Article

Three-Level Distributed Real-Time Monitoring of Construction near Underground Infrastructure Using a Combined Intelligent Method

Biao Zhou [1,2,*], Yingbin Gui [1,2], Xiaojian Wang [1,2] and Xiongyao Xie [1,2]

1 Key Laboratory of Geotechnical & Underground Engineering, Ministry of Education, Tongji University, Shanghai 200092, China; 1932597@tongji.edu.cn (Y.G.); seanwxj@tongji.edu.cn (X.W.); xiexiongyao@tongji.edu.cn (X.X.)
2 Department of Geotechnical Engineering, Tongji University, Shanghai 200092, China
* Correspondence: zhoubiao@tongji.edu.cn

Abstract: With the rapid development of underground infrastructure and the uncertainty of its location, the possibility of damage due to nearby construction has increased. Thus, for the early warning of dangerous construction behaviors around underground facilities, this paper proposes a novel real-time distributed monitoring method with three levels, comprised of the terminal node, relay node, and server. Corresponding to these three monitoring levels, a vibration-based intelligent solution for recognizing the construction source is presented and compared with the traditional method. First, the blind source separation method was used to separate collected signals into a limited number of monitoring object sources; this helped to minimize the number of required classification categories and reduce the recognition uncertainty caused by signal mixing. Then, the mutual information (MI) method was used to select suitable vibration features, which were used as the input matrix for the resulting intelligent recognition. Finally, the construction behaviors were identified at the server based on returned features. Guided by this method, a sample dataset including pile-driving, train-operation, and environment-vibration signals was constructed and combined with a multi-layer perceptron (MLP) and a long short-term memory (LSTM) network. The effects of blind source separation and the MI method are discussed in depth in this paper.

Keywords: underground infrastructure; construction monitoring; real-time; mutual information; blind source separation

1. Introduction

As the density of subway and underground pipe networks increases with the rapid development of urban construction, the possibility of damage to these infrastructures owing to nearby construction also increases. Pile foundation construction is among the most common construction behaviors that can affect underground infrastructure. In addition, the uncertainty of the location of existing infrastructures and the randomness of construction behavior has caused several accidents in the past. Indeed, there have been several incidents in the Chinese cities of Chengdu, Shenzhen, and Shanghai, where subway tunnel structural damage has occurred as a result of nearby pile construction, including a pile head hitting an operating train, causing permanent tunnel damage. Damage to underground pipelines is much more common and has resulted in gas pipe explosions in Qingdao, China. Therefore, the development of a real-time monitoring system for nearby construction is of great significance to prevent the destruction of important underground facilities.

However, this is a considerably challenging task. First, most underground facilities have become increasingly networked. For instance, subway tunnels are extremely long linear structures, with the length of a single line typically exceeding 10 km; these lines also pass through different areas. Hence, given the limited monitoring points, realizing early warnings for the damage caused by pile driving and other surrounding construction behaviors is

remarkably difficult. In the past, underground infrastructure monitoring has been primarily concerned with structural deformation or leakage, and related settlement monitoring systems have been developed using the static level or optical fiber test methods [1,2]. However, these methods can only cover significantly small areas, and it is difficult to conduct real-time monitoring and realize early warnings for nearby constructions. As construction behaviors, such as pile driving, result in structural soil and tunnel vibrations, a number of vibration or sound source perception and detection methods have been widely applied with seismic [3], noise [4], and acoustic emission (AE) [5] applications; the vibration detection approach can provide the real-time monitoring of nearby construction over a large area.

Second, apart from the vibrations caused by nearby construction, underground spaces comprise various other vibration sources; these mainly include train-induced vibrations and various environmental noises. The interaction and superposition of such vibration signals can generate new recognition categories, which, in turn, can considerably increase the difficulty in recognition. Furthermore, most previously proposed sound- and vibration-based monitoring methods are dependent on factors such as wave propagation properties [6,7] and modal characteristics [8–10]. However, in-depth analyses of such characteristics require complex and time-consuming calculations such as waveform-based seismic location methods [11] including full waveform inversion, which depends on complex inversion or reverse imaging algorithms and requires powerful computing resources. Furthermore, previous studies have shown that vibration responses and features are characterized by strong nonlinearity and uncertainty due to the structure and soil interference, especially when applying vibration-based methods in underground infrastructures [12,13]. The aforementioned characteristics and features may change depending on the variations in train speeds, soil properties, tunnel structures, and construction locations; hence, it is difficult to determine specific evaluation ranges for automatic judgments without human intervention. As a result, traditional signal processing methods are unable to directly conduct continuous and stable information perception and online monitoring.

Recently, a number of deep learning networks have been rapidly developed and applied in medical and mechanical fields, such as disease and human activity pattern recognition [14–16], and optical and medical signal diagnosis [17]. Deep learning networks have also been employed in applications closer to the focus of this study, such as bridge monitoring data analysis [18], truck loading detection [19], indoor fall detection for the elderly [20], and construction monitoring using sound classification [21]. However, these methods have all encountered various challenges, including the complexities of potential classification types and the need for large sample sets. Considering the difficulties in sample collection and the cases involving signal mixtures, it is necessary to develop new models or employ additional algorithms.

Meanwhile, rapid progress in blind source separation methods [22–25] has made it possible to develop more efficient models for use in vibration signal classification and recognition. Among these methods, mutual information (MI) analysis can be used to determine the degree of uncertainty associated with the assigned classification category by using several quantified features, such as information entropy [26]. Indeed, Shannon's MI-based theories [27] have been widely employed in applications such as the selection of multi-channel electroencephalogram features [28]. Furthermore, the combination of MI-based methods and intelligent algorithms has been demonstrated to extend the depth and breadth of feature mining; this approach has achieved continuous progress in recognition accuracy [29,30].

Therefore, considering the challenges associated with intelligent recognition owing to vibration mixing and the difficulties in signal acquisition, a novel real-time monitoring method is proposed. The blind source separation, MI analysis, and deep learning algorithms were combined to develop a limited-sample driven vibration perception and nearby construction classification method. This method can achieve real-time monitoring and early warnings for surrounding construction near underground facilities. Notably, the use of vibration recognition differs from the use of sound recognition in that the former can also reflect the

physical characteristics of underground media, and as such, it can be extended in the future to monitor excavations and other construction behaviors, in addition to pile driving.

2. Proposed Vibration Recognition Method

2.1. Outline of Proposed Method

As pile foundation construction is the most common construction behavior affecting underground infrastructure, this study investigated the vibration monitoring method for pile foundation construction near an operating subway tunnel. As shown in Figure 1, based on the distributed layout principle, a novel real-time monitoring method is proposed for the perception of surrounding construction behaviors and early warnings near super-long tunnel structures. This approach mainly consists of three levels, as follows:

- Terminal node for data acquisition and processing

To establish a sensor network covering the entire tunnel network, a super-long subway tunnel is divided into several monitoring sections. At each section, double sensors are symmetrically arranged along the tunnel structure; the section intervals are determined according to the sensing range of the acceleration sensors. For pile-driving vibration monitoring, this interval is typically 200 m; therefore, for a 1 km-long tunnel section, four or five test sections are sufficient. Accordingly, the total number of sensors can be less than ten by arranging two sensors at each section. If necessary, a processing module can be installed on the terminal node for signal sampling and signal processing, in order to reduce data transmission.

- Relay node for mutual analysis and feature screening

In each test section, a relay node is arranged as a terminal server; the connection with the terminal node can be wired or wireless. Moreover, the MI of the sampled signals between terminal nodes is further explored; time or frequency features will be processed here to meet the demands of source recognition. This will also further reduce the data transmission from the relay node to the server.
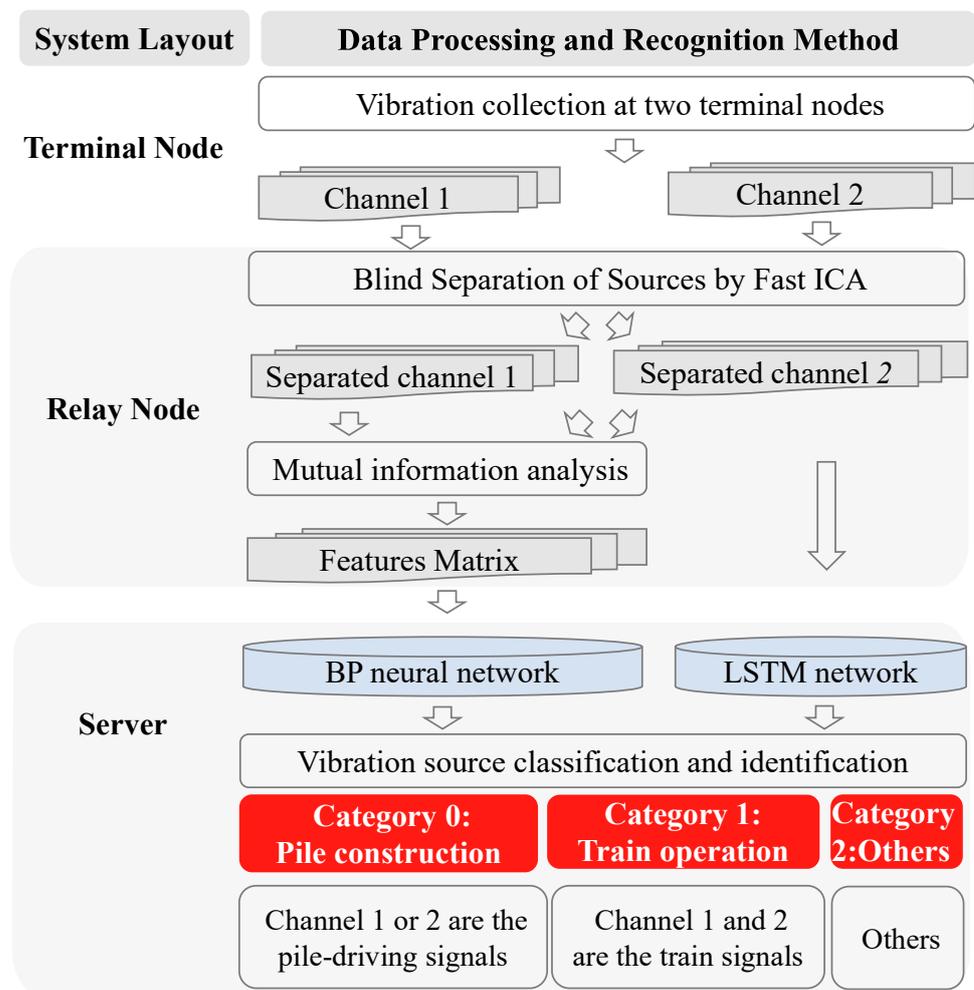
- Intelligent recognition of construction behavior at server

