

Detection of damage locations and damage steps in pile foundations using acoustic emissions with deep learning technology

Alipujiang JIERULA^{a,b}, Tae-Min OH^{*}, Shuhong WANG^a, Joon-Hyun LEE^b, Hyunwoo KIM^d, Jong-Won LEE^c

^a School of Resource & Civil Engineering, Northeastern University, Shenyang 110819, China

^b School of Mechanical Engineering, Pusan National University, Busan 46241, Korea

^c Department of Civil and Environmental Engineering, Pusan National University, Busan 46241, Korea

^d Korea Institute of Geoscience and Mineral Resource (KIGAM), Daejeon 34132, Korea

^{*} Corresponding author: E-mail: geotaemin@pusan.ac.kr

© Higher Education Press 2021

ABSTRACT The aim of this study is to propose a new detection method for determining the damage locations in pile foundations based on deep learning using acoustic emission data. First, the damage location is simulated using a back propagation neural network deep learning model with an acoustic emission data set acquired from pile hit experiments. In particular, the damage location is identified using two parameters: the pile location (P_L) and the distance from the pile cap (D_S). This study investigates the influences of various acoustic emission parameters, numbers of sensors, sensor installation locations, and the time difference on the prediction accuracy of P_L and D_S . In addition, correlations between the damage location and acoustic emission parameters are investigated. Second, the damage step condition is determined using a classification model with an acoustic emission data set acquired from uniaxial compressive strength experiments. Finally, a new damage detection and evaluation method for pile foundations is proposed. This new method is capable of continuously detecting and evaluating the damage of pile foundations in service.

KEYWORDS pile foundations, damage location, acoustic emission, deep learning, damage step

1 Introduction

Group pile foundations are used in modern infrastructure construction to enhance the ground bearing capacity [1–4]. Owing to the complexity of underground environments, unexpected damage or failure can occur during the service life of a structure. It is thus very important to implement effective monitoring of the health status of pile foundations and detect potential or actual damage locations, such as cracks and impacts [5]. However, it is more difficult to check foundations for damage or failure compared with upper structures, as foundations are always located underground [6]. Thus, source location detection and damage evaluations of pile foundations are urgent issues for engineers and researchers.

Acoustic emission (AE) refers to the elastic waves that

are released through structural media when microcracks occur. AE is a structural health monitoring (SHM) technique [7] and is widely utilized for nondestructive testing of in-service concrete structures [8]. Compared with other nondestructive testing methods, AE techniques have been applied during loading, which has allowed for convenient detection of in-service structures in recent years [9].

Deep learning is a class of machine learning that has been widely used for diagnostic purposes in recent years [10,11]. Deep learning involves several aspects, such as data analysis, image classification, and voice recognition [12–14]. Deep learning has also been applied to solve partial differential equations (PDEs) [15–17]. Recently, several studies have investigated the application of deep learning for AE source locations with convolutional neural networks (CNNs). Ebrahimkhanlou et al. [18] proposed a general deep learning model for the source location and

characteristics of AE in plate-like structures. Ebrahimkhanlou and Salamone [19] proposed two different deep learning models to detect the AE of source locations in metallic plates by considering geometric features. However, existing studies have only concentrated on metallic plate structures and cannot be applied for the detection of damage locations in in-service pile foundations.

The damage step status should be monitored for repair assessments of damaged structures. Several studies have been conducted to identify damage steps in terms of AE data. Jiao et al. [20] performed model tests on reinforced concrete beams with AE and studied the correlations between the cumulative crack width and AE parameters. Logoń [21] conducted tests on quasi-brittle composites to identify the destruction process using AE and a sound spectrum. It was necessary to clarify the damage steps separately to understand the failure mechanisms. Deep learning methods have also been used to verify the damage steps in recent studies. Xu et al. [22] investigated the identification of seismic damage and source locations using a CNN with 400 raw images. The images were captured using a digital camera at different locations with varying degrees of damage. Furthermore, Jang et al. [23] proposed a new detection method for concrete cracks using CNNs with hybrid images. The hybrid images improved the crack detection accuracy by combining visible and infrared information. Silva et al. [24] presented a deep principal component analysis (PCA) approach for concrete damage identification. However, the correlations between AE parameters and damage steps have yet to be studied.

To investigate how the damage location and damage step correlate with AE parameters, this study carried out pile hit experiments and uniaxial compressive strength (UCS) experiments. Pile hit experiments were conducted to simulate the damage location, and compressive uniaxial experiments were performed to simulate the damage steps of concrete. Subsequently, a new damage location detection and damage step evaluation method was proposed using the analyzed test data with deep learning.

The remainder of this paper is organized as follows. First, a damage location prediction method based on a back propagation (BP) neural network deep learning model is developed and verified using the pile hit experiment data. Second, a damage step prediction model is developed based on the classification learner app using the UCS experiment data. Finally, a new damage detection and evaluation method for pile foundations is proposed.

2 Damage location prediction

2.1 Data collection from pile hit experiments

Figure 1 shows the experimental setup of the pile hit test, which was performed to obtain damage location simulation data for deep learning. The specimen was a scale model of

a pile foundation, including a pile cap, an artificial bedrock and six piles. The dimensions of the specimen are shown in Fig. 1. The three-dimensional sizes of the pile cap and artificial bottom rock bed were 144 cm × 96 cm × 25 cm, and the depth of the piles was 100 cm. Six piles were deployed in two rows and three columns on the artificial bottom rock bed. The row and column spacings between the centers of the piles were 48 cm. The diameter of the piles was 16 cm (refer to Fig. 1(g)). Ten AE sensors were installed on the specimen, with six sensors installed on the pile cap and four sensors installed on the artificial bedrock, as shown in Figs. 1(g) and 1(h).

Six hit points were selected on the pile at 20 cm intervals from the pile bottom to top, with each point including four different hit directions separated by 90° angles (Fig. 1(c)). Each point was hit with a small iron hammer five times in each hit direction (Fig. 1(f)).

The pile hit experiment data were recorded using a Micro-II Digital AE system (Fig. 1(d)). The threshold (trigger level for recording AE signals) was set to 60 dB after pretesting, which could effectively prevent surrounding noise interference.

2.2 Deep learning data set

The pile foundation specimen included six piles, and each pile was hit in four directions. Each direction included six hit points with varying depths, and each target point was hit five times. With each hit, the AE signals were received by the ten sensors installed on the specimen, with each sensor receiving one AE signal. In total, 7200 AE signal data points were detected (Fig. 2(a)). An AE signal has 17 characteristics (called AE parameters), which are defined in Table 1. The final size of the deep learning data set was 17 × 7200, of which a 17 × 7188 data set was randomly selected as the sample set for training, and a 17 × 12 data set was randomly selected as the validation set.

2.3 Back propagation neural network deep learning model

BP neural networks are adopted in artificial neural networks [25]. These are a type of supervised learning used to minimize errors, and the algorithm is trained by the BP algorithm [26]. In the BP algorithm, signals propagate forward and errors propagate in reverse during the learning process [27]. The BP network structure consists of at least three layers: an input layer, hidden layer, and output layer, as shown in Fig. 2(b). The BP network learning model for this study was built using the neural network toolbox in the MATLAB R2019 software.

The process diagram of the BP network simulation model used in this study is shown in Fig. 2(c). The input sample data set was obtained from the pile hit experiments. The size of this data set was 17 × 7188 (7188 columns of data with 17 AE parameters).

The outputs were considered based on three target cases.

