



City Research Online

City, University of London Institutional Repository

Citation: Yu, A., Liu, X., Fu, F., Chen, X. & Zhang, Y. (2022). Acoustic Emission Signal Denoising of Bridge Structures using SOM Neural Network Machine Learning. *Journal of Performance of Constructed Facilities*,

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/28434/>

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

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

Aiping Yu¹, Xiangtai Liu² Feng Fu³ C.Eng, F.ASCE. · Xuandong Chen⁴, Yan Zhang⁵

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 test 183 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 Yu¹, Professor, College of Civil and Architectural Engineering, Guilin University of Technology, Guilin 541004. apyu@glut.edu.cn;

2. Xiangtai Liu¹, 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

35 detection technique (Dunegan 1969), and there is no strict requirement for the size and detection
36 range of the detected object. Moreover, non-contact detection can be achieved, such as detection
37 in high temperature and corrosion that cannot be approached by human beings.

38 This technology has been widely used in the damage detection of homogeneous materials,
39 such as damages in mechanical bearing (Nguyen et al. 2018), metal tool (Rmili et al. 2016;
40 Bhuiyan et al. 2014), metal container corrosion detection (Li et al. 2015), aeronautical metal
41 materials (Holford et al. 2017), metal track detection (Zhang et al. 2017), glass (Njuhovic et al.
42 2014), carbon fiber (McCrory et al. 2015) and rock (Liu et al. 2020). However, there is little
43 application in reinforced concrete materials (Colombo et al. 2003; Lamonaca et al. 2012;
44 Noorsuhada 2016). The main reason is the complexity and diversity of environmental noise signals
45 causing ambiguities when identify the noise signals and structural damage signals, which makes
46 it difficult to effectively identify and characterize damage signals.

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

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

55 Artificial intelligence technology is one of the means of signal recognition. Its algorithms are
56 mainly divided into traditional machine learning algorithms (Fu,2020,2018) and neural network

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

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

79 Therefore, in this paper, the AE signals of the bridges in operation under certain specific
80 loading state were tested in this experiment. Based on the signal of hardware filtering technology
81 and spatial identification technology, as well as the wavelet packet energy analysis (Guo et al.
82 2020, 2021) and wavelet packet entropy analysis (Safty and El-Zonkoly 2008; Yin et al. 2004) the
83 characteristic frequency bands were extracted from simulated acoustic emission signal and noise
84 signal. Finally, the signal was clustered by using Konhonen's self-organizing feature map and
85 neural network (SOM neural network) (Kohonen 1998) to establish an acoustic emission detection
86 and recognition algorithm, which provided new ideas and methods for solving the noise reduction
87 problem of bridge acoustic emission damage signal. It is expected to solve the difficulty to
88 effectively identify and characterize structural damage due to the confusion between noise signals
89 and structural damage signals.

90 **2. Test set up**

91 *2.1 Introduction of the prototype bridge*

92 The prototype bridge is a single flyover, and the superstructure is a four-span simply
93 supported hollow beam bridge with a span of 10 m + 20 m + 20 m + 10 m. The main beam with a
94 span of 20 m is a prestressed reinforced concrete hollow beam, and the main beam with a span of
95 10 m is a common reinforced concrete hollow beam. The net width of the bridge deck is 11m +
96 2×0.5m (anti-collision wall), and the bridge design grade is grade I.

97 *2.2 Detection scheme*

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

101 test machinery.

102 *2.3 sensor installation*

103

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

112

113 *2.4 Simulated Acoustic Emission Signal and Noise Source*

114

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

120

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

123

124 2) Signal in driving condition : the test vehicle passes the bridge at the speeds of 10 km / h,
125 20 km / h and 30 km / h to collect the AE signals generated by the vertical vibration and friction
126 of the fourth span bridge.

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

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

134

135 **3 Filtering and Spatial Identification of Noise Signal**

136 *3.1 Data filtering*

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

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

146 generated by the environment.

147 *3.2 Spatial Identification*

148 Spatial identification technology is to place two types of sensors (monitoring sensors and
149 guard sensors) at different locations to eliminate noise signals. In the test, the main sensor (No. 1-
150 7 sensor) is placed at the bottom of the test beam, and the guard sensor (No. 8-9 sensor) is placed
151 at the side of the tested area to shield the interference noise signal produced when driving in both
152 direction. The three-dimensional schematic layout of the sensor is shown in Fig. 4. The noise signal
153 positioning maps before and after setting the guard sensor are shown in Fig. 5 (a) and Fig. 5 (b).