Finally, the processed data at the relay node are sent to the server via wireless transmission or optical fibers; the intelligent algorithm is then used to identify construction behaviors and provide warnings, if required.
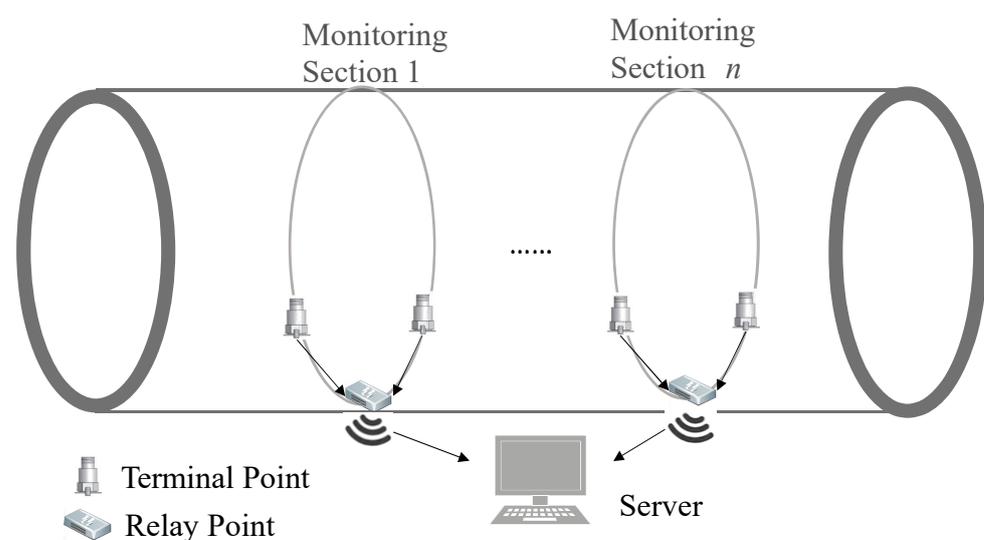
Figure 1. Vibration-based real-time monitoring system for construction behaviors surrounding super-long tunnel structures.

Corresponding to the three levels of the monitoring system, an intelligent solution for data processing is proposed herein, as shown in Figure 2.

System Layout	Data Processing and Recognition Method

Terminal Node

Vibration collection at two terminal nodes

⇩

Channel 1　　　Channel 2

⇩　　　⇩

Blind Separation of Sources by Fast ICA

Relay Node

⇧⇩

Separated channel 1　　　Separated channel *2*

⇩　　　⇧⇩

Mutual information analysis

⇩

Features Matrix　　⇩

⇩

Server

BP neural network　　　LSTM network

⇩　　　⇩

Vibration source classification and identification

Category 0: Pile construction	**Category 1: Train operation**	**Category 2:Others**
Channel 1 or 2 are the pile-driving signals	Channel 1 and 2 are the train signals	Others

Figure 2. Flow chart describing the two evaluated methods for the intelligent recognition of nearby construction behavior.

- Signal collection and sampling at terminal node

To meet the requirements of the signal acquisition for subsequent blind source separation, the number of sensors needs to be greater than the number of signal types to be identified, i.e., more than two. Here, by taking two single-channel sensors as an example, two channels of signals can be obtained at the terminal point, and the signals will be further sampled by time windows of 10 s. Then, if the peak value of the sampled signal exceeds the environment levels, the data will be transmitted to the relay point.

- Blind source separation and MI analysis at relay node

After the sampled signal enters the relay node, the blind source separation technology is first applied to separate it into two channel signals with a limited number of monitoring object sources, including pile-driving vibration, train-induced vibration, and others; it will minimize the number of required classification categories. The details are introduced in Sections 2.2 and 3.2.

Meanwhile, to improve the robustness and training efficiency of the intelligent recognition algorithm, the MI is applied at the relay node to analyze the correlation between the vibration features and the recognition objects; features with high correlation are then used to form a feature matrix that is transmitted back to the server. This matrix is employed by the intelligent network to reduce the input matrix dimensions and improve the system robustness and efficiency.

- Vibration source classification and construction behavior intelligent recognition at server

Finally, the separated two channels signals are transmitted back to the server, and the intelligent algorithm is used to determine the vibration signal category. Category 1 is the train operation case, with both channels of data determined to be train-induced signals. The second category is the determination of whether any channel data is determined to be pile-driving signals; in this case, the early warning is announced at once. If any channel does not contain a train response and a pile-driving signal, this situation will be classified as "others".

To highlight the advantages of combining the MI algorithm, the MLP network is used for feature and signal source recognition; this approach is here termed as the MI-MLP method. Most traditional methods adopt similar solution methodologies, where the separated signal is directly transmitted to the server. For a comparison of the proposed and existing methods, the commonly used long short-term memory (LSTM) method is employed to determine the vibration signal categories.

2.2. Vibration Signal Blind Source Separation

In the scenario investigated in this study, the main vibration types include train operation-induced vibration and pile foundation construction-induced vibration. Therefore, an independent component analysis (ICA) algorithm based on negative entropy is applied, as it can effectively separate relatively independent signals according to their sources [31]. As negative entropy can provide non-Gaussian measurements of random variables, it can realize improved computational efficiency and robustness. Therefore, the FastICA algorithm based on negative entropy was employed in this study, as follows [32]:

$$J(y) = H(y_{Gauss}) - H(y) \tag{1}$$

where $J(y)$ represents the negative entropy of the random variable y, $H(y)$ represents the entropy value of the random variable y, and y_{Gauss} represents a random variable with the same variance as that of y that conforms to the Gaussian distribution. According to the definition of entropy, the solution of negative entropy must be able to calculate the probability density of random variables. In actual use, the mean negative entropy value can be approximately expressed as

$$J(y) = \{E[G(y)] - E[G(y_{Gauss})]\}^2 \tag{2}$$

where G is a non-quadratic function, generally defined by

$$G_1(y) = \frac{1}{k} \log[\cosh(ky)], \ 1 \leq k \leq 2 \tag{3}$$

$$G_2(y) = -\exp(\frac{y^2}{2}) \tag{4}$$

The goal of the FastICA algorithm is to determine the $y = Wx$ value that maximizes Equation (2). This optimization employs the Newton iterative method using the following procedure:

1. Generate a random unmixing matrix W_0 that satisfies $\|W_0\| = 1$;
2. Assume $W_{k+1} = E[zG(W_k^T z)] - E[G'(W_k^T z)]W_k$ and iterate k;
3. Standardize and make $W_{k+1} = W_{k+1}/\|W_{k+1}\|$;
4. When the convergence condition is verified according to $|W_{k+1} - W| < \varepsilon$, end the cycle step, and end the algorithm after outputting the final unmixing matrix to obtain the solution of independent component $y = Wx$; otherwise, return to step 2 and continue with the iterations.

2.3. Mutual Information Analysis Method

The MI method has proven to be an effective tool to measure interdependence by calculating the shared information between two random variables. In the proposed method,

it is used to select independent features to form the input matrix of the MLP neural network. According to the MI analysis theory, the MI value between two random variables X and Y is defined as

$$I(X,Y) = \sum_{y_i \in Y} \sum_{x_i \in X_j} p(X_j = x_i, Y = y_i) \log_2 \left(\frac{p(X_j = x_i, Y = y_i)}{p(X_j = x_i)p(Y = y_i)} \right) \tag{5}$$

where $p(X_j = x_i)$ denotes the probability that X_j is equal to x_i, $p(Y = y_i)$ denotes the probability that Y is equal to y_i, and $p(X_j = x_i, Y = y_i)$ indicates the probability of both variables simultaneously matching their counterparts.

