Acoustic Emission Signal Denoising of Bridge Structures using SOM Neural Network Machine Learning

Aiping Yu[1], Xiangtai Liu[2] Feng Fu[3] C.Eng, F.ASCE.， Xuandong Chen [4], Yan Zhang[5]

ABSTRACT

Identification Noise signal is one of the challenging problems in the health monitoring of bridge structure using acoustic emission monitoring and identification technology. Hardware filtering technology and spatial identification technologies are the most common method in identifying of the signals from the defect of the bridge, which have great limitations due to the presence of environmental noise. Therefore, this paper focus on the AE noise signal from a bridge in operation state and other specific loading state, which is diagnosed in the hardware filtering technology, spatial identification and SOM neural network, to obtain the new noise recognition methods. It is found that the first two methods can indeed filter the noise signal, but the filtering rate can only reach about 50 %, and can barely filter strong noise signal. The SOM neural network had strong self-recognition ability. The classification accuracy of simulated AE signals is 90 % and 100 % respectively. The trained network is used to test183 sample signals, the defect signal detection accuracy reaches 76 % and 78.8 %, therefore, the noise signal filtering effect is significantly improved.

Keywords: Noise, SOM neural network, Wavelet packet energy analysis, Wavelet packet entropy analysis

1. Aiping Yu [1], Professor，College of Civil and Architectural Engineering, Guilin University of Technology, Guilin 541004. apyu@glut.edu.cn；

2. Xiangtai Liu [1], Master Degree Candidate，College of Civil and Architectural Engineering, Guilin University of Technology, Guilin 541004. 1271229848@qq.com;

3. Feng Fu, Associate Professor -Department of Civil Engineering, School of Mathematics, Computer Science & Engineering, City, University of London, Northampton Square, London, EC1V 0HB,U.K. Corresponding author), feng.fu.1@city.ac.uk

4.Xuandong Chen lecturer，College of Civil and Architectural Engineering, Guilin University of Technology, Guilin 541004. chenxuandong@glut.edu.cn;

5. Yan Zhang Professor， College of Civil and Architectural Engineering, Guilin University of Technology, Guilin 541004. zhangyan@glut.edu.cn;

1. Introduction

Acoustic emission signal processing is a passive, non-destructive and real-time dynamic detection technique (Dunegan 1969), and there is no strict requirement for the size and detection range of the detected object. Moreover, non-contact detection can be achieved, such as detection in high temperature and corrosion that cannot be approached by human beings.

This technology has been widely used in the damage detection of homogeneous materials, such as damages in mechanical bearing (Nguyen et al. 2018), metal tool (Rmili et al. 2016; Bhuiyan et al. 2014), metal container corrosion detection (Li et al. 2015), aeronautical metal materials (Holford et al. 2017), metal track detection (Zhang et al. 2017), glass (Njuhovic et al. 2014), carbon fiber (McCrory et al. 2015) and rock (Liu et al. 2020). However, there is little application in reinforced concrete materials (Colombo et al. 2003; Lamonaca et al. 2012; Noorsuhada 2016). The main reason is the complexity and diversity of environmental noise signals causing ambiguities when identify the noise signals and structural damage signals, which makes it difficult to effectively identify and characterize damage signals.

In the current practical engineering, most noises signal are filtered by hardware filtering technology and spatial identification (Noorsuhada et al. 2011; Kalyanasundaram et al. 2007; Li and Ou 2007), which has a positive effect on the identification of damage signals, but it has great limitations and obvious defects for more complex environmental noise.

Therefore, scholars began to extract and identify the key characteristic parameters of damage signal and noise signal (Dijck et al. 2009; Kacimi and Laurens 2009; Surgeon and Wevers 1999; Fu et al. 2011; Bianchi et al. 2015; Velayudham et al. 2005; Deng et al. 2009). It has been found that the extraction of signal characteristic parameters can better identify and eliminate noise signals.

Artificial intelligence technology is one of the means of signal recognition. Its algorithms are mainly divided into traditional machine learning algorithms (Fu,2020,2018) and neural network algorithms (Cai et al , 2020,2019), such as support vector machine Noble (2006), Gaussian process regression (HU et al. 2010, Wang et al.(2022), long short-term memory network Moon et al.(2022), and so on. In recent years, with the rapid development of neural network algorithms, researchers have studied the identification of damage signals and noise signals through artificial neural networks (Zafar et al 2017, Ekici et al ,2008) based on spectrum, energy and entropy of wavelet packets. With the development of artificial neural network in signal processing, the way of distinguishing based on the neural network of wavelet packet energy and entropy has been addressed by various researchers in the other areas of signal processing such as, track defects detection, power systems, mechanical engineering. In the field of mechanical engineering, (Luo et al.) extract the energy eigenvector of the signal failure die using the wavelet packet analysis technique, and the energy percentage is taken as the characteristic parameters. Then a BP neural network recognition model was established. The BP neural network recognition model can quickly identify new sample data with an accuracy rate of 95 %. This new technology enabled an more accurate identification method of acoustic emission signals and assessing the degree of structural or material damage

From previous studies, it can also be found that it is feasible for some researchers to try to identify signals and remove noise signals through parameter analysis and neural networks. However, most of the existing research focuses on homogeneous materials, and concrete materials are multiphase heterogeneous materials, which makes the damage signal become complex and diverse in the process of propagation, which will lead to differences in signal recognition between concrete materials and homogeneous materials. Therefore, this has not been investigated in the past for background noise reduction of concrete structures such as building structures and bridges.

