Research Article

An Improved High-Intelligence Method of Gas and Oil Pipeline Prewarning System in Real Soil Environment

Fang Wang, 1 Jichuan Xing 2, Jinxin Li, 1 Feng Zhao, 2 and Shufeng Zhang 3

1 Optoelectronic Department, Beijing Institute of Technology, Beijing 100081, China
2 Hilong Petroleum Information Technology Company, Hilong Group of Companies, Shanghai 201103, China
3 College of Computing, Southeast University, Nanjing 211189, China

Correspondence should be addressed to Jichuan Xing; michaelhsing@bit.edu.cn

Received 8 February 2021; Revised 22 February 2021; Accepted 18 September 2021; Published 19 October 2021

Academic Editor: Yufei Wang

Copyright © 2021 Fang Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the development of technology, the total extent of global pipeline transportation is also increased. However, the traditional long-distance optical fiber prewarning system has poor real-time performance and high false alarm rate when recognizing events threatening pipeline safety. The same vibration signal would vary greatly when collected in different soil environments and this problem would reduce the signal recognition accuracy of the prewarning system. In this paper, we studied this effect theoretically and analyzed soil vibration signals under different soil conditions. Then we studied the signal acquisition problem of long-distance gas and oil pipeline prewarning system in real soil environment. Ultimately, an improved high-intelligence method was proposed for optimization. This method is based on the real application environment, which is more suitable for the recognition of optical fiber vibration signals. Through experiments, the method yielded high recognition accuracy of above 95%. The results indicate that the method can significantly improve signal acquisition and recognition and has good adaptability and real-time performance in the real soil environment.

1. Introduction

Pipeline integrity is vital to the transmission of gas and oil [1–4]. Timely discovering the intrusion events around the pipelines and preventing pipeline leakage is the major direction of current researches. The optical fiber prewarning system (OFPS) is mainly used in underground cable and pipeline transportation safety. In comparison to other methods, distributed optical fiber sensing technology is used in the OFPS with higher positioning accuracy and sensitivity and lower investment cost [5–8]. Mainly, the OFPS includes two components of Pattern Recognition System (PRS) and Distributed Acoustic sensing System (DAS). DAS is a distributed acoustic wave sensing system enabling remote monitoring, and PRS is utilized to recognize and classify various events from the DAS-collected data. In this paper, the DAS based on the Φ-Optical Time-domain Reflectometry (Φ-OTDR) principle is utilized to collect the soil vibration signal [9–11]. Normally, gas and oil pipelines are laid in a complex geographical environment. The same vibration signals will vary greatly when collected in different soil environments and this problem can reduce the signal recognition accuracy of the OFPS. The signal acquisition of the OFPS in the real soil environment is complicated and different from the simple environment in the laboratory [12, 13]. Figure 1 represents the OFPS in the real soil environment. The whole process is similar to a black box from the occurrence to the identification and classification of the vibration signals caused by the intrusion events [14, 15].

Currently, commonly used gas and oil pipeline prewarning technologies mainly include optical fiber sensing and seismic wave detection. Seismic wave detection technology is mainly used in places where there is no communication optical cable. The monitoring distance of a single seismic wave sensor is 150–200 m. Since the installation and maintenance of each sensor requires excavation, which leads to higher use and maintenance costs, this technology is only suitable for monitoring of key pipe...
Soil Vibration Signals

1. Soil Environmental Factors
2. Optical Fiber Sensing
3. Other Special Factors

Figure 1: OFPS in the real soil environment.

sections [16]. Therefore, the long-distance gas and oil pipeline prewarning system mainly uses optical fiber sensing technology. According to the different optical principles used, the optical fiber sensing pipeline prewarning technology can be divided into four types: Mach–Zehnder, Sagnac, Bragg fiber grating, and Φ-OTDR. The Mach–Zehnder technology requires a 3-core optical fiber, and the sensitivity of the system is limited; when multiple vibration events occur simultaneously, the system cannot distinguish and identify [17]. The structure of the Sagnac technology system is relatively complex, and the length of the delay fiber is required to be much larger than the length of the sensing fiber of the interferometer, which increases the difficulty of application [18]. The technology of Bragg fiber grating is complicated, and the installation and maintenance cost are relatively high [19]. The Φ-OTDR method occupies less fiber resources, has higher sensitivity, can detect weak vibration signals, and can monitor and locate multiple vibration events that occur simultaneously at multiple points along the pipeline. Thus, the most commonly used today is the Φ-OTDR technology. In recent years, some researches provide techniques to represent the system recognition accuracy of Φ-OTDR [20]. Various Φ-OTDR-based vibration sensor schemes have been suggested to increment the dynamic range (DR) and the signal-to-noise (SNR) of the vibration measurement by controlling or reducing the noise present in the system [21–23]. For a given spatial resolution (SR), increasing the optical peak power input to the fiber via an erbium doped fiber amplifier (EDFA) increases the DR and the SNR of the system. Other techniques have emerged to break totally this dilemma in the Φ-OTDR system and enabled improving significantly the SR [24, 25]. Such methods mainly involve the linear frequency modulation pulse compression and the chirp pulse amplification. In these researches, only the optical fiber sensing system itself is enhanced; however, the recognition algorithm is not improved.