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

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

163 **4 Wavelet packet based feature extraction**

164 *4.1 Basic principles*

165 wavelet packet transform (WPT) has become one of the most widely used signal analysis
166 methods because of its multi-resolution and ability to characterize the local characteristics of
167 signals in both time domain and frequency domain. Wavelet entropy analysis is a new method to

168 measure the complex sequence of the signal, which is the combination of wavelet transform and
169 information entropy. In signal processing, it not only has the advantages of changeable resolution
170 and no signal -stationary requirements, but also can statistically analyze the complexity of entropy
171 on signal, which can be used to detect the local characteristics of non-stationary signals.

172

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

$$177 \quad \varphi(t) = \sqrt{2} \sum_k h(k) \varphi(2t - k) \quad (1)$$

$$178 \quad \psi(t) = \sqrt{2} \sum_k g(k) \varphi(2t - k) \quad (2)$$

179 Where: k is the translation amplitude. $h(k)$ is the low-pass filter corresponding to the scale
180 function $\varphi(t)$. $g(k)$ is a high-pass filter corresponding to the wavelet function $\psi(t)$.

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

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

$$186 \quad E_{i,j} = \int |S_{i,j}(t)|^2 dt = \sum_{k=1}^n |x_{i,j}(k)|^2 \quad (3)$$

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

189

190 Therefore, the total energy of the whole signal is :

$$191 \quad E = \sum_{j=0}^{2^l-1} E_{i,j} \quad (4)$$

192 Wavelet packet energy coefficient:

$$193 \quad P_{i,j} = \frac{E_{i,j}}{E} \quad (5)$$

194

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

$$197 \quad H_{i,j} = - \sum_{j=0}^n P_{i,j} \log P_{i,j} \quad (6)$$

198 *4.2 Wavelet packet decomposition*

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

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

211 *4.3 Energy Analysis of Wavelet packet Coefficient*

212 The results of calculating the characteristic energy of each decomposed signal in each
213 frequency band are shown in Fig. 6, and the energy proportion of each signal is shown in Table 3.
214 It can be seen that the frequency components of simulated AE signals with different propagation
215 distances are quite different. The frequency band range of the sound source is wide, and both high
216 and low frequency information exist. The energy is mainly concentrated in the 6 ~ 8 high frequency
217 band, accounting for 72.2 % of the total energy. The high frequency energy components such as
218 14 ~ 16 account for 10.4 % of the total energy, and the 2 ~ 4 band accounts for 14.2 % of the total
219 energy. With the propagation of signals, the high-frequency components of 6th, 7th and 8th
220 continuously decay. When propagating to 1.2 m, the energy is mainly concentrated in the first,
221 second and fourth bands, accounting for 87.4 % of the total energy. The energy of noise in other
222 working conditions is mainly concentrated in the 1st, 2nd and 4th low which is similar to the
223 energy of simulated AE signal at 1.2 m and can be greater than 87.4 %. So the noise signals can
224 be distinguished by the main frequency band distribution.

225 *4.4 Entropy Analysis of Wavelet packet Coefficient*

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

233 in the other conditions is similar to that of the simulated AE signal at 1.2 m, which is also
234 concentrated in the first, second and fourth low frequency bands, accounting for more than 84.4 %.

235 **5 SOM Neural Network Machine learning**

236 *5.1 SOM Clustering Principle*

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

250 *5.2 SOM neural network training*

251 **5.2.1 Standard Sample Design**

252 The acoustic emission signals collected in test of flyover are clustered. 20 groups of simulated
253 AE signals and 100 groups of noise signals are selected as standard sample set input, in total of
254 120 groups of sample feature vectors. Training input mode is:

255
$$P_k = (P_1^k, P_2^k, \dots, P_{16}^k) \quad k = 1, 2, \dots, 120$$

256 **5.2.2 Network structure design**

257 1) Number of neurons in input layer

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

262 2) Number of neurons in competition layer

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

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

272 Create a SOM neural network using the NEWSOM function The code is : net = NEWSOM
273 (minmax (P),[6 6]). Among them, minmax (P) specifies the maximum and minimum values of
274 input vector elements, and 6×6 denotes the structure of 6×6 competition layer of the network. The
275 network is trained and simulated by function train and simulation function sim. The size of learning
276 rate ^[35] and the number of training steps will affect the clustering performance of the network

277 (training time and convergence rate). When the learning rate ($\eta(t)$) is 1.0 and the number of steps(T)
278 is 500, it is clear to observe the clustering results.