Therefore, in this paper, the AE signals of the bridges in operation under certain specific loading state were tested in this experiment. Based on the signal of hardware filtering technology and spatial identification technology, as well as the wavelet packet energy analysis (Guo et al. 2020, 2021) and wavelet packet entropy analysis (Safty and El-Zonkoly 2008; Yin et al. 2004) the characteristic frequency bands were extracted from simulated acoustic emission signal and noise signal. Finally, the signal was clustered by using Konhonen's self-organizing feature map and neural network (SOM neural network) (Kohonen 1998) to establish an acoustic emission detection and recognition algorithm, which provided new ideas and methods for solving the noise reduction problem of bridge acoustic emission damage signal. It is expected to solve the difficulty to effectively identify and characterize structural damage due to the confusion between noise signals and structural damage signals.

2. **Test set up**

2.1 Introduction of the prototype bridge

The prototype bridge is a single flyover, and the superstructure is a four-span simply supported hollow beam bridge with a span of 10 m + 20 m + 20 m + 10 m. The main beam with a span of 20 m is a prestressed reinforced concrete hollow beam, and the main beam with a span of 10 m is a common reinforced concrete hollow beam. The net width of the bridge deck is 11m + 2×0.5m (anti-collision wall), and the bridge design grade is grade I.

2.2 Detection scheme

The signal acquisition instrument is full digital Sensor Highway III (SH-III) acoustic emission device manufactured by American Physical Acoustics (PAC) as shown in Fig. 1. Acoustic emission parameter settings are shown in Table 1. Two 360kN truck are used for loading test machinery.

2.3 sensor installation

In order to facilitate the monitoring of structural stress changes and considering the convenience of sensor installation and arrangement. The monitoring position chosen are the most unfavorable position of shear force when the bridge is under load (the 2-2 section of the middle span of the 4th span). Fig. 2 shows the overall elevation of the bridge. When the sensor is installed, the concrete surface is smoothed by grinding the bottom of the bridge, and the Vaseline glue is uniformly coated on the sensor to stabilized sensors at the bottom of the bridge. Before monitoring, the Pencil Lead Break Testing (PLBT)(Lopes et al. 2018) was used to identify whether the sensor coupling was good.

2.4 Simulated Acoustic Emission Signal and Noise Source

Simulated AE signal is produced using Pencil Lead Break Testing. To perform the tests, a mechanical pencil with 3 mm of length and 0.5 mm of diameter were mounted on the bridge which maintained 30°angle between the pencil and the bridge surface. The noise source mainly monitors the sound source of static load condition, driving condition, jumping condition and braking condition. The environment of various noise sources is as follows :

1) Signal under static load condition : AE signals generated by the tiny vertical vibration of the bridge under environmental vibration are collected under conditions such as no driving.

2) Signal in driving condition : the test vehicle passes the bridge at the speeds of 10 km / h, 20 km / h and 30 km / h to collect the AE signals generated by the vertical vibration and friction of the fourth span bridge.

3) Signal of vehicles bumping condition : a single vehicle with 360 kN self- weight was used to simulate the bad state of vehicle bumping by let it pass a wedge with a height of 10 cm in the middle of the fourth span bridge, and the AE signals generated by vertical vibration and friction under the bad state of vehicle load were collected.

4) Signal of vehicles braking condition : an emergency braking test was conducted with a single 360 kN vehicle at the middle of the fourth span bridge to collect AE signals generated by vibration and friction of the bridge.

3 Filtering and Spatial Identification of Noise Signal

3.1 Data filtering

Select the appropriate filter in acoustic emission system, that is, select the appropriate 'window' to suppress noise. At present, the most commonly used engineering is to set the appropriate amplitude threshold, the noise below the threshold will be isolated by the detection system.

In the static load condition detection, a large number of high frequency continuous noise signals are collected when the amplitude threshold is set to 20 dB, As shown in Fig. 3. It can be clearly seen that the amplitude distribution of the interference noise signal is mainly concentrated below 43 dB, and only a few amplitude reaches more than 43 dB. In the subsequent data processing, the amplitude threshold is increased to 43dB, almost filtering out the interference noise signal generated by the environment.

3.2 Spatial Identification

Spatial identification technology is to place two types of sensors (monitoring sensors and guard sensors) at different locations to eliminate noise signals. In the test, the main sensor (No. 1-7 sensor) is placed at the bottom of the test beam, and the guard sensor (No. 8-9 sensor) is placed at the side of the tested area to shield the interference noise signal produced when driving in both direction. The three-dimensional schematic layout of the sensor is shown in Fig. 4. The noise signal positioning maps before and after setting the guard sensor are shown in Fig. 5 (a) and Fig. 5 (b).

It can be seen from Fig. 5 (a) that the number of events is 32, indicating the randomness of the noise signal. The number of events in Fig. 5 (b) is 16. The latter eliminates 50% noise and improves the signal-to-noise ratio. It shows that the system can shield the noise signal after setting the guard sensor.

In summary, the use of data filtering and spatial identification technology can effectively eliminate the interference of noise. However, the characteristics parameter of some strong interference noise signals is intertwined with the characteristics parameter of simulated AE signals, which cannot accurately filter out the noise signals. Therefore, it is necessary to further use SOM neural network to cluster the signals and establish an effective pattern recognition method.