Artificial intelligence algorithms are applied in various fields of industry [26]. There are also some different algorithms in the field of optical fiber sensing. Different algorithms lead to significant differences in the accuracy of external intrusion event detection. Most of the methods used in the previous researches are hard to be applied. First, in some researches the signal used to input into the neural network (NN) is mainly about the feature vectors constructed by decomposing the signal spectra from the multilayer wavelets. The wavelet decomposition used to extract the features in [27, 28] generates a large number of wavelet coefficients, which increases the complexity of network training. This process of extracting signal features takes a lot of time. Second, the NN in [29–31] used a single NN, which is hard to effectively learn the signal characteristics collected by the DAS. Besides, the existing experiments in [32–35] only utilized the data being generated very close to the sensed area (tens or hundreds of meters), which had a small sensing area or were at a location with a clean background environment. This makes it almost impossible to verify the recognition accuracy of the recognition model with dozens of kilometers’ long distance and with unknown background noises. Thus, these systems are not able to fully meet the needs of the indicators, and soil signal processing is still not perfect. No article exists focusing on the role of the combination of soil environment and optical fiber sensing in the OFPS.

In this paper, we studied the signal acquisition problem of long-distance gas and oil pipeline prewarning system in real soil environment. Accordingly, an improved high-intelligence method was proposed for optimization. This method can solve the problem that the same vibration signal would vary greatly when collected in different soil environments, which would reduce the signal recognition accuracy of the prewarning system. This method is based on the real application environment, which has practical significance for industrial applications. The improved high-intelligence method is a new neural network method, which has the advantages of LSTM and CNN. LSTM is good at processing time-related data, and CNN can abstract and extract features from data at multiple time points to reduce invalid data. These advantages determine that the improved NN is more suitable for the recognition of optical fiber vibration signals than other traditional methods. It is a new attempt in the field of optical fiber prewarning. Through experiments, it is verified that this improved NN has good adaptability and real-time performance. We collected three types of event signals in different soil environments for training and recognition, and it is proved that the improved high-intelligence method has good adaptability, and the average recognition rate of three types of event signals is more than 95%. Since this method has good adaptability and real-time performance, it can be popularized in the optical fiber prewarning systems all over the world. Furthermore, the results of the study have high practical value for monitoring the safety of gas and oil pipelines.

2. Optical Fiber Prewarning System

The OTDR system uses a low-coherence light source, and the phase information of the light wave is masked. But the phase change caused by external disturbance is more sensitive. In order to take advantage of the high sensitivity of phase shift, a Φ-OTDR system was developed. The structure of the Φ-OTDR system is similar to that of the OTDR, but with a narrow linewidth laser light source, the interference phenomenon of the backward Rayleigh scattered light within the optical pulse width is enhanced, and the sensitivity of the system is greatly improved [36–38]. The OFPS based on the Φ-OTDR principle includes two main parts of PRS and DAS. Figure 2 represents the OFPS processing flow. The DAS’s main function is to gather soil vibration signals near the pipeline in real time for monitoring invasion events.
Within the DAS, a narrow line width laser source mainly emits the laser yielding the result of the interference of reflected Rayleigh scattered light. The events near the optical fiber can cause soil vibration signals. The pressure created by soil vibrations will act on the optical fiber sensing system, and then as a result of the photoelastic effect of the fiber, the refractive index of the fiber at the corresponding position will also alter. Subsequently, it will change the phase of light at the same point. Owing to the interference, the intensity of the backward Rayleigh scattered light will change with the phase. Therefore, when an intrusion event happens near the prewarning system, the intruding position intensity on the fiber varies from that before occurrence of the intrusion event, but the intensity at other positions remains unchanged. Vibration signals are collected by the DAS, and then they will be input into the PRS. The PRS is able to give the location and the type of the intrusion event.