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

282 **5.2.3 SOM neural network training algorithm steps**

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

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

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

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

$$291 \quad q(t) = \arg \min_j \|P_k(t) - w_j(t)\| \quad (7)$$

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

$$293 \quad w_j(t+1) = \begin{cases} w_j(t) + \eta(t)(P_k(t) - w_j(t)) & j \in N_q(t) \\ w_j(t) & j \notin N_q(t) \end{cases} \quad (8)$$

294 $\eta(t)$ is the learning rate parameter, and $0 < \eta(t) < 1$, $N_q(t)$ is the adjacent region of the
295 winning neuron q , both of which are decreasing functions with the increase of time t .

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

298 (7) Update learning rate and neighborhood radius.

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

301 5.3 SOM neural network testing

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

306 5.3.1 Test sample design

307 33 groups of simulated AE signals and 150 groups of noise signals in AE test are selected as
308 test samples set , in total are 183 groups of test sample feature vectors. Each sample includes 16
309 evaluation indexes, and the network is tested by test samples. Test input mode is :

$$310 X_k = (X_1^k, X_2^k, \dots, X_{16}^k) \quad k = 1, 2, \dots, 183$$

311 5.3.2 Mapping of SOM neural network clustering results

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

317 6 Clustering results analysis of AE signal pattern recognition based on SOM neural 318 network

319 6.1 Wavelet Energy

320 The wavelet energy coefficient of the standard sample is extracted as the input feature vector,

321 and the SOM neural network will conduct automatic classification training. Finally, the test dataset
322 is used for the training. The clustering formula is shown in Equation 9. The clustering results are
323 shown in Fig. 9 and the clustering summary table is shown in Table 5.

324

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

325

326 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
327 each character C (noise signal neuron), N is the total number of input noise signals.

328

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

342 6.2 entropy

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

347

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

348

349 ,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
350 signals in each character I (simulated AE signal neurons), N is the total number of input simulated AE signals.
351

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

363 **7 Conclusion**

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

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

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

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

385 components of signals decay continuously, When the signal was spread to 1.2m, the frequency
386 gradually concentrated in the 1, 2, 4 bands. The energy and entropy of the noise signal are mainly
387 concentrated in the 1, 2, 4 frequency bands, which are similar to those of the simulated damage AE
388 signal at 1.2 m, but the frequency characteristics between AE simulation signal and noise signal
389 are obviously different, so the feature vector can be formed by extracting the characteristic
390 frequency band of signal to establish the identification mode of AE signal.

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

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

407

408 **Data Availability Statement**

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

411

412 **Acknowledgments**

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

417 **References**

418 Bafroui, H. H., A. Ohadi. (2014). “Application of wavelet energy and Shannon entropy for
419 feature extraction in gearbox fault detection under varying speed conditions.” *Neurocomputing*.
420 133: 437-445.

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

424 Bianchi,D., E. Mayrhofer, M. Gröschl, G. Betz, and A. Vernes. (2015). “Wavelet packet
425 transform for detection of single events in acoustic emission signals.” *Mech. Syst. Signal. Pr.* 64:
426 441-451.

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

428 GA-BPNN Machine Learning, Journal of Performance of Constructed Facilities 34 (6), 04020105

429 Cai B, Xu LF, Fu F (2019) Shear resistance prediction of post-fire reinforced concrete beams
430 using artificial neural network, International Journal of Concrete Structures and Materials 13 (1),
431 1-13

432 Choi, Ahnryul, Jae-Moon Lee, and Joung Hwan Mun. (2013) Ground reaction forces
433 predicted by using artificial neural network during asymmetric movements, Int J Precis Eng Man.
434 14.3 475-483.

435 Colombo, Ing. S., I. G. Main, and M. C. Forde. (2003). "Assessing Damage of Reinforced
436 Concrete Beam using b -value Analysis of Acoustic Emission Signals." *J. Mater. Civil. Eng.* 15
437 (3): 280-286.