2.4. Intelligent Recognition Method

2.4.1. MLP Neural Network

Two intelligent signal source recognition methods were applied and compared in this study. With regard to the first method, as shown in Figure 3, the MLP recognition network used in this study is a fully connected MLP neural network employing error back propagation; it was constructed based on the TensorFlow module of Python. The sigmoid function is used as the activation function, and the cross-entropy function is used as the loss function. The input p of the network is an m-dimensional feature vector obtained through the MI. The weight w_i in the hidden layer is an $m \times n$ matrix, while the bias b_i is an n-dimensional feature vector.
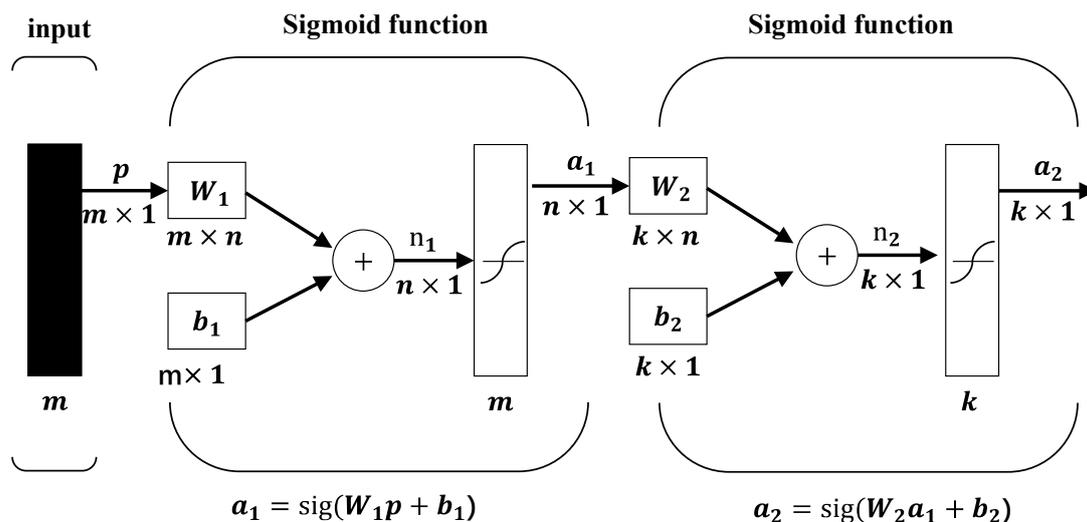

$$a_1 = \text{sig}(W_1 p + b_1) \qquad\qquad a_2 = \text{sig}(W_2 a_1 + b_2)$$

Figure 3. MLP neural network model structure.

2.4.2. LSTM Method

The LSTM [33], illustrated in Figure 4, offers the advantage of being able to solve the gradient instability problem associated with recurrent neural networks, thus making it superior to the conventional MLP model. This is accomplished by adding a new cell σ and a gating system in memory module A [34]. Accordingly, the training of the LSTM employs the back propagation through time algorithm, which propagates the error term back along time and layers [35,36].
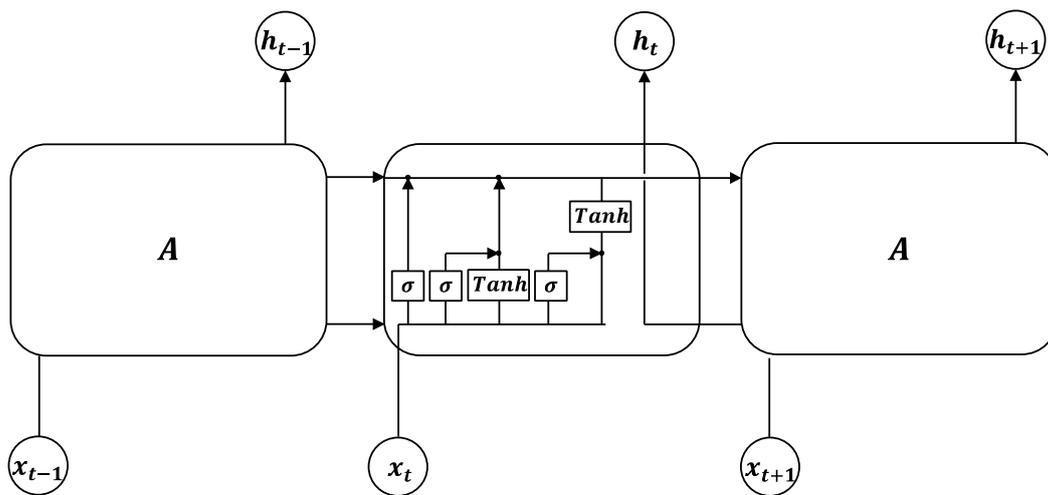
Figure 4. Internal structure of the LSTM computing layer.

3. Vibration Signal Separation and Data Preparation

3.1. Vibration Measurement and Data Preparation

For obtaining the vibration signals required for the subsequent feature analysis and intelligent recognition, considering the difficulty in acquiring mixed vibration signals from an operational subway tunnel, field tests were only conducted at two sites in order to collect train-induced and pile-driving vibration signals.

3.1.1. Pile-Driving-Induced Signals

One of the field tests was conducted in a power tunnel to collect the vibration signals induced by a nearby pile foundation construction. This tunnel was a pipe jacking tunnel with an outer diameter of 3.2 m and an inner diameter of 2.7 m. In the nearby construction, a steel pipe pile was being driven using the resonance hammer method. Six accelerometers were installed on the interior wall of three tunnel sections in order to collect the tunnel vibration responses in the radial direction at a sampling frequency of 1000 Hz. Overall, after being sampled with a time window of 10 s, 14,404 groups of pile-driving vibration signals with different distances from the power tunnel were obtained via this field test.

3.1.2. Train-Operation-Induced Signals

Another field test was used to collect train-operation-induced vibrations; this test was conducted in Shanghai metro line 12, which is a shield tunnel with an outer diameter of 6.2 m and an inner diameter of 5.7 m. Eight accelerometers were arranged on the track bed and side walls of four different tunnel sections at intervals of 12 m. The accelerometers collected data at 1000 Hz for three months. Finally, the collected data were sampled using a time window of 10 s, and a total of 29,592 groups of signals were prepared for the following intelligent recognition.

3.1.3. Others

For training and testing the classification of data other than pile-driving and train-vibration signals, the environment vibration data collected at the above two sites and other power and metro line tunnels were used, including the vibration response caused by ground traffic and human activities in the tunnel. Finally, 10,286 groups of signals were prepared.

3.2. Signal Separation Using FastICA and Comparison

As described in Section 2, the FastICA method was employed in this study to separate the mixed signals and thus minimize the number of classification categories. In this section, to verify the effect of separation, the mixed and single-type signals were prepared for

separation. The mixed signals were obtained by combining the two types of signals with different weights.

3.2.1. The Separation of Train Operation and Pile-Driving Mixed Signals

Since it is difficult to obtain the measured mixed signal, the pile-driving- and train-operation-induced signals were therefore combined with different weights to form two-channel mixed signals; the mixture and separation results are shown in Figure 5.
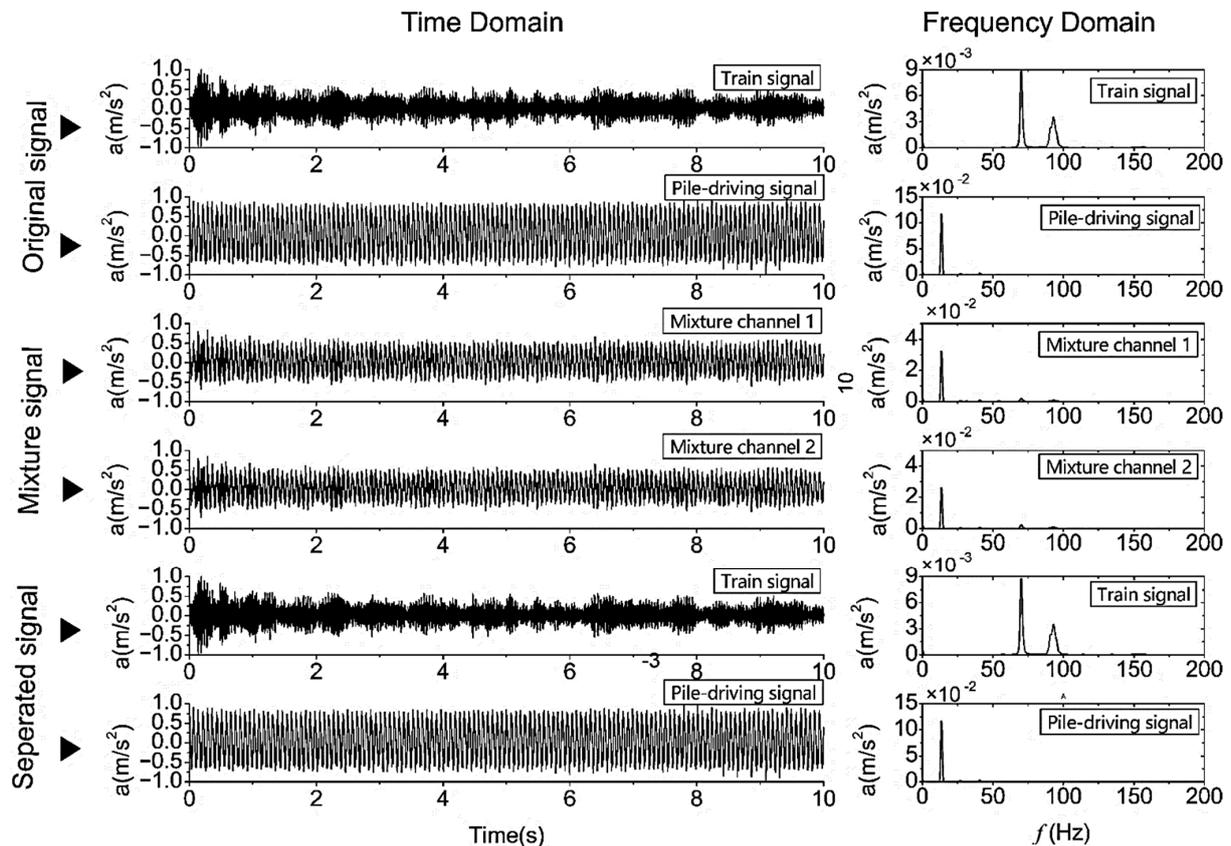