4 Wavelet packet based feature extraction

4.1 Basic principles

wavelet packet transform (WPT) has become one of the most widely used signal analysis methods because of its multi-resolution and ability to characterize the local characteristics of signals in both time domain and frequency domain. Wavelet entropy analysis is a new method to measure the complex sequence of the signal, which is the combination of wavelet transform and information entropy. In signal processing, it not only has the advantages of changeable resolution and no signal -stationary requirements, but also can statistically analyze the complexity of entropy on signal, which can be used to detect the local characteristics of non-stationary signals.

Wavelet packet decomposition methods first decomposes the input signal into high frequency and low frequency dataset through orthogonal wavelet bases, and then decomposes the two datasets of the signal to obtain the next high and low frequency datasets. In the process interaction, the scale function φ (t) and the wavelet function ψ (t) satisfy the below equation:

$$\varphi(t) = \sqrt{2}\sum_{k} \mathrm{h}(k)\varphi(2t-k) \tag{1}$$

$$\psi(t) = \sqrt{2}\sum_{k} g(k)\varphi(2t-k) \tag{2}$$

Where: k is the translation amplitude. h (k) is the low-pass filter corresponding to the scale function φ (t). g (k) is a high-pass filter corresponding to the wavelet function ψ (t).

After the original signal is decomposed by i-layer wavelet packet, the characteristic signal composed of 2i frequency bands from low frequency to high frequency in the i-layer is obtained. The decomposed wavelet packet coefficients are reconstructed to extract the signals in each frequency band.

The energy E corresponding to the j-band signal in layer i is :

$$E_{\mathrm{i,j}} = \int \left|S_{i,j}(t)\right|^2 dt = \sum_{k=1}^{n} \left|x_{i,j}(k)\right|^2 \tag{3}$$

Where , $x_{i,j}$ (k) denotes the amplitude of the discrete points of the reconstructed signal $S_{i,j}$, and n is the number of discrete points of the reconstructed signal.

Therefore, the total energy of the whole signal is :

$$E = \sum_{j=0}^{2^i-1} E_{i,j} \tag{4}$$

Wavelet packet energy coefficient:

$$P_{i,j} = \frac{E_{i,j}}{E} \tag{5}$$

According to the basic theory of information entropy, the wavelet packet characteristic entropy is defined as :

$$H_{i,j} = -\sum_{j=0}^{n} P_{i,j} \log P_{i,j} \tag{6}$$

4.2 Wavelet packet decomposition

The wavelet energy and wavelet entropy are used to study the characteristics of AE signals. The sampling frequency of the AE signal is 1MHz. According to the sampling theorem, the Nyquist frequency is 512kHz, the wavelet basis function and the number of wavelet packet decomposition layers can determine the optimal solution according to the norm entropy (Lp). Thus, the db6 wavelet basis function is selected to decompose the AE signal into five levels of wavelet packet. The signal is decomposed into 25 sub-bands, and each band width is 16kHz. Therefore, this decomposition basically meets the requirements of acoustic emission time domain waveform signal frequency band division.

According to the relevant literature(Wen 2015), the frequency of AE signal and noise signal in concrete is mostly less than 150 KHz, so this paper extracts the first 16 frequency components of the fifth layer from low frequency to high frequency, which can basically reflect the characteristics of each signal. The frequency ranges are shown in Table 2.

4.3 Energy Analysis of Wavelet packet Coefficient

The results of calculating the characteristic energy of each decomposed signal in each frequency band are shown in Fig. 6, and the energy proportion of each signal is shown in Table 3. It can be seen that the frequency components of simulated AE signals with different propagation distances are quite different. The frequency band range of the sound source is wide, and both high and low frequency information exist. The energy is mainly concentrated in the 6 ~ 8 high frequency band, accounting for 72.2 % of the total energy. The high frequency energy components such as 14 ~ 16 account for 10.4 % of the total energy, and the 2 ~ 4 band accounts for 14.2 % of the total energy. With the propagation of signals, the high-frequency components of 6th, 7th and 8th continuously decay. When propagating to 1.2 m, the energy is mainly concentrated in the first, second and fourth bands, accounting for 87.4 % of the total energy. The energy of noise in other working conditions is mainly concentrated in the 1st, 2nd and 4th low which is similar to the energy of simulated AE signal at 1.2 m and can be greater than 87.4 %. So the noise signals can be distinguished by the main frequency band distribution.

4.4 Entropy Analysis of Wavelet packet Coefficient

The wavelet entropy is calculated by using the energy of decomposed signals in each frequency band, and the results are shown in Fig. 7. The proportion of entropy of each signal is shown in Table 4. It can be seen that, the above series of signal wavelet entropy coefficient wavelet energy coefficient follows the same pattern. The entropy of the sound source is mainly concentrated in the 6th-8th high frequency band, accounting for 79.1 %. As the propagation distance of the simulated AE signal increases, when it reaches 1.2 m, the entropy gradually concentrates in the first, second and fourth low frequency bands. The entropy of the noise signal in the other conditions is similar to that of the simulated AE signal at 1.2 m, which is also concentrated in the first, second and fourth low frequency bands, accounting for more than 84.4 %.