Figure 3 shows the structure of the soil utilized in this paper. A narrow line width laser is used in the DAS system as the light source to output continuous light with a wavelength of 1550 nm. In order to modulate the emitted light, the acoustooptic modulator is used to obtain a laser pulse, which is then amplified by an EDFA. The system utilizes two-way pumping Raman amplification to enhance system performance and amplify signal light at different locations in the fiber. A Raman laser is used by the amplifying part of the system as the pump source. The center wavelength of the pump source is 1450 nm with the maximum output power of 0.5 W. Through a 3 dB coupler and a wavelength division multiplexer (WDM), the amplified light pulse enters the sensing fiber. The sensing fiber is the common single-mode communication fiber. The amplified optical signal is filtered into the fiber grating (FBG) through the optical circulator, and the noise resultant from amplification was removed. Then, the intrusion signal is collected by the photodetector. Ultimately, the collected data is passed to the PRS for calculation and analysis.

3. Theoretical Analysis

3.1. Soil Vibration. Soil is a three-phase discontinuous material since the air, liquid, and soil particles are three different materials. The skeleton of soil is composed of loose soil particles, and each soil particle is filled with space between which the air and liquid exist. However, these three media are not closely connected. Thus, when we analyze the indicators of soil physical properties, soil can be regarded as a discontinuous dispersion system [39].

Figure 4 shows the structure of the soil. As shown in the figure, the structure is divided into air, water, and soil particles from top to bottom. \( m_a \) and \( V_a \) represent the mass and volume of the air; \( m_w \) and \( V_w \) represent the mass and volume of the liquid; \( m_i \) and \( V_i \) represent the mass and volume of soil particles. The total mass and volume of soil are expressed by \( m \) and \( V \), respectively. Since the air has a mass of 0, the total mass of the soil is equal to the sum of the mass of liquid and the mass of soil particles; the volume of soil interstitial \( V_i \) can be regarded as the sum of the volume of liquid and the volume of air. The nature of the geotechnical soil is determined by these parameters. Thus, the forces on the skeleton of soil are different from the forces on the fluid. The forces on the soil skeleton can cause the deformation of both the soil itself and the fluid within the soil. In addition, according to different physical states, the density of medium can be divided into solid density and fluid density.

When the movement of soil particles, liquid, and air in the soil is not synchronized, the viscous force between them also weakens the elasticity of vibration. The range of action of vibration impact force on the cable is wide in the working process of the distributed optical fiber sensing system, and the effect of soil vibration signal is distributed in a section of fiber optic cable near the impact force. We used the theory of elastic half-space in our study. In this theory of elastic, there are three different types of waves including Primary Wave (P wave), Secondary Wave (S wave), and another elastic wave, Rayleigh wave (R wave). The vibration wave generated by the earth’s surface is transmitted in the form of the R wave.

In the elastic half-space model, Poisson’s ratio of foundation soil \( \mu \) and the modulus of rigidity \( G \) can be written as

\[
\mu = \frac{v_p^2 - 2v_s^2}{2(v_p^2 - v_s^2)},
\]

\[
G = \rho v_s^3,
\]

where \( v_s \) and \( v_p \) represent wave velocity of S wave and P wave in elastic half-space theory, respectively. Poisson’s ratio and shear modulus of soil are affected by the change of wave velocity. Under the vertical harmonic force of vibration source, at each point \( A_1 \), the soil vibration signal amplitudes can be expressed as

\[
A_1 = \frac{F}{4\pi G} \left[ \frac{2(1 - \mu)}{\sqrt{x^2 + Z^2} + \left(\frac{Z}{x^2 + Z^2}\right)^{3/2}} \right],
\]

where \( F \) shows the concentrated vertical harmonic force, \( Z \) is the buried depth of fiber, and \( x \) represents the abscissa of any point on the fiber.

3.2. Influence of Different Soil Types. When the soil types are different, the basic parameters of the soil will also alter in the same season in a certain area. Table 1 represents the
empirical values of two physical parameters of several common soil types in Shanghai city of China. The two physical parameters are very essential as the necessary dynamic parameter in the analysis of soil seismic response and an indispensable content in site vibration evaluation. \( \mu \) presents Poisson's ratio of foundation soil, and \( G \) denotes the modulus of rigidity.