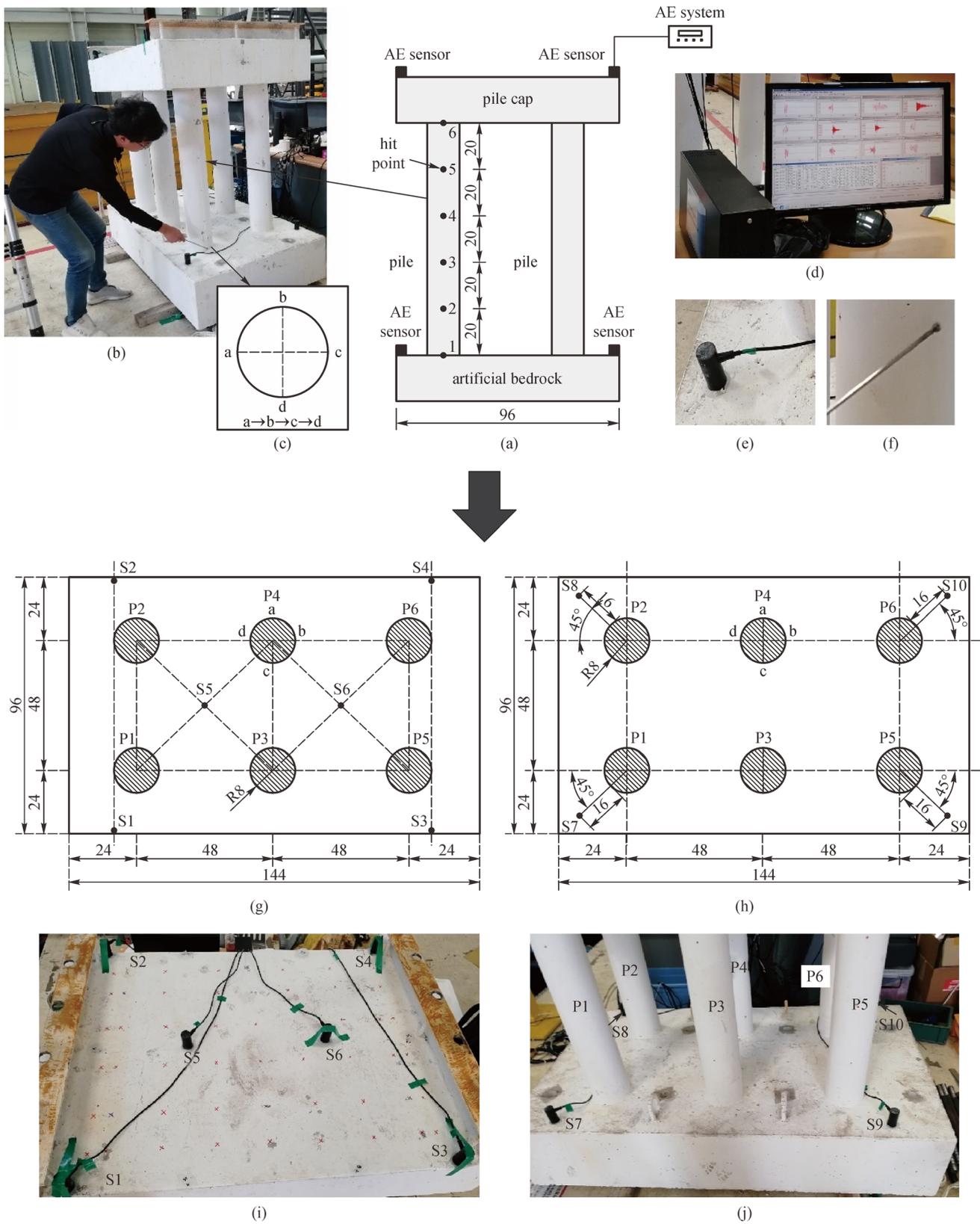


Fig. 1 Experimental setup for the pile hit tests. (a) Diagrammatic sketch of test setup (cm); (b) specimen and testing; (c) hit directions; (d) AE system; (e) AE sensor; (f) hammer; (g) diagrammatic sketch of top sensors installation (cm); (h) diagrammatic sketch of bottom sensors installation (cm); (i) installation of top sensors; (j) installation of bottom sensors.

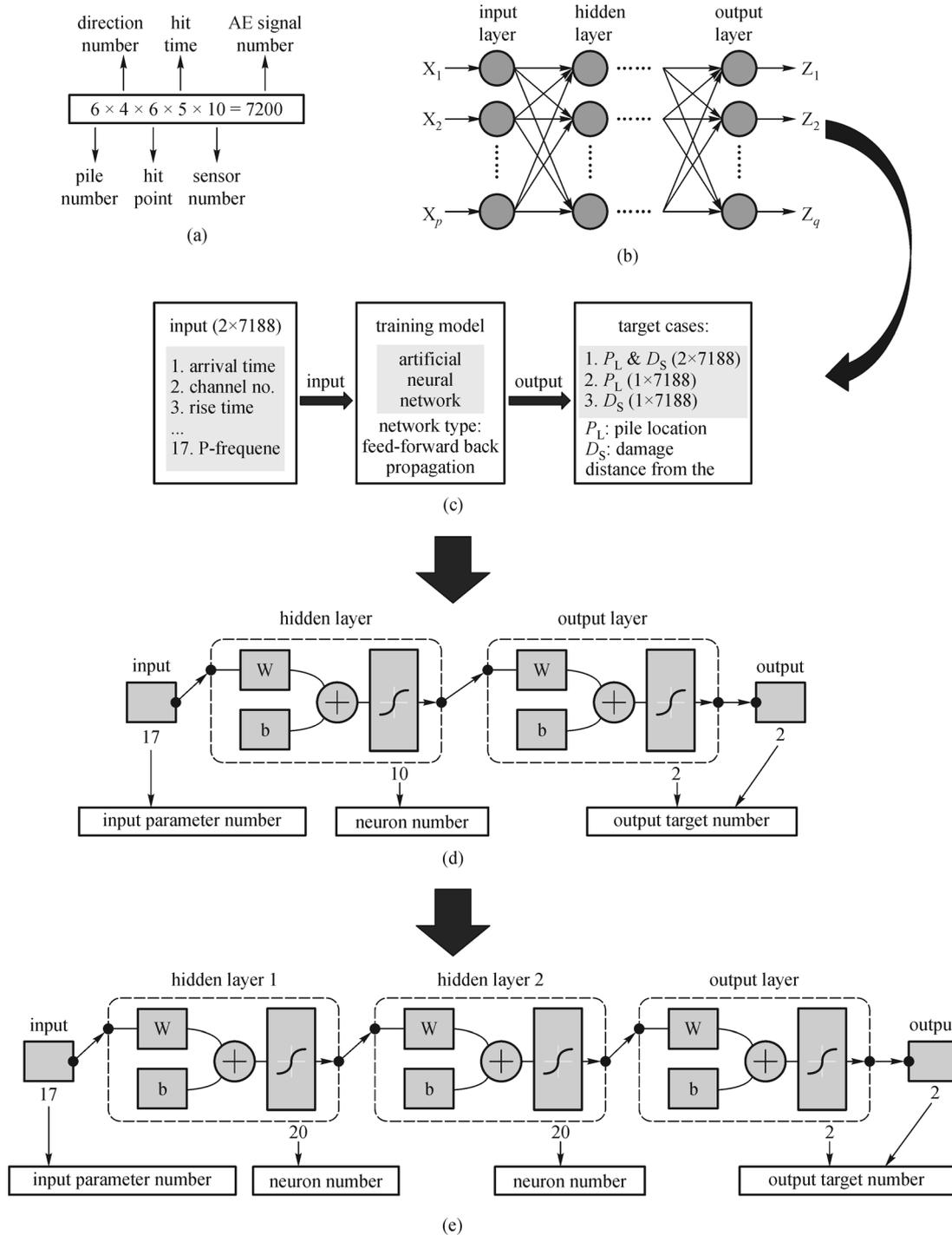


Fig. 2 BP neural network modeling. (a) Total number of data points obtained in the AE test; (b) topology of the BP neural network; (c) learning process of the BP neural network; (d) schematic concept of the three-layer BP neural network model; (e) schematic concept of the four-layer BP neural network model.

For Case 1, a data set with a size of $2 (P_L \text{ and } D_S) \times 7188$ was used to predict both P_L and D_S , where P_L represents the pile location, and D_S represents the damage distance from the pile cap. P_L was set as the output for Case 2, and D_S was considered in Case 3. All three case targets were output independently by three different training networks

with the same input sample data set. The network structure built in this study is shown in Fig. 2(d).