5 SOM Neural Network Machine learning

5.1 SOM Clustering Principle

SOM is a competitive artificial neural network with self-organization, self-learning and lateral association ability proposed by Professor Kohonen of Helsinki University of Technology in 1981. Firstly, it is a single-layer neural network composed of input layer and output layer and realizes the orderly mapping of high-dimensional data distribution to regular shape low-dimensional grids (generally two-dimensional). Secondly, the output layer is a one-dimensional or two-dimensional regular lattice grid composed of logical units, and there is a short-range lateral feedback between each unit in a certain neighborhood, and the feedback intensity varies with the distance. Therefore, the adjacent neurons stimulate to each other, while the slightly distant neurons inhibit from each other, and the farther neurons have a weaker incentive effect. Finally, through competitive learning, the input vector continuously adjusts the connection weight to make it closer to a certain type of input vector. The final similar input vectors can be clustered at a node and separated from the dissimilar input vectors. The signal recognition process using SOM network structure is shown in Fig. 8.

5.2 SOM neural network training

5.2.1Standard Sample Design

The acoustic emission signals collected in test of flyover are clustered. 20 groups of simulated AE signals and 100 groups of noise signals are selected as standard sample set input, in total of 120 groups of sample feature vectors. Training input mode is:

$$P_k = (P_1^k, P_2^k, \cdots, P_{16}^k) \qquad \mathrm{k} = 1, 2, \cdots, 120$$

5.2.2 Network structure design

1) Number of neurons in input layer

The simulated AE signal and noise signal are concentrated below 300 kHz. The wavelet characteristic parameters of 0 ~ 16 band (0 ~ 256 kHz) can fully reflect the characteristics of each band of the signal, and the data is rich and reliable. So the number of input layer neurons in SOM neural network (N)is 16.

2) Number of neurons in competition layer

The selection of the number of neurons in the competition layer will affect the performance of the network. If the number is too large, it will increase the amount of calculation and reduce the learning speed of the network. If the number is too small, it is possible to misclassify signals of different modes. The sample size here is small, and according to the parameter recommendation of literature (Silva et al 2019, Khanzadeh,2018,) supplemented by the clustering performance observation of network structure adjustment. Finally, the structure with a competitive layer of 6×6 is selected, which can achieve better clustering results. So, the number of competition layer neurons in SOM neural network (M) is 36.

3) Determine function, select learning efficiency, set training steps

Create a SOM neural network using the NEWSOM function The code is : net = NEWSOM (minmax (P),[6 6]). Among them, minmax (P) specifies the maximum and minimum values of input vector elements, and 6×6 denotes the structure of 6×6 competition layer of the network. The network is trained and simulated by function train and simulation function sim. The size of learning rate [35] and the number of training steps will affect the clustering performance of the network (training time and convergence rate). When the learning rate (η(t))is 1.0 and the number of steps(T) is 500, it is clear to observe the clustering results.

In the is paper, **hyperparameters are empirically determined, however, according to** In Liang et al 2009, Zhang et al 2012, The Bayesian optimization can be used to determine the critical hyperparameters.

5.2.3 SOM neural network training algorithm steps

(1) Determine SOM network topology, the number of neurons in input layer and competition layer.

(2) Set t = 0, random initialization weight vector w_j (0) (j = 1, 2, ..., M), M is the number of neurons in the competition layer.

(3) The network is randomly provided with an input vector $P_k(t)=(P_1^k, P_2^k, \ldots, P_N^k)^T$, (k=1, 2, …, L), where L is the total number of data set vectors to be input.

(4) Calculate the distance between the current input vector and the neurons in the competition layer, and select the neurons with the smallest distance as the winning neurons q (t)

$$q(t) = \arg\min_{j} \left\| P_k(t) - w_j(t) \right\| \tag{7}$$

(5) Adjust the weighted vector of the winning neuron and its neighborhood to

$$w_j(t+1) = \begin{cases} \dfrac{w_j(t) + \eta(t)(P_k(t) - w_j(t))}{w_j(t)} & \begin{matrix} j \in N_q(t) \\ j \notin N_q(t) \end{matrix} \end{cases} \tag{8}$$

η(t) is the learning rate parameter, and 0 < η (t) < 1, N_q (t) is the adjacent region of the winning neuron q, both of which are decreasing functions with the increase of time t.

(6) Determine whether all the input vectors are provided to the network, if it is transferred to the next step, otherwise return step (3).

(7) Update learning rate and neighborhood radius.

(8) If the total number of iterations reaches T, the algorithm ends, otherwise step (3) is returned.

5.3 SOM neural network testing

The trained network clusters the input test sample data, which is called the network recall process. If the training input mode P_k wins at node j, then when the test input model X_k is similar to the training input mode P_k, node j will be more likely to win, that is, the category attribute of the test input mode X_k is identified.

5.3.1 Test sample design

33 groups of simulated AE signals and 150 groups of noise signals in AE test are selected as test samples set , in total are 183 groups of test sample feature vectors. Each sample includes 16 evaluation indexes, and the network is tested by test samples. Test input mode is :

$$X_k = (X_1^k, X_2^k, \cdots, X_{16}^k) \qquad \mathrm{k} = 1, 2, \cdots, 183$$

5.3.2 Mapping of SOM neural network clustering results

For each input signal, the output plane array of SOM neural network has a specific neuron sensitive to it. This input-output mapping relationship is very clear in the output characteristic plane array. Fig. 9 is the signal clustering result map. Character I maps simulate AE signals neurons, character C maps neurons that interfere with noise signals, and character subscripts are cluster serial numbers.

