emission signals can be captured. Two channels are used for simultaneous acquisition. Figure 5 shows the block diagram of the steel strand acoustic emission monitoring system. Figure 6 shows the working interface of computer acquisition software. From the measurement results, the frequency range of typical broken wire signals is 0~300 kHz with peak point spectrum range of 40~60 kHz, i.e., see Figure 7. The frequency range of typical non-broken wire signals is 0~200 kHz with peak point spectrum range of 44~65 kHz, i.e., see Figure 8.



Figure 1. Tensile testbed of steel strand.

Table 1. Parameters of the acoustic emission sensors.

Model	Diameter (mm)	Height (mm)	Frequency Range (kHz)	Center Frequency (kHz)	Temperature (°C)
WG50	19	15	100~1000	500	−20~120
SR150M	19	15	60~400	150	−20~120
SR40M	22	36.8	15~70	40	−20~120



Figure 2. Prefabricate a notch on the steel strand.



Figure 3. Acoustic emission sensors.



Figure 4. Schematic picture of the tensile load application process.

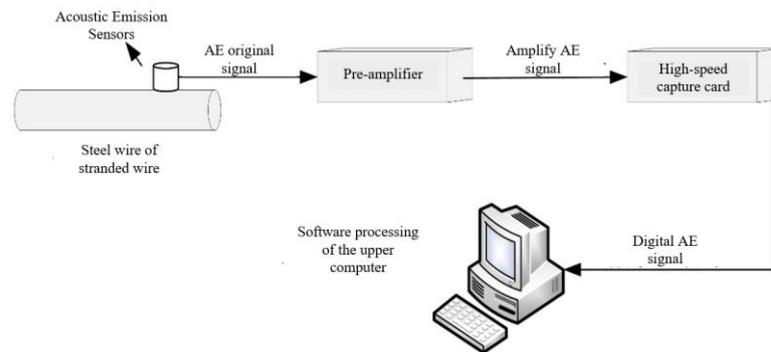


Figure 5. Steel strand acoustic emission monitoring system.

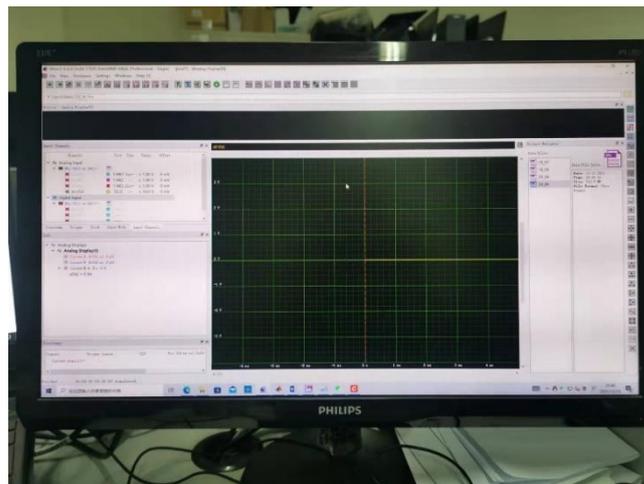


Figure 6. Working interface of upper computer acquisition software.

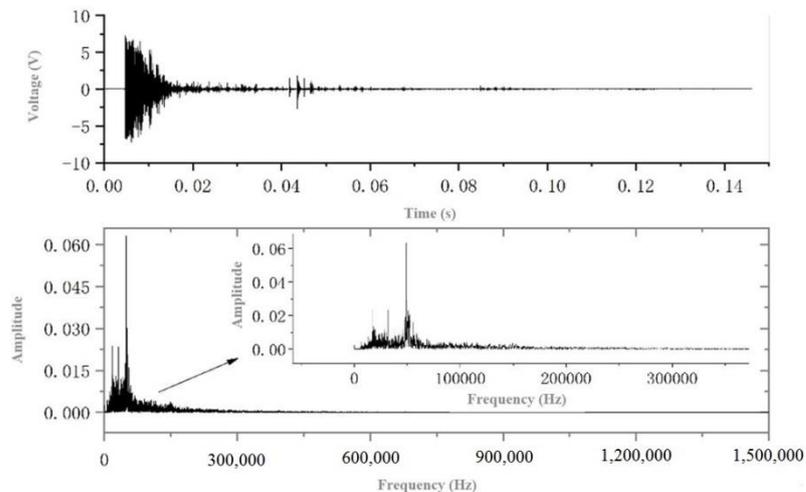


Figure 7. Spectrogram of broken wire acoustic emission signal.

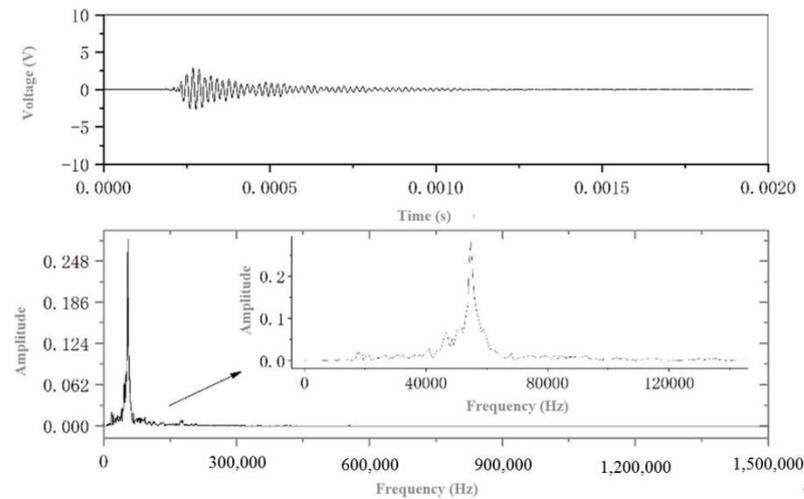


Figure 8. Spectrogram of non-broken wire acoustic emission signal.