Figure 5. Example of mixed signals and separation using the FastICA method.

From Figure 5, two signals of train-operation-induced and pile-driving signals were combined by two random weights between 0.35 and 0.5, forming two signal mixtures, mixture channel 1 and mixture channel 2. Then, by using the FastICA-based separation algorithm, it found that only two separated signals can be obtained from the mixed signals.

As shown in Figure 5, before and after signal separation, the time domain waveform and frequency components are well preserved. It shows that the mixed signal will be well restored to the source signal, so as to reduce the number of classifications.

3.2.2. The Separation of Single-Type Signals

More commonly, the input two-channel signals are of a single type, either the train-operation-induced or the pile-driving-induced signals. Because the time and frequency characteristics of train-operation-induced signals are more sophisticate than those of pile-driving signals, for illustrating the separation of a single-type signal, two channels of train-operation-induced signals will be used as an example. The results are shown in Figure 6. It was also found that the waveform and frequency components are well preserved before and after separation. This means that, for a single signal, blind source separation can retain the characteristics of the original signal.
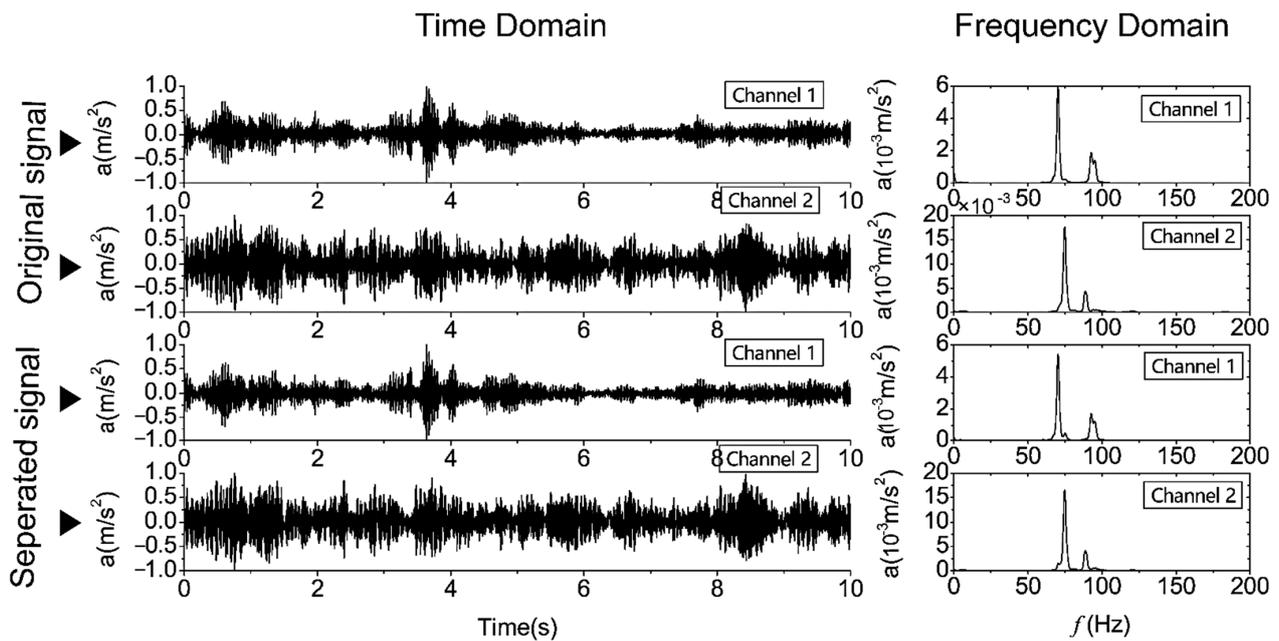
Figure 6. Example of two-channel single-type signal separation using the FastICA method.

For further comparison of the advantages of the blind source separation method, the HHT method was used [37]. The results in terms of performing empirical mode decomposition (EMD) on the mixed signal are presented in Figure 7; ten IMFs can be identified using EMD. Additionally, the time-domain characteristics are significantly different from those of the original signal prior to synthesis.
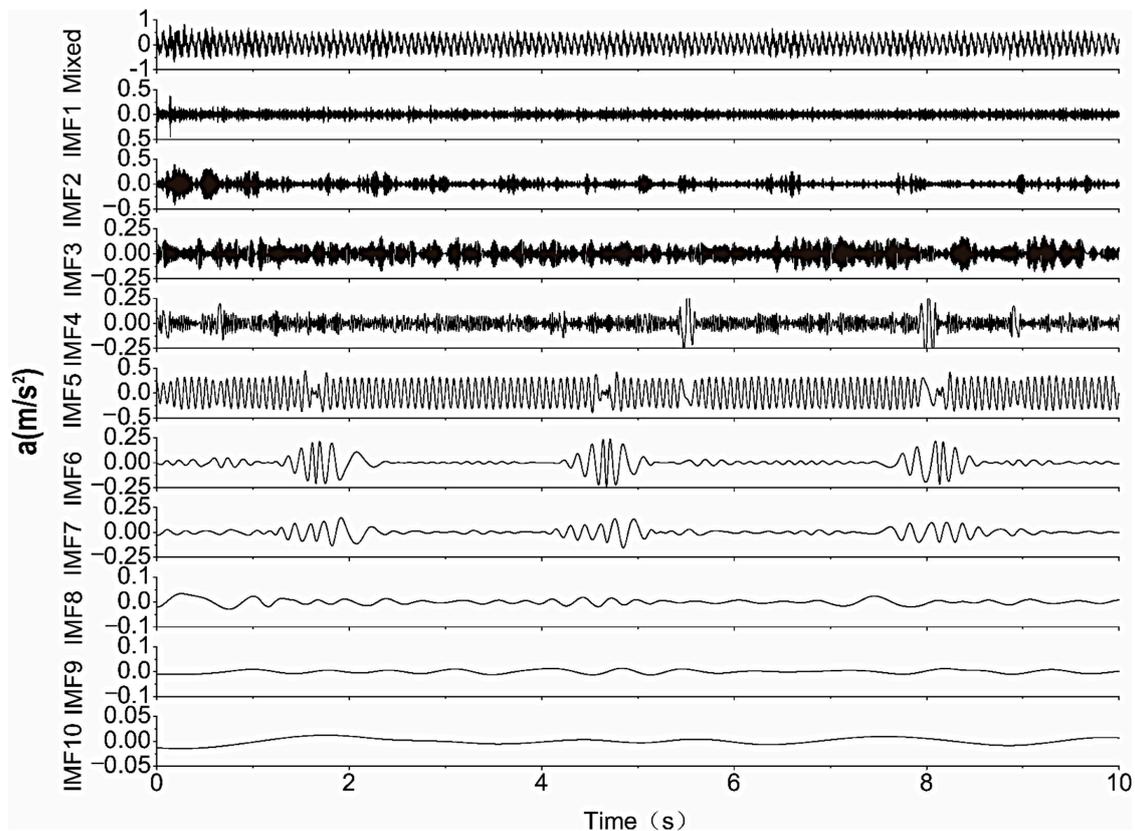

Figure 7. IMF components by performing EMD for the mixed signal.

As shown above, the aim of signal separation is to limit the types of typical signals in order to reduce the number of intelligent recognition categories. Additionally, the blind source separation discussed above can separate signal sources from different collecting channels, not only for the mixed signals, but also for the single type signals, without creating any new signals.

3.3. Dataset Preparation and Features Analysis

3.3.1. Sample Preparation and Dataset Construction

Based on the original signals collected at the field sites, three parts of the data were prepared for the construction of training and testing datasets, including the pile-driving signals, train operational signals, and others.

As shown in Table 1, 13,184 groups of pile-driving signals, 23,832 groups of train operational signals, and 8229 groups of environment signals collected from the field sites were used for the training dataset.