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

441 Dijck, G.V., M. Wevers, and M. V. Hulle. (2009). "Wavelet Packet Decomposition for the
442 Identification of Corrosion Type from Acoustic Emission Signals." *Int. J. Wavelets. Multi.* 7 (4):
443 513-534.

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

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

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

450 Buildings using machine learning, *Journal of Constructional Steel Research*

451 Fu, F. (2018). *Design and Analysis of Tall and Complex Structures*. Butterworth-Heinemann.

452 ISBN 978-0-08-101121-8.

453 Fu, P., W. L. Li, L. Q. Zhu. (2011). “Cutting Tool Wear Monitoring Based on Wavelet

454 Denoising and Fractal Theory.” *Appl. Mech. Mater.* 48-49: 349-352.

455 Guo, J., C. J. Hu, M. J. Zhu, and Y. Q. Ni. (2021). “Monitoring-based evaluation of dynamic

456 characteristics of a long span suspension bridge under typhoons.” *J. CIV. Struct. Health Monit.* 11:

457 397-410.

458 Guo, J., D. Hang, X. Zhu. (2020). “Prediction of Crack Propagation in U-Rib Components

459 Based on the Markov Chain.” *J. Bridge. Eng.* 25 (10): 04020089.

460 Holford, K. M., M. J. Eaton, J. J. Hensman, R. Pullin, S. L. Evans, N. Dervilis, and K.

461 Worden. (2017). “A new methodology for automating acoustic emission detection of metallic

462 fatigue fractures in highly demanding aerospace environments: an overview.” *Proc. Aerosp. Sci.*

463 90: 1-11.

464 Hu I C, Liao H J, (2010) Method of up-slope mitigation priority for Alishan mountain road

465 in Taiwan, *J Perform Constr Fac.* 24 (4) 373-381.

466 Kacimi, S., S. Laurens. (2009). “The correlation dimension: A robust chaotic feature for

467 classifying acoustic emission signals generated in construction materials.” *J. Appl. Phys.* 106:

468 024909.

469 Kalyanasundaram, P., C. K. Mukhopadhyay, and S. V. SubbaRao. (2007). *Practical Acoustic*

470 *Emission*. Indian Society for Non-Destructive Testing – National Certification Board Series.

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

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

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

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

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

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

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

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

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

491 Lopes, B. G., F. A. Alexandre, W. N. Lopes, P. R. Aguiar, E. C. Bianchi, and M. A. A. Viera.
492 (2018). "Study on the effect of the temperature in Acoustic Emission Sensor by the Pencil Lead
493 Break Test." *2018 13th IEEE International Conference on Industry Applications (INDUSCON)*.

494 1226-1229.

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

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

501 Moon S, Chung S, Chi S, (2020) Bridge damage recognition from inspection reports using
502 NER based on recurrent neural network with active learning, *J Perform Constr Fac.* 34 (6)
503 04020119.

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

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

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

510 Noorsuhada, M. N. (2016). "An overview on fatigue damage assessment of reinforced
511 concrete structures with the aid of acoustic emission technique." *Constr. Build. Mater.* 112: 424-
512 439.

513 Nor, N. M., N. M. Bunnori, A. Ibrahim, S. Shahidan, and S. N. M. Saliyah. (2011). "An
514 observation of noise intervention into acoustic emission signal on concrete structure." *2011 IEEE
515 7th International Colloquium on Signal Processing and its Applications.*

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

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

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

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

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

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

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

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

535 Yang Z, Yu Z. (2012)Grinding wheel wear monitoring based on wavelet analysis and support
536 vector machine[J]. *The International Journal of Advanced Manufacturing Technology*, 2012, 62(1):
537 107-121.)

538 Yin, X., Y. He, Z. Peng, and F. Chu. (2004). "Wavelet entropy and its application in state trend

539 analysis (in Chinese).” *China J. Vibra. Eng.* 17 (2): 49-53.

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

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

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

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

551

552

553 **Table**

554

555

Table 1. Acoustic emission detection parameter settings

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

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

556

557 **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

558

559

560 **Table 3.** statistics tables of Energy in each frequency band

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

Note: “—” means no statistics

563

564

565 **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

566

567

568 **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

569

570

571

572