According to Table 1, when the buried soil is different, Poisson’s ratio of the soil is mainly within the range of 0.20–0.35, with little change. This parameter has little effect. However, the modulus of rigidity varies extensively. This will cause alterations in the soil vibration signal and affect the recognition accuracy of the OFPS. The vibration amplitude of the vibration signal of the soil is simulated by using equation (2), and the experimental parameters are shown in Table 2.

Figure 5 represents the mathematical relationship between the modulus of rigidity and the vibration amplitude of the soil vibration signal. It shows that the vibration amplitude of the soil’s vibration signal gradually decreases by increasing the modulus of rigidity.

3.3. Analysis of Real Signals. Real soil vibration signals under different soil conditions were collected and analyzed. The distributed optical fiber sensing system used in the experiment is 22 km in Shanghai, with a buried depth of 1.2 m. The 22 km optical fiber is 12-core armored optical cable, which is directly buried under the surface without any packaging. The response time is 60 s or less, and the sampling frequency of the system is 750 Hz. We utilized a small tamping machine in the same gear to generate vibration signals and collected soil vibration signals from different sensor locations at the same temperature. The signals collected here are mechanical vibration signals. The soil types at the locations where the signals are collected are gravel soil, sandy soil, and clay with high moisture content. Figure 6 shows the signals of mechanical vibration under three certain conditions.

By comparison, the same vibration signal is changed after collecting by the OFPS in the case of different soil types, and the time-domain characteristics of some signals are quite different compared to others, such as the signals in clay with high moisture. By analyzing the collected soil vibration signals, the system gives the event recognition probability without the NN method. The specific information is shown in Table 3. According to Table 3, the system’s recognition probability of event occurrence varies greatly under different soil types before the improved neural network method is used. And the recognition probability is less than 85% and the recognition effect of the system is relatively poor.

<table>
<thead>
<tr>
<th>Physical parameters</th>
<th>Gravel soil</th>
<th>Sand</th>
<th>Cohesive soil</th>
<th>Sand clay</th>
</tr>
</thead>
<tbody>
<tr>
<td>( G )</td>
<td>54–65</td>
<td>31–42</td>
<td>16–59</td>
<td>18–39</td>
</tr>
<tr>
<td>( \mu )</td>
<td>0.20</td>
<td>0.20–0.25</td>
<td>0.25–0.35</td>
<td>0.25–0.3</td>
</tr>
</tbody>
</table>

Table 1: The empirical values of two physical parameters of different soil types.

Figure 3: The DAS signal acquisition system.

Figure 4: The structure of soil.

3.3. Analysis of Real Signals. Real soil vibration signals under different soil conditions were collected and analyzed. The distributed optical fiber sensing system used in the experiment is 22 km in Shanghai, with a buried depth of 1.2 m. The 22 km optical fiber is 12-core armored optical cable, which is directly buried under the surface without any packaging. The response time is 60 s or less, and the sampling frequency of the system is 750 Hz. We utilized a small tamping machine in the same gear to generate vibration signals and collected soil vibration signals from different sensor locations at the same temperature. The signals collected here are mechanical vibration signals. The soil types at the locations where the signals are collected are gravel soil, sandy soil, and clay with high moisture content. Figure 6 shows the signals of mechanical vibration under three certain conditions.

By comparison, the same vibration signal is changed after collecting by the OFPS in the case of different soil types, and the time-domain characteristics of some signals are quite different compared to others, such as the signals in clay with high moisture. By analyzing the collected soil vibration signals, the system gives the event recognition probability without the NN method. The specific information is shown in Table 3. According to Table 3, the system’s recognition probability of event occurrence varies greatly under different soil types before the improved neural network method is used. And the recognition probability is less than 85% and the recognition effect of the system is relatively poor.
4. Signal Recognition Methods

4.1. Improved Neural Network (Method 1). An improved NN method is proposed to solve the above problem in this paper. The architecture of the neural network mainly integrates the advantages of the network such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) in modeling, which can obtain the characteristic information of the signal data more effectively [40, 41].

