

Received May 4, 2022, accepted May 19, 2022, date of publication May 23, 2022, date of current version June 6, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3177595

Leak Identification Based on CS-ResNet Under Different Leakage Apertures for Water-Supply Pipeline

LIN MEI^{1,2}, JUN ZHOU^{1,2}, SHUAIYONG LI³, MENGQIAN CAI³, AND TONG LI^{1,2}

¹Chongqing Special Equipment Inspection and Research Institute, Chongqing 401121, China

²Key Laboratory of Electromechanical Equipment Security, Western Complex Environment for State Market Regulation, Chongqing 401121, China

³Key Laboratory of Industrial Internet of Things and Networked Control, Ministry of Education, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Corresponding author: Shuaiyong Li (lishuaiyong@cqupt.edu.cn)

This work was supported in part by the National Natural Science Foundation Project of China under Grant 61703066; in part by the Natural Science Foundation Project of Chongqing under Grant cstc2018jcyjAX0536; and in part by the Chongqing Special Project of Technology Innovation and Application Development under Grant cstc2018jcsx-cyztzxX0028, Grant cstc2019jcsx-fxydX0042, and Grant cstc2019jcsx-zdztzxX0053.

ABSTRACT Considering the problem of difficulty in transmission and storage due to a large amount of data in the water-supply network monitoring system based on a wireless sensor network (WSN), we propose a sparse representation of the water-supply network monitoring data by using compressed sensing (CS) method. At the same time, aiming at the problem of low leakage identification accuracy caused by information loss under compressed sensing, we propose a leak identification method for a water-supply pipe network based on compressed sensing and deep residual neural network (ResNet). Firstly, under the condition that the observation matrix ensures the integrity of signal information, the compressed sensing theory is used to compress and observe leakage signals to obtain observation data, to reduce the redundant information and volume of the data. At the same time, the observation data is preprocessed to realize the transformation of a one-dimensional signal to a two-dimensional matrix. Then the residual neural network is trained by using the two-dimensional data to realize the automatic, efficient, and accurate leak identification under different leakage apertures. Experimental results show that the proposed method can obtain relatively high accuracy and greatly reduce the training time of ResNet by using compressed data. When the Compression rate (CR) is 70% and the observation matrix is a Gaussian random matrix, the average accuracy is 96.67% and the training time is only 50% compared to uncompressed data. The research work provides a new intelligent leak identification under different leak apertures using WSN and has important application prospects in saving water resources.


INDEX TERMS Pipeline leakage, compressed sensing, residual neural network, observation matrix.

I. INTRODUCTION

A survey report shows that 43% of water is wasted due to leakage in pipelines [1]. If a pipeline leak is not detected, the leakage will generate a significant waste of water resources. Therefore, the research work of leak identification in the water-supply pipeline has been widely concerned at home and abroad. In recent years, wireless sensor network (WSN) has been developed rapidly and attracts more and more attention in the field of pipeline monitoring, which organizes and

combines tens of thousands of sensor nodes freely through wireless communication technology. The leak identification based on WSN can efficiently discover the leak events to reduce the waste of water resources and economic losses [2]–[4], which is popularly applied in leak monitoring of water-supply pipelines at present.

Nowadays, several leak identification schemes have emerged, such as time-domain reflectometry (TDR) [5], [6], real-time transient modeling [7], negative pressure wave-based method [8]–[10], transient pressure wave-based method [11]–[16] and acoustic-based methods [17], [18]. The TDR-based method is suitable and effective which requires

The associate editor coordinating the review of this manuscript and approving it for publication was Wen Chen .

only an already-installed metal pipeline as a prerequisite. It is necessary to build a theoretical model for real-time transient modeling, which lacks adaptability leading to low accuracy of leak identification and has large computation. The negative pressure wave is not obvious in the case of small leakage under complicated background noise, which may result in a larger error of leak identification. The transient pressure wave-based method has the desirable merits of high efficiency, low cost, and non-intrusion which requires additional excitation signals [11]. The acoustic-based method has high sensitivity and accuracy of leak identification [19], which has great advantages in leak identification in water-supply pipelines [20]–[23].

The most popular leak identification methods by acoustic wave adopt machine learning such as k-nearest neighbors, support vector machines (SVMs) [24], and artificial neural networks [25]. In the classification process, the extraction of suitable features of leak acoustic signal is considered as an important factor for increasing performance. Recently, deep learning has become the most successful machine learning method and attracted more attention in leak identification [26], [27]. In particular, convolutional neural network (CNN) has become popular in condition monitoring [28], which can automatically extract features from raw data for fault identification and classification. Ince *et al.* [29] and Abdeljaber *et al.* [30] proposed fast and accurate condition monitoring, acoustic-based structural health monitoring, and early fault identification systems using 1D-CNN. Similarly, the goal of the approach in [31] was to autonomously learn useful features for bearing fault detection in rotating machinery based on raw 1-D signals. The autonomous feature learning approach was compared with the handcrafted features approach using the same data to objectively quantify performance, with the former producing an accuracy of 93.61% and the latter 87.25%.

Leak identification based on wireless acoustic sensor network (WASN) has widely applied in water-supply pipelines and generates a mass of data, which results in a lot of training time of CNN model and has a great impact on the real-time performance of monitoring system. The compressive sensing (CS) first proposed in [32] attracts increasing attentions. CS provides a promising technique to reconstruct a signal from a small number of measurements by exploiting the sparseness nature of the signal [33], [34], which has shown significant success in the processing of big data [35], synthetic aperture radar imaging [36], and WASN based monitoring system [37]. Hence, the combination of CS and CNN for leak identification is expected to solve the problem of poor real-time performance caused by a large training time for monitoring of water-supply pipelines based on WASN.

In this paper, we propose a CS-ResNet framework for multi-class leak identification under leakage apertures in WASN based water-supply pipelines. Our contributions are as follows.

- 1) Under the condition that the observation matrix ensures the integrity of the signal information, the compressed

sensing theory is used to compress and observe the leakage signal of the water-supply pipeline to obtain the observation data, so as to reduce the redundant information and volume of the data.

- 2) The observation data is preprocessed to realize the transformation of one-dimensional signal to two-dimensional matrix data.
- 3) The residual neural network is trained by using the observed data, to realize the automatic, efficient, and accurate identification of the leakage under different apertures of the water-supply pipeline.
- 4) The experiments are conducted to verify the effectiveness of the proposed method.

The rest of the paper is organized as follows. In Section II, we provide a brief overview and derivation of theoretical models. In Section III, we give an identification method of the leakage of water-supply pipeline based on residual neural network in compressed sensing domain. In Section IV, the experimental study on leak identification of water-supply pipeline is conducted. In section V, we end the paper with some concluding remarks.

II. BASIC THEORY

A. COMPRESSED SENSING THEORY

The main idea of compressed sensing theory is that if the signal has sparsity, the signal can be projected into a space with a very low dimension relative to the original signal through a matrix (observation matrix) irrelevant to sparsity, so as to obtain the observation signal with a very small length relative to the original signal. The reconstruction algorithm can be used to reconstruct the low-dimensional observation signal into the original signal. Its principle is as follows:

Suppose the signal x with signal length N is linearly measured in an observation matrix $\Phi \in R^{M \times N} (M \ll N)$ to obtain the observation value $y \in R^M$, as follows:

$$y = \Phi x \quad (1)$$

This process can be regarded as the linear projection of the signal on the observation matrix.

