A Novel DAS Signal Recognition Method Based on Spatiotemporal Information Extraction With 1DCNNs-BiLSTM Network

HUIJUAN WU [1], **(Member, IEEE), MINGRU YANG**[1]**, SIQI YANG**[1]**, HAO LU**[1]**, CHAOQUN WANG**[1]**, AND YUNJIANG RAO** [1,2]**, (Fellow, IEEE)**

[1]Key Laboratory of Optical Fiber Sensing and Communications, Ministry of Education, University of Electronic Science and Technology of China, Chengdu 611731, China

[2]Research Center for Optical Fiber Sensing, Zhejiang Laboratory, Hangzhou 310000, China

Corresponding author: Huijuan Wu (hjwu@uestc.edu.cn)

ABSTRACT Extracting more and more accurate information to understand the detected vibration or acoustic targets better, has always been an important goal in signal recognition for Distributed Acoustic Sensor(DAS) with optical fiber. In this paper, we use one-dimensional Convolution Neural Networks(1D-CNNs) to extract the detailed temporal structure information at each signal node and utilize a bidirectional Long Short Term Memory(BiLSTM) network to dig out the spatial relationship among the different signal nodes, and then propose a novel identification method by treating the spatial- and temporal- information in a different way, which is denoted as the 1DCNNs-BiLSTM model. The experimental results on the field data show better recognition performance can be achieved in the safety monitoring of the buried optical communication cable in urban with DAS. It helps to improve the recognition rate further compared with the other deep-learning methods frequently or possibly used for DAS signal recognition, such as the 1D-CNNs with a single temporal feature extraction, and 1DCNN-CNN and 2D-CNN models with simultaneous spatiotemporal feature learning. To the best of our knowledge, it is the first time to simultaneously extract and utilize the detailed temporal structure feature and the overall spatial connection through a customized deep learning network.

INDEX TERMS Vibration identification, spatiotemporal information extraction, DAS, CNN, LSTM.

I. INTRODUCTION

Distributed Acoustic Sensor (DAS) based on the phase-sensitive optical time-domain reflectometry (Φ-OTDR) technology [1]–[4], provides a highly sensitive and cost-effective vibration or sound sensing way for our environment in long distance or wide range. It has been extensively applied in many safety monitoring fields, such as natural disaster prediction including seismic waves [5], the third party damage warning for oil/gas/water/heating pipelines [6], [7], and perimeter security [1], [8] and large structure health monitoring, etc. Especially, it is a promising technology in maintenance of urban communication cables [9]. In recent

years, the laid volume of the communication optical cables is dramatically growing with the increasing construction demand of 5G, Internet of things, industrial Internet, and data center. By utilizing the existing fiber cables, the vibration or sound signals generated by various destructive sources on the ground can be sensed and located by DAS with high sensitivity and precision. However, intelligent recognition and deep understanding of real threats such as mechanical operations, manual diggings and other vibration sources as the traffic interferences along the fiber is still a challenging problem, which results in frequent nuisance alarms in practical applications. In fact, to improve its smart sensing ability, a lot of useful signal processing work has been involved, which includes conventional machine learning methods with fixed hand-crated feature

H. Wu *et al.*: Novel DAS Signal Recognition Method Based on Spatiotemporal Information Extraction With 1DCNNs-BiLSTM Network

IEEE *Access*

extraction [6], [7], [10], [11], and deep learning methods with one dimensional(1-D) [12] or two dimensional (2-D) Convolution Neural Networks (CNNs) [13]–[15] to extract the hidden distinguishable feature automatically.

However, most of the identification methods for DAS still mainly rely on the temporal feature extraction of the vibration signals at each sampling node at different fiber locations, while they ignore the spatial relationship among them. Some methods try to use traditional image processing methods [11] to detect and identify some certain 2-D patterns in the sampled space-time matrix. But fixed hand-crated feature extraction is time-consuming and laborious, and relying heavily on the expert knowledge. Thus in this paper, to improve the recognition ability in the safety monitoring of the buried optical communication cable, it is explored a novel identification method by using a deep learning structure with an array of 1-D Convolution Neural Networks(CNNs) combined with the Bi-directional Long Short-Term Memory(BiLSTM) model, which is shortened as 1DCNNs-BiLSTM. In this algorithm, CNN is used to automatically extract the temporal structure feature of the signals at each acquisition node on fiber, and the BiLSTM network is composed of forward and backward LSTMs, which are designed to mine the internal spatial relationship among the temporal signals at different nodes from right to left and from left to right in succeed.