It should be noted that a large number of samples may be required for the training of a machine learning (deep learning) model. Owing to the slow generation of acoustic emission signals from the testbed, it will be difficult to have sufficient samples within an acceptable time and cost. At the moment, the total measured samples from the testbed are 249 samples of broken wire signals and 363 samples of non-broken wire signals. Clearly, such a number of samples will not be sufficient for some machine learning models, e.g., deep learning. The required number of training samples varies largely with applications and machine learning methods. From the rule of thumb, however, it is commonly suggested that the required training samples should be at least 10 times the number of model parameters, and may be better up to a hundred times if deep learning methods are used. In Section 3, 22 model parameters are proposed for machine learning. Hence, more than a thousand training samples should be used to satisfy the requirements of all machine learning methods. Although techniques such as transfer learning have been developed to reduce the required samples for training, it complicates the modelling process. In a most simple way, transfer learning first trains machine learning models with other data sets. Then, the models are further trained with measured acoustic emission signal samples [25,26]. It will work well if you already have data sets being similar to the measured acoustic emission signal samples. Usually, the generation of the sample can be rather time-consuming and expensive in the lab environment. In practical measurement, there are also commonly insufficient samples acquired due to the field restrictions and the expensive setup. To simplify the application of machine learning (deep learning) modelling, we therefore propose using a synthetic data approach to solve the problem. With a proper algorithm, we can generate a sufficient number of simulated samples based on the limited number of measured acoustic emission signals from the testbed. For the synthetic data approach being applicable; however, a thorough understanding of the acoustic emission signals is required.

2.2. Generating the Synthetic Data of AE Signals

We generated the synthetic data using the algorithm shown in Figure 9. Using the broken wire signals generation as an example, the following steps are needed:

- (1) Group all the broken wire signal lengths and record their lengths.
- (2) Make a grouping of signals with different lengths. For each *length*, find all the broken wire signals that are close to this *length*, and here the distance tolerance is considered close within 5%. For the signals with length less than *length*, the mean value is added at the end to make up the length, and for the signals with length more than *length*, the excess part at the end of the signal is truncated to achieve the same signal length. However, for the above-mentioned signals that change the original length, *Flag* = 0 is marked, while the signal whose length is not changed is marked *Flag* = 1.

- (3) In a length group, two signals *data1* and *data2* are randomly selected, intercept the oscillation part of the two signals (signal amplitude exceeds 3 standard deviations from the mean) and compare the length of the oscillation part, if they do not agree, continue to randomly select two signals, and if the length is close, continue processing.
- (4) The oscillation part is uniformly divided into *N* segments. A segment is randomly selected to complete the exchange of the corresponding signal segments in the oscillation parts of the two signals *data1* and *data2*. The signal with *Flag* = 1 in the newly generated simulation signal is saved.
- (5) Repeat steps (3) and (4) for each length group until the number of simulated signals in that length group reaches the preset value.
- (6) Repeat steps (2), (3), (4), and (5) for each different length.

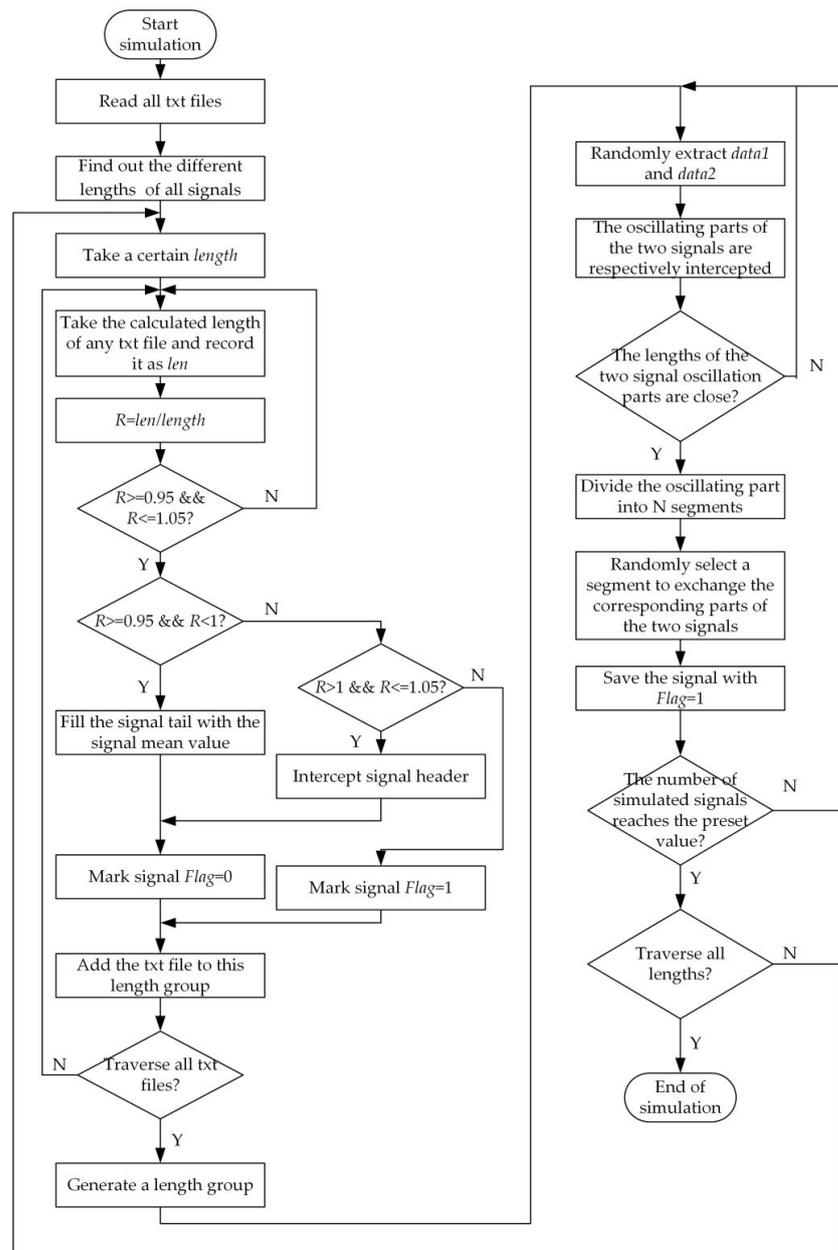


Figure 9. Simulation algorithm.