The feed-forward BP network learning model was selected in this study, and its specific details are listed in Table 2. The model included 14 different training functions. To select the best training function to achieve

Table 1 AE parameters with statistical values of the data set

no.	AE parameter	definition	minimum	maximum	mean	median	standard deviation
1	arrival time (μs)	time at which an AE wave reaches the sensor	1.2	39.5	17.6	17.6	9.6
2	channel no.	also referred to as the sensor number; different sensors are connected to the AE system by different channels	1.0	10.0	5.0	6.0	2.8
3	rise time (μs)	time between the first signal crossing the threshold and the maximum amplitude	60.0	15960.0	1754.3	1570.0	1279.5
4	counts	number of oscillations of the signal over the threshold	11.0	364.0	125.0	126.0	38.0
5	energy	area under the signal detection envelope	421.0	65535.0	16161.0	14245.0	10849.0
6	duration (μs)	time between the first and last threshold crossing	4140.0	56950.0	20619.0	19880.0	6746.0
7	amplitude (dB)	maximum voltage of the AE waveform	65.0	99.0	91.3	93.0	6.7
8	<i>RMS</i> (V)	root mean square (<i>RMS</i>) value of the detected signal	0.4×10^{-3}	0.7	0.2	0.1	0.1
9	<i>ASL</i> (dB)	average signal level (<i>ASL</i>) of the detected signal	18.0	79.0	64.6	66.0	7.2
10	counts to peak	number of threshold crossings from the first to maximum voltage	1.0	118.0	18.5	17.0	10.4
11	signal strength ($\text{pV}\cdot\text{s}$)	time integral of the absolute signal voltage	2.6×10^6	4.6×10^8	1.0×10^8	8.9×10^7	6.8×10^7
12	absolute energy (aJ)	time integral of the square of the unamplified signal voltage, expressed in attojoules	2.5×10^5	1.7×10^9	2.0×10^8	1.2×10^8	2.3×10^8
13	average frequency (kHz)	ratio of the counts to duration, divided by 1000	2.0	12.0	6.2	6.0	1.5
14	reverberation frequency (MHz)	(counts-counts to peak) divided by (duration-rise time)	1.0	11.0	5.3	5.0	1.5
15	initiation frequency (MHz)	counts to peak divided by the rise time	2.0	26.0	10.9	11.0	3.1
16	frequency centroid (kHz)	centroid of the power spectrum	5.0	21.0	10.2	10.0	2.3
17	peak frequency (kHz)	greatest power point of the spectrum	1.0	17.0	6.6	4.0	4.8

Table 2 Specific details of the BP network

network type	feed-forward BP
adaption learning function	LEARNDM
performance function	<i>MSE</i>
number of layers	2
transfer function	TANSIG

high accuracy, all of the functions were tested in the BP network learning model; Table 3 lists the regression values (*R*-values) for each function (a higher regression value indicates a higher accuracy of the neural network simulation model). The TRAINGLM training function, which had the highest *R*-value, was selected to build the BP network.

The TRAINGLM function is an artificial neural network training function in which the weight and basis values are updated via the Levenberg–Marquardt optimization. This function is regarded as the fastest BP algorithm for neural networks. Although it requires more memory than other

algorithms, it is strongly recommended to use TRAINGLM as the preferred supervision algorithm.

The number of hidden layers and number of neurons in the hidden layers will affect the prediction accuracy of a given BP network learning model. To investigate the best combination of hidden layers and neurons, the *R*-values were determined for varying numbers of hidden layers (1 or 2) and neurons (10, 20, or 30) with the TRAINGLM training function. Table 4 indicates that the combination of two hidden layers with 20 neurons in each layer provided the highest regression value (0.8076). The final BP network structure established in this study is shown in Fig. 2(e).

2.4 Verification of the damage location

2.4.1 Result validation

The BP neural network deep learning model was built according to Fig. 2(e) and trained using the test data set (size: 17×7188). The output of this training model was

designed as target case 1 (output: both P_L and D_S). The regression value of the training model was 0.8076. Twelve data groups (size: 17×12) were applied to validate the trained model, and the P_L and D_S values were predicted, as listed in Table 5. P_L and D_S were subsequently extracted from the output data separately, and a comparison of the predicted and actual values is shown in Fig. 3. It is clear that the predicted values of D_S are in good agreement with the actual values; however, this is not the case for the predicted P_L values, in which only three data groups (groups 5, 7, and 8) are in good agreement. As such, P_L cannot be accurately predicted by the trained BP network deep learning model.

Table 3 Comparison of different training functions for the BP network

No.	training function	regression value
1	TRAINBFG	0.7077
2	TRAINBR	0.7088
3	TRAINCGB	0.7028
4	TRAINCGF	0.6977
5	TRAINCGP	0.7038
6	TRAINGD	0.6902
7	TRAINGDM	0.6932
8	TRAINGDA	0.6787
9	TRAINGDX	0.6849
10	TRAIINGLM	0.7579
11	TRAINOSS	0.4741
12	TRAINR	0.6845
13	TRAINRP	0.7066
14	TRAINSSG	0.7114

Table 4 Comparison of varying numbers of hidden layers and neurons

hidden layers	10	20	30
1	0.7579	0.7825	0.7835
2	0.7737	0.8076	0.8051

2.4.2 Influences of individual parameters on the regression values

The AE data comprise 17 parameters. However, the influence of each parameter has yet to be verified to determine the weighting that will result in the highest regression value for the developed model. It is crucial to explore the impact weight of each parameter to achieve the BP network learning model with the best performance for damage location prediction. To determine the sensitivity of each AE parameter in the built network model, a single parameter was removed from the input data set to generate new simulation data sets (refer to Table 6).

The 14 cases with different input parameters are listed in Table 6; test case 1 was represents the group containing all parameters. The ‘rise time’ was removed from test case 2, and in a similar manner, ‘counts’ was removed from test case 3. Subsequently, different parameters were removed from each of test cases 4 to 14 in the same manner. In test case 12, the five types of frequency parameters (‘average frequency’, ‘reverberation frequency’, ‘initiation fre-

Table 5 Validation results for P_L and D_S

data group	$P_L^{(a)}$ (pile no.)		$D_S^{(b)}$ (cm)	
	actual	predicted	actual	predicted
1	1.00	3.02	20.00	5.38
2	1.00	2.04	80.00	80.55
3	2.00	2.85	40.00	33.80
4	2.00	4.19	100.00	95.89
5	3.00	3.17	0.00	0.33
6	3.00	3.68	60.00	57.34
7	4.00	4.31	20.00	21.44
8	4.00	3.98	40.00	29.01
9	5.00	2.71	80.00	87.48
10	5.00	1.05	100.00	99.98
11	6.00	5.25	0.00	1.10
12	6.00	3.83	60.00	50.18

Notes: a) P_L : pile location; b) D_S : damage distance from the pile cap.

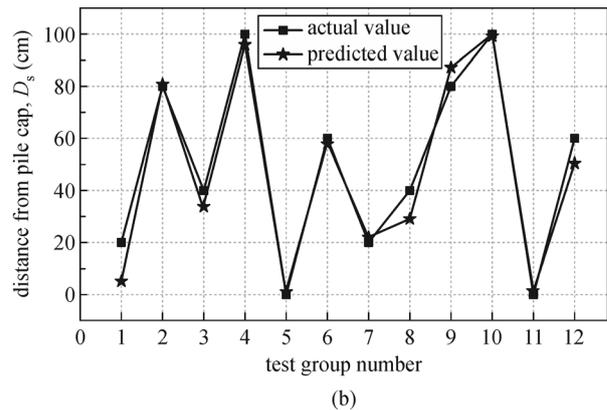
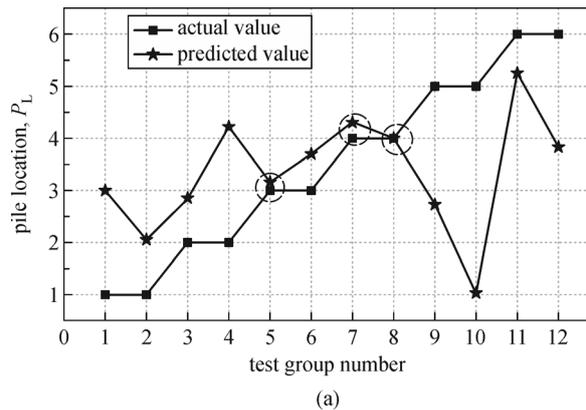


Fig. 3 Model validation results for P_L and D_S . (a) Validation results of P_L ; (b) validation results of D_S .