First, the intrusion signals gathered by the system are labeled and standardized. Standardization of data refers to scaling the data to fall into a small specific interval. To ensure the reliability of the results, it is essential to standardize the data and utilize the standardized data for analysis before

<table>
<thead>
<tr>
<th>Force on the fiber optic cable ( F ) (kN)</th>
<th>Horizontal coordinates of fiber optic cable ( x ) (m)</th>
<th>Buried depth of fiber optic cable ( Z ) (m)</th>
<th>Poisson’s ratio ( \mu )</th>
<th>Numerical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0.25</td>
<td>0.012</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.5</td>
<td>0.25</td>
<td>0.010</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.5</td>
<td>0.25</td>
<td>0.008</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0.5</td>
<td>0.25</td>
<td>0.006</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0.5</td>
<td>0.25</td>
<td>0.004</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0.5</td>
<td>0.25</td>
<td>0.002</td>
</tr>
</tbody>
</table>

**Table 2**: The numerical value of the simulation parameter.

**Figure 5**: The relationship of vibration amplitude of soil vibration signal with the modulus of rigidity.

**Figure 6**: The signals under different conditions: (a) gravel soil, (b) sandy soil, and (c) clay with a high moisture.
modeling the neural network. The Max Abs Scaler is the standardization technique explored in this study adjusting each feature within the range of $[-1, 1]$. It is divided by the maximum absolute value within each feature. The main mathematical basis is shown in

$$x' = \frac{x}{\max}$$

(3)

In the formula, $x$ and $x'$ are the signal matrix before and after normalization, and max is the maximum value of a certain row or column of the signal matrix. Signal matrix is a kind of matrix whose element is soil vibration signal amplitude in each row. The elements are arranged by the same time interval. The specific situation depends on whether the normalization is the row compression mode or the column compression mode. This method has the characteristics of not destroying the data structure and not moving and gathering data, so it will not destroy any sparseness, and it is suitable for signals collected by optical fiber systems [42].

The standardized signals will be input into the improved NN for deep learning. The signals first are input into the CNNs. Through convolution, the CNNs effectively avoid the complex data reconstruction in signal feature extraction and classification. The special structure of weight sharing greatly reduces the parameters of the model and reduces the complexity. CNN mainly includes convolutional layer and pooling layer. The convolutional layer recognizes different shallow features of the input data by using multiple filters. The pooling layer is a filtering process. The pooling layer can reduce the connections between the convolutional layers, thereby reducing the computational burden and reducing the complexity of the operation. CNN can effectively reduce the frequency variation. LSTM can model the signal time series characteristics, connect CNN to reduce the signal frequency domain variation, and then insert the output results of CNN into several LSTM layers for time series modeling. Ultimately, the output results of LSTM are inserted into the fully connected layer for analysis, and the feature space will be mapped to the output layer. This layer is more easily categorized to create the feature representation that is more easily separated. Therefore, the recognition model established by this method has better recognition and classification characteristics than a single neural network. The illustrated neural network structure is mainly constructed by combining the features of the CNN network and the LSTM network comprising three CNN layers, one linear layer (dim red), one LSTM layer, and a fully connected layer in Figure 7(a). First, the intrusion signals gathered by the system are labeled and standardized. The Max Abs Scaler is the standardization technique explored in this study. The time of each labeled signal is 4 seconds, and the frequency is 750 Hz. The signals first are input into the CNNs. There are three CNNs in this method in Figure 7(b). The CNN of each layer includes a frequency-time filter $(1 \times 5)$ and a maximum pooling layer of 2 and a step size of 2. Since the output size of CNN is the number of feature maps $\times$ time $\times$ frequency, the size is very large. Adding a linear layer can effectively reduce the parameters without loss of accuracy by following the CNN layer. In experiments, it is found that it is appropriate to reduce the dimensionality so that the linear layer has 256 outputs. After passing through the CNN layer, a linear layer with 256 outputs is used to construct the input of the LSTM layer. By modeling the frequency, the output is passed to the LSTM layer to model the signal’s time series. The flow of LSTM is shown in Figure 7(c). The utilized LSTM layer included 64 units and 256 hidden units. Ultimately, the results of a fully connected layer are used as the output of the signal. With the abovementioned steps, the system can perform deep learning modeling simply on the collected intrusion signals to accelerate the creation of a neural network model. Then, the system is able to identify and classify the gathered intrusion signals through the model.

4.2. Method for Comparison (Method 2). In order to compare the effectiveness of Method 1, Method 2 is proposed. A 5-layer Deep Neural Network (DNN) is used (Figure 8) including one input layer, three hidden layers, and one output layer. For the three hidden layers, a hyperbolic tangent sigmoid function generally is selected as the activation function, in which five knots are set in the first hidden layer, four knots are set in the second hidden layer, and three knots are set in the three hidden layers based on an experimental parameters test. The output layer with a linear activation function consists of three knots.