The idea of the proposed simulated signal generate algorithm is rather simple. Two similar broken (or non-broken) wire signals are selected and segmented. After exchanging a pair of segments at a random position of the signals, two simulated broken (or non-broken)

wire signals are generated with little differences to their originals. The process will repeat until sufficient number of simulated signals are generated. Note that the signal length distribution of the simulated signals will be the same as that of measure signals if all signal pair selections are purely random. Moreover, one can dynamically adjust the signal length distribution of simulated signals by setting proper number of simulated signals in each length group as shown in step 5 of above algorithm list.

After finishing the signal simulation, wavelet packet decomposition is performed on the simulated broken wire signals and the simulated non-broken wire signals. The j th layer wavelet packet decomposition of the signal $S(t)$ can get 2^j sub-bands, and the energy of the k th sub-band in the j th layer can be calculated as:

$$e_{j,k} = \sum_{i=0}^N [d_{j,k}(i)]^2 \tag{1}$$

In the above equation, N denotes the length of the signal and $d_{j,k}$ denotes the wavelet coefficient. The energy share of the k th sub-band of the j th layer can be calculated as:

$$R_k = \frac{e_{j,k}}{\sum_{k=0}^{2^j-1} e_{j,k}} \tag{2}$$

The difference between the simulated signal and the measured signal can be quantified by the Euclidean distance after wavelet packet decomposition. After analyzing the simulated signals, the distance filtering threshold of the simulated signals is set as 20. With the time domain waveform manually screening and Euclidean distance filtering, a total of 832 simulated broken wire signals and 832 simulated non-broken wire signals are obtained. The comparison between the two types of simulated signals and the measured signals is shown in Figures 10 and 11. Note that the simulated signals are rather close to the measured signals. To better utilize the limited number of measured signals, it will be used as the test set of the coming machine learning models. The simulated signals are mainly for the model training purpose.

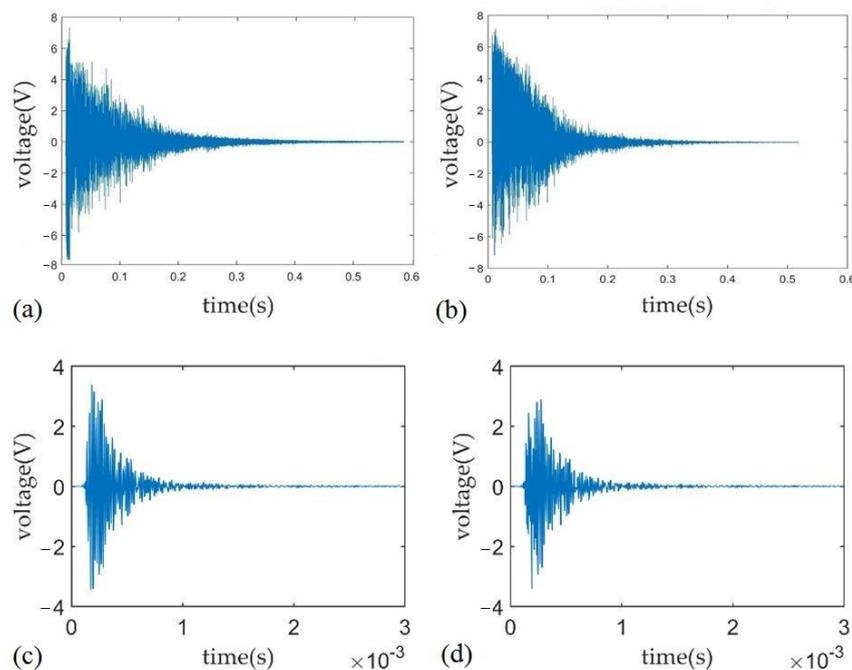


Figure 10. Comparison between simulated signals and measured signals (type 1). (a) Simulated broken wire signal, (b) Broken wire signal, (c) Simulated non-broken wire signal, (d) Non-broken wire signal.

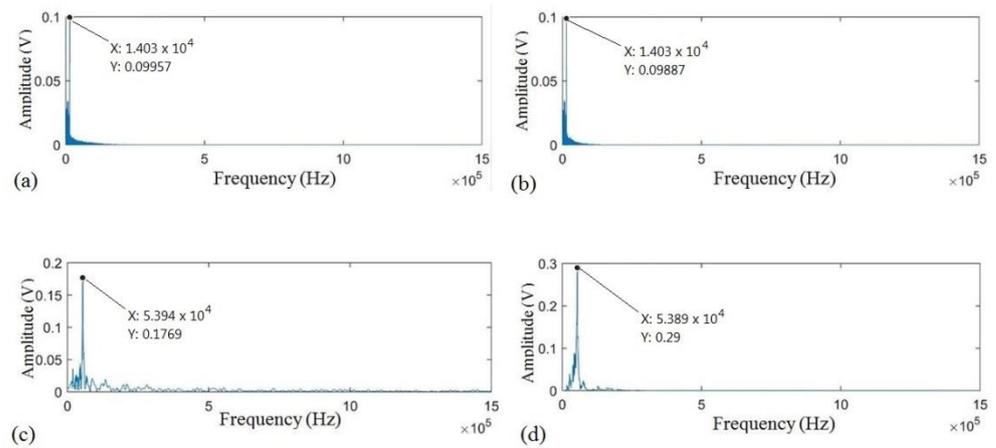


Figure 11. Comparison between simulated signals and measured signals (type 2). (a) Simulated broken wire signal, (b) Broken wire signal, (c) Simulated non-broken wire signal, (d) Non-broken wire signal.