Table 6 Comparison of the test groups to evaluate parameter sensitivity

test case	input parameters	data set size	P_L			D_S		
			regression value	$MSE^{(a)}$	$RMSE^{(b)}$	regression value	MSE	$RMSE$
1	all	17×7188	0.5801	1.91	1.38	0.9814	43.87	6.62
2	remove (rise time)	16×7188	0.5950	1.93	1.39	0.9801	42.55	6.52
3	remove (counts)	16×7188	0.5628	1.97	1.40	0.9817	44.18	6.69
4	remove (energy)	16×7188	0.5272	2.01	1.42	0.9811	44.06	6.64
5	remove (duration)	16×7188	0.5857	1.92	1.39	0.9810	41.38	6.43
6	remove (amplitude)	16×7188	0.5554	1.98	1.41	0.9824	42.02	6.48
7	remove (RMS)	16×7188	0.5868	1.92	1.38	0.9809	44.14	6.64
8	remove (ASL)	16×7188	0.5750	1.96	1.40	0.9828	44.27	6.53
9	remove (counts to peak)	16×7188	0.5777	1.95	1.39	0.9821	41.38	6.43
10	remove (signal strength)	16×7188	0.6132	1.86	1.36	0.9800	44.53	6.67
11	remove (absolute energy)	16×7188	0.5858	1.92	1.39	0.9820	43.03	6.5594
12	remove frequencies	12×7188	0.5251	2.16	1.47	0.9804	44.78	6.69
13	remove (arrival time)	16×7188	0.5555	2.05	1.43	0.6931	641.46	25.32
14	remove (channel no.)	16×7188	0.5396	2.09	1.45	0.9808	43.52	6.60

Notes: a) MSE : mean square error; b) $RMSE$: root mean square error.

quency’, ‘frequency centroid’, and ‘peak frequency’) were removed together, as they have the same frequency characteristics and were thus regarded as a single parameter.

Figures 4(a) and 4(b) shows the regression plot for Case 1. The 7188 data groups in Case 1 were separated into three sets: a training set (70%), validation set (15%), and testing set (15%). Subsequently, a regression plot was generated for each set. In the regression plot of D_S , the regression values of the three sets were approximately the same ($R \approx 0.9800$). In the regression plot of P_L , the training set ($R = 0.6346$) performed better than the validation set ($R = 0.5009$) and test set ($R = 0.4749$).

In the training process for the neural network, more epochs of training were effective for reducing the error. However, as overfitting or underfitting problems occur in the neural network, the error in the validation set may begin to increase. In the default setup of the neural network applied in this study, the training stopped after six continual increases in the validation error. The best validation performance of the neural network was determined at the epoch with the minimum validation error. Figures 4(c) and 4(d) shows the performance of the neural network for Case 1. The training of the neural network for the prediction of D_S was stopped at epoch 32, and the best validation performance was determined at epoch 26. The training of the neural network for the prediction of P_L was stopped at epoch 31, and the best

validation performance was determined at epoch 25.

The mean square error (MSE) and root mean square error ($RMSE$) were used to evaluate the regression model in this study. MSE is a measure of the distance between a fitting line and a data point. $RMSE$ is the root of MSE . Smaller MSE and $RMSE$ values indicate a higher accuracy of the neural network simulation model. MSE and $RMSE$ are defined as follows [28]:

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2, \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \quad (2)$$

where $\hat{y} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ (predicted value), $y = \{y_1, y_2, \dots, y_n\}$ (true value), and $n =$ number of data groups.

The test results show that, with the exception of Case 13, the regression values for the prediction of D_S remained mostly high (≈ 0.9800 , refer to Fig. 5). For Case 13, the regression value of D_S dropped to 0.69 when the ‘arrival time’ was removed from the full data set; the MSE increased from 43.87 to 641.46, and the $RMSE$ increased from 6.62 to 25.32. This implies that the ‘arrival time’ is an important and sensitive parameter for D_S , more so than the other parameters. The regression values for P_L vary within a small range.

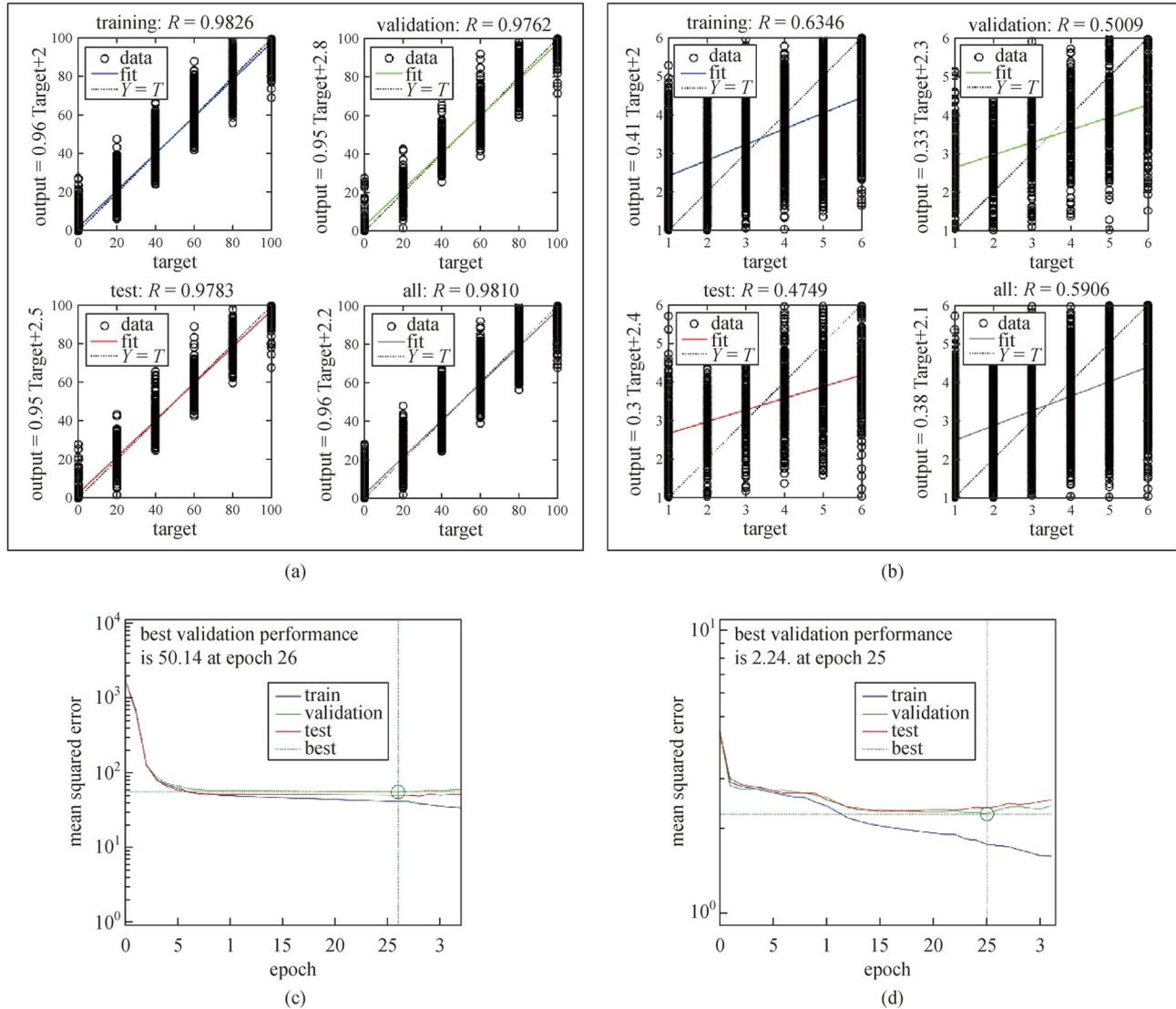


Fig. 4 Regression and performance plot for Case 1. (a) Regression plot for predicting D_S ; (b) regression plot for predicting P_L ; (c) ANN training performance for predicting D_S ; (d) ANN training performance for predicting P_L .

To analyze the possibility of using a single parameter to predict P_L or D_S , the BP neural network deep learning model was trained using each of the individual parameters. The BP neural network deep learning model was built as described previously, and a selected single parameter was set as the input parameter. In contrast to the previous tests presented in Table 6, only a single parameter was input for each case in this model with the same output target case (Table 7). Case 1 represents the data set containing all parameters for comparison, whereas the other cases represent the single input parameters; for example, only the ‘rise time’ was input in Case 2.