5. Model Building and Recognition Test Analysis

5.1. Model Building and Recognition Test Analysis. In this section, the data collected in Shanghai was used to train the neural network model. First, the training sets and the test sets of the recognition model were prepared. We utilized signals from three different soil types to train and test the recognition accuracy of the neural network model. Then, the collected signals with time-domain characteristics were processed and labeled into 4 s sample data for each segment. The time-domain signals of labeled samples collected from three different soil types are shown in Figure 6. The division ratio of the training sets and the test sets is 7:2. The specific experimental data is shown in Table 4. In total, there are 2100 sets of data for training including 700 sets of data collected from gravel soil at 2.99 km, 700 sets of data collected from sandy soil at 10.21 km, and 700 sets of data collected from clay with high moisture content at 18.86 km for training neural networks to get recognition model. After training the model for all data, the test sets are selected which are different from the training sets to test the recognition effect of the model. A total of 900 sets of new

<table>
<thead>
<tr>
<th>Table 3: The event recognition probability under different soil types.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of soil</strong></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Gravel soil</td>
</tr>
<tr>
<td>Sandy soil</td>
</tr>
<tr>
<td>Clay</td>
</tr>
</tbody>
</table>
data are utilized as the test sets to test the recognition effect of the model including 300 sets from gravel soil, 300 sets from sandy soil, and 300 sets from clay.

When training, all training datasets are trained once for one epoch. The fitting degree of the neural network to the data is determined by the number of the epochs. Numerous epochs may cause overfitting of the data; otherwise, underfitting the data is caused. The number is determined by the validation accuracy of the neural network. Provided that the network self-test accuracy reaches over 99%, the training can be stopped to avoid overfitting caused by excessive epoch. Figure 9 represents the relationship between epoch and validation accuracy used in the neural network in the first method indicating that when the epoch reaches 1183 times, validation accuracy meets the requirement. Training loss is also close to 0 based on Figure 9.

After training the model, the test sets are selected to examine the recognition effect of the model. According to Table 5, the signal recognition average accuracy of the OFPS obtained by Method 1 is 95.1%, and the recognition average

![Diagram](image-url)

Figure 7: The improved neural network. (a) The flow of an improved neural network. (b) The convolutional layers. (c) The LSTM flow.
accuracy of the OFPS obtained by Method 2 is 89.8%. Based on the comparison, it is observed that the testing effect of Method 1 is better and the recognition accuracy is over 95%.

5.2. The Adaptability of the Method. The adaptability of the method was studied by implementing the test and verifying of classification and identification effect of the optical fiber warning system in the actual environment. The adaptability of the method referring to the data collected by the system in different regions has good recognition and classification capabilities. Thus, the signals of the three events were collected including manual digging (MG), mechanical excavation (ME), and noninvasion (NI) in different regions. The distributed optical fiber sensing system used in the experiment is 22 km in Shanghai, with a buried depth of 1.2 m. The response time is 60 s or less, and the sampling frequency of the system is 750 Hz. The signals of manual digging (MG) were obtained by digging the soil with a shovel. The signals of mechanical excavation (ME) were obtained by using a tamping machine to vigorously excavate the soil. Finally, we selected the places where no incident occurred to collect the signals of noninvasion (NI). This experiment is aimed to ensure the diversity and complexity of the collected data and to recognize the vibration signal at a remote location of the system. The vibration signal samples of three typical

<table>
<thead>
<tr>
<th>Type of soil</th>
<th>Number of training sets</th>
<th>Collection location of training sets (km)</th>
<th>Number of test sets</th>
<th>Collection location of test sets (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravel soil</td>
<td>700</td>
<td>2.99</td>
<td>300</td>
<td>2.99</td>
</tr>
<tr>
<td>Sandy soil</td>
<td>700</td>
<td>10.21</td>
<td>300</td>
<td>10.21</td>
</tr>
<tr>
<td>Clay</td>
<td>700</td>
<td>18.68</td>
<td>300</td>
<td>18.68</td>
</tr>
</tbody>
</table>

Table 4: The first experiment data.