The rest of the paper is organized as follows: Section II presents related work about the temporal and spatiotemporal information extraction methods and challenges in DAS signal processing at present; Section III, by taking the application of safety monitoring for a long distance underground communication cable as an example, it introduces the proposed 1DCNNs-BiLSTM algorithm, in which it can extract not only the structural feature of the temporal signals at each fiber location, but also the spatial correlation among the signals at different locations and give out recognition results of the space-time matrix at the end of the network; and the experimental results with real field test data are presented in Section IV; finally, the conclusion is given in Section V.

II. RELATED WORK

At present, intelligent recognition and deep understanding of real threats and other nonthreatening sources caused by various production and life noises along the buried fiber is a challenging problem, which still results in inaccurate recognition and frequent nuisance alarms in long distance monitoring applications. In order to propel the practical applications of DAS, more and more research is devoted to its signal processing methods, such as the vibration or acoustic source identification.

At the first stage (before 2015), most work is devoted to get better SNR for reliable detection and location of the events, such as by correlation detection and moving average [16], 2-D image edge operator [17], Wavelet transform [18] and Hilbert-Huang transform [19], and other signal denoising and separation techniques [20]. These methods are helpful to depress the noises caused by the frequency

shift and other unstable system factors, and the stationary background to some extent. But in real applications, it is found that the nuisance alarm rate (NAR) mainly comes from poor understanding of the vibration or sound especially in complicated time-varying noisy environments. Thus at the second stage (from 2015-2017), more researchers focus on exploring various feature extraction and proper identification methods. The feature includes the magnitude [7], level crossing rate [21], and periodic gait characteristics [22] in time domain, and the frequency energy distribution of its Fast Fourier Transform(FFT) spectra [23], and the morphological features in space-time domain [11], and the spectra obtained by short-time Fourier Transform(STFT) [24], Wavelet or Wavelet packet energy spectra [25], [26], and Mel-Frequency Ceptral Coefficient (MFCC) [27] in time-frequency domain and etc. And the classification models include artificial neural networks(ANN) [6], [25], [26], Gaussian mixture model (GMM) [24], [27], Support Vector Machine(SVM) [7] and Relevance Vector Machine (RVM) [11] and etc. All the feature engineering methods enhance the perception ability further, but they rely heavily on the expert knowledge. And developing an applicable recognition algorithm is time-consuming and laborious, and the model has poor transfer ability in different application environments. Then along with rapid development of the deep learning algorithms and their successful applications in image processing [28], [29], speech recognition [30] and fault diagnosis [31] and etc., the vibration source identification of DAS enters a new phase at the third stage. In [15], [13], and [14], the hidden feature of different types of DAS signals is automatically extracted and identified by using CNN network. To extract more useful information, Tejedor J and et al proposed to utilize the temporal contextual connection by integrating the sequential feature vectors in a multi-layer perceptron (MLP) network in [32] and [10]. In further, we proposed a knowledge mining method based on the hidden Markov models (HMMs) [33] to extract the dynamic time sequence feature and its evolution information, and then identify the sequential state process of typical events.

However, generally most of the present identification methods for DAS still rely on the temporal feature extraction of the signals obtained at each sampling node on fiber, while they ignore the spatial information which is the relationship among the signals at different nodes. In [11], it considers the spatiotemporal structure of DAS signals, but it utilizes fixed hand-crated feature extraction for the 2-D space-time matrix, which is laborious. At present, 1-D CNN and 2-D CNN are still frequently used networks while they are proposed for the one-dimensional DAS signal recognition at every fiber node. And until recently, the spatiotemporal feature of DAS signals has not been fully considered or automatically excavated. On the other hand, a new deep hybrid learning model of CNN-LSTM is proposed and sequence connection and spatiotemporal information are excavated by using this network for video action recognition in [34], and transportation flow prediction in [35]. In [35], CNN is used to extract the spatial

information while the bidirectional LSTM (BiLSTM) is used to extract the temporal information, and in [34] the BiLSTM is configured in a two-layer structure to extract the hidden connection among the image sequences. In this paper, to fully utilize both the structural information in the time sequences and the spatial distribution mode among the sequences influenced by the vibration source, it borrows the idea in [34] and [35] but designs a customized deep learning structure of 1DCNNs-BiLSTM for DAS according to its spatiotemporal signal structure. A separate 1-D CNN network is used to extract the detailed temporal information of the signal at each spatial sampling node; while a one-layer BiLSTM is used to automatically dig out the spatial connection among the node signals in order to give full play to their advantages respectively. Purposes of these two sub-networks are different and customized in our application.