Table 1. Division of sample sets for intelligent model.

	Training Set (Group)	Test Set (Group)
Pile-driving signals	13,184 (original signal)	1220 (separation signal)
Train operational signals	23,832 (original signal)	5760 (separation signal)
Others (noise, etc.)	8229 (original signal)	2057 (original signal)

In total, 1220 groups of pile-driving signals and 5760 groups of train operational data were used to conduct the signal separation, and the re-separated data were used for the test set. Among them, 1220 groups of pile-driving data and train operational data were used to form the mixed signals, and 4540 groups of train operational data were employed as two-channel data to achieve single-type signal separation.

3.3.2. Typical Time and Frequency Domain Features

In order to explore the time and frequency domain characteristic distribution of the above dataset, and as a basis of the following MI analysis, sixteen time and frequency domain features were selected and are listed in Table 2, where x_p is the peak value of the vibration, N is the length of the signal sampling window, and S and E are the beginning and end of the sampling signal, respectively.

Table 2. Selected features for MI analysis and feature compression.

Feature	Equation	Feature	Equation
Mean value	$\mu_x(t) = \frac{1}{N}\sum_{i=S}^{E} x_i = E(X)$	Kurtosis factor	$KF = \frac{\sum_{i=S}^{E} x_i^4}{\sqrt{\sum_{i=S}^{E} x_i^2}}$
Standard deviation	$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=s}^{E}(x_i - \mu_x)^2}$	Pulse factor	$C_{pul} = \frac{x_p}{\mu_x}$
Kurtosis	$K = \frac{1}{N}\sum_{i=S}^{E}\left(\frac{x_i - \mu_x}{\sigma_x}\right)^4$	Clearance factor	$C_{clea} = \frac{x_p}{\left(\frac{1}{N}\sqrt{|x_i|}\right)^2}$
Root mean square	$x_{rms} = \sqrt{\frac{x_S^2 + x_{S+1}^2 + \ldots + x_E^2}{E-S+1}}$	Waveform center	$\langle t \rangle = \frac{1}{N}\sum_{i=S}^{E}(i \cdot |x_i|^2)$
Wave form factor	$C_{Ws} = \frac{x_{rms}}{\mu_x}$	Time width	$\sigma_t = \sqrt{\sum_{i=S}^{E}(i - \langle t \rangle)^2 |x_i|^2}$
Peak factor	$C_f = \frac{x_p}{x_{rms}}$	Mean square frequency	$MSF = \frac{\int_0^\infty f^2 FFT(x,t)df}{\int_0^\infty FFT(x,t)df}$
Center frequency	$FC = \frac{\int_0^\infty f FFT(x,t)df}{\int_0^\infty FFT(x,t)df}$	Root mean square frequency	$RMSF = \sqrt{MSF}$
Frequency variance	$VF = \frac{\int_0^\infty (f-FC)^2 FFT(x,t)df}{\int_0^\infty FFT(x,t)df}$	Frequency standard deviation	$RVF = \sqrt{VF}$

Additionally, the box chart of the statistical distribution of each feature on three types of target signal data is shown in Figure 8.
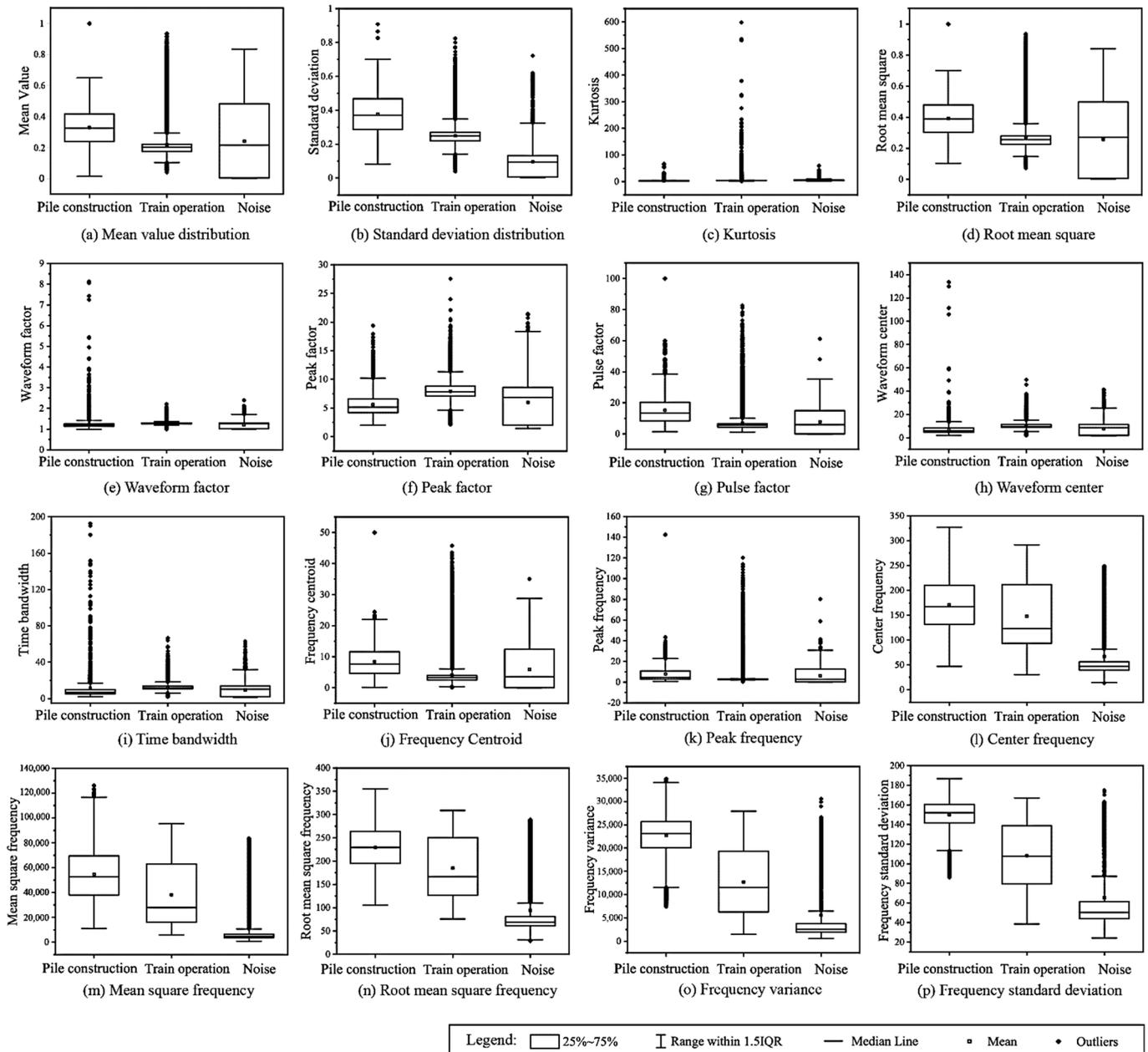

Figure 8. Statistical distribution of each feature on two types of target signal data.

As can be seen from Figure 8, for the pile-driving signals collected from sites at different distances from the sensor, and the train-induced vibration signals obtained at different times, the features show obvious discreteness, and have overlapping distribution areas, representing a challenge for the MLP network. Meanwhile, it was also found that the discreteness of the frequency-domain features is less than that of the time-domain features. Especially for the train operation signals, the feature distribution in the frequency domain is more stable than that in the time domain, and the increased use of the frequency features will increase the recognition accuracy.

4. Application of Combined MI-MLP Neural Network Method

In this section, the intelligent recognition of vibration signal sources using the combined MI-MLP neural network is discussed based on the data prepared, as described in Section 3. This process mainly consists of the MI analysis of vibration features, the feature selection and input matrix formation, and the training and testing of the MLP neural network.

4.1. MI Analysis and Feature Selection

To measure the correlation between the feature and the corresponding vibration category, the variable X can be treated as the feature vector and Y can be considered as the corresponding vibration-type label. When using X_n groups of training and validation set samples, as listed in Table 1, there exists a feature vector X_j with the dimensions $X_n \times 1$. A corresponding label vector representing the vibration types is assembled by assigning a category label to each row of the feature vector, namely $Y = \{y_1, y_2, y_3, \ldots y_i, \ldots, y_{Xn}\}$, where y_i is 0 when x_i is the pile-construction-induced response and 1 when x_i is the train-operation-induced response.