<table>
<thead>
<tr>
<th>Type of soil</th>
<th>Method 1 (%)</th>
<th>Method 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravel soil</td>
<td>96.5</td>
<td>91.6</td>
</tr>
<tr>
<td>Sandy soil</td>
<td>95.4</td>
<td>89.4</td>
</tr>
<tr>
<td>Clay</td>
<td>92.9</td>
<td>88.9</td>
</tr>
<tr>
<td>Average</td>
<td>95.1</td>
<td>89.8</td>
</tr>
</tbody>
</table>

Table 5: The signal recognition accuracy in the first experiment.

![5-layer DNN network](image)

Figure 8: The 5-layer DNN network.

![Validation Accuracy](image)

![Training Loss](image)

Figure 9: The relationship between epoch and validation accuracy and training loss.

![Graph](image)
intrusion events in this experiment are collected at different positions to guarantee the validity and authenticity of the sample set. Similarly, the collected signals with time-domain characteristics are also processed and labeled into 4 s sample data. As shown in Figure 10, they are the time-domain signals of labeled samples gathered.

The specific information of the input training and test sets of the neural network is shown in Table 6. In total, there are 16733 sets of data for training including 5528 sets of the NI events data, 5674 sets of the MG events data, and 5531 sets of the ME events data applicable for training neural networks to build recognition models.

The test results are provided in Table 7. In the second experiment, the signal recognition accuracy of the OFPS obtained by Method 1 is 95.6%, and the average accuracy obtained by Method 2 is 90.4%. By comparing the data in Table 7, it can be observed that the accuracy of Method 1 remained above 95% and this improved NN has good adaptability in the actual environment and can be popularized in long-distance OFPS.

The specific information of the input training and test sets of the neural network is shown in Table 6. In total, there are 16733 sets of data for training including 5528 sets of the NI events data, 5674 sets of the MG events data, and 5531 sets of the ME events data applicable for training neural networks to build recognition models.

The test results are provided in Table 7. In the second experiment, the signal recognition accuracy of the OFPS obtained by Method 1 is 95.6%, and the average accuracy obtained by Method 2 is 90.4%. By comparing the data in Table 7, it can be observed that the accuracy of Method 1 remained above 95% and this improved NN has good adaptability in the actual environment and can be popularized in long-distance OFPS.

5.3. The Real-Time Capability of the Method. Through the two experiments above, it can be seen that this new method can accurately classify signals, indicating that this improved NN is feasible. Therefore, we will run this identification part in real time in the pipeline prewarning system. Figure 11 shows the process of the pipeline prewarning system for real-time detection. The signal collection part of the pipeline prewarning system collects the data of each optical fiber sensor on the optical fiber every four seconds. When a harmful intrusion event such as mechanical excavation occurs near the pipeline, it will cause the fiber optic sensor to respond. After the data is collected, it will be transmitted to the neural network for identification and classification, and the predicted probability of the occurrence of the intrusion event at that point will be given. The data used for model training of the neural network is collected in different soil environments, which will make the training set more complete. The data used for testing is the real-time data collected by the pipeline prewarning system every four seconds. This can better verify the real-time performance of the system.

To ensure the integrity of the analysis, the predictive ability of the system for the occurrence of interference events was compared between Method 1 and Method 2. The abscissa in the next two figures is the length of the whole optical fiber, and the ordinate is the probability of the predicted event. Figure 12 represents the recognition probability of each point on the whole fiber length when occurring in the manual mining event. The experimental data are the data collected on the whole optical fiber length during manual mining at 6.52 km. At this point, no other intrusions exist elsewhere on the fiber. It can be seen in Figure 12(a) that the system can estimate the manual mining events at 6.52 km with Method 2. However, the recognition rate is less than 80%, and there is a false alarm at other locations to a certain extent. After using the improved NN, it is observed (Figure 12(b)) that the recognition rate at 6.81 km exceeds 95%, and there is no false alarm at other locations. The overall recognition rate of the system model is enhanced.

![Figure 10: The signal samples. (a) The NI event. (b) The MG event. (c) The ME event.](image-url)
Similarly, Figure 13 represents the recognition probability of each point on the whole fiber length when occurring in the mechanical mining event. The experimental data are the data collected on the whole fiber length during mechanical excavation at 5.57 km. No other intrusions exist at any other point along the length of the fiber. Considering the large intensity and wide influence range of mechanical excavation events, it is observed that the occurrence of events is predicted within a range of about 5.57 km. The comparison of the two figures shows that, before using the improved NN, the recognition rate of the system for mechanical excavation was also about 80%, and the false alarm was reported at other locations. After analyzing the data via the improved NN, it is observed that the recognition rate of the event exceeds 95%, and no events occur in other locations. This is consistent with the actual situation.