Now that the signal is observed and measured, it is necessary to consider reconstructing the signal x from the observed value y . Since $M \ll N$, (1) has an infinite solution, and it is difficult to reconstruct the signal x . If the signal x is sparse, x is k -sparse, and y and Φ also meet certain conditions, it is proved by the theory that x can be reconstructed by solving the optimization problem shown in (2), expressed as:

$$x' = \arg \min \|x\|_0 \text{ s.t. } y = \Phi x \quad (2)$$

where $\|\cdot\|_0$ is the l_0 -norm of the vector. The compressed sensing theory of E. J. Candes points out that to reconstruct k -sparse signal x from high precision, the dimension of observation value y must satisfy (3), and the observation matrix Φ must satisfy the restricted isometric property (RIP).

$$M = O(\kappa \lg(N)) \quad (3)$$

The sparse representation theory of signal indicates that the signal x can be sparsely represented by a certain transformation Ψ or a certain over-complete dictionary Ψ , then the signal x can be represented by the sparse signal s , which is:

$$x = \Psi s \tag{4}$$

Since (1) can sparsely represent the signal x , then:

$$y = \Phi x = \Phi \Psi s = A s \tag{5}$$

where $A = \Phi \Psi, A \in R^{M \times N}$ is called perception matrix, as shown in Figure 1.

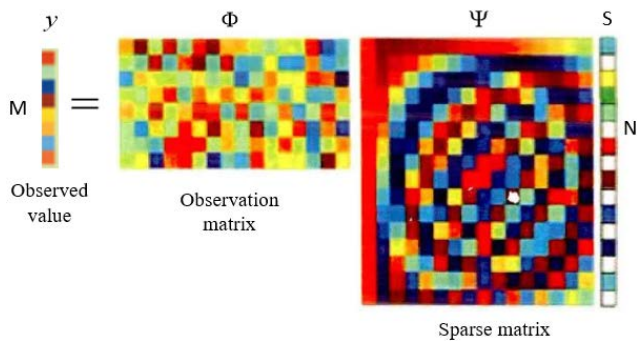


FIGURE 1. Compressed sensing measurement process.

Equation (5) is an underdetermined equation, and to reconstruct the linear system of equations with a unique sparse solution from, the constraint of sparsity needs to be added. Equation (5) can be regarded as the linear observation value of the sparse signal under the perception matrix A , and at the same time A satisfies the RIP condition, the problem of reconstructing the signal x is transformed into solving the optimal norm to reconstruct the sparse signal, expressed as:

$$s' = \arg \min \|s\|_0 \text{ s.t. } y = A s \tag{6}$$

The reconstructed sparse signal s' is obtained, and the reconstructed signal x is obtained by sparse inverse transformation of (7).

$$\hat{x} = \Psi s' \tag{7}$$

CS theory breaks through the frequency limitation of the Nyquist sampling theorem, and at the same time realizes the acquisition and compression of the signal, using a small amount of observed sampling data to retain a good approximation to the original signal, which can achieve the reconstruction of the original signal. Compression rate: the ratio of the length of the original signal minus the length of the observed value to the length of the original signal, which is used to measure the degree of compression and acquisition of the original signal. It is defined as:

$$CR = \frac{N - M}{N} \times 100\% \tag{8}$$

B. COMPRESSION ACQUISITION

In the compressed sensing theory, if the signal has sparsity, the observed value can be obtained by linear projection of the signal through the observation matrix, as shown in equation (1). The data dimension of the observed value is different from that of the original signal, and the dimension of the observed value is much lower than that of the actual signal, so as to reduce the dimension of the data. In the design of the observation matrix, the RIP conditions must be obeyed while ensuring the integrity of the original signal information, which realizes a high-precision reconstruction of the original signal.

On the premise that the observation matrix ensures the integrity of the signal information, if the signal x has sparsity, the signal x is linearly projected under the observation matrix to obtain the low-dimensional observation value y , which realizes the dimensionality reduction of data from N to M dimensions, and its compressed collection and dimensionality reduction method as shown in Figure 2.

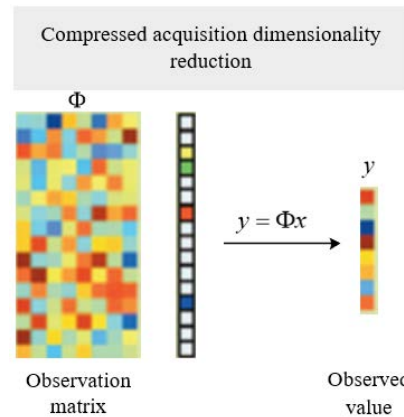


FIGURE 2. Dimension reduction of compressed acquisition.

Since the observation matrix ensures the integrity of the signal information, the obtained observation value contains a lot of useful information of the original signal. Combining CS with the convolution neural network model, CS reduces the volume of data and greatly simplifies the computational complexity in model training, thus improving the recognition efficiency of the convolutional neural network model. The compression rate of CS is a performance index to measure the degree of data dimensionality reduction. The higher the CR, the less information of the original signal contained in the observation value, which has an impact on the accuracy of the convolutional neural network model in the compressed sensing domain.

C. OBSERVATION MATRIX

The observation matrix $\Phi \in R^{M \times N} (M \ll N)$, also called the measurement matrix, is an important part of the compressed sensing framework. The observation matrix is used to observe the original signal to obtain the observation value, and the observation matrix is also required for signal reconstruction from the observation value. Therefore, the observation

matrix plays a very important role in the acquisition of the observation value and the signal reconstruction. In the compressed sensing theory, the observation matrix must be incoherent with the sparse matrix of the signal, that is, the matrix Φ cannot be expressed linearly by the matrix Ψ . The commonly used observation matrices include Gaussian random observation matrix, Bernoulli random observation matrix, local Hadamard observation matrix, and Fourier observation matrix.

1) GAUSSIAN RANDOM OBSERVATION MATRIX

In compressed sensing, the Gaussian random observation matrix is the most widely used observation matrix. Its design method is to randomly generate a matrix Φ of size $M \times N$, and each element of the matrix Φ must obey the Gaussian distribution of the mean value of 0 and the variance of $\frac{1}{\sqrt{M}}$, that is:

$$\Phi_{i,j} \sim N(0, \frac{1}{\sqrt{M}}) \tag{9}$$

The matrix has strong randomness. Because the matrix is not related to the sparse matrix of most orthogonal bases and orthogonal dictionaries, when the signal length is N , the sparsity is K , and the observation value under the Gaussian random observation matrix satisfies $M > cK \lg(N/K)$, the matrix satisfies the RIP property with great probability, so the original signal can be reconstructed accurately.

2) BERNOULLI RANDOM OBSERVATION MATRIX

The Bernoulli random observation matrix is also one of the commonly used observation matrices in compressed sensing framework. Its design method is to design a matrix Φ with the size of $M \times N$, and make each element in the matrix Φ obey Bernoulli distribution, that is:

$$\Phi_{i,j} = \begin{cases} +\frac{1}{\sqrt{M}}P = \frac{1}{2} \\ -\frac{1}{\sqrt{M}}P = \frac{1}{2} \end{cases} \tag{10}$$

The matrix has strong randomness like the Gaussian random matrix, and it is proved by theory that this matrix satisfies the RIP property. Compared with the Gaussian random matrix, the element of Bernoulli random observation matrix is ± 1 , so it is easier to implement and store in practical application.

3) LOCAL HADAMARD OBSERVATION MATRIX

The local Hadamard matrix can also be used as the observation matrix in compressed sensing. Its design method is to generate a Hadamard matrix with the size of $N \times N$, and then randomly select M row vectors from the matrix to construct an observation matrix Φ with the size of $M \times N$. The observation matrix is a local Hadamard matrix of size $M \times N$ consisting of M row vectors selected from the Hadamard matrix. At the same time, since the Hadamard matrix is orthogonal, the local Hadamard has partial orthogonality and non-correlation, which makes the less observations are needed to reconstruct the signal under the observation matrix.