6 Clustering results analysis of AE signal pattern recognition based on SOM neural network

6.1 Wavelet Energy

The wavelet energy coefficient of the standard sample is extracted as the input feature vector, and the SOM neural network will conduct automatic classification training. Finally, the test dataset is used for the training. The clustering formula is shown in Equation 9. The clustering results are shown in Fig. 9 and the clustering summary table is shown in Table 5.

$$N_C\% = \frac{\sum_{i=1}^{k} N_{Ci}}{N} \tag{9}$$

Where , $N_C\%$ is the percentage of the noise signals, $\sum_{i=1}^{k} N_{Ci}$ is the sum of the number of noise signals in each character C (noise signal neuron), N is the total number of input noise signals.

It can be seen from Table 5 and Fig. 10 that SOM neural network has strong self-recognition ability, the classification accuracy of simulated AE signals in the standard sample set reaches 90 %. When predicting the signal dataset in the test, it also shows a similar classification trend, and the accuracy reaches 76 %. However, the simulated AE signals and noise signals are mixed to some extent, reflecting the diversity and complexity of the signal mode. The main reasons for the errors are as follows : Firstly, the simulated AE signal is a point source, the signals received by different distance sensors are transmitted and attenuated, and have different modes. Secondly, the noise signals generated by different sound sources such as vibration and friction also have different modes. For the noise signals whose feature energy is obviously concentrated in the second and fourth bands and is not easy to cross with other classes, the classification accuracy of the standard sample set reaches 95 %, while the classification accuracy of the test sample set reaches 94 %. In summary, SOM neural network shows good classification ability and can clearly identify AE signals from a large number of interference noise signals.

6.2 entropy

The wavelet energy entropy of each frequency band of the signal in the standard sample dataset and the test sample dataset is used as the feature vector for similar clustering. The clustering formula is shown in Equation 10. The clustering results are summarized in Table 6. Fig. 11 is the output of the signal clustering result.

$$N_I\% = \frac{\sum_{i=1}^{k} N_{Ii}}{N} \quad (10)$$

,Where $N_I\%$ is the percentage of analog AE signals, $\sum_{i=1}^{k} N_{Ii}$ is the sum of the number of simulated AE signals in each character I (simulated AE signal neurons), N is the total number of input simulated AE signals.

Table 6 and Fig. 11 show that: first of all, noise signal recognition can reach more than 99 %, simulated AE signal classification accuracy in standard sample dataset of is 100 %. The trained network is used for the test the sample signals, and the correct rate reaches 78. 8 %. The reason is similar to wavelet energy coefficient analysis. Secondly, some simulated AE signals are mistakenly identified as noise signals by SOM neural network, indicating that the characteristics between different types are sometimes very similar and there is a certain degree of confusion. However, the network has good classification ability and strong generalization ability, which can clearly identify AE signals from a large number of noise signals. Finally, compared with the wavelet energy coefficient analysis method, the wavelet entropy coefficient analysis method has stronger clustering ability for signals. The reason is that wavelet entropy is the representation of the complexity of wavelet energy, which can better represent the characteristics of the signal.

7 Conclusion

The effectiveness of data filtering and spatial discrimination technology on noise signal recognition is analyzed in this paper based using Pencil Lead Break Testing on AE signal and noise AE signal of the bridge in operation state and other specific loading state, and the study further used wavelet packet energy and wavelet packet entropy analysis to extract the characteristic frequency bands of AE signal and interference noise signal, to train and identify defect signals using SOM neural networks. A more effective method to denoising noise signal was developed, the main conclusions are as follows:

1. When threshold is greater than 43dB, it can eliminate most of the environment and other interference noise signal. However, the noise signal and damage signal are often intertwined in the field detection. When the threshold is low, it often contains many unnecessary noise signals. When the threshold is high, all the noise is filtered, but the system will filter out the AE signal with low amplitude. Therefore, in practical engineering detection, Data filtering needs to rely on the experienced engineers, and is not that reliable

2. Under the same conditions, by using the guard sensor, 50 % of the noise signal is eliminated and the signal to noise ratio is improved. It shows that the noise signal can be well shielded by setting the guard sensor. Although spatial filtering can eliminate the noise signal from the distance, it cannot process the noise signal in the detection area. Therefore, it is necessary to conduct in-depth research on the noise characteristics in the field detection to find a more suitable noise processing method.

3. Energy and entropy analysis have similar laws. The simulated AE signal source is mainly concentrated in the 6-8 frequency band. With the propagation of the signal, the high frequency

components of signals decay continuously, When the signal was spread to 1.2m, the frequency gradually concentrated in the1, 2, 4 bands. The energy and entropy of the noise signal are mainly concentrated in the1, 2, 4 frequency bands, which are similar to those of the simulated damage AE signal at 1.2 m, but the frequency characteristics between AE simulation signal and noise signal are obviously different, so the feature vector can be formed by extracting the characteristic frequency band of signal to establish the identification mode of AE signal.

4. The classification accuracy of simulated AE signals obtained by wavelet packet energy and entropy analysis method reaches 90 % and 100 %, respectively. The trained NN is used for the test set signal, and the accuracy reaches 76 % and 78.8 %. The error is caused by the characteristic cross between the effective AE signal and the interference noise signal, and the input data are similar. However, the network has good classification ability and can clearly identify AE signals from a large number of interference noise signals. It shows that compared with the defects of commonly used methods (hardware filtering and spatial identification technology) in practical engineering, this method can more accurately identify and separate noise signals, reduce the distortion of damage acoustic emission signal caused by noise environment, and make more accurate and effective acoustic emission identification and characterization of structural damage.