The validation results show that the regression value of P_L for Case 12 (i.e., frequency) was 0.35, whereas the values for the other cases were approximately 0.1, except for Case 1 (Table 7). This implies that the frequencies slightly affected P_L . The regression value of D_S was

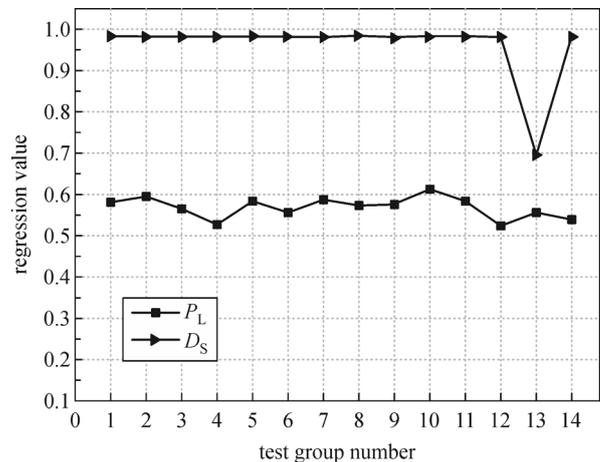


Fig. 5 Regression values for the prediction of D_S and P_L according to the test group number (refer to Table 6 for definitions of the test group numbers).

Table 7 Regression values with single parameters

test case	input parameters	data set size	output regression	
			P_L	D_S
1	all	17×7188	0.5801	0.9814
2	rise time	1×7188	0.0965	0.0868
3	counts	1×7188	0.0739	0.1140
4	energy	1×7188	0.1062	0.1257
5	duration	1×7188	0.1695	0.1448
6	amplitude	1×7188	0.0782	0.1179
7	<i>RMS</i>	1×7188	0.1158	0.0946
8	<i>ASL</i>	1×7188	0.1168	0.0984
9	counts to peak	1×7188	0.0692	0.0913
10	signal strength	1×7188	0.1114	0.1257
11	absolute energy	1×7188	0.1131	0.0940
12	frequency (A^a - R^b - I^c - C^d - P^e)	5×7188	0.3500	0.3291
13	arrival time	1×7188	0.1454	0.9760
14	channel no.	1×7188	0.0016	0.0012

Notes: a) A: average frequency; b) R: reverberation frequency; c) I: initiation frequency; d) C: frequency centroid; e) P: peak frequency.

similar to that in the previous test for Cases 1 and 13; however, the other cases exhibited low accuracy for D_S . The ‘arrival time’ was a key parameter for the prediction of D_S ; in other words, D_S can be predicted by a BP neural network deep learning model using the ‘arrival time’ of an AE signal. In contrast, there were no key parameters for the prediction of P_L , and even the use of all parameters cannot satisfy the accuracy requirements, as the regression value was limited to only 0.58.

2.4.3 Influences of the installation location and number of sensors

The influences of the installation location and number of sensors were explored using BP neural network deep

learning. Ten AE sensors were installed on the pile foundation specimen. Six sensors were installed on the top surface of the pile cap, which are defined as the ‘top sensors’. The remaining four sensors were installed on the artificial bedrock and are referred to as ‘bottom sensors’ (Fig. 1(j)).

As summarized in Table 8, the data sets gathered from each sensor were prepared as ten cases. Each data group was composed of 720 columns of data with 16 parameters. As there is no ‘channel no.’ in the single-sensor data sets, the number of parameters was reduced to 16 compared with the data sets for all sensors.

The ten data sets from the individual sensors were used to generate a new BP neural network deep learning model in the same manner as described previously. The neural

Table 8 Comparison of the results for test groups comprising individual channels

data set	sensor number	data size	P_L			D_S		
			regression value	<i>MSE</i>	<i>RMSE</i>	regression value	<i>MSE</i>	<i>RMSE</i>
	all channels	17×7188	0.5801	1.91	1.38	0.9814	43.87	6.62
S1	sensor 1	16×720	0.6762	1.65	1.28	0.9862	39.91	9.32
S2	sensor 2	16×720	0.6501	1.69	1.30	0.9832	37.49	6.12
S3	sensor 3	16×720	0.6209	1.79	1.34	0.9815	44.41	6.66
S4	sensor 4	16×720	0.6635	1.64	1.28	0.9840	38.57	6.21
S5	sensor 5	16×720	0.6917	1.54	1.24	0.9858	33.01	5.75
S6	sensor 6	16×720	0.6695	1.62	1.27	0.9853	34.69	5.89
S7	sensor 7	16×720	0.7926	1.08	1.04	0.9849	35.45	5.95
S8	sensor 8	16×720	0.7854	1.12	1.06	0.9847	36.24	6.02
S9	sensor 9	16×720	0.7957	1.09	1.05	0.9848	38.23	6.18
S10	sensor 10	16×720	0.7976	1.06	1.03	0.9833	39.19	6.26

network evaluation metrics, i.e., the regression, MSE , and $RMSE$ values, are listed in Table 8. The regression values for the prediction of P_L were significantly greater for the four bottom sensors: the average values for the top and bottom sensors were 0.6620 and 0.7928, respectively. The MSE and $RMSE$ for the prediction of P_L were significantly lower for the four bottom sensors: the average MSE values for the top and bottom sensors were 1.65 and 1.09, respectively, and the average $RMSE$ values for the top and bottom sensors were 1.29 and 1.05, respectively. For D_S , the average values of the evaluation metrics were maintained at the same level in both the top and bottom sensors. These results demonstrate that the installation location and number of sensors influence P_L but do not affect D_S (Fig. 6). Regardless of the number or location of the sensors used, the regression value of D_S remained almost constant.

Given these results, it seems likely that a single-sensor data set is sufficient for the prediction of D_S using the BP network deep learning model. For P_L , data sets from the individual channels resulted in varying degrees of improvement in terms of the regression value. The four bottom sensors installed on the artificial bottom bedrock resulted in significant improvements, with the regression value reaching 0.7976. The remaining six top sensors installed on the pile cap also improved the regression value, but were not sufficient to meet the prediction requirements.

2.4.4 Comparison of ‘time difference’ and ‘arrival time’

In this study, the ‘arrival time’ data were used for BP neural network deep learning. However, ‘arrival time’ does not refer to the actual time of the signal wave, as it includes the waiting time of each hit. Previous studies often used the ‘time difference’ of different sensors to calculate the location of the AE source [29,30]. To verify the influences

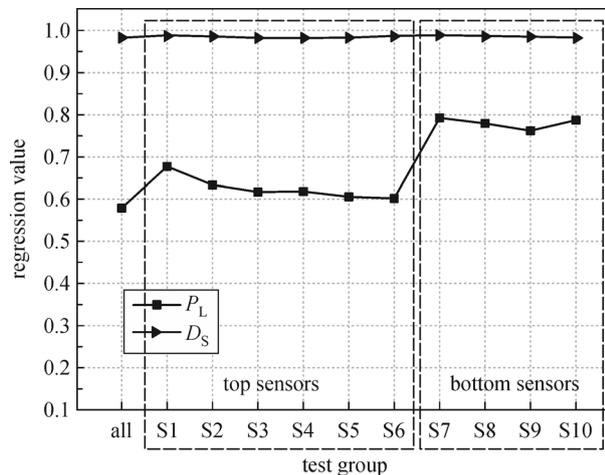


Fig. 6 Change in the regression value with different sensor locations (refer to Table 8 for definitions of the test groups).

of the ‘arrival time’ and ‘time difference’ on the regression values of the BP deep learning model, a comparison study was conducted.

The six top sensors were selected for this comparison study, and a deep learning data set was prepared. The size of the new data set was 17×4320 (4320 columns of data with 17 parameters). The ‘arrival time’ was used in the original group, and the ‘time difference’ was used in the comparison group. As the AE signals were received by the six sensors, the first ‘arrival time’ was set as the base time, and the ‘time difference’ was calculated for each sensor according to the base time. Here, the first ‘arrival time’ was set to zero.

The results of this comparison study shows that when the ‘time difference’ was used instead of the ‘arrival time’, the regression value of P_L increased from 0.5362 to 0.7472, whereas the regression value of D_S decreased from 0.9823 to 0.7198. This implies that using the ‘time difference’ can improve the accuracy of the BP neural network deep learning model for the detection of P_L . However, it has the opposite influence on the detection of D_S . Thus, separate BP neural network deep learning models should be built to detect P_L and D_S . The ‘arrival time’ should be used to detect D_S , and the ‘time difference’ should be used to detect P_L .