The MI values of the sixteen feature vectors and the corresponding label vector are shown in Figure 9; from the figure, it is evident that the frequency domain features are most relevant to the vibration signal classification. Finally, the MI values of ten features were found to be greater than 0.4 and were accordingly selected to form the input matrix of the MLP neural network.
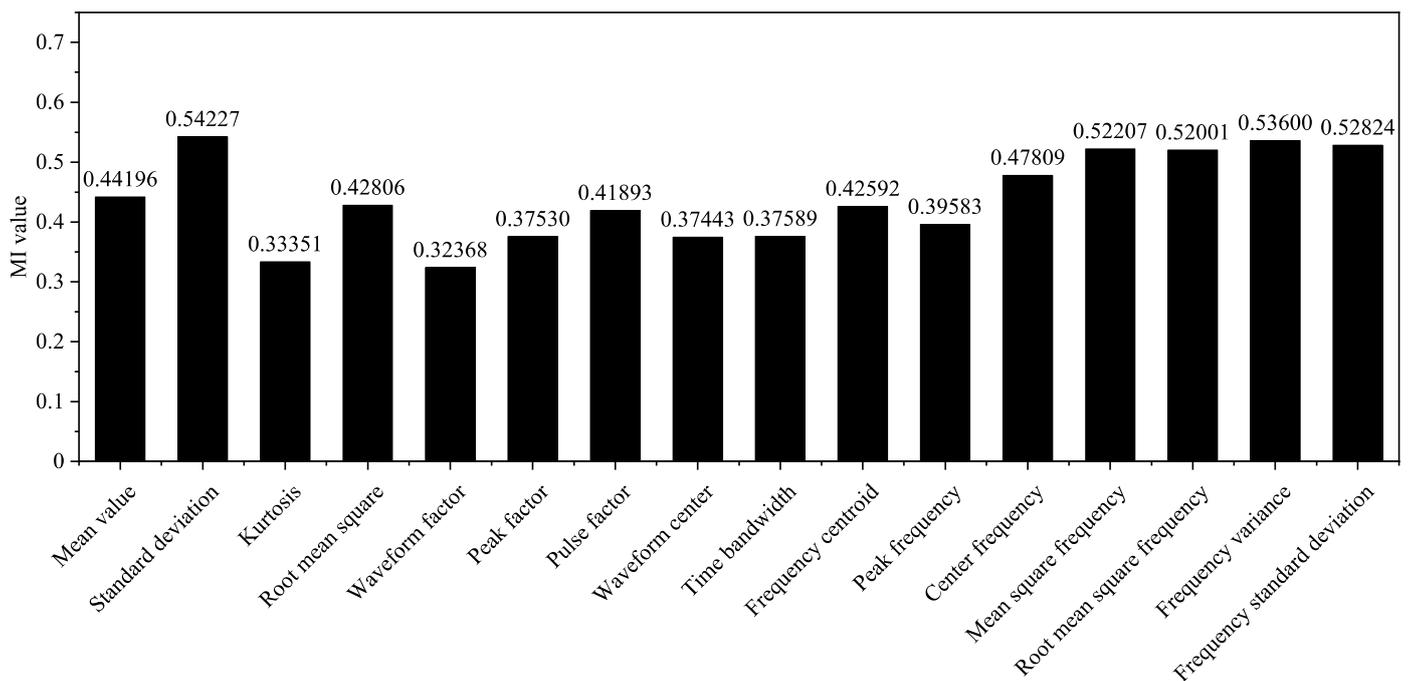

Figure 9. MI values between different features and the corresponding signal types.

4.2. Training and Testing of the MLP Neural Network

Based on the method described in Section 5.1, 600 feature vectors were used as the input for the MLP network, instead of the 600 groups of vibration signals. The training set and test set were then constructed by randomly selecting values from these vectors at a ratio of 8:2.

Hyperparameter searching was performed when preparing the data set, as the selection of appropriate hyperparameters can ensure a smooth training process and improve the final model training quality. The most widely used hyperparameter automatic searching algorithms include the grid search method and the random search method. In the grid search algorithm, the specified parameter array grows from different sequences and then enters the training model to finally select the optimal set of parameters according

to their accuracies. This approach is more suitable for cases involving fewer parameters; as the number of hyperparameters increases, the computational complexity of the grid search method increases exponentially. In contrast to the grid search method, the random search method does not need to go through all possible combinations of hyperparameters; instead, this approach defines their distributions. As a result, it has a lower calculation complexity and, consequently, a greater efficiency when searching for a large number of hyperparameters. Accordingly, this approach was employed in this study.

When using the random search method, the search routes begin from a random starting parameter combination and converge to a final parameter value according to the probability distribution of each parameter. For this purpose, in this study, a function was coded in Python that considered the hyperparameter distributions to be uniform, and the ratio of model accuracy to training time was used as the evaluation index. The search parameters and results are shown in Table 3.

Table 3. Setting and results of hyperparameter searching for the MLP neural network.

Search Parameters			
Search range of learning rate	[0, 1)	Iteration step for learning rate	0.01
Search range of calculation layers	[1, 3)	Number of calculation layers in each search step	1
Search range of hidden layer nodes	[1, 100)	Number of hidden layer nodes in each search step	10
Search results			
Optimal learning rate	0.1	Optimal number of hidden layer nodes	10
Optimal number of computing layers	2	Optimal accuracy	94.68%

The MLP neural network shown in Figure 3 was trained and tested using the ten feature vectors selected using MI analysis, as discussed in Section 4.1. Additionally, because of the time and frequency domain features used, there is a large gap between the two types of features in terms of values. Therefore, normalization was employed to the feature matrix before inputting it into the training and testing network.

The changes in the loss function with the number of MLP network iterations using the hyperparameter search results (Table 3) are shown in Figure 10a. Clearly, the loss function decreased rapidly over the first fifty epochs before converging to 0.39 in the training set and 0.23 in the test set. Additionally, from Figure 10b, it can be observed that the final accuracy of the network structure when using the training set was near 87.17%, and near 94.68% when using the test set.

(a) Loss of training and test set

(b) Accuracy of training and test set
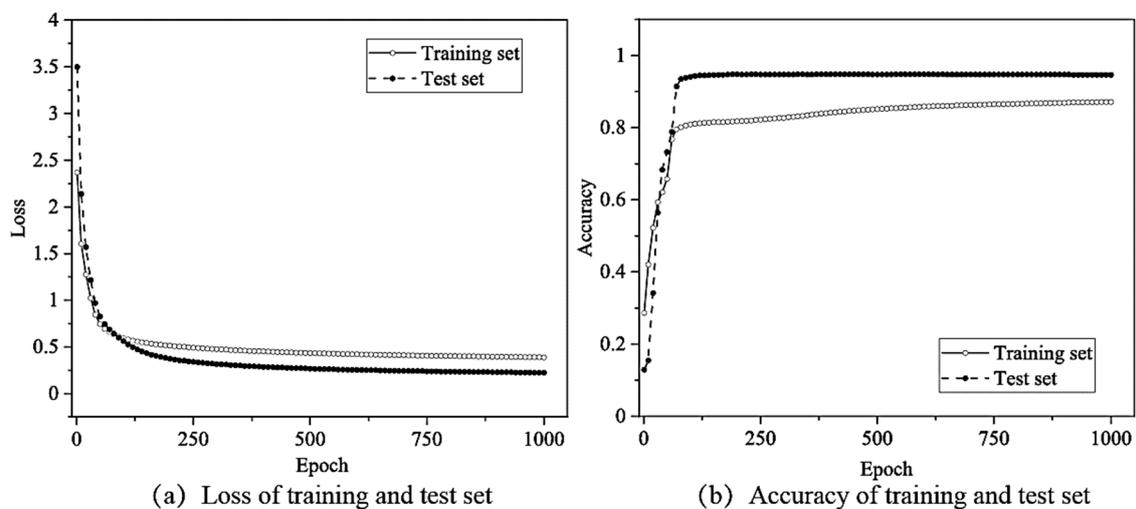
Figure 10. Loss function and accuracy changes according to epoch when using the MLP neural network.

4.3. The Effect of MI Analysis and Feature Selection

The feature matrix after MI analysis and the original signal samples were separately used as the input of the MLP model to evaluate the effects of feature selection on the loss and accuracy of the training and validation sets (Figure 11).
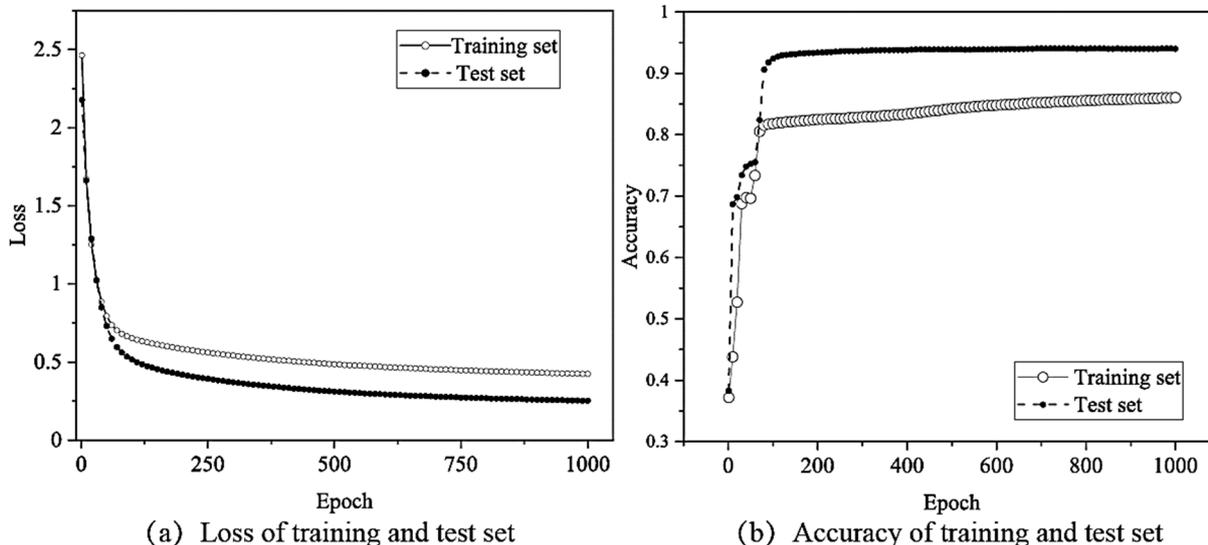

(a) Loss of training and test set (b) Accuracy of training and test set

Figure 11. Loss and accuracy changes by MLP neural network without feature selection.