3. Acoustic Emission Signal Feature Extraction

In order to create a comprehensive picture of the signals, this study performed feature extraction of the wire-break and non-wire-break signals from various perspectives, i.e., time domain, frequency domain, and time-frequency domain analysis. The features extracted in the time domain include the acoustic emission parameters of the signal and the statistical parameters of the signal waveform. They are shown in Table 2.

Table 2. Time domain characteristic parameter.

Feature Parameter	Computing Formula (Definition)	Feature Description
Mean square root (V)	$y_1 = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$	Signal RMS value, reflecting the ability of the signal to send power.
Variance (V ²)	$y_2 = \frac{1}{N-1} \sum_{i=1}^{N-1} (x_i - \bar{x})^2$	Characterizes the extent to which the signal deviates from the mean.
Rectification average value (V)	$y_3 = \frac{1}{N} \sum_{i=1}^N x_i $	Indicates the central tendency of the signal oscillation amplitude.
Peak factor	$y_4 = \frac{\max(x_i)}{\sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}}$	Characterizes the sharpness of the signal waveform.
Waveform factor	$y_5 = N \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}}{\sum_{i=1}^N x_i }$	Characterizes the change in signal waveform.
Energy	$y_6 = \sum_{i=1}^N x_i^2$	Reflects the intensity of the acoustic emission event.
Risetime	The length of time between the moment when the acoustic emission signal first crosses the ringing threshold and the moment when the maximum amplitude of the acoustic emission signal is located.	—
Duration	Length of time from the moment when the acoustic emission signal first crosses the ringing threshold to the moment when the acoustic emission signal last crosses the ringing threshold.	—

Table 2. Cont.

Feature Parameter	Computing Formula (Definition)	Feature Description
Ringing count	Each oscillation pulse that crosses the ringing threshold in the acoustic emission signal is called a ringing, and the total number of all oscillation pulses that cross the ringing threshold.	Reflects the strength of the signal.
Peak value	Maximum amplitude of acoustic emission oscillation pulse	—
Descending time	The length of time from the moment when the maximum amplitude of the acoustic emission signal is located to the moment when the acoustic emission signal last crosses the ringing threshold.	—

1. N in the table indicates the length of the signal; 2. x_i in the table indicates the amplitude of the signal at time i .

The frequency domain feature extraction is based on Fast Fourier Transform (FFT). After transforming the discrete acoustic emission signals from the time domain to the frequency domain, frequency domain feature parameters can be easily obtained. The parameters of the frequency domain feature extraction are shown in Table 3.

Table 3. Frequency domain characteristic parameter.

Feature Parameter	Computing Formula (Definition)	Feature Description
Centroid frequency (kHz)	$y_1 = \frac{\sum_{i=1}^{N/2} f_i p_i}{\sum_{i=1}^{N/2} p_i}$	Characterizes the frequency at the center of gravity of the entire spectrum.
Spectral entropy	$y_2 = - \sum_{i=1}^{N/2} \frac{ p_i \frac{2}{N}}{\sum_{i=1}^{N/2} p_i \frac{2}{N}} \log_2 \left(\frac{ p_i \frac{2}{N}}{\sum_{i=1}^{N/2} p_i \frac{2}{N}} \right)$	Characterizes the uncertainty of the spectrum and reflects the spectrum information.
Frequency standard deviation (kHz)	$y_3 = \sqrt{\frac{\sum_{i=1}^{N/2} (f_i - y_1)^2 p_i}{\sum_{i=1}^{N/2} p_i}}$	Describes the fluctuations of spectral energy changes.
RMS frequency (kHz)	$y_4 = \sqrt{\frac{\sum_{i=1}^{N/2} f_i^2 p_i}{\sum_{i=1}^{N/2} p_i}}$	Describes the shift of the main band in the spectrum.
Frequency variance (kHz ²)	$y_5 = \frac{\sum_{i=1}^{N/2} (f_i - y_1)^2 p_i}{\sum_{i=1}^{N/2} p_i}$	Reflects spectral energy distribution.
Mean square frequency (kHz ²)	$y_6 = \frac{\sum_{i=1}^{N/2} f_i^2 p_i}{\sum_{i=1}^{N/2} p_i}$	Reflects the change in the main band in the spectrum.

1. f_i in the table indicates the frequency at i in the frequency spectrum; 2. N in the table indicates the number of FFT points; 3. p_i in the table indicates the absolute value of spectrum amplitude.

The time-frequency analysis method of acoustic emission signal uses continuous wavelet transform. The essence of continuous wavelet transform is the process of wavelet function to process the measured signal under different time domain and frequency domain windows corresponding to different scale factors. When the scale decreases, the time domain window becomes narrower and the frequency domain window becomes wider, so that the high frequency components of the signal can be extracted, while when the scale increases, the time domain window becomes wider and the frequency domain window becomes narrower, so that the low frequency components of the signal can be extracted. Different scales correspond to a filter set, which can realize the filtering of different fre-

quency components, and different translation factors correspond to different positions of the time domain window, which can realize the processing of different time positions of the signal. Therefore, the continuous wavelet transform is an adaptive time-frequency analysis method, which can realize multi-resolution analysis of the signal.

The continuous wavelet transform of the signal is based on the Morlet function and the decomposition scale is 7. The wavelet coefficients of each scale obtained by the continuous wavelet transform are further calculated as the energy share of each scale, and the energy share is used as the wavelet characteristics of the signal. Using the broken wire signal as an example, let the total decomposition scale of the broken wire signal be s , and the wavelet coefficients of each scale decomposition be $coef$, the further energy of the signal at each scale j can be calculated as:

$$E_j = \sum_{i=1}^N |coef_i|^2 \tag{3}$$

where N denotes the signal length, the proportion of signal energy E_j per scale is:

$$R = \frac{E_j}{\sum_{j=1}^s E_j} \tag{4}$$

After the computational analysis, the wavelet features of the broken wire signal and the non-broken wire signal differ significantly on the second, third, fourth, sixth, and seventh scales. These five scales are used as the features extracted from the signal time-frequency analysis.