3 Damage step prediction

3.1 Data collection from failure experiments

To obtain experimental data for the damage step prediction with deep learning, a UCS test was performed using concrete specimens. As illustrated in Fig. 7, the test setup consisted of a compressive testing system and an AE system. Six sensors were installed on the cylindrical specimen, and the installation locations of the sensors are shown in Fig. 7(a). The cylindrical specimens (dimensions: 5 cm in diameter \times 10 cm in height) were placed on the loading cell and acted on by the compressive press with a loading rate of 0.15 mm/min (Fig. 7(c)).

Table 9 lists the physical properties of the specimens with different mixing ratios. The samples were prepared in three groups according to the mixing ratio (water: cement: sand): Group A, Group B, and Group C had mixing ratios of 1:2:0, 1:2:2, and 1:2:4, respectively. Each group included five different specimens with the features of UCS and P-waves.

3.2 Data set for the classification learner

The data set for the classification learner was built by combining all of the data acquired by the six installed sensors. This data set contained all of the parameters except for the ‘arrival time’ and ‘channel no.’ as these two parameters had no influence on the damage step classifica-

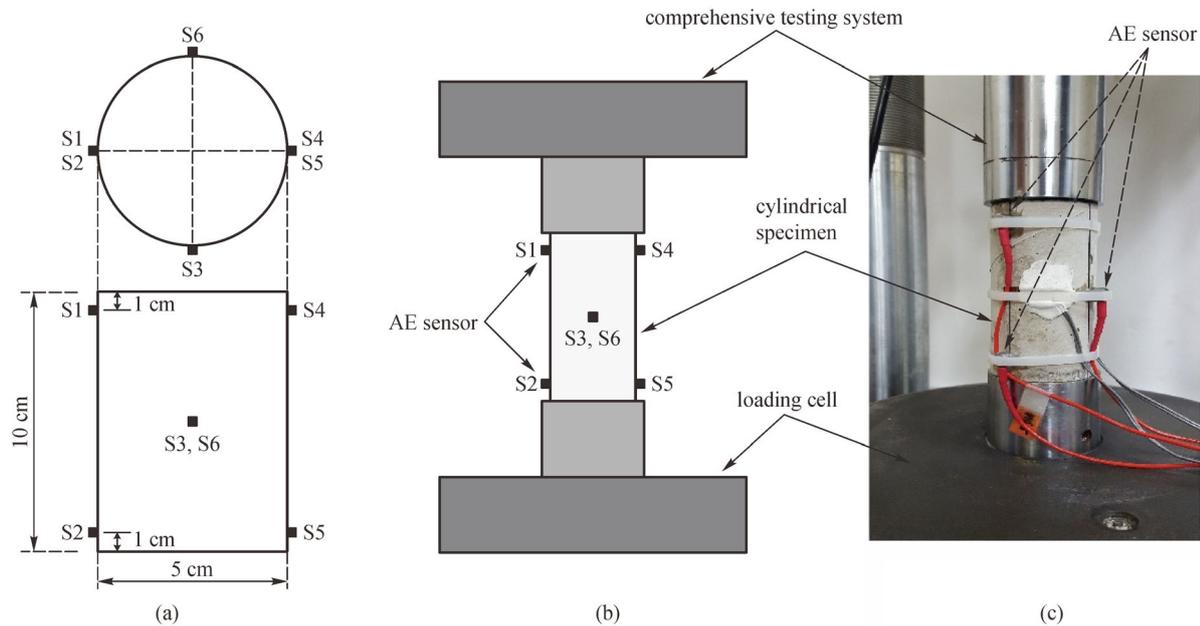


Fig. 7 Illustration of the UCS tests used to obtain AE signals. (a) Sensor installation; (b) test setup; (c) testing.

Table 9 Physical properties of the concrete specimens

group	sample	UCS (MPa)	P-wave velocity (m/s)	density (g/cm ³)	mixing ratio (water:cement:sand)
A	a-1	25.0	2861	1.59	W:C:S = 1:2:0
	a-2	26.8	2951	1.57	
	a-3	20.9	2171	1.67	
	a-4	28.8	4353	1.67	
	a-5	12.4	3800	1.71	
B	b-1	26.4	4519	2.02	W:C:S = 1:2:2
	b-2	19.3	4977	1.98	
	b-3	26.7	4895	1.96	
	b-4	28.7	5057	1.99	
	b-5	39.5	4949	2.02	
C	c-1	23.2	4910	2.03	W:C:S = 1:2:4
	c-2	21.9	4493	2.07	
	c-3	6.0	3968	1.95	
	c-4	15.3	5057	1.97	
	c-5	18.4	4811	1.99	

tion. Instead of ‘arrival time’ and ‘channel no.’, the UCS and P-wave velocity data were added to the classification data set to identify different types of concrete.

The data set was 166000×18 in size, i.e., it contained 166000 rows of data with 18 columns. Columns 1 to 17 are the observations to classify, and column 18 is the labeled layer. Each row represents one sample, and each column represents a different parameter. The labeled layer defines the damage step. The damage step was identified by the

stress ratio (applied stress/failure stress) acquired during the test (Table 10) [31].

Dimension reduction is effective for improving the classification accuracy of high-dimensional data. Neighborhood component analysis (NCA) is an effective dimension reduction method for high-dimensional data [32,33]. In this study, NCA was applied to reduce the dimensions of the data set.

Figure 8 shows the average values of the normalized

Table 10 Damage step identification

stress ratio	damage step	phenomenon
0%–35%	I	no crack generation
35%–97%	II	crack generation and expansion
97%–100%	III	structure failure

weights for the 17 AE parameters computed 50 times using the NCA. The normalized weights of nine AE parameters were greater than 0.01. Thus, the following parameters were determined to be effective parameters for the damage step: the ‘UCS’ (0.39), ‘P-wave velocity’ (0.52), ‘amplitude’ (0.65), ‘average frequency’ (0.29), ‘ASL’ (1), ‘reverberation frequency’ (0.54), ‘initiation frequency’ (0.25), ‘frequency centroid’ (0.68), and ‘peak frequency’ (0.42). The normalized weights of eight AE parameters were less than 0.01. These were determined to be ineffective parameters for the damage step: the ‘rise time’, ‘counts’, ‘energy’, ‘duration’, ‘RMS’, ‘counts to peak’, ‘signal strength’, and ‘absolute energy’. The eight ineffective parameters were removed from the data set. As

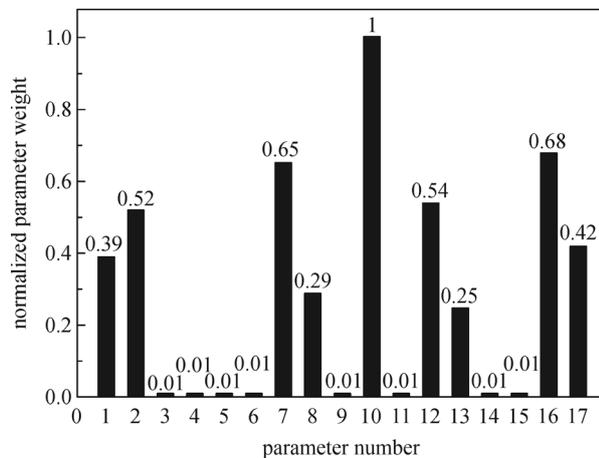


Fig. 8 Normalized weights of 17 AE parameters (1. UCS (MPa); 2. P-wave velocity (m/s); 3. rise time; 4. counts; 5. energy; 6. duration; 7. amplitude; 8. average frequency; 9. RMS; 10. ASL; 11. counts to peak; 12. reverberation frequency; 13. initiation frequency; 14. signal strength; 15. absolute energy; 16. frequency centroid; 17. peak frequency).

a consequence, the dimension of the data set was reduced from 17 to 9 parameters.

3.3 Prediction model

A classification learner is a type of supervised machine learning that offers learnable algorithms for classifying new observations based on given labeled data. The accuracy of the result can be provided directly by the classification learner, and the confusion matrix plot and receiver operating characteristic (ROC) curve can also be explored [34]. The ROC is a metric used to evaluate the output quality of the classifier. The ROC area refers to the area enclosed by the ROC curve, X axis, and Y axis. The range of the ROC area is [0,1]. The output quality of the classification depends on the ROC area: a higher value of the ROC area indicates a higher classification output quality.