The relevant machine learning on denoise is rare. Most research is to use ML to predict the acoustic emission signals, such as use Long Short-Term Memory (LSTM) (Zhang et al ,2018) network and support vector machine learning models (Yang et al, 2012). The reason we choose SOM neural network is because that it can use Wavelet packet energy analysis, Wavelet packet entropy to analyze the defect signal. Further research on other learning model will be continued in our future project.

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

The first author would like to acknowledge the financial support provided by the National Natural Science Foundation of China (Grant No. 51968014) , Key R&D projects in the Guangxi Autonomous Region (Grant No. AA20302006) and Guangxi Key Laboratory of New Energy and Building Energy Saving (Grant No. 19-J-21-20).

References

Bafroui, H. H., A. Ohadi. (2014). “Application of wavelet energy and Shannon entropy for feature extraction in gearbox fault detection under varying speed conditions.” *Neurocomputing.* 133: 437-445.

Bhuiyan, M. S. H., I. A. Choudhury and M. Dahari. (2014). “Monitoring the tool wear, surface roughness and chip formation occurrences using multiple sensors in turning.” *J. Manuf. Syst.* 33 (4): 476-487.

Bianchi,D., E. Mayrhofer, M. Gröschl, G. Betz, and A. Vernes. (2015). “Wavelet packet transform for detection of single events in acoustic emission signals.” *Mech. Syst. Signal. Pr.* 64: 441-451.

Cai B, Pan G, Fu F (2020) Prediction of the Postfire Flexural Capacity of RC Beam Using

GA-BPNN Machine Learning,J ournal of Performance of Constructed Facilities 34 (6), 04020105

Cai B, Xu LF, Fu F (2019) Shear resistance prediction of post-fire reinforced concrete beams using artificial neural network,International Journal of Concrete Structures and Materials 13 (1), 1-13

Choi, Ahnryul, Jae-Moon Lee, and Joung Hwan Mun. (2013) Ground reaction forces predicted by using artificial neural network during asymmetric movements, Int J Precis Eng Man. 14.3 475-483.

Colombo, Ing. S., I. G. Main, and M. C. Forde. (2003). "Assessing Damage of Reinforced Concrete Beam using b -value Analysis of Acoustic Emission Signals." *J. Mater. Civil. Eng.* 15 (3): 280-286.

Deng, A., L. Zhao, Y. Bao, and Q. Gao. (2009). "Acoustic emission recognition based on wavelet entropy in noisy environment (in Chinese)." *China J. Southeast University (Natural Science Edition).* 39 (06): 1151-1155.

Dijck, G.V., M. Wevers, and M. V. Hulle. (2009). "Wavelet Packet Decomposition for the Identification of Corrosion Type from Acoustic Emission Signals." *Int. J. Wavelets. Multi.* 7 (4): 513-534.

Dunegan, H., and D. Harris. (1969). "Acoustic emission-a new nondestructive testing tool." *Ultrasonics.* 7.3: 160-166.

Ekici, S., S. Yildirim, M. Poyraz. (2008). "Energy and entropy-based feature extraction for locating fault on transmission lines by using neural network and wavelet packet decomposition." *Expert. Syst. Appl.* 34 (4): 2937-2944.

Fu F (2020), Fire induced Progressive Collapse Potential assessment of Steel Framed

Buildings using machine learning, Journal of Constructional Steel Research

Fu, F. (2018). Design and Analysis of Tall and Complex Structures. Butterworth-Heinemann. ISBN 978-0-08-101121-8.

Fu, P., W. L. Li, L. Q. Zhu. (2011). “Cutting Tool Wear Monitoring Based on Wavelet Denoising and Fractal Theory.” *Appl. Mech. Mater.* 48-49: 349-352.

Guo, J., C. J. Hu, M. J. Zhu, and Y. Q. Ni. (2021). “Monitoring-based evaluation of dynamic characteristics of a long span suspension bridge under typhoons.” *J. CIV. Struct. Health Monit.* 11: 397-410.

Guo, J., D. Hang, X. Zhu. (2020). “Prediction of Crack Propagation in U-Rib Components Based on the Markov Chain.” *J. Bridge. Eng.* 25 (10): 04020089.

Holford, K. M., M. J. Eaton, J. J. Hensman, R. Pullin, S. L. Evans, N. Dervilis, and K. Worden. (2017). “A new methodology for automating acoustic emission detection of metallic fatigue fractures in highly demanding aerospace environments: an overview.” *Proc. Aerosp. Sci.* 90: 1-11.

Hu I C, Liao H J, (2010) Method of up-slope mitigation priority for Alishan mountain road in Taiwan, J Perform Constr Fac. 24 (4) 373-381.

Kacimi, S., S. Laurens. (2009). “The correlation dimension: A robust chaotic feature for classifying acoustic emission signals generated in construction materials.” *J. Appl. Phys.* 106: 024909.

Kalyanasundaram, P., C. K. Mukhopadhyay, and S. V. SubbaRao. (2007). *Practical Acoustic Emission.* Indian Society for Non-Destructive Testing – National Certification Board Series.