Based on the above multi-angle feature extraction, a comprehensive feature vector characterizing the acoustic emission signal can be constructed, and this feature vector is the basis for generating samples to establish the broken wire signal recognition model.

4. Long Short-Term Memory (LSTM)

From the discussion in Section 2, it is apparent that there are mainly two kinds of acoustic emission signals, i.e., broken wire signals and non-broken wire signals. Non-broken wire signals are the acoustic emission signals generated from the wires under certain pulling forces, but the wires are still non-broken. Apart from providing alarms by detecting the wire broken signals, a machine learning model may be possible to provide the health status of wires. Determining the health status of wires is similar to estimating the probability of a wire broken signal to happen within a predefined time period. By analyzing currently received signals, clearly, it will be much more complicated than only detecting the wire broken signals. Such applications require more powerful machine learning models and deep learning capability will be needed. In this study, long short-term memory (LSTM) was chosen as a typical powerful machine learning model which also demands a large number of training samples.

LSTM is an artificial neural network often used in the fields of artificial intelligence and deep learning [27]. It is a chain structure containing a large number of repetitive neural network modules, three gates (input, forgetting, and output gates), and the same memory cells as the hidden state. The internal structure of a LSTM cell is shown in Figure 12.

Let the number of cells in the hidden layer be h . Given a small batch of input samples $X_t \in \mathbb{R}^{n \times d}$ at time t , the number of samples n , the number of inputs d , and the hidden state $H_{t-1} \in \mathbb{R}^{n \times h}$ at the previous time $t - 1$, the status of input gate, forgetting gate, and output gate at time t can be calculated as:

$$\begin{cases} I_t = \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i) \\ F_t = \sigma(X_t W_{xf} + H_{t-1} W_{hf} + b_f) \\ O_t = \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o) \end{cases} \tag{5}$$

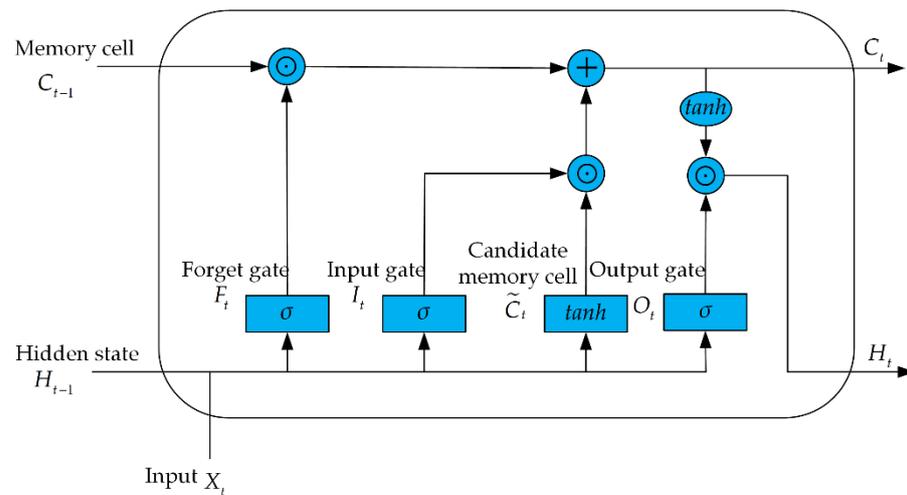


Figure 12. LSTM unit structure.

In the above equation, $W_{xi}, W_{xf}, W_{xo} \in \mathbb{R}^{n \times h}$, W_{hi}, W_{hf} and $W_{ho} \in \mathbb{R}^{h \times h}$ are weights and b_i, b_f and $b_o \in \mathbb{R}^{1 \times h}$ are biases. The candidate memory cell uses a different activation function than the three gates, and the tanh function it uses can obtain an output in the range of $[-1, 1]$. From the above figure, the candidate memory cell output for time t is:

$$\tilde{C}_t = \tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c) \tag{6}$$

In the above equation, $W_{xc} \in \mathbb{R}^{d \times h}$ and $W_{hc} \in \mathbb{R}^{h \times h}$ denote the weights and $b_c \in \mathbb{R}^{1 \times h}$ denotes the bias parameter. The memory cell C_t at the current time t carries the information of the memory cell at the previous time step and the candidate memory cell at the current time step, C_t can be calculated as:

$$C_t = F_t \odot C_{t-1} + I \odot \tilde{C}_t \tag{7}$$

Combined with the above analysis, the main role of the forgetting gate in the figure is to control whether the information in the memory cell C_{t-1} of the previous time step can be passed to the memory cell of the current time step t . The main role of the input gate is to control how the information from the input X_t of the current time step t flows through the candidate memory cell to the memory cell of the current time step. If the output of the forgetting gate is kept as 1 and the output of the input gate is 0, the information in the past memory cells will be passed to the current time step over time, which is similar to a conveyor belt. Such a network design can cope with the gradient decay problem in RNN networks and better capture the dependencies in a time series where the time steps are far away from each other.

The working process of LSTM can be simply understood as follows: given the input value X_t at the current time step, useful information will be filtered through the candidate memory cells under the control of the input gate for the current memory cell update, while the forgetting gate will control whether the information passed from the previous cell flows into the current memory cell, and the two parts of the retained valuable information, i.e., the updated memory, will be passed to the next LSTM cell module. The output gate controls whether the information in the memory cell is passed to the hidden state for use in the output layer, and H_t is also connected to the next LSTM cell module. The interaction and control of the three gates achieves a longer-term memory of the input information.

The structure of the LSTM model built in this paper is shown in Figure 13. For the previously extracted acoustic emission signal features, the main role of the input layer is to import the feature vector of each acoustic emission signal into the LSTM network. The LSTM hidden layer is responsible for further analysis of the feature vector of the input batch samples and passing the valuable information to the fully connected layer, which

is mainly responsible for converting the dimension of the LSTM output vector into the dimension of the model label vector. The fully-connected layer is mainly responsible for converting the dimensionality of the LSTM output vector into the dimensionality of the model label vector, so that the loss function can be calculated. The final Softmax layer is mainly responsible for mapping the category scores' output from the fully connected layer to a positive range, and then normalizing them to (0, 1) to obtain the probability of each category. The category to which each acoustic emission signal sample belongs is finally obtained from the output layer.