4) PARTIAL FOURIER OBSERVATION MATRIX

The Partial orthogonal matrix can also be used as the observation matrix in compressed sensing, such as partial Fourier matrix, and it is proved theoretically that partial orthogonal matrix also satisfies the RIP condition. The design method of partial orthogonal matrix is to generate an orthogonal matrix H of size $N \times N$, then randomly select M row vectors from the matrix H to form a matrix of size $M \times N$, and finally the matrix of size $M \times N$ is normalized by column vector to obtain the observation matrix Φ . Under this observation matrix, to accurately reconstruct the original signal, its sparsity must meet the following requirements:

$$\kappa \leq c \frac{1}{\mu} \frac{M}{(\log(N))^6} \tag{11}$$

where $\mu = \sqrt{M} = \max(|H_{i,j}|)$, when $\mu=1$, the partial orthogonal matrix becomes the partial Fourier observation matrix.

D. RESIDUAL NEURAL NETWORK

There are many models based on convolutional neural network, such as LeNet, AlexNet, VGG16, Inception. At present, with the improvement of computer computing power, the number of convolution network layers of the model is increasing, but there is a problem that the effect of the model is not very good when the number of convolution network layers is increasing. He Kaiming’s team proposed the residual neural network structure in 2016, which can improve this problem. The residual neural network solves the gradient dispersion problem by making up the weight of the shallow network for the weight of the deep network, and has a very strong representation ability.

1) RESIDUAL ELEMENT

The residual element is the main component of the residual neural network, which mainly includes two-layer residual element and three-layer residual element, as shown in Figures 3 and 4. A layer of identity mapping X is added to the residual neural network, which can increase the depth of the neural network without degrading the model effect.

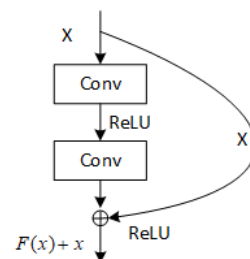


FIGURE 3. Two-layer residual element.

The two-layer residual element has fewer network layers and is generally used for the model with less demand for network layers, which can reduce the number of parameters for model training and recognition, to reduce the amount of

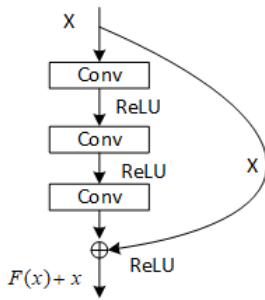


FIGURE 4. Three-layer residual element.

calculation of the model. The three-layer residual element has more network layers, which is mostly used in models with more demand for network layers. In this paper, the two-layer residual element is used to construct the convolutional neural network.

2) CONVOLUTION LAYE

The convolution layer is one of the basic unit structures of convolution neural network and is a network structure for feature extraction of input data. The convolution layer uses the convolution kernel to perform sliding convolution operation on the local area of the input data to extract the corresponding features. The feature of weight sharing of the convolutional layer can reduce its network parameters, which is beneficial to avoid over-fitting caused by too many parameters. The formula of convolution operation is shown in (12).

$$y^{l(i,j)} = M_i^l x^{l(r^j)} = \sum_{j'=0}^{W-1} M_i^{l(j')} x^{l(j+j')} \quad (12)$$

Where $x^{l(r^j)}$ is the i -th convoluted local region of the l -th layer, $M_i^{l(j')}$ is the j -th weight of the i -th convolution kernel of the l -th layer, and W is the length of the convolution kernel.

III. IDENTIFICATION METHOD OF LEAKAGE OF WATER-SUPPLY PIPELINE BASED ON RESIDUAL NEURAL NETWORK IN COMPRESSED SENSING DOMAIN

The observed value contains a lot of useful information of the original signal, which can be directly used to identify the leakage of the water-supply pipeline. Because the residual neural network model has better performance, this paper proposes a method for identifying the leakage of the water-supply pipeline based on CS and the residual neural network (CS-ResNet).

A. FRAMEWORK OF PROPOSED METHOD

The flow framework of the method is shown in Figure 5. The method is mainly divided into four parts: compression acquisition, data preprocessing, residual neural network training, and water-supply pipeline leakage identification.

Compression acquisition: At the leakage of the water-supply pipeline, the leakage vibration signal is acquired, and

the observation value is obtained by observing the acquired signal under an observation matrix.

Data preprocessing: Use observations to form an observation dataset. Then data preprocessing is performed to convert the one-dimensional water-supply pipeline leakage signal into a two-dimensional positive matrix format.

Residual neural network training: The preprocessed data is constructed into a data set, and the data set is divided into training set and test set at a ratio of 4:1. Then the residual neural network is trained to update the parameters and errors in the network to achieve the optimal model, and the water-supply pipeline leakage identification model is obtained.

Leakage identification: The leakage data after data preprocessing is sent to the leakage identification model based on CS-ResNet for identification, and the result of the leakage identification of the water-supply pipeline can be obtained.

B. RESIDUAL NEURAL NETWORK MODEL

In this paper, according to the length and signal characteristics of the leakage vibration signal data set of the water-supply pipeline, and the number of network layers of residual neural network should not be too deep, the two-layer residual element is selected to construct an 18-layer residual neural network model (ResNet18) as the residual neural network model of the leakage recognition experiment. Its network structure is shown in Figure 6.

C. DATA PREPROCESSING

Generally, the collected water-supply pipeline leakage vibration signal is a one-dimensional time-domain signal, which is difficult to be processed directly by convolutional neural network. Therefore, it is necessary to preprocess the water-supply pipeline leakage vibration signal data sets with different leakage apertures to convert the original water-supply pipeline leakage vibration signal into a matrix of size $n \times n$.

In general, the environment of water-supply pipeline leakage is very complex. In the collected water-supply pipeline leakage vibration signals, the median difference of the data is relatively large, and the number of outliers and outliers is large, which will adversely affect the identification accuracy and convergence performance of the leakage identification model. Therefore, the data need to be standardized and normalized.

The standardized expression of standard deviation is shown in (13).

$$x' = \frac{x - u}{\sigma} \quad (13)$$

where u is the mean value of the data sample, σ is the standard deviation of the data sample.

The normalized expression is shown in (14).

$$x'' = \frac{x' - x_{\min}}{x_{\max} - x_{\min}} \quad (14)$$

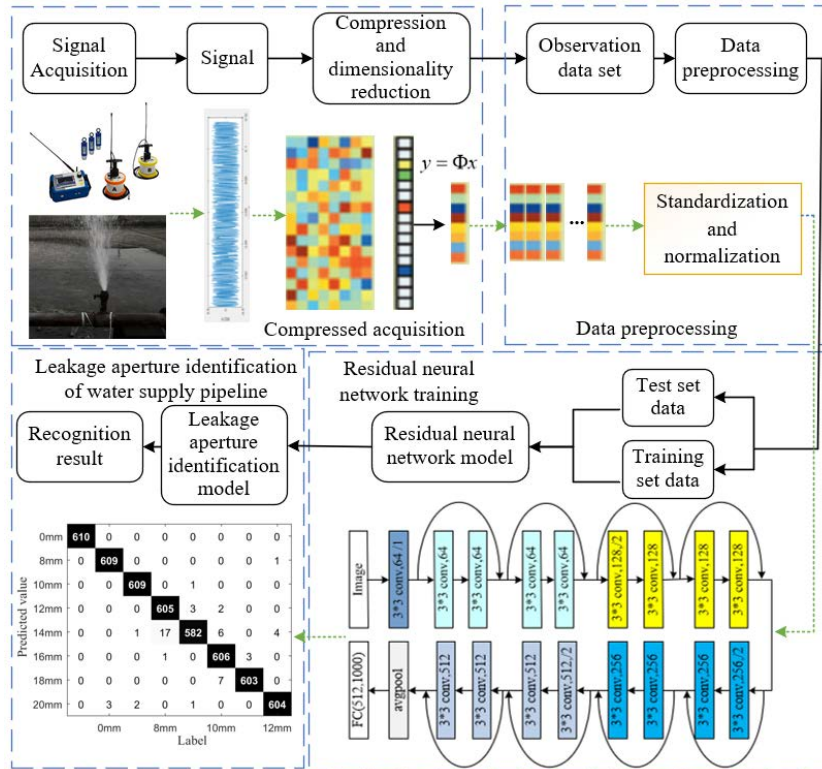


FIGURE 5. The framework of the leakage identification method of water-supply pipeline based on CS-ResNet.