Different classifiers have different performance for the prediction of the damage step. Table 11 lists the average value of the classification accuracy and ROC area for six different classifiers after performing the classification 50 times. The classification accuracy of the decision tree, support vector machine (SVM), and ensemble classifier were greater than 70%. The ensemble classifier has the highest accuracy (78.2%) compared with the other classifiers. The classification accuracies of the naive Bayes classifier, nearest neighbor classifier, and discriminant analysis were lower than 70%. Consequently, the ensemble classifier was determined to be an effective classifier for predicting the damage step.

3.4 Verification of the damage step

The classification accuracy of the ensemble classifier was 78.2% with 5-fold cross-validation. The confusion matrix of the ensemble classifier and the prediction results are shown in Fig. 9. The positive predictive value of damage step I was 83%, which is a reliable value that is of great importance in SHM because damage step I is the critical step for fixing structures. The positive predictive values of damage steps II and III were 75% and 77%, respectively. These values, although not ideal, can offer reference values for evaluating the damage steps, as both are greater than

Table 11 Comparison of different classifiers

classifier	tree	accuracy	ROC area
decision tree	fine tree	73.4%	0.86
SVM	quadratic SVM	72.7%	0.85
naive Bayes classifier	Gaussian naive Bayes	59.2%	0.74
nearest neighbor classifier	fine KNN ^{a)}	63.4%	0.72
ensemble classifier	bagged tree	78.2%	0.90
discriminant analysis	quadratic discriminant	64.2%	0.79

Notes: a) KNN: k-nearest neighbors.

		1	2	3
true class	1	83%	13%	5%
	2	15%	75%	18%
	3	2%	12%	77%
positive predictive value		83%	75%	77%
false discovery rate		17%	25%	23%
		1	2	3
		predicted class		

Fig. 9 Confusion matrix and prediction results.

50%. In summary, the trained classification learner can be used to evaluate the damage step of concrete structures; in particular, the classification learner can effectively evaluate damage step I.

4 Discussion: detection and evaluation method for pile foundations

Owing to the complexity of underground environments, pile foundations are vulnerable to unexpected damage. Thus, it is essential to perform continuous monitoring of such structures during their service lives. This is particularly important as pile foundations remain underground, and thus it is difficult to excavate the surrounding soil to monitor the foundations during their service life. This study provides a practical detection method that utilizes deep learning technology.

The detection method includes two steps (Fig. 10). The first step is a damage-simulating phase in which data are collected for varying damage locations and damage steps of the target structure. This step is carried out during the construction of pile foundations. The damage location simulation test can be carried out using hammer hits or an ultrasonic generator, and all potential source locations must be covered. Damage step simulation tests are carried out by conducting failure tests using target concrete specimens. Subsequently, the data sets for deep learning are produced after the two aforementioned simulation tests.

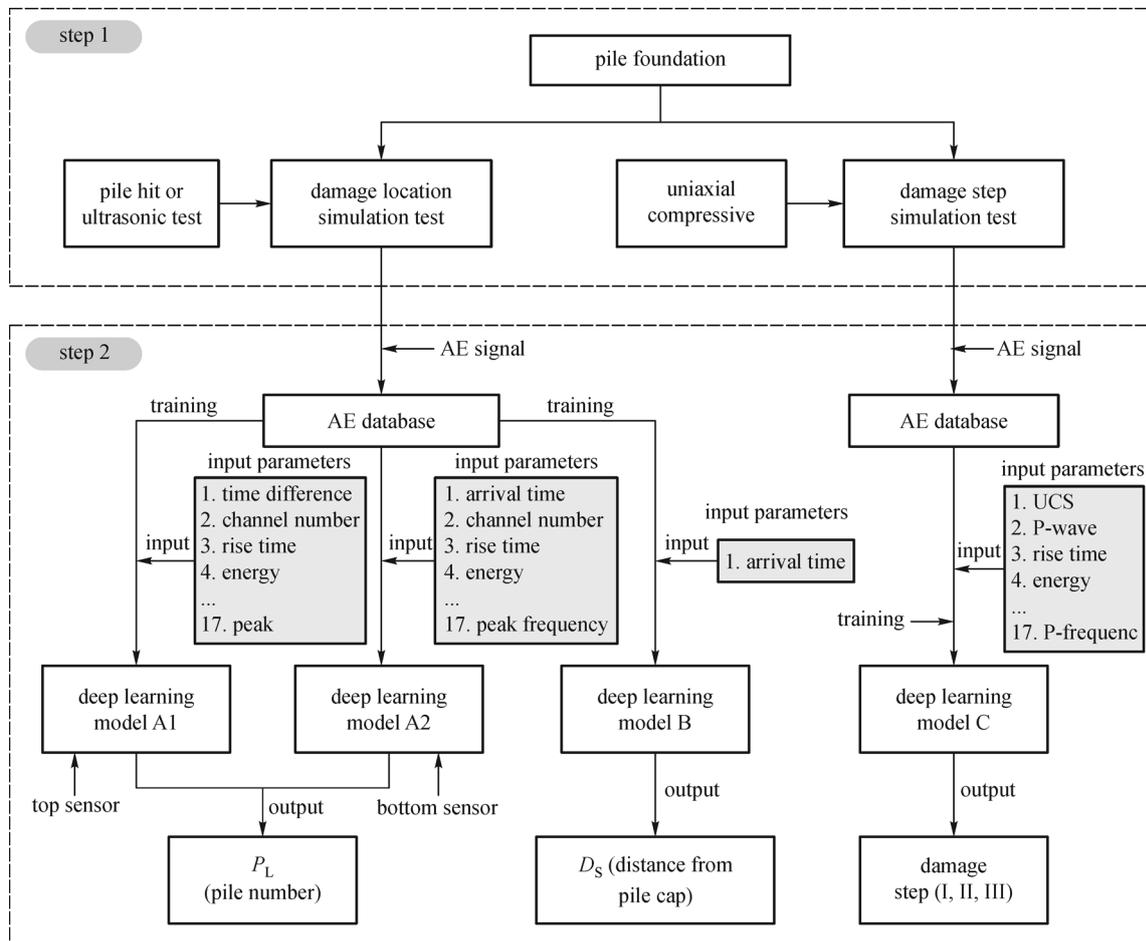


Fig. 10 Conceptual diagram of the process for detecting the damage locations and steps of pile foundations.

The second step is a deep learning training phase in which BP neural network deep learning models are used to predict P_L and D_S . Different BP networks are trained using the data sets obtained in the previous step.

Models A1 and A2 are provided for the prediction of P_L . The difference between models A1 and A2 is that the data set of the former is collected from the top sensors with ‘time difference’ as an input parameter, whereas the data set of model A2 is acquired from the bottom sensors and uses the ‘arrival time’ as an input parameter. Model B is used for the prediction of D_S . In this model, only the ‘arrival time’ is used as the input parameter, and there are no prerequisites for the sensor installation location. Model C is provided for the damage step evaluation.

The trained models can be used directly to detect the damage location of in-service pile foundations without the need for excavation, and can easily evaluate the damage using the classification learning model. The accuracy of D_S (model B) is very high and always satisfies the detection requirement. The accuracy of P_L cannot reach such accuracy levels but can offer reference values for health monitoring. The accuracy levels could be improved by using high-quality sensors or modified deep learning models.

5 Conclusions

This study performed experimental and analytical investigations to propose a new method for detecting and evaluating damage in pile foundations. The main findings are as follows.

1) The damage location was identified using the pile location (P_L) and the distance from the pile cap (D_S). It was found that P_L and D_S should not be predicted together by a single trained BP neural network deep learning model; specifically, two separate data set models are required to predict P_L and D_S .

2) The ‘arrival time’ is a key parameter for detecting D_S ; the BP neural network deep learning model achieves the same regression value regardless of whether all parameters are used or only ‘arrival time’ is used. Thus, using only ‘arrival time’ is sufficient for predicting D_S . For P_L , there is no key parameter that can be used to predict the value, unlike for D_S . Using the ‘time difference’ instead of the ‘arrival time’ significantly increases the regression value of P_L , but decreases the regression value of D_S . Thus, the ‘arrival time’ should be replaced by the ‘time difference’ when predicting P_L .