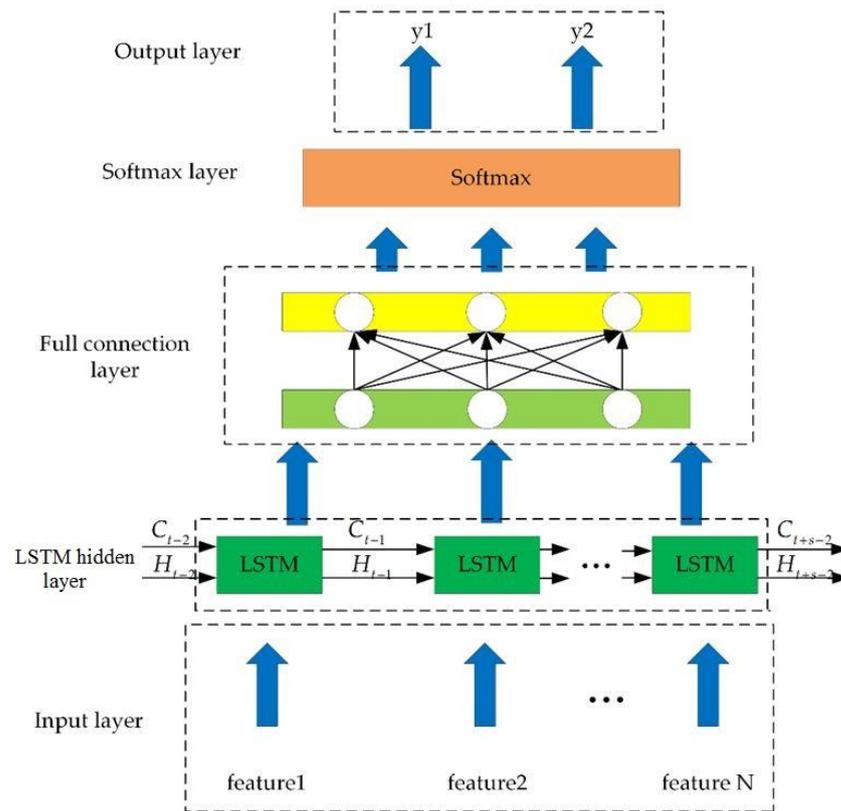


Figure 13. LSTM model structure.

The deep learning model needs to set the relevant parameters of the model before training. According to the previously established LSTM structure, in order to avoid overfitting and increase the number of operations, the hidden layers of LSTM should not be stacked with multiple layers. The model structure parameters are as follows: the input layer has a total of 22 dimensions; the LSTM hidden layer contains 10 network module units, and the activation functions used include sigmoid and tanh functions; a fully connected layer includes 2 neurons for dimensional conversion; the output of the Softmax classification layer is connected to 2 signal categories. By tuning the model parameters, the LSTM model training parameters are shown in Table 4.

Table 4. LSTM model training parameters.

Model Parameter	Parameter Selection Results
Optimization algorithm	Adam
Hardware resources	CPU
Initial learning rate	0.01
Learning rate decline factor	0.0001
Learning rate decline cycle	5
Number of training rounds	10
Batchsize	16

5. Performance Evaluation

Three metrics (precision, recall rate, and F1-score) are used to evaluate the classification ability of the long short-term memory (LSTM) model using the simulated signals for training and measured signals for test set (Table 5). Apart from the LSTM model we discussed in Section 3, a set of machine learning models are used to demonstrate the feasibility of using the proposed synthetic data approach to solve the problem of insufficient training samples. They are support vector machine (SVM), particle swarm optimized support vector machine (PSO-SVM), multilayer perceptron, k-nearest neighbors (KNN), decision tree, and Naive Bayes models. The precision performance between machine learning models with simulated signals for training and measured signals for test sets are also compared. To further compare the performance between the machine learning models, they are used to classify the simulated signals with noise of SNR 30dB.

Table 5. Confusion matrix.

	Prediction of Broken Wire Signal (Positive)	Prediction of non-Broken Wire Signal (Negative)
Measure broken wire signal (positive)	TP (true positive)	FN (false negative)
Measured non-broken wire signal (negative)	FP (false positive)	TN (true negative)

Precision (i.e., recognition accuracy) is the most common performance index in verifying machine learning models. It is defined as

$$precision = \frac{TP}{TP + FP} \tag{8}$$

and that of the non-broken wire signal is

$$precision = \frac{TN}{TN + FN} \tag{9}$$

The definition of *TP*, *TN*, *FP*, and *FN* are listed in Table 5.

The recall rate, also known as the check-all rate, is a measure of coverage, and can measure how many broken/non-broken wire signals are identified accurately, using the broken wire signal as an example. It is calculated as follows:

$$R = \frac{TP}{TP + FN'} \tag{10}$$

and that of the non-broken wire signals is

$$R = \frac{TN}{TN + FN}. \tag{11}$$

The F-Score can be calculated by combining the precision *P* and recall rate *R* metrics as follows

$$F = \frac{(\alpha^2 + 1)P * R}{\alpha^2(P + R)}. \tag{12}$$

When $\alpha = 1$, it is the F1-score index. Owing to the limited number of measured signals (249 and 363 measured samples of broken and non-broken wire signals, respectively), the algorithm in Figure 9 has been used to generate simulated broken and non-broken wire signals (each having 832 samples) to train the LSTM model described in Section 3 and the measured signals are used for test set. Table 6 shows the performance of LSTM on detecting both broken-wire and non-broken wire signals under such arrangement. One can observe that the LSTM model has rather good performance with the three metrics shown in Table 6.

Table 6. The performance of LSTM (simulated signal for training, and measure signals for test set).