The change in the loss function and recognition accuracy during model training, obtained when using the MI analysis and feature selection, are shown in Figure 9. Additionally, the training results without feature selection are shown in Figure 12 (all the 16 features were used). The training loss converged slowly, only reaching convergence after 1000 epochs; the resulting accuracy was only 86.05%. Similarly, the loss of the test set converged after nearly 1000 epochs, and the accuracy only reached 94.03%, decreasing by 0.65% before feature selection. Thus, the use of feature selection can improve the training speed and recognition accuracy of the vibration signal source recognition method.

Figure 12. Confusion matrix for vibration source recognition by the MLP neural network without feature selection. (**a**) MI + MLP; (**b**) MLP.

Additionally, as shown in Figure 12, the confusion matrix also shows that the recognition accuracy of all the categories will decrease if the MI and feature selection are not used, and the total accuracy decreases by 0.65%. Therefore, feature selection by the MI method can improve the recognition accuracy.

5. Comparison with LSTM Network-Based Recognition Method
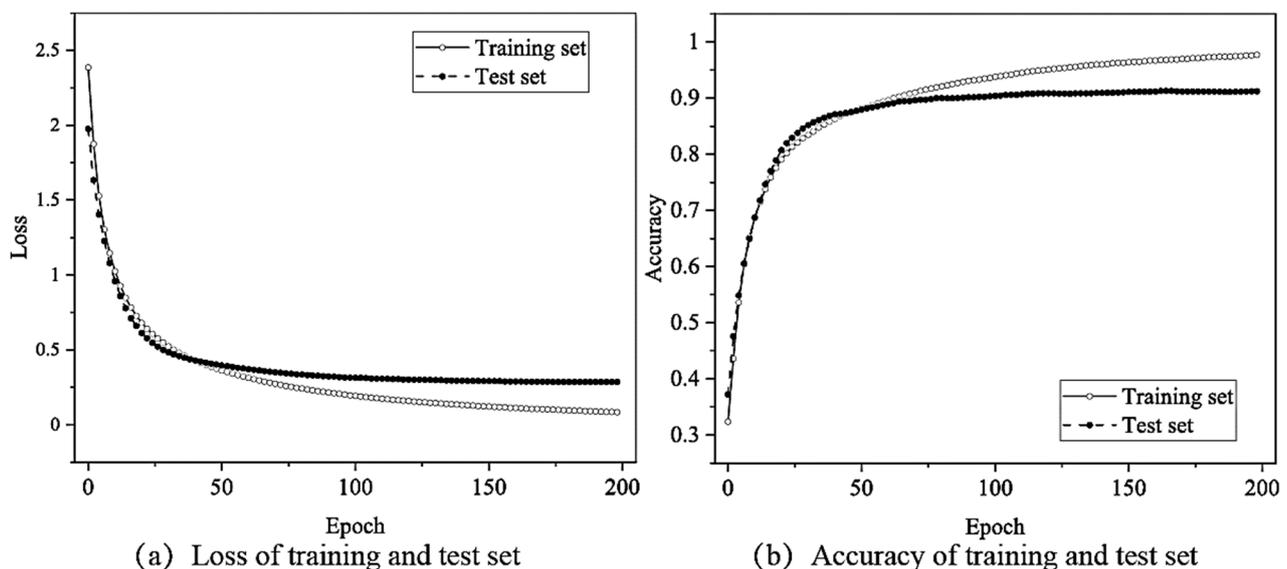
5.1. Hyperparameter Searching

Hyperparameter searching was then performed using the random search method, similar to that performed for the MI-MLP method, as described in Section 4; the results are shown in Table 4.

Table 4. Setting and results of LSTM network hyperparameter searching.

Search Parameters			
Search range of learning rate	[0, 1)	Iteration step for learning rate	0.01
Search range of calculation layers	[1, 3)	Number of calculation layers in each search step	1
Search range of hidden layer nodes	[1, 100)	Number of hidden layer nodes in each search step	10
Search results			
Optimal number of computing layers	2	Optimal accuracy	91.25%
Optimal number of hidden layer nodes	50		

5.2. Training and Testing of LSTM Network

The LSTM network was applied to identify the vibration signals, as described in Section 2.4 and shown in Figure 3. It should be noted that because the LSTM network offers advantages in identifying discrete time series, the training and validation signal samples were divided into ten time steps and directly used as the network input (Figure 13).

(a) Loss of training and test set (b) Accuracy of training and test set

Figure 13. Change in the loss function and the accuracy of the training and validation sets according to epoch when using the LSTM network.

Based on the hyperparameter search results, as listed in Table 4, the changes in the loss function with the number of iteration epochs were determined, as shown in Figure 14a. It can be observed that the loss function decreases rapidly over the first twenty epochs, before finally converging to 0.08 for the training set and approximately 0.28 for the test set. From Figure 7, it can be observed that the final accuracy of the network structure over the training set was 97.7% and 91.25% over the test set.

(a) MI-MLP method

Recognize category

	Object (1,0,0)	Object (0,1,0)	Object (0,0,1)	
(1,0,0)	1088 / 12.04%	132 / 1.46%	0 / 0.00%	89.18% / 10.82%
(0,1,0)	6 / 0.07%	5753 / 63.66%	1 / 0.01%	99.88% / 0.12%
(0,0,1)	120 / 1.33%	222 / 2.46%	1715 / 18.98%	83.37% / 16.63%
	89.62% / 10.38%	94.20% / 5.80%	99.94% / 0.06%	94.68% / 5.32%

Object category (1,0,0) (0,1,0) (0,0,1)

(b) LSTM

Recognize category

	Object (1,0,0)	Object (0,1,0)	Object (0,0,1)	
(1,0,0)	1153 / 12.76%	56 / 0.62%	11 / 0.12%	94.51% / 5.49%
(0,1,0)	82 / 0.91%	5293 / 58.56%	385 / 4.27%	91.88% / 8.12%
(0,0,1)	38 / 0.42%	218 / 2.41%	1801 / 19.93%	87.55% / 12.45%
	90.57% / 9.43%	95.08% / 4.92%	81.94% / 18.06%	91.25% / 8.75%

Object category (1,0,0) (0,1,0) (0,0,1)

Figure 14. Confusion matrix for vibration source recognition by the MLP neural network and LSTM. (**a**) MI-MLP method; (**b**) LSTM.

5.3. Method Comparison

Meanwhile, the details of the recognition results are shown in the confusion matrix in Figure 14. The object category (1, 0, 0) (pile-driving-induced vibration recognition) achieved an accuracy of 94.51%, correctly identifying 54 of the 75 signals, higher than that achieved by the MLP + MI method; object category (0, 1, 0) (train operation-induced vibration recognition) achieved an accuracy of 91.88%, correctly identifying 5293 of the 5760 signals; and the total accuracy of the LSTM-based vibration signal recognition was 91.25%, which is lower than the recognition accuracy obtained when using the MI-MLP method.

The LSTM has been proven to be a particularly efficient network and good at time series prediction; therefore, the recognition accuracy for pile-driving and environment vibration is higher than that of the MLP network. However, for the train operational signals, as shown in Figure 8, the frequency-domain feature distribution concentration is much better than that of the time-domain feature, and future compression and selection by the MI method greatly improve the recognition accuracy. Meanwhile, the training of the LSTM network took 525.35 s, much longer than that of the MLP network, which was 20.02 s. This indicates that the use of suitable feature selection can considerably improve the training efficiency of vibration source recognition.

6. Conclusions

This paper proposed a method for the real-time monitoring and early warning of surrounding construction behaviors near underground facilities. Considering the challenges associated with intelligent recognition owing to vibration mixing and signal acquisition difficulties, the MI analysis and blind source separation methods were employed. The conclusions of this work can be detailed as follows:

- The recognition accuracies obtained by the LSTM and MI-MLP methods when using the test set were both greater than 90%, which can meet the needs of practical applications. This demonstrates the feasibility of real-time construction monitoring with suitable recognition robustness.

- Blind source separation can decompose mixed signals into a limited number of identifiable samples, which can help minimize the number of classification categories caused by signal mixing and improve recognition accuracy.
- Feature compression and selection after MI analysis can increase the training efficiency and recognition accuracy. Especially for signals with a centralized distribution of frequency domain features, the recognition accuracy of the MLP network will have a better performance than that of a well-known time series prediction LSTM model.