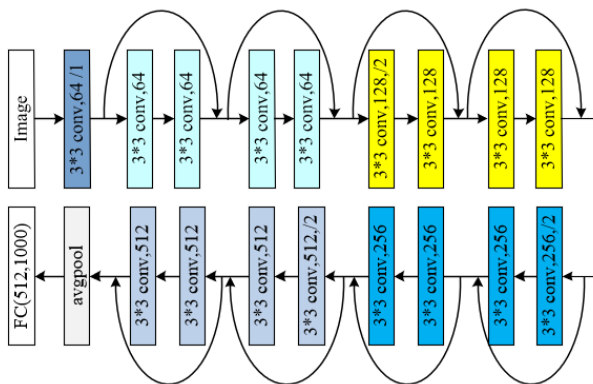


FIGURE 6. The network structure of ResNet18.

where x'' is the normalized data sample, x_{\min} is the minimum value of the data sample, and x_{\max} is the maximum value of the data sample.

After normalization, matrix mapping is performed on each data, each data sample is transformed into a matrix of size $n \times n$. And the matrix dimension must be greater than the number of data features. If the data length is not enough, a zero filling operation is taken for the deficiency of the matrix.

D. TRAINING OF RESIDUAL NEURAL NETWORK

To obtain accurate output, it is necessary to optimize the network model and update the model parameters.

The convolutional neural network has the same training method as a general neural network. It needs to be trained step by step according to loss function, gradient descent algorithm and error backpropagation to make it reach the optimal model. The training process of residual neural network is shown in Figure 7.

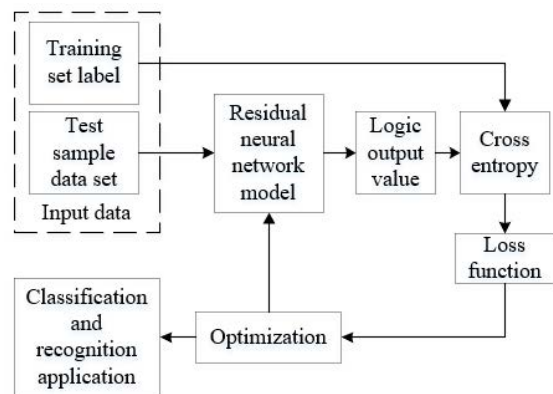


FIGURE 7. The training process of convolutional neural network.

IV. EXPERIMENTAL STUDY ON LEAK IDENTIFICATION OF WATER-SUPPLY PIPELINE

To verify the performance of the leakage identification method of the water-supply pipeline based on CS-ResNet, and the influence of compression rate and observation matrix

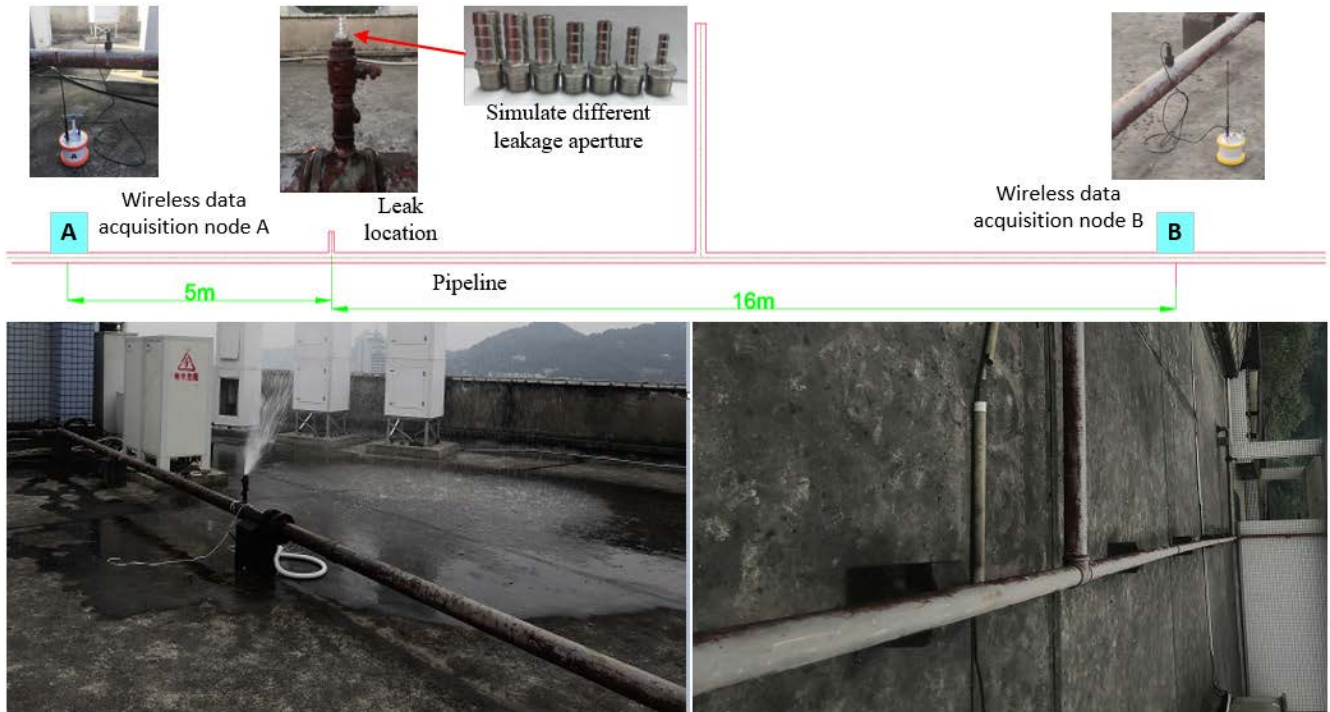


FIGURE 8. Experimental site of leakage of water-supply pipeline with different diameters.



FIGURE 9. Size of 304 stainless steel hexagonal pagoda with 4 points.

on the method, ResNet18 residual neural network model is used in this experiment. Different compression rates and observation matrices are used for compression collection, different observation value data sets are constructed, and the leakage aperture identification model of the water-supply pipeline is trained. The trained model is used to identify the leakage of the water-supply pipeline under different apertures, and the identification accuracy of test set is used to measure performance of the model. All of the computations are conducted using a computer with 2.9 GHz Intel processor, a maximum memory of 16 GB and GeForce GTX 1660 graphics, and all of the programs are written in Tensorflow 2.1 environment.

A. SIGNAL ACQUISITION OF LEAKAGE OF WATER-SUPPLY PIPELINE

To verify the performance of the method proposed in this paper, an experimental site that simulates the leakage of water-supply pipes with different apertures is built based

on laboratory equipment. As shown in Figure 8, the leakage of water-supply pipelines with different diameters is simulated by installing straight water pipe joints with different diameters on the fire pipe, and the data acquisition system of SebaKMT composite correlator is used for data acquisition. Among them, according to reference [19] and the experimental analysis results of leakage signal characteristics, the dominant frequency band of leakage vibration signal energy is 200Hz-1500Hz. At the same time, the leakage vibration signal is relatively weak as a whole, and is susceptible to noise interference, resulting in a low SNR. A high-sensitivity accelerometer is required to obtain the leakage signal. The higher the sensitivity, the stronger the correlation of the collected signals, and the smaller the leak detection positioning error. Therefore, the high-sensitivity Lens acceleration sensor LC0115 selected in this paper, its detailed specifications are shown in Table 1:

In the process of data acquisition, the pipeline parameters of the fire pipe: the pressure is 0.2MPa, the material is steel,

TABLE 1. Specifications of lc0115.