Metrics	Broken Wire Signals	Non-Broken Wire Signals
Recognition accuracy	98.02%	99.44%
Call rate	99.2%	98.62%
F1-score	0.986	0.9903

To further verify the effectiveness of the sample generation algorithm proposed in Section 2, the performance differences between machine learning models were compared. Table 7 shows the comparison of performance of several machine learning models using the simulated signals for training and measured signals for test set. From Table 7, one can observe that all machine learning models have similar good performance under such arrangement. Hence, we arranged another experiment such that simulated signals with Gaussian white noise (SNR = 30 dB) were used as the test set. Table 8 shows the performance of the machine learning models under such arrangement. From Table 8, one can observe that some machine learning models such as SVM and decision trees will have large performance degradation in such a situation. It implies that more factors should be considered when choosing a machine learning model, e.g., the required number of training samples and performance under different situations. Table 9 shows the F1-score performance of the machine learning models under both test set arrangement. One can observe similar phenomena. Nevertheless, all machine learning models will have good performance in the tests only if training samples are appropriate and their number is sufficiently large. It demonstrates the usefulness of the proposed simulated signal generation algorithm. A limitation of the investigated case is that cross validation is not performed due to a limited number of data samples. Future studies should incorporate cross validation when using the proposed approach. A larger number of data samples are encouraged for the variation.

Table 7. Recognition accuracy performance (simulated signal for training, and measure signals for test set).

Model	Recognition Accuracy of Broken Wire Signal
LSTM	99.20%
SVM	97.99%
PSO-SVM	98.39%
Multilayer perceptron	98.39%
KNN	98.39%
Decision tree	99.20%
Naive Bayes	99.20%

Table 8. Recognition accuracy performance (simulated signal for training, and simulated signals with Gaussian white noise (SNR = 30 dB) for test set).

Model	Recognition Accuracy of Broken Wire Signal
LSTM	100%
SVM	94.23%
PSO-SVM	98.92%
Multilayer perceptron	99.40%
KNN	98.08%
Decision tree	93.99%
Naive Bayes	98.68%

Table 9. F1-score of broken wire signals (simulated signal for training, and simulated signals with Gaussian white noise (SNR = 30dB) for test set).

Model	F1-Score	F1-Score (SNR = 30dB)
LSTM	0.9861	0.9846
SVM	0.9799	0.9584
PSO-SVM	0.9819	0.9827
Multilayer perceptron	0.9819	0.9833
KNN	0.9819	0.9784
Decision tree	0.9802	0.9572
Naive Bayes	0.9763	0.9780

6. Conclusions

Acoustic emission (AE) is a dynamic nondestructive testing method that is increasingly used in the local monitoring of bridge cables. In this paper, a testbed is described for generating the acoustic emission signals for signal identification testing with machine learning models. Owing to the limited number of measured signals being available, an algorithm is proposed to simulate acoustic emission signals for model training. A multi-angle feature extraction method was used to extract the acoustic emission signals and construct a comprehensive feature vector to characterize the acoustic emission signals. Seven ML models were trained with the simulated acoustic emission signals. As all machine learning models (including LSTM) provide desired performance, it shows the approach of simulated acoustic emission signals to be favorable. A limitation of the study is that the model has not been applied to real field data. Future studies are encouraged to test the model in practice.

Another limitation of the proposed simulated signal generation algorithm is that it relies too much on professional expertise. One example is the hand-crafted statistical features. Using powerful machine learning models such as LSTM may help us to extract overlooked features and consolidate the required model parameters. Consequently, it will reduce the model training time and improve the performance of machine learning models. To obtain the proper feature extraction, however, a sufficient large number of raw samples are required. As the raw sample generation is costly and only a limited number of samples (249 broken wire signals and 363 non-broken wire signals) have been recorded, it is difficult to currently carry out serious work on feature extraction. Nevertheless, we have planned to refine the proposed simulated signal generation algorithm.

Author Contributions: Conceptualization, G.L., H.D., Y.L., C.-Y.L. and C.-C.L.; methodology, G.L. and H.D.; software, G.L. and H.D.; validation, G.L. and H.D.; formal analysis, G.L. and H.D.; investigation, G.L. and H.D.; resources, G.L. and H.D.; data curation, G.L. and H.D.; writing—original draft preparation, G.L. and H.D.; writing—review and editing, G.L., H.D., Y.L., C.-Y.L. and C.-C.L.; visualization, G.L. and H.D.; supervision, G.L. and C.-C.L.; project administration, G.L. and C.-C.L.; funding acquisition, G.L. and C.-C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Coordinative Innovation Plan of Guangdong Province, Weihai Science and Technology Development Plan by the National Natural Science Foundation of China, grant number 41904158, the Technology Developing Project of Shenzhen, grant number Key 20180126, the China Postdoctoral Science Foundation, grant number 2019M652385, Shandong post-doctoral innovation project, grant number key 202002004, the Young Scholars Program of Shandong University, Weihai (20820201005), the Hong Kong SAR, RGC Faculty Development Scheme (Project No. UGC/FDS16/E04/21); the RGC Research Matching Grant Scheme (Project No. 2021/3008); and HKMU Faculty Advancement Fund.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: We would like to thank the editors and the anonymous reviewers for their insightful comments and constructive suggestions.