3) The regression value of D_S is unaffected by the number of sensors used. However, the number of sensors slightly influences P_L ; a single sensor can increase the regression value of P_L .

4) The installation locations of the sensors did not influence the regression value of D_S . On the other hand, the regression value of P_L increases significantly when using

the bottom sensors. This implies that bottom sensors are required to detect P_L with high accuracy.

5) The trained classification learner can be used to evaluate the damage steps of concrete structures.

Acknowledgements This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIT) (No. NRF-2019R1G1A1100517), the Fundamental Research Funds for the Central Universities (N170108029), the National Natural Science Foundation of China (Grant Nos. U1602232 and 51474050), and China Government Scholarship (201806080061); all of the above-mentioned funding sources and kind help are gratefully acknowledged.

References

1. Mao W, Aoyama S, Goto S, Towhata I. Acoustic emission characteristics of subsoil subjected to vertical pile loading in sand. *Journal of Applied Geophysics*, 2015, 119: 119–127
2. Mao W W, Aoyama S, Towhata I. Feasibility study of using acoustic emission signals for investigation of pile spacing effect on group pile behavior. *Applied Acoustics*, 2018, 139: 189–202
3. Mao W W, Towhata I, Aoyama S, Goto S. Grain crushing under pile tip explored by acoustic emission. *Geotechnical Engineering*, 2016, 47(4): 164–175
4. Mao W W, Aoyama S, Goto S, Towhata I. Behaviour and frequency characteristics of acoustic emissions from sandy ground under model pile penetration. *Near Surface Geophysics*, 2016, 14(6): 515–525
5. Shehadeh M F, Elbatran A H, Mehanna A, Steel J A, Reuben R L. Evaluation of acoustic emission source location in long steel pipes for continuous and semi-continuous sources. *Journal of Nondestructive Evaluation*, 2019, 38(2):40
6. Mao W, Yang Y, Lin W. An acoustic emission characterization of the failure process of shallow foundation resting on sandy soils. *Ultrasonics*, 2019, 93: 107–111
7. Madarshahian R, Ziehl P, Caicedo J M. Acoustic emission Bayesian source location: Onset time challenge. *Mechanical Systems and Signal Processing*, 2019, 123: 483–495
8. Ohtsu M. The history and development of acoustic emission in concrete engineering. *Magazine of Concrete Research*, 1996, 48(177): 321–330
9. Grosse C U. Acoustic emission (AE) evaluation of reinforced concrete structures. In: Maierhofer C, Reinhardt H-W, Dobmann G, eds. *Non-Destructive Evaluation of Reinforced Concrete Structures*. Oxford: Woodhead Publishing, 2010, 185–214
10. Li X Y, Li J L, Qu Y Z, He D. Gear pitting fault diagnosis using integrated CNN and GRU network with both vibration and acoustic emission signals. *Applied Sciences-Basel*, 2019, 9(4): 768
11. He M, He D. Deep learning based approach for bearing fault diagnosis. *IEEE Transactions on Industry Applications*, 2017, 53(3): 3057–3065
12. Guo W C, Li B Z, Shen S G, Zhou Q Z. An intelligent grinding burn detection system based on two-stage feature selection and stacked sparse autoencoder. *International Journal of Advanced Manufacturing Technology*, 2019, 103(5–8): 2837–2847
13. Sohaib M, Kim J M. Data driven leakage detection and classification

- of a boiler tube. *Applied Sciences-Basel*, 2019, 9(12): 2450
14. Carter A, Briens L. An application of deep learning to detect process upset during pharmaceutical manufacturing using passive acoustic emissions. *International Journal of Pharmaceutics*, 2018, 552(1–2): 235–240
 15. Samaniego E, Anitescu C, Goswami S, Nguyen-Thanh V M, Guo H, Hamdia K, Zhuang X, Rabczuk T. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. *Computer Methods in Applied Mechanics and Engineering*, 2020, 362: 112790
 16. Guo H, Zhuang X, Rabczuk T. A deep collocation method for the bending analysis of Kirchhoff plate. *Computers, Materials & Continua*, 2019, 59(2): 433–456
 17. Anitescu C, Atroshchenko E, Alajlan N, Rabczuk T. Artificial Neural Network methods for the solution of second order boundary value problems. *Computers, Materials & Continua*, 2019, 59(1): 345–359
 18. Ebrahimkhanlou A, Dubuc B, Salamone S. A generalizable deep learning framework for localizing and characterizing acoustic emission sources in riveted metallic panels. *Mechanical Systems and Signal Processing*, 2019, 130: 248–272
 19. Ebrahimkhanlou A, Salamone S. Single-sensor acoustic emission source localization in plate-like structures using deep learning. *Aerospace (Basel, Switzerland)*, 2018, 5(2): 50
 20. Jiao Y, Zhang Y, Shan W, Han Q, Zhao Y, Liu S. Damage fracture characterization of reinforced concrete beam subjected to four-point bending with parametric analysis of static, dynamic, and acoustic properties. *Structural Health Monitoring*, 2020, 19(4): 1202–1218
 21. Logoń D. Identification of the destruction process in quasi brittle concrete with dispersed fibers based on acoustic emission and sound spectrum. *Materials (Basel)*, 2019, 12(14): 2266
 22. Xu Y, Wei S, Bao Y, Li H. Automatic seismic damage identification of reinforced concrete columns from images by a region-based deep convolutional neural network. *Structural Control and Health Monitoring*, 2019, 26(3): e2313
 23. Jang K, Kim N, An Y K. Deep learning-based autonomous concrete crack evaluation through hybrid image scanning. *Structural Health Monitoring*, 2019, 18(5–6): 1722–1737
 24. Silva M, Santos A, Santos R, Figueiredo E, Sales C, Costa J C W A. Deep principal component analysis: An enhanced approach for structural damage identification. *Structural Health Monitoring*, 2018, 18(5–6): 1444–1463
 25. Liu Y, Zhu J J, Roberts N, Chen K M, Yan Y L, Mo S R, Gu P, Xing H Y. Recovery of saturated signal waveform acquired from high-energy particles with artificial neural networks. *Nuclear Science and Techniques*, 2019, 30(10): 148
 26. Tao H, Liao X, Zhao D, Gong X, Cassidy D P. Delineation of soil contaminant plumes at a co-contaminated site using BP neural networks and geostatistics. *Geoderma*, 2019, 354: 113878
 27. Zhu C, Zhang J, Liu Y, Ma D, Li M, Xiang B. Comparison of GA-BP and PSO-BP neural network models with initial BP model for rainfall-induced landslides risk assessment in regional scale: A case study in Sichuan, China. *Natural Hazards*, 2020, 100(1): 173–204
 28. Kim C H, Kim Y C. Application of Artificial Neural Network over nickel-based catalyst for Combined Steam-Carbon Dioxide of Methane Reforming (CSDRM). *Journal of Nanoscience and Nanotechnology*, 2020, 20(9): 5716–5719
 29. Zhou Z, Zhou J, Dong L, Cai X, Rui Y, Ke C. Experimental study on the location of an acoustic emission source considering refraction in different media. *Scientific Reports*, 2017, 7(1): 7472
 30. Mu W L, Zou Z X, Sun H L, Liu G J, Wang S J. Research on the time difference of arrival location method of an acoustic emission source based on visible graph modelling. *Insight (American Society of Ophthalmic Registered Nurses)*, 2018, 60(12): 697–701
 31. Lee J W, Kim H, Oh T M. Acoustic emission characteristics during uniaxial compressive loading for concrete specimens according to sand content ratio. *KSCE Journal of Civil Engineering*, 2020, 24(9): 2808–2823
 32. Aydemir E, Tuncer T, Dogan S. A Tunable-Q wavelet transform and quadruple symmetric pattern based EEG signal classification method. *Medical Hypotheses*, 2020, 134: 109519
 33. Ayyıldız H, Arslan Tuncer S. Determination of the effect of red blood cell parameters in the discrimination of iron deficiency anemia and beta thalassemia via Neighborhood Component Analysis Feature Selection-Based machine learning. *Chemometrics and Intelligent Laboratory Systems*, 2020, 196: 103886
 34. Bonah E, Huang X, Yi R, Aheto J H, Osae R, Golly M. Electronic nose classification and differentiation of bacterial foodborne pathogens based on support vector machine optimized with particle swarm optimization algorithm. *Journal of Food Process Engineering*, 2019, 42(6): e13236