According to Sections 5.2 and 5.3, the improved NN has better adaptability and real-time capability. The reason why the improved NN method works is as follows: LSTM is good at processing time-related data; CNN can abstract and extract features from data at multiple time points to reduce invalid data. CNN has the ability to process time information, and the ability of this abstract feature of CNN is higher than that of LSTM. The CNN layer can extract hidden information in the time dimension and then pass higher-quality and high-concentration features to the LSTM layer. These characteristics determine that the improved NN is more suitable for the recognition of optical fiber vibration signals than Method 2 and other methods. However, this method also has some disadvantages. First, this improved NN has only been tested in the field of fiber vibration signal recognition and has not been tried in other speech signal recognition. Besides, this method requires relatively high GPU performance, so a high-performance GPU is required for calculation. Since the optical fiber early warning system requires real-time detection, a high-performance computer is also necessary.

<table>
<thead>
<tr>
<th>Type of event</th>
<th>Number of training sets</th>
<th>Collection location of training sets (km)</th>
<th>Number of test sets</th>
<th>Collection location of test sets (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NI</td>
<td>5528</td>
<td>2.99</td>
<td>2000</td>
<td>14.96</td>
</tr>
<tr>
<td>MG</td>
<td>5674</td>
<td>10.21</td>
<td>2000</td>
<td>16.86</td>
</tr>
<tr>
<td>ME</td>
<td>5531</td>
<td>18.68</td>
<td>2000</td>
<td>6.24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of event</th>
<th>Method 1 (%)</th>
<th>Method 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NI</td>
<td>97.5</td>
<td>93.6</td>
</tr>
<tr>
<td>MG</td>
<td>94.3</td>
<td>89.7</td>
</tr>
<tr>
<td>ME</td>
<td>95.2</td>
<td>88.3</td>
</tr>
<tr>
<td>Average accuracy</td>
<td>95.6</td>
<td>90.4</td>
</tr>
</tbody>
</table>

Table 6: The second experiment data.

Table 7: The signal recognition accuracy in the second experiment.
6. Conclusion

In this paper, we found that the same vibration signals would vary greatly when collected in different soil environments and this problem would reduce the signal recognition accuracy of the OFPS. This phenomenon was theoretically studied, and soil vibration signals under different soil conditions were collected and analyzed. A novel highly intelligent method was proposed to improve signal acquisition of gas and oil pipeline prewarning system in the complex soil environment. First, we analyzed the effects of the real soil environment. The two principles of elastic half-space theory and optical fiber sensing theory were creatively integrated. The results indicate that this problem will affect the recognition rate of the OFPS. Therefore, we utilized an improved NN based on LSTM and CNN which is mainly attempted in the field of optical fiber signal recognition. Based on the investigation, some remarkable conclusions can be drawn as follows:

1. The improved high-intelligence method is a new neural network method, which has the advantages of LSTM and CNN. LSTM is good at processing time-related data, and CNN can abstract and extract features from data at multiple time points to reduce invalid data. These advantages determine that the improved NN is more suitable for the recognition of optical fiber vibration signals than other traditional methods. It is a new attempt in the field of optical fiber prewarning. Through experiments, it is verified that this improved NN has good adaptability and real-time performance.

2. We collected three types of event signals in different soil environments for training and recognition, and
it is proved that the improved high-intelligence method has good adaptability, and the average recognition rate of three types of event signals is more than 95%.

(3) This improved method was used to conduct real-time online monitoring of the pipeline prewarning system in Shanghai. Through the experiments of two types of intrusion events, it is proved that the method has good real-time performance and can quickly give the location and the probability of the event, and the prediction rate of the event has reached about 95%.

Since this method has good adaptability and real-time performance, it can be popularized in the optical fiber prewarning systems all over the world, which is helpful to maintain the safety of gas and oil pipelines and has far-reaching significance for the intelligence of gas and oil transportation systems. Furthermore, the results of the study have high practical value for monitoring the safety of gas and oil pipelines.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

The authors gratefully acknowledge the support given by HILONG Group of Companies. This work was supported by Shanghai Municipal Economic and Informatization Commission (Grant no. 2018-RGZN-02015).

References