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

References

1. Dong, Z.S.; Xian, Y.F. Analysis of integrated technology of intelligent health monitoring system for large bridge structures. *Technol. Innov. Appl.* **2020**, *12*, 146–147.
2. Wang, S.N.; Sun, F.J.; Zhao, W.X.; Cheng, J.T. Exploration of intelligent monitoring of highway bridges. *Eng. Technol. Res.* **2019**, *4*, 100–101. (In Chinese)
3. Editorial Board of the Chinese Journal of Highways. A review of academic research on bridge engineering in China 2021. *Chin. J. Highw.* **2021**, *34*, 1–97.
4. Tong, G.W.; Zhou, X.D.; Huang, L.Y.; Chen, C.Y.; Xu, H.W.; Yang, L. Acoustic emission detection signal analysis and source localization method research. *Instrum. Technol. Sens.* **2021**, *5*, 96–100. (In Chinese)
5. Li, D.S.; Yang, W.; Zhang, W.Y. Cluster analysis of stress corrosion mechanisms for steel wires used in bridge cables through acoustic emission particle swarm optimization. *Ultrasonics* **2017**, *77*, 22–31. [[CrossRef](#)]
6. Hu, P.Y.; Li, S.L.; Jiang, N.; Yan, Y. Investigation of the impressed current cathodic protection method for the cable parallel wires in the rainwater electrolyte based on acoustic emission method. *Constr. Build. Mater.* **2019**, *229*, 116918. [[CrossRef](#)]
7. Li, D.S.; Tan, M.L.; Zhang, S.F.; Ou, J. Stress corrosion damage evolution analysis and mechanism identification for prestressed steel strands using acoustic emission technique. *Struct. Control. Health Monit.* **2018**, *25*, e2189. [[CrossRef](#)]
8. Huang, H.; Wu, Y.H.; Xu, P.; Chen, X. Corrosion assessment of high-strength steel wires based on guided-wave modal analysis. *Sci. Technol. Eng.* **2020**, *20*, 11332–11338.
9. Drummond, G.; Watson, J.F.; Acarnley, P.P. Acoustic emission from wire ropes during proof load and fatigue testing. *NDT E Int.* **2006**, *40*, 94–101. [[CrossRef](#)]
10. Li, H.; Huang, Y.; Ou, J.P. Estimation and warning of fatigue damage of FRP stay cables based on acoustic emission techniques and fractal theory. *Comput.-Aided Civ. Infrastruct. Eng.* **2011**, *26*, 500–512. [[CrossRef](#)]
11. Li, D.S.; Hu, Q.; Ou, J.P. Fatigue damage evolution and monitoring of carbon fiber reinforced polymer bridge cable by acoustic emission technique. *Int. J. Distrib. Sens. Netw.* **2013**, *8*, 282139. [[CrossRef](#)]
12. Qu, H.Y.; Li, T.T.; Chen, G.D. Bridge cable fracture detection with acoustic emission test. In Proceedings of the NDE/NDT for Highways & Bridges: SMT 2016, Portland, OR, USA, 29 August 2016; pp. 38–41.
13. Zhang, W.Y.; Yang, W.; Zhang, W.; Wu, N.; Li, D.S. Xiangyou Mulanxi bridge suspender damage acoustic emission monitoring. *Fujian Constr. Technol.* **2014**, *5*, 21–24. (In Chinese)
14. Ma, M.L.; Li, H.; Chen, W.L.; Huang, Y.; Han, W. Acoustic emission characters of glass fibre reinforced polymer stay cable. *J. Comput. Theor. Nanosci.* **2012**, *9*, 1357–1363. [[CrossRef](#)]
15. Deng, Y.; Ding, Y.L.; Li, A.Q. Experimental study on damage evolution of steel strand based on acoustic emission rate process theory. *J. Southeast Univ. Nat. Sci. Ed.* **2010**, *40*, 1238–1242.
16. Li, D.S.; Ou, J.P. Acoustic emission characteristics and damage evolution model of steel strands in tensile test. *J. Highw. Transp. Res. Dev. Engl. Ed.* **2008**, *3*, 1238–1242. [[CrossRef](#)]
17. Cao, L.L.; Li, A.Q.; Deng, Y.; Ding, Y.L. Application of acoustic emission and wavelet packet analysis in damage condition monitoring. *Vib. Test. Diagn.* **2012**, *32*, 591–595, 688–689. (In Chinese)
18. Nair, A.; Cai, C.S. Acoustic monitoring of bridges: Review and case studies. *Eng. Struct.* **2010**, *32*, 1704–1714. [[CrossRef](#)]
19. Son, H.; Pham, V.T.; Jang, Y.; Kim, S.E. Damage localization and severity assessment of a cable-stayed bridge using a message passing neural network. *Sensors* **2021**, *21*, 3118. [[CrossRef](#)]
20. Feng, J.; Gao, K.; Gao, W.; Liao, Y.; Wu, G. Machine learning-based bridge cable damage detection under stochastic effects of corrosion and fire. *Eng. Struct.* **2022**, *264*, 114421. [[CrossRef](#)]
21. Han, G.; Oh, T.M.; Kim, H.; Song, K.I.; Kim, Y.; Kwon, T.H. Determination of Crack Signals Using the Deep Learning Technique Based on a 1D Convolutional Neural Network for Smart Detection of Structural Damage Cracking. *J. Korean Soc. Hazard Mitig.* **2019**, *19*, 1–7. [[CrossRef](#)]
22. Wu, Y.; Li, S. Damage degree evaluation of masonry using optimized SVM-based acoustic emission monitoring and rate process theory. *Measurement* **2022**, *190*, 110729. [[CrossRef](#)]
23. Xin, H.; Cheng, L.; Diender, R.; Veljkovic, M. Fracture acoustic emission signals identification of stay cables in bridge engineering application using deep transfer learning and wavelet analysis. *Adv. Bridge Eng.* **2020**, *1*, 6. [[CrossRef](#)]
24. Li, G.; Zhao, Z.; Li, Y.; Li, C.-Y.; Lee, C.-C. Preprocessing Acoustic Emission Signal of Broken Wires in Bridge Cables. *Appl. Sci.* **2022**, *12*, 6727. [[CrossRef](#)]
25. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A Comprehensive Survey on Transfer Learning. *Proc. IEEE* **2021**, *109*, 43–76. [[CrossRef](#)]
26. Lu, J.; Behbood, V.; Hao, P.; Zuo, H.; Xue, S.; Zhang, G. Transfer learning using computational intelligence: A survey. *Knowl.-Based Syst.* **2015**, *80*, 14–23. [[CrossRef](#)]
27. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]