Khanzadeh, M., Rao, P., Jafari-Marandi, R., Smith, B. K., Tschopp, M. A., & Bian, L, (2018).

Quantifying geometric accuracy with unsupervised machine learning: Using self-organizing map on fused filament fabrication additive manufacturing parts, J Manuf Sci E-T Asme. 140.3

Kohonen, T. 1998. “The self-organizing map.” *Neurocomputing.* 21 (1-3): 1-6.

Lamonaca, F., A. Carrozzini, D. Grimaldi, and R. S. Olivito. (2012). “Acoustic emission monitoring of damage concrete structures by multi-triggered acquisition system.” *2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings.* 1630-1634.

Li, D. S., J. P. Ou. (2007). “Application of Acoustic Emission Technology in Damage Monitoring of Arch Bridge Hangers (in Chinese).” *China J. Shenyang Jianzhu University (Natural Science).* 23 (1) :4-9.

Li, L., F. Chu. (2011). “Feature extraction of AE characteristics in offshore structure model using Hilbert-Huang transform.” *Measurement.* 41 (1): 46-54.

Li, Y. B., Y. X. Zhang, H. Y. Zhu, R. X. Yan, Y. Y. Liu, L. Y. Sun, and Z. M. Zeng. (2015). “Recognition Algorithm of Acoustic Emission Signals Based on Conditional Random Field Model in Storage Tank Floor Inspection Using Inner Detector.” *Shock. Vib.* 2015: 1-9.

Liang X, (2019) Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization, Computer-Aided Civil and Infrastructure Engineering 34(5):415-430

Liu, X. L., Z. Liu, X. B. Li, F. Q. Gong, and K. Du. (2020). “Experimental study on the effect of strain rate on rock acoustic emission characteristics.” *Int. J. Rock Mech. Min.* 133: 104420.

Lopes, B. G., F. A. Alexandre, W. N. Lopes, P. R. Aguiar, E. C. Bianchi, and M. A. A. Viera. (2018). “Study on the effect of the temperature in Acoustic Emission Sensor by the Pencil Lead Break Test.” *2018 13th IEEE International Conference on Industry Applications (INDUSCON).*

1226-1229.

Luo, Z. G., X. Wang, J. Li, B. B. Fan, and X. D. Guo. (2008). “Study of Punch Die Condition Discrimination Based on Wavelet Packet and Genetic Neural Network.” *Advances in Neural Networks - ISNN 2008.*https://doi.org/10.1007/978-3-540-87734-9_55.

McCrory, J. P., S. K. Al-Jumaili, D. Crivelli, M. R. Pearson, M. J. Eaton, C. Featherston, M. Guagliano, K. Holford, and R. Pullin. (2015). “Damage classification in carbon fibre composites using acoustic emission: A comparison of three techniques.” *Compos. Part. B-Eng.* 68: 424-430.

Moon S, Chung S, Chi S, (2020) Bridge damage recognition from inspection reports using NER based on recurrent neural network with active learning, J Perform Constr Fac. 34 (6) 04020119.

Nguyen, H. N., C. H. Kim and J. M. Kim. (2018). “Effective Prediction of Bearing Fault Degradation under Different Crack Sizes Using a Deep Neural Network.” *Appl. Sci.* 8: 2332.

Njuhovic, E., M. Bräu, F. Wolff-Fabris, K. Starzynski, V. Altstädt. (2014). “Identification of interface failure mechanisms of metallized glass fibre reinforced composites using acoustic emission analysis.” *Compos. Part. B-Eng.* 66: 443-452.

Noble W S (2006), What is a support vector machine? Nat. Biotechnol. 24 (12) 1565-1567.

Noorsuhada, M. N. (2016). “An overview on fatigue damage assessment of reinforced concrete structures with the aid of acoustic emission technique.” *Constr. Build. Mater.* 112: 424-439.

Nor, N. M., N. M. Bunnori, A. Ibrahim, S. Shahidan, and S. N. M. Saliah. (2011). “An observation of noise intervention into acoustic emission signal on concrete structure.” *2011 IEEE 7th International Colloquium on Signal Processing and its Applications.*

Rmili, W., A. Ouahabi, R. Serra, and R. Leroy. (2016). “An automatic system based on vibratory analysis for cutting tool wear monitoring.” *Measurement.* 77: 117-123.

Safty, S. El., A. El-Zonkoly. (2008). “Applying wavelet entropy principle in fault classification.” *Int. J. Elec. Power.* 31 (10): 604-607.

Silva, Rui G., and Steven J. Wilcox, (2019) Feature evaluation and selection for condition monitoring using a self-organizing map and spatial statistics, AI EDAM. 33.1 1-10.

Surgeon, M., M. Wevers. (1999). “Modal analysis of acoustic emission signals from CFRP laminates.” *NDT & E Int.* 32 (6): 311-322

Velayudham, A., R. Krishnamurthy, T. Soundarapandian. (2005). “Acoustic emission based drill condition monitoring during drilling of glass/phenolic polymeric composite using wavelet packet transform.” *Mat. Sci. Eng. A-Struct.* 412 (1-2): 141-145.
Volume 147 Issue 1

Wang H, Zhang Y M, Mao J X, (2022) Sparse Gaussian process regression for multi-step ahead forecasting of wind gusts combining numerical weather predictions and on-site measurements, J. Wind. Eng. Ind. Aerodyn, 220 104873.