Model	Sensitivity (mV/g)	Measurement range (g)	Frequency Range (Hz/±10%)	Resonant frequency (kHz)	Resolution (g)	Impact resistance (g)
LC0115	5000	1	0.1-1500	7	0.000004	200

the inner diameter is 100mm, and the propagation speed of the leakage vibration signal in the steel pipeline can be obtained as 1250m/s according to the sound velocity empirical model. According to the reference [19] and the characteristics of the leakage vibration signal of the water supply pipeline, the dominant frequency range of the signal energy is mainly concentrated between 200Hz-1500Hz. It can be known from the sampling law that a sampling frequency that is more than twice the maximum frequency can meet the requirements of signal acquisition without distortion. In addition, combined with the parameter requirements of the AD converter of the selected data acquisition device, the maximum can only be set to 6554Hz. Therefore, we use an anti-aliasing filter with a cutoff frequency of 1500Hz at the data acquisition end to filter out signals greater than 1500Hz, and use a sampling rate of 6554Hz to satisfy Nyquist's law to ensure that the collected data will not be distorted. The number of sampling points is set to 4096 for the following two reasons, which can not only meet the storage requirements of the data acquisition terminal register, but also meet the requirements of the effective detection distance of the leakage vibration signal of the water supply pipeline under different circumstances. 1) The number of sampling points is set according to the capacity of the register of the data acquisition terminal used to ensure that the data will not overflow or be lost during storage. 2) Determine the number of sampling points according to the length of the pipeline under test. The effective detection length of the leak detection method proposed in this paper is within 200 meters for buried pipelines, and the effective detection length for overhead pipelines is within 500 meters. The vibration signal of pipeline leakage beyond this detection length is so weak that the leakage cannot be accurately identified. The length of the pipeline used in this experiment is 50 meters. The maximum time delay of 4096 sampling points selected in this paper is $4096 \times (1/6554) = 0.625s$. According to the diameter of the experimental pipeline in this paper is 100mm, the propagation speed of the leakage vibration signal is 1250m/s, and the maximum length of the pipeline that can be detected is $1250m/s \times 0.625s = 781.25m$. The straight water pipe joint adopts the 4-point 304 stainless steel hexagonal pagoda. Its size is shown in Figure 9. There are seven different diameters: 4 points -8mm, 4 points -10mm, 4 points -12mm, 4 points -14mm, 4 points -16mm, 4 points -18mm, 4 points -20mm.

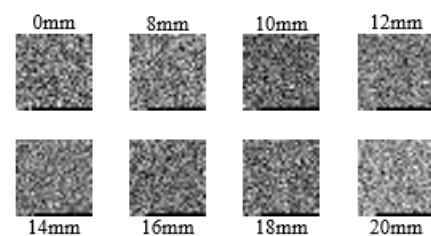
As shown in Figure 8, the wireless data acquisition node A is attached to the pipeline 5 meters to the left of the leak through its magnetic base, and node B is attached to the

pipeline 16 meters to the right of the leak in the same way. The acquisition pipeline is 21 meters long, and the wireless data acquisition node B has a sub-pipeline. Seven different 304 stainless steel hexagonal pagodas are installed at the leakage place to collect the leakage acoustic and vibration signals of the water-supply pipeline, and one leakage-free signal is collected to form eight different types of data. The wireless data acquisition nodes A and B collect 3050 samples of each type of data respectively, and the signal length of each sample is 4096. The data set collected by the wireless data acquisition node B is used in this experiment. The data set is divided into training data set and test data set at a ratio of 4:1, and the number of samples in the training data set is 19520 and the number of samples in the test data set is 4880.

B. OBSERVATION DATA SET AND DATA PREPROCESSING

To verify the performance of the water-supply pipeline leakage identification method based on CS-ResNet, each data sample in the data set is observed and compressed under a certain observation matrix and CR to obtain the corresponding observation data sample, to realize the compressed acquisition of the water-supply pipeline leakage signal. Finally, the observation data set samples of different leakage signals are obtained, where each type of data is still 3050 and the number of overall samples is 24400.

According to Section 2.3, data preprocessing is performed for each data sample in the observation data set. When the observation matrix is a Gaussian random matrix and CR is 80%, the data length of the sample before the observation is 4096, and the length of the observation value after observation compression is 819. After the data sample of the observation value is preprocessed, a group of data is randomly selected and transformed into gray image format through matrix mapping.

**FIGURE 10.** Visualization of observations with a CR of 80%.

During the conversion, if the data length is not enough to be converted into a similar positive matrix form, it is completed through the zero filling operation to obtain the gray atlas, as shown in Figure 10, and the size of each picture is 29×29 . The image format of leakage data under different apertures after the compressed acquisition is quite different, while the grayscale images of the leakage signal under different apertures without compression acquisition in Figure 11 have little difference, and the grayscale images of different leakage apertures also have a certain similarity. From the perspective of visual processing, the compressed collected observations

are sent to the resnet18 model for training, which can achieve better accuracy.

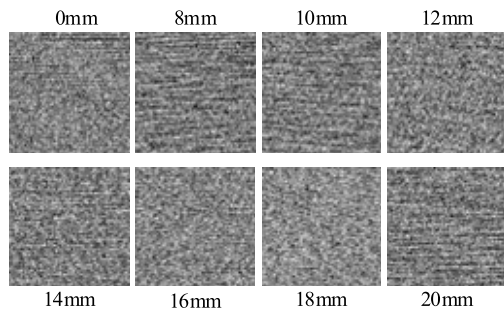


FIGURE 11. Visualization of different leakage data.

C. PERFORMANCE UNDER DIFFERENT COMPRESSION RATIOS

To discuss the influence of CR on the performance of the network model, the observation compression collection with CR of 50%, 60%, 70%, 80%, and 90% is carried out respectively when the observation matrix is a Gaussian random matrix, and the observation data sets with different CR are constructed. Then, the data sets are sent to ResNet18 model for training and recognition, in which the number of iterations during ResNet18 model training is 100 and batch_size is 64, the optimization algorithm is Adam algorithm, and the loss function is cross-entropy loss function, which are consistent. The observed value data set of each compression rate is used to do five experiments, and the identification accuracy of the five experiments is averaged.

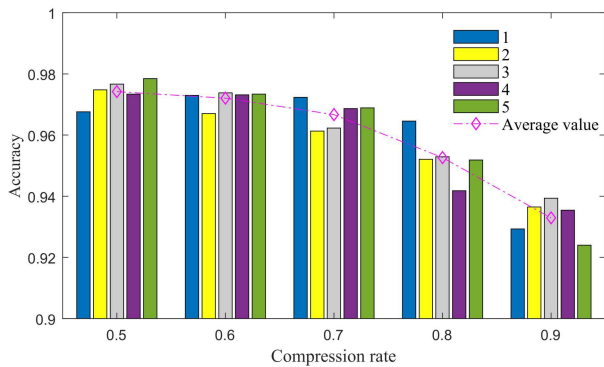


FIGURE 12. Accuracy of five times training recognition under different compression rates.

TABLE 2. Average time of five training under different compression rates.

CR (%)	Training time (s)
50	1436
60	1251
70	1102
80	803
90	651

The accuracy rate of five experimental training identification with CR of 50% - 90% is shown in Figure 12.

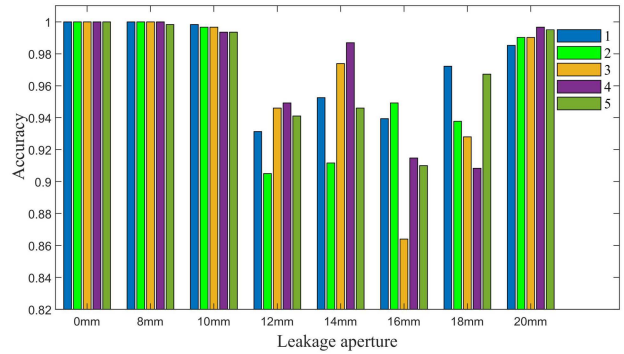


FIGURE 13. Accuracy of five times training recognition for each leakage aperture when CR is 70%.