The results of this study can be used to develop a real-time vibration monitoring system for underground infrastructure that is capable of identifying the construction activity source of a particular vibration. This will help ensure the safety of underground infrastructure, such as transit and pipeline networks, near construction sites. Although the proposed method was evaluated in this study using pile-driving vibrations, it should be capable of identifying a wide range of construction activities. Therefore, future research should focus on applying the proposed method to identify various construction activities based on measured vibrations.

Author Contributions: Supervision, project administration, and methodology, B.Z.; validation, data curation, and writing—original draft, Y.G.; conceptualization, investigation, and resources, X.W.; formal analysis and writing—review and editing, X.X. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Xie, X.; Wang, Q.; Shahrour, I.; Li, J.; Zhou, B. A real-time interaction platform for settlement control during shield tunnelling construction. *Autom. Constr.* **2018**, *94*, 154–167. [CrossRef]
2. Gómez, J.; Casas, J.R.; Villalba, S. Structural Health Monitoring with Distributed Optical Fiber Sensors of tunnel lining affected by nearby construction activity. *Autom. Constr.* **2020**, *117*, 103261. [CrossRef]
3. Ross, Z.E.; Trugman, D.T.; Hauksson, E.; Shearer, P.M. Searching for hidden earthquakes in Southern California. *Science* **2019**, *364*, 767–771. [CrossRef]
4. Umlauft, J.; Korn, M. 3-D fluid channel location from noise tremors using matched field processing. *Geophys. J. Int.* **2019**, *219*, 1550–1561. [CrossRef]
5. Zhang, H.; Eaton, D.; Rodríguez-Pradilla, G.; Jia, S.Q. Source-Mechanism Analysis and Stress Inversion for Hydraulic-Fracturing-Induced Event Sequences near Fox Creek, Alberta. *Bull. Seismol. Soc. Am.* **2019**, *109*, 636–651. [CrossRef]
6. Ventura, C.E.; Liam Finn, W.D.; Lord, J.-F.; Fujita, N. Dynamic characteristics of a base isolated building from ambient vibration measurements and low level earthquake shaking. *Soil Dyn. Earthq. Eng.* **2003**, *23*, 313–322. [CrossRef]
7. Newton, C.; Snieder, R. Estimating Intrinsic Attenuation of a Building Using Deconvolution Interferometry and Time Reversal. *Bull. Seismol. Soc. Am.* **2012**, *102*, 2200–2208. [CrossRef]
8. Lorenzoni, F.; Casarin, F.; Caldon, M.; Islami, K.; Modena, C. Uncertainty quantification in structural health monitoring: Applications on cultural heritage buildings. *Mech. Syst. Signal Process.* **2016**, *66–67*, 268–281. [CrossRef]
9. HoThu, H.; Mita, A. Assessment and evaluation of damage detection method based on modal frequency changes. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2013*; Lynch, J.P., Yun, C.-B., Wang, K.-W., Eds.; SPIE: Washington, DC, USA, 2013; Volume 8692, p. 86923M.
10. Spence, S.M.J.; Bernardini, E.; Guo, Y.; Kareem, A.; Gioffrè, M. Natural frequency coalescing and amplitude dependent damping in the wind-excited response of tall buildings. *Probabilistic Eng. Mech.* **2014**, *35*, 108–117. [CrossRef]
11. Li, L.; Tan, J.; Schwarz, B.; Staněk, F.; Poiata, N.; Shi, P.; Diekmann, L.; Eisner, L.; Gajewski, D. Recent Advances and Challenges of Waveform-Based Seismic Location Methods at Multiple Scales. *Rev. Geophys.* **2020**, *58*, e2019RG000667. [CrossRef]
12. Zhou, B.; Zhang, F.; Xie, X. Vibration Characteristics of Underground Structure and Surrounding Soil Underneath High Speed Railway Based on Field Vibration Tests. *Shock Vib.* **2018**, *2018*, 3526952. [CrossRef]
13. Zhou, B.; Xie, X.; Wang, X. The Tunnel Structural Mode Frequency Characteristics Identification and Analysis Based on a Modified Stochastic Subspace Identification Method. *Shock Vib.* **2018**, *2018*, 6595841. [CrossRef]

14. Zhao, A.; Qi, L.; Dong, J.; Yu, H. Dual channel LSTM based multi-feature extraction in gait for diagnosis of Neurodegenerative diseases. *Knowl.-Based Syst.* **2018**, *145*, 91–97. [CrossRef]

15. Xu, Y.; Lu, L.; Xu, Z.; He, J.; Zhou, J.; Zhang, C. Dual-channel CNN for efficient abnormal behavior identification through crowd feature engineering. *Mach. Vis. Appl.* **2019**, *30*, 945–958. [CrossRef]

16. Jung, M.; Chi, S. Human activity classification based on sound recognition and residual convolutional neural network. *Autom. Constr.* **2020**, *114*, 103177. [CrossRef]

17. Sezgin, N. A new hand finger movements' classification system based on bicoherence analysis of two-channel surface EMG signals. *Neural Comput. Appl.* **2019**, *31*, 3327–3337. [CrossRef]

18. Tang, Z.; Chen, Z.; Bao, Y.; Li, H. Convolutional neural network-based data anomaly detection method using multiple information for structural health monitoring. *Struct. Control Health Monit.* **2019**, *26*, e2296. [CrossRef]

19. Bai, Y.; He, P.; Zhao, Y.; Ma, S.; Mi, H.; Wei, Z.; Li, Z. Real-time online detection of trucks loading via genetic neural network. *Autom. Constr.* **2020**, *120*, 103354. [CrossRef]

20. Clemente, J.; Li, F.; Valero, M.; Song, W. Smart Seismic Sensing for Indoor Fall Detection, Location, and Notification. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 524–532. [CrossRef]

21. Liu, Z.; Li, S. A sound monitoring system for prevention of underground pipeline damage caused by construction. *Autom. Constr.* **2020**, *113*, 103125. [CrossRef]

22. Musafere, F.; Sadhu, A.; Liu, K. Towards damage detection using blind source separation integrated with time-varying auto-regressive modeling. *Smart Mater. Struct.* **2015**, *25*, 15013. [CrossRef]

23. Cavalcante, C.C.; Filho, D.Z.; Romano, J.M.T. Multiuser processing using blind source separation methods. *Eur. Trans. Telecommun.* **2008**, *19*, 827–836. [CrossRef]

24. Ma, B.; Zhang, T. Single-channel blind source separation for vibration signals based on TVF-EMD and improved SCA. *IET Signal Process.* **2020**, *14*, 259–268. [CrossRef]

25. Nordhausen, K.; Fischer, G.; Filzmoser, P. Blind Source Separation for Compositional Time Series. *Math. Geosci.* **2020**, *53*, 905–924. [CrossRef]

26. Shannon, C.E. A Mathematical Theory of Communication. *SIGMOBILE Mob. Comput. Commun. Rev.* **2001**, *5*, 3–55. [CrossRef]

27. Thuraisingham, R.A. Estimating Electroencephalograph Network Parameters Using Mutual Information. *Brain Connect.* **2018**, *8*, 311–317. [CrossRef]

28. Afshani, F.; Shalbaf, A.; Shalbaf, R.; Sleigh, J. Frontal–temporal functional connectivity of EEG signal by standardized permutation mutual information during anesthesia. *Cogn. Neurodyn.* **2019**, *13*, 531–540. [CrossRef]

29. Ren, F.; Dong, Y.; Wang, W. Emotion recognition based on physiological signals using brain asymmetry index and echo state network. *Neural Comput. Appl.* **2019**, *31*, 4491–4501. [CrossRef]

30. Tafreshi, T.F.; Daliri, M.R.; Ghodousi, M. Functional and effective connectivity based features of EEG signals for object recognition. *Cogn. Neurodyn.* **2019**, *13*, 555–566. [CrossRef]

31. Hinton, G.E.; Osindero, S.; Teh, Y.-W. A Fast Learning Algorithm for Deep Belief Nets. *Neural Comput.* **2006**, *18*, 1527–1554. [CrossRef]

32. Himberg, J.; Hyvärinen, A.; Esposito, F. Validating the independent components of neuroimaging time series via clustering and visualization. *NeuroImage* **2004**, *22*, 1214–1222. [CrossRef] [PubMed]

33. Wang, K.P. *Methodologies and Applications of Weak Signal Detection Based on Blind Source Separation*; Chongqing University: Chongqing, China, 2014.

34. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]

35. Gers, F.A. Learning to forget: Continual prediction with LSTM. In Proceedings of the 9th International Conference on Artificial Neural Networks: ICANN '99, Edinburgh, UK, 7–10 September 1999; IEEE: Piscataway NJ, USA, 1999; Volume 1999, pp. 850–855.

36. Li, A.J.; Khoo, S.; Lyamin, A.V.; Wang, Y. Rock slope stability analyses using extreme learning neural network and terminal steepest descent algorithm. *Autom. Constr.* **2016**, *65*, 42–50. [CrossRef]

37. Goncalves, P.; Rilling, G.; Flandrin, P. On empirical mode decomposition and its algorithms. In Proceedings of the IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, Baltimore, MD, USA, 3–6 June 2003; Volume 3, pp. 8–11.