Wang, W. H., Hong, G. S., Wong, Y. S., & Zhu, K. P, (2007) Sensor fusion for online tool condition monitoring in milling, Int J Prod Res. 45.21 5095-5116.

Wen, X. (2015). *The neural network is realized by MATLAB (in Chinese).* Beijing: National Defense Industry Press.

Yang Z, Yu Z. (2012)Grinding wheel wear monitoring based on wavelet analysis and support vector machine[J]. The International Journal of Advanced Manufacturing Technology, 2012, 62(1): 107-121.)

Yin, X., Y. He, Z. Peng, and F. Chu. (2004). “Wavelet entropy and its application in state trend analysis (in Chinese).” *China J. Vibra. Eng.* 17 (2): 49-53.

Zafar, T., K. Kamal, Z. Sheikh, S. Mathavan, U. Ali, and H. Hashmi. (2017). “A neural network based approach for background noise reduction in airborne acoustic emission of a machining process.” *J. Mech. Sci. Technol.* 31 (7): 3171-3182.

Zhang X, Zou Z, Wang K, et al (2018). A new rail crack detection method using LSTM network for actual application based on AE technology[J]. Applied Acoustics, 2018, 142: 78-86.)

Zhang Y, Hao Wang, Jian-Xiao Mao; Zi-Dong Xu; and Yu-Feng Zhang,(2021)Probabilistic Framework with Bayesian Optimization for Predicting Typhoon-Induced Dynamic Responses of a Long-Span Bridge, Journal of Structural Engineering,

Zhang, X., K. Wang, Y. Wang, Y. Shen, and H. Hu. (2017). “An improved method of rail health monitoring based on CNN and multiple acoustic emission events.” *2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC).*

Table

Table 1. Acoustic emission detection parameter settings

parameter	settings
Sensor model	R6
Pre-amplifier	40dB
Acquisition threshold	43dB

Sample rate	1M
Pre-trigger	100μs
Length	1k
PDT	300μs
HDT	600μs
HLT	1000μs
Positioning wave speed	2700m·s^{-1}
Number of sensors	9
Layout	flat

Table 2. frequency ranges

frequency band	1	2	3	4	5	6	7	8
Frequency interval(kHz)	0～16	16～32	32～48	48～64	64～80	80～96	96～112	112～128
frequency band	9	10	11	12	13	14	15	16
Frequency interval(kHz)	128～144	144～160	160～176	176～192	192～208	208～224	224～240	240～256

Table 3. statistics tables of Energy in each frequency band

Frequency band	**0m**	**0.6m**	**1.2m**	**Static load**	**10km/h**	**20km/h**	**30km/h**	**Bounce state**	**Braking state**	**Operational status**
1、2、4	——	——	87.44%	91.41%	91.93%	93.37%	88.11%	87.64%	92.10%	92.90%

1	0.52%	2.53%	——	——	——	——	——	——	——	——
3	——	——	5.18%	5.99%	3.89%	3.76%	6.76%	6.62%	4.50%	2.20%
2-4	14.20%	83.99%	——	——	——	——	——	——	——	——
5	1.12%	0.60%	0.23%	0.12%	0.14%	0.38%	0.41%	0.09%	0.03%	2.64%
6-8	72.19%	12.17%	6.45%	2.48%	3.95%	2.49%	4.67%	5.57%	3.29%	2.26%
9-13	1.61%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%	0.00%	0.00%
14-16	10.36%	0.70%	0.69%	0.00%	0.08%	0.00%	0.06%	0.04%	0.08%	0.00%
Total	1	1	1	1	1	1	1	1	1	1

Note:"——" means no statistics

Table 4. Entropy statistics table of each frequency band

Frequency band	**0m**	**0.6m**	**1.2m**	**Static load**	**10km/h**	**20km/h**	**30km/h**	**Bounce state**	**Braking state**	**Operational status**
1、2、4	——	——	86.71%	91.00%	92.56%	92.06%	84.59%	84.38%	87.87%	91.25%
1	1.31%	3.25%	——	——	——	——	——	——	——	——
3	——	——	4.90%	3.84%	3.24%	4.04%	8.69%	8.02%	7.02%	4.50%
2-4	10.09%	68.02%	——	——	——	——	——	——	——	——
5	1.08%	1.09%	0.27%	0.09%	0.00%	0.25%	0.52%	0.08%	0.14%	0.12%
6-8	79.09%	25.05%	7.42%	4.97%	4.04%	3.65%	6.20%	7.27%	4.57%	3.99%
9-13	1.77%	0.43%	0.00%	0.00%	0.00%	0.00%	0.00%	0.05%	0.12%	0.06%
14-16	6.66%	2.16%	0.70%	0.09%	0.16%	0.00%	0.00%	0.19%	0.28%	0.09%
Total	1	1	1	1	1	1	1	1	1	1

Note:"——" means no statistics

Table 5. Cluster results

Class	Sample set			Test set		
	Simulated AE signal	Noise signal	Cluster accuracy （%）	Simulated AE signal	Noise signal	Cluster accuracy （%）
Simulated AE signal	18	2	90	25	8	76
Noise signal	5	95	95	9	141	94

Table 6. Cluster results

Class	Sample set			Test set		
	Simulated AE signal	Noise signal	Cluster accuracy （%）	Simulated AE signal	Noise signal	Cluster accuracy （%）
Simulated AE signal	20	0	100	26	7	78.8

Noise signal	1	99	99	1	149	99.4