The accuracy rate of water-supply pipeline leakage identification decreases slowly as the increase of CR. Because the higher the CR, the leakage characteristic information contained in the observed value of compressed collected leakage signal is relatively low, but it does not have a great impact on the accuracy of leakage identification, and the overall accuracy is still above 92%. As shown in Table 2, the training time of CS-ResNet18 model under different CR decreases as the increase of CR. when CR is 90%, the data length of the sample set is only one-tenth of the original data, and its training time is only 651s, which also has high accuracy. CS-ResNet18 has excellent performance under a high compression rate. If the leakage identification of the water-supply pipeline does not require high accuracy, the use of a high compression rate can reduce the sample points of data, thereby greatly reducing the training time of the model and quickly identifying leakage. At the same time, it can reduce the pressure of network communication and data storage, which is conducive to the load and real-time performance of the water-supply pipeline monitoring system network.

D. PERFORMANCE AT FIXED COMPRESSION RATE

When CR is 70% and the observation matrix is a Gaussian random observation matrix, the compressed acquisition data set is compared with the data set without compressed acquisition. The parameters of their ResNet18 model training are the same, which is consistent with the above. When the leakage signal is compressed and collected, the recognition accuracy rate of each type of water-supply pipeline leakage aperture trained for five times is shown in Figure 13, and the 8-dimensional confusion matrix output after a set of data training and recognition is shown in Figure 14. When the leakage signal is not compressed and observed, the recognition accuracy of each type of water-supply pipeline leakage aperture trained for five times is shown in Figure 15.

From Figure 13 and Figure 15, in the five training of each leakage aperture signal under the condition that the leakage signal is compressed and not compressed, when the leakage aperture is 0mm and 8mm, the recognition rate of the leakage aperture reaches 100%, and the leakage aperture is 12mm, 14mm, 16mm, 18mm, the recognition

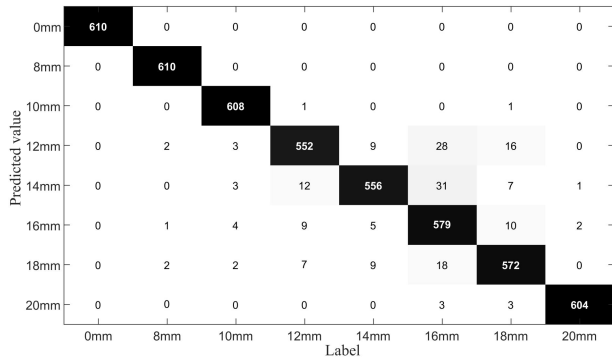


FIGURE 14. Confusion matrix under compressed acquisition.

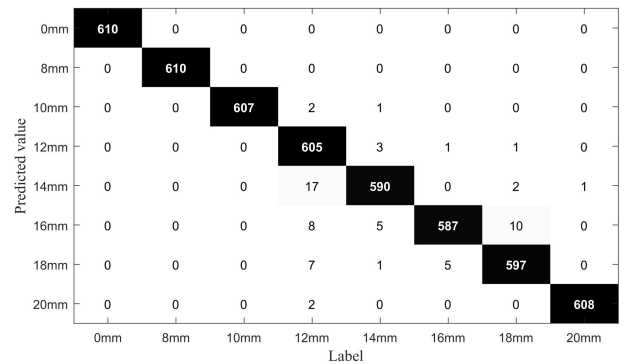


FIGURE 16. Confusion matrix under compressed acquisition.

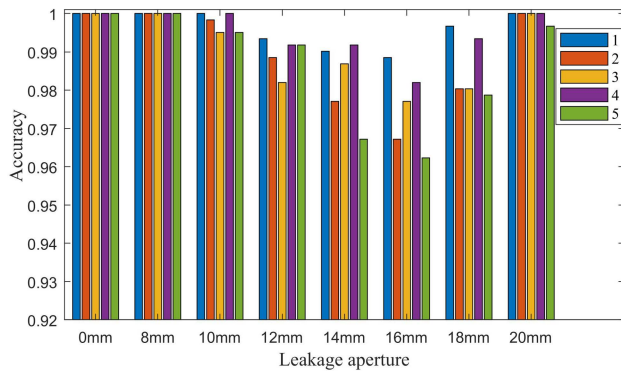


FIGURE 15. Accuracy of five times training recognition for each leakage aperture under uncompressed acquisition.

rate of the leakage aperture is relatively low. From the perspective of the confusion matrix, there is a small amount of aliasing in these four different apertures, which leads to misidentification as other leakage apertures. It can be seen from Figure 14 and Figure 16 that there is more aliasing after compressed acquisition. This is because the observation value of compressed acquisition lacks a small amount of information relative to the original signal, but does not cause important information to be lost, so the accuracy of recognition is still very high. As shown in Table 3, the training time of the model is greatly reduced, and the training time is only 50% of the uncompressed collection, which greatly shortens the training time of the model.

TABLE 3. Performance comparison between compressed acquisition and uncompressed acquisition.

Method	Accuracy (%)	Training time (s)
Compressed acquisition	96.67	1102
Uncompressed acquisition	99.13	2226

E. PERFORMANCE UNDER DIFFERENT OBSERVATION MATRICES

To analyze the influence of the observation values obtained under different observation matrices on performance of leakage identification model, the observation compression

collection of the observation matrices of Gaussian random matrix, Bernoulli random matrix, local Hadamard matrix and partial Fourier matrix is carried out respectively under the compression rate of 80%, and the data sets of different observation matrices are constructed. The data sets of different observation matrices are sent to the ResNet18 model for five times of training and recognition. The number of iterations, optimization algorithms, loss functions, and batch_size of the ResNet18 model during model training are consistent with the above.

When performing linear projection, the observation matrices with different structures have a great impact on the amount of useful information contained in the observations, which has a high impact on the accuracy of the convolution neural network model. Under Gaussian random observation matrix, Bernoulli random observation matrix, local Hadamard observation matrix and partial Fourier observation matrix, the vibration signals of the water-supply pipeline leakage are compressed and observed respectively.

A sample of acoustic vibration signal of the water-supply pipeline leakage is randomly selected from the experimental data of the water-supply pipeline leakage under different apertures. The time-frequency diagram of this group of signals is shown in Figure 17, and its amplitude range is about $-3-3$.

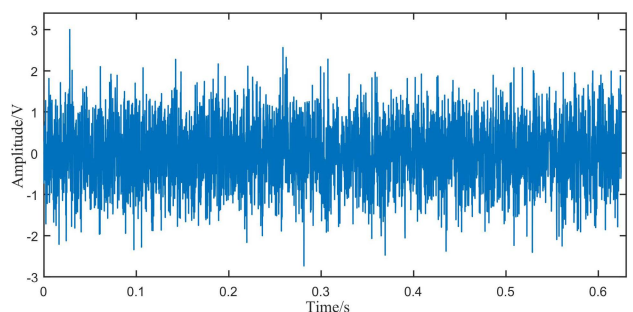


FIGURE 17. Time-frequency of acoustic vibration signal of water-supply pipeline leakage.

The sample is compressed with $CR = 50\%$ under Gaussian random observation matrix, Bernoulli random observation matrix, local Hadamard observation matrix and partial

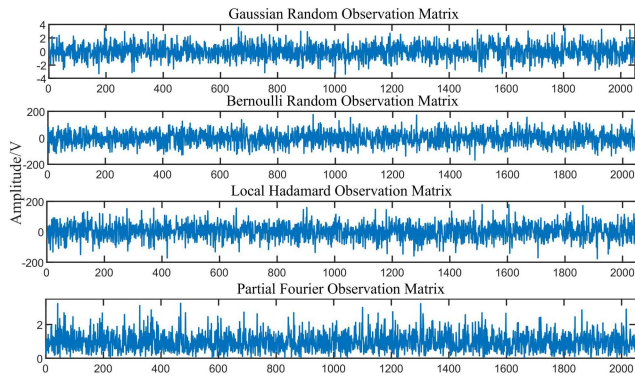


FIGURE 18. Observation values under different observation matrices.

Fourier observation matrix to obtain the observed value. The observed values under different observation matrices are shown in Figure 18. Among them, the Gaussian random matrix and the partial Fourier observation matrix have little change to the amplitude range of the signal. The amplitude range of the observed values under the Gaussian random matrix is about $-4-4$ and the amplitude range of the observed values under the partial Fourier observation matrix is about $0-3$. While the Bernoulli random observation matrix and the local Hadamard observation matrix greatly change the amplitude range of the signal, the amplitude range of the observed values under them is about $-200-200$. The difference in the amplitude range of the observed values of Bernoulli random observation matrix, local Hadamard observation matrix, Gaussian random matrix and partial Fourier observation matrix is caused by their different construction methods.

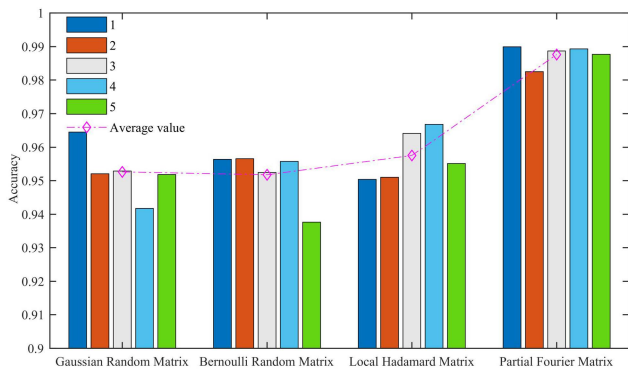


FIGURE 19. Accuracy of five times training recognition under different observation matrices.

Then, the data sets of different observation matrices are respectively sent to the residual neural network model. After five training and recognition experiments, the mean value is calculated, and the accuracy of the model under different observation matrices is shown in Figure 19. When the observation matrix is Gaussian random matrix, Bernoulli random matrix and local Hadamard matrix, the accuracy of the model is about 95%. While under partial Fourier matrix, the accuracy of the five experiments is higher than

that of other observation matrices, and the average accuracy is 98.77%.

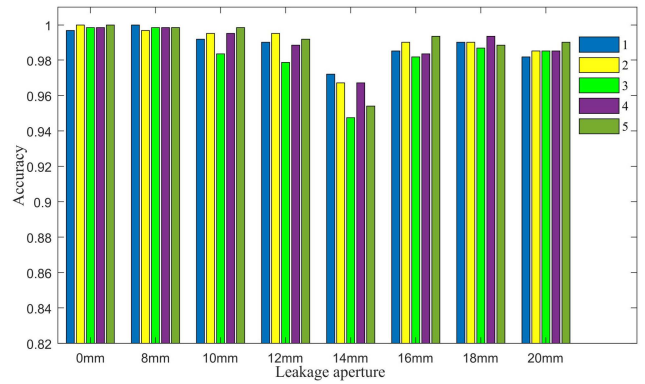


FIGURE 20. Accuracy of five times training and recognition for each leakage aperture under partial Fourier observation matrix.

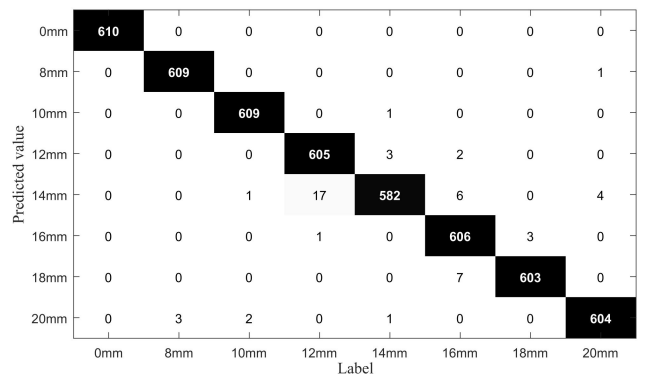


FIGURE 21. Confusion matrix under partial Fourier observation matrix.

When the observation matrix is a partial Fourier matrix, the accuracy of the five-time training and recognition of the leakage under different apertures of each water-supply pipe is shown in Figure 20, and the 8-dimensional confusion matrix output after a set of data training and recognition is shown in Figure 21. Compared with the previous experiment, the recognition rate is 100% when the leakage aperture is 0mm and 8mm, while the recognition rates of 12mm, 14mm, 16mm, and 18mm are relatively low but higher than the recognition rate under the Gaussian random matrix. From the confusion matrix in Figure 21, the aliasing of these four different apertures is less than that under Gaussian random matrix, which is close to that without compressed acquisition.

V. CONCLUSION

Aiming at the problems of too large data volume and too much information redundancy based on Nyquist data acquisition, which is not conducive to the load and real-time performance of the water-supply pipeline monitoring system network, this paper proposes a leakage identification method based on CS and the residual neural network. Under the condition of the observation matrix to ensure the integrity of the signal information, the compressed collection is

performed to reduce the redundant information and volume of the data. Firstly, it analyzes the compressed collection of the water-supply pipeline leakage signal, different observation matrices and residual neural network, and then explains the method proposed in this paper, and finally verifies the performance of the method through experiments, as well as the influence of CR and observation matrix on the performance of the model. Experiments show that:

- 1) The accuracy rate of the water-supply pipeline leakage identification decreases slowly with the increase of CR. But the overall accuracy rate is still above 92%, and the model training time is greatly reduced.
- 2) When the CR is 70% and the observation matrix is a Gaussian random matrix, the average accuracy rate reaches 96.67%, and the training time is only 50% of the uncompressed acquisition, which greatly reduces the training time of the model.
- 3) When the observation matrix is a partial Fourier matrix, the average accuracy of the leakage identification of the water-supply pipeline is 98.77%.

Therefore, the proposed method can achieve relatively high recognition accuracy using compressed observations, while greatly reducing the time of model training.

REFERENCES

- [1] A. F. Colombo, P. Lee, and B. W. Karney, "A selective literature review of transient-based leak detection methods," *J. Hydro-Environ. Res.*, vol. 2, no. 4, pp. 212–227, Apr. 2009.
- [2] B. Zhou, A. Liu, X. Wang, Y. She, and V. Lau, "Compressive sensing-based multiple-leak identification for smart water supply systems," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1228–1241, Apr. 2018.
- [3] B. Zhou, A. Liu, and V. K. N. Lau, "Sparse variational Bayesian inference for water pipeline systems with parameter uncertainties," *IEEE Access*, vol. 6, pp. 49664–49678, 2018.
- [4] H.-F. Duan, "Transient frequency response based leak detection in water supply pipeline systems with branched and looped junctions," *J. Hydroinform.*, vol. 19, no. 1, pp. 17–30, 2017.
- [5] A. Cataldo, G. Cannazza, E. De Benedetto, and N. Giaquinto, "A new method for detecting leaks in underground water pipelines," *IEEE Sensors J.*, vol. 12, no. 6, pp. 1660–1667, Jun. 2012.
- [6] J. Buerck, S. Roth, K. Kraemer, and H. Mathieu, "OTDR fiber-optical chemical sensor system for detection and location of hydrocarbon leakage," *J. Hazardous Mater.*, vol. 102, no. 1, pp. 13–28, Aug. 2003.
- [7] C. Verde, L. Molina, and L. Torres, "Parameterized transient model of a pipeline for multiple leaks location," *J. Loss Prevention Process Ind.*, vol. 29, pp. 177–185, May 2014.
- [8] C. Ge, G. Wang, and H. Ye, "Analysis of the smallest detectable leakage flow rate of negative pressure wave-based leak detection systems for liquid pipelines," *Comput. Chem. Eng.*, vol. 32, no. 8, pp. 1669–1680, Aug. 2008.
- [9] J. Hu, L. Zhang, and W. Liang, "Detection of small leakage from long transportation pipeline with complex noise," *J. Loss Prevention Process Ind.*, vol. 24, no. 4, pp. 449–457, Jul. 2011.
- [10] W. Liang and L. Zhang, "A wave change analysis (WCA) method for pipeline leak detection using Gaussian mixture model," *J. Loss Prevention Process Ind.*, vol. 25, no. 1, pp. 60–69, Jan. 2012.
- [11] T.-C. Chen, H.-F. Duan, and P. J. Lee, "Transient wave-based methods for anomaly detection in fluid pipes: A review," *Mech. Syst. Signal Process.*, vol. 160, Nov. 2021, Art. no. 107874.
- [12] H.-F. Duan, B. Pan, M. Wang, L. Chen, F. Zheng, and Y. Zhang, "State-of-the-art review on the transient flow modeling and utilization for urban water supply system (UWSS) management," *J. Water Supply, Res. Technol.-Aqua*, vol. 69, no. 8, pp. 858–893, Dec. 2020.
- [13] A. Keramat and H.-F. Duan, "Spectral based pipeline leak detection using a single spatial measurement," *Mech. Syst. Signal Process.*, vol. 161, Dec. 2021, Art. no. 107940.
- [14] J. Sun, R. Wang, and H.-F. Duan, "Multiple-fault detection in water pipelines using transient-based time-frequency analysis," *J. Hydroinform.*, vol. 18, no. 6, pp. 975–989, Nov. 2016.
- [15] H.-F. Duan, "Uncertainty analysis of transient flow modeling and transient-based leak detection in elastic water pipeline systems," *Water Resour. Manage.*, vol. 29, no. 14, pp. 5413–5427, Nov. 2015.
- [16] H.-F. Duan, "Accuracy and sensitivity evaluation of TFR method for leak detection in multiple-pipeline water supply systems," *Water Resour. Manage.*, vol. 32, no. 6, pp. 2147–2164, Apr. 2018.
- [17] X. Yu, W. Liang, L. Zhang, H. Jin, and J. Qiu, "Dual-tree complex wavelet transform and SVD based acoustic noise reduction and its application in leak detection for natural gas pipeline," *Mech. Syst. Signal Process.*, vols. 72–73, pp. 266–285, 2016.
- [18] L. Cui-Wei, L. Yu-Xing, F. Jun-Tao, and L. Guang-Xiao, "Experimental study on acoustic propagation-characteristics-based leak location method for natural gas pipelines," *Process Saf. Environ. Protection*, vol. 96, pp. 43–60, Jul. 2015.
- [19] Y. Gao, M. J. Brennan, P. F. Joseph, J. M. Muggleton, and O. Hunaidi, "On the selection of acoustic/vibration sensors for leak detection in plastic water pipes," *J. Sound Vib.*, vol. 283, nos. 3–5, pp. 927–941, May 2005.
- [20] C. Liu, Y. Li, L. Meng, W. Wang, and F. Zhang, "Study on leak-acoustics generation mechanism for natural gas pipelines," *J. Loss Prevention Process Ind.*, vol. 32, pp. 174–181, Nov. 2014.
- [21] P.-S. Murvay and I. Silea, "A survey on gas leak detection and localization techniques," *J. Loss Prevention Process Ind.*, vol. 25, no. 6, pp. 966–973, Nov. 2012.
- [22] W. Liang, L. Zhang, Q. Xu, and C. Yan, "Gas pipeline leakage detection based on acoustic technology," *Eng. Failure Anal.*, vol. 31, pp. 1–7, Jul. 2013.
- [23] L. Weiguang, W. Xiaodong, W. Haiyan, M. Changli, and W. Fenwei, "A dual-sensor-based method to recognize pipeline leakage and interference signals," *J. Loss Prevention Process Ind.*, vol. 38, pp. 79–86, Nov. 2015.
- [24] Q. Zhang, Z. Y. Wu, M. Zhao, J. Qi, Y. Huang, and H. Zhao, "Leakage zone identification in large-scale water distribution systems using multiclass support vector machines," *J. Water Resour. Planning Manage.*, vol. 142, no. 11, pp. 1–15, Nov. 2016.
- [25] C. C. D. Santos and A. J. P. Filho, "Water demand forecasting model for the metropolitan area of São Paulo, Brazil," *Water Resour. Manage.*, vol. 28, no. 13, pp. 4401–4414, Aug. 2014, doi: 10.1007/s11269-014-0743-7.
- [26] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015, doi: 10.1016/j.neunet.2014.09.003.
- [27] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, Apr. 2017, doi: 10.1016/j.neucom.2016.12.038.
- [28] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, Mar. 2016, doi: 10.1109/TBME.2015.2468589.
- [29] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016, doi: 10.1109/TIE.2016.2582729.
- [30] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks," *J. Sound Vib.*, vol. 388, pp. 154–170, Feb. 2017, doi: 10.1016/j.jsv.2016.10.043.
- [31] O. Janssens, V. Slavkovic, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. van de Walle, and S. Van Hoecke, "Convolutional neural network based fault detection for rotating machinery," *J. Sound Vib.*, vol. 337, pp. 331–345, Sep. 2016, doi: 10.1016/j.jsv.2016.05.027.
- [32] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [33] G. Oliveri, M. Salucci, N. Anselmi, and A. Massa, "Compressive sensing as applied to inverse problems for imaging: Theory, applications, current trends, and open challenges," *IEEE Antennas Propag. Mag.*, vol. 59, no. 5, pp. 34–46, Oct. 2017.
- [34] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [35] J. Chen, Y. Gao, C. Ma, and Y. Kuo, "Compressive sensing image reconstruction based on multiple regulation constraints," *Circuits, Syst., Signal Process.*, vol. 36, no. 4, pp. 1621–1638, Apr. 2017.

- [36] J. Wu, F. Liu, L. C. Jiao, and X. Wang, "Compressive sensing SAR image reconstruction based on Bayesian framework and evolutionary computation," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1904–1911, Jul. 2011.
- [37] J. L. Paredes, G. R. Arce, and Z. Wang, "Ultra-wideband compressed sensing: Channel estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 3, pp. 383–395, Oct. 2007.



SHUAIYONG LI was born in Luohe, Henan, China. He received the Ph.D. degree from Chongqing University, China, in 2014. Since 2015, he has been with the Chongqing University of Posts and Telecommunications, where he is currently a Professor. His current research interests include fault diagnosis of dynamic systems, information acquisition and processing, industrial Internet of Things, and leak detection and location in fluid-filled pipelines.



LIN MEI was born in Mianyang, Sichuan, China. He received the B.Eng. and Ph.D. degrees from Chongqing University, China, in 2010 and 2015, respectively. Since 2015, he has been with the Chongqing Special Equipment Inspection and Research Institute, where he is currently a Researcher. His current research interests include corrosion protection system of buried pipeline, integrity management of long distance pipeline, and risk assessment on gas pipelines.



MENGQIAN CAI was born in Zhumadian, Henan, China. She received the B.Eng. degree from Xinyang Normal University, in 2020. She is currently pursuing the M.Sc. degree with the Department of Automation, Chongqing University of Posts and Telecommunications. Her main research interests include information acquisition and processing, industrial Internet of Things, and leak detection and location in fluid-filled pipelines.



JUN ZHOU was born in Bishan, Chongqing, China. He received the B.Eng. degree from Sichuan Institute of Light and Chemical Engineering, China, in 1985. Since 2006, he has been with the Chongqing Special Equipment Inspection and Research Institute, where he is currently the Director of the Pressure Pipeline Center. His current research interests include corrosion protection system of buried pipeline, integrity management of long distance pipeline, and risk assessment on gas pipelines.



TONG LI was born in Xi 'an, Shaanxi, China. He received the master's degree from Xi 'an Technological University, in 2012. Since 2013, he has been working at the Chongqing Special Equipment Inspection Research Institute and currently serves as a Responsible Engineer. His research interests include corrosion protection system for buried pipelines, integrity management of long-distance pipelines, and risk assessment of gas pipelines.

• • •