III. METHODOLOGY

A. SYSTEM STRUCTURE OF DAS AND THE PROPOSED RECOGNITION SCHEME

Taking the application of safety monitoring for long-distance underground communication cable as an example, the system structure of DAS and the distributed recognition scheme are demonstrated as in Fig. 1. The hardware consists of three parts: the probe fiber, the optical signal demodulator of DAS, and the signal processing unit. The Φ-OTDR linearly demodulated by a 3×3 coupler [2] is used in the demodulator. The probe fiber takes a spare core of the communication cable laid along the pipelines under the urban ground and is used to sense/detect the ambient disturbing events which causes vibration. Each section of the fiber is equivalent to a sensor sampling node in space. These distributed nodes cooperate to pick up the vibration signals on the whole line. The system originally returns a space-time signal matrix, which is defined as

$$\{XX = x_{ts}(t = 1, 2, \ldots, T; s = 1, 2, \ldots S)\} \quad (1)$$

In the matrix in (1), the row index t represents the time and T is the time length; and the column index s denotes the spatial sampling node and S is the spatial width. And the spatial interval of each two nodes is ΔS, and the temporal interval is $\Delta T = 1/f_s$, in which f_s is the sampling frequency. One column data represents the temporal signal collected at a sampling node, which is always the basis of event recognition in most of the related work.

Actually, any a machinery excavation or manual operation influences not just a single column, but some columns in a range. Most of the present work identify the temporal signal at each node respectively and ignores the internal spatial correlation of the signals among the nodes. Thus in this paper the acquisition matrix XX is segmented into small frames of event centric signal matrix, which are taken as a space-time recognition area. And the signal matrix in the recognition area is input into the following recognition network as shown in Fig. 1.
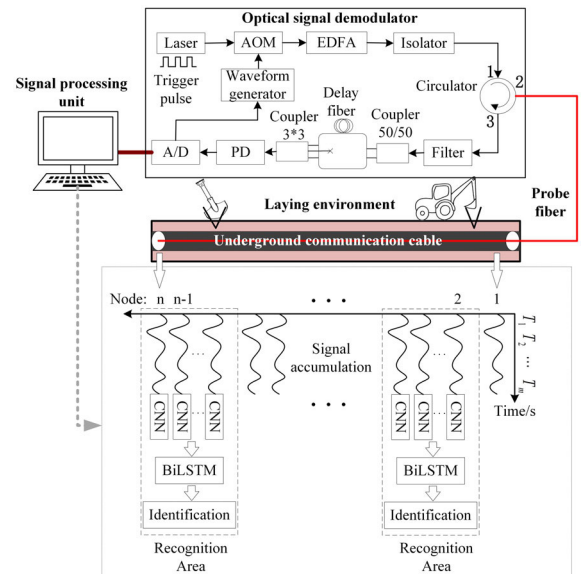

FIGURE 1. System structure of DAS and the proposed recognition scheme in the safety monitoring of long-distance underground communication cable.

In Fig. 1, a customized hybrid model of 1DCNNs-BiLSTM is designed according to the specific space-time structure of the DAS sensing signals, in which there are three modules: firstly, there is only one signal per node, and the signal at one node is preprocessed and input into a separate 1D-CNN network, which corresponds to one cell of the LSTM to extract the local and global structural feature; then the extracted 1D-CNN feature vectors are fed in parallel into a bidirectional LSTM network to continue mining the spatial association among the signals at different nodes in the event frame; finally, the extracted spatio-temporal feature sets are stacked and input to the full connection layer of the whole network to identify the event type. In this way, the space-time information, such as the characteristics of spatial distribution as well as the temporal structure feature of the vibration signals can be both fully utilized, while they are treated differently according to the different contribution of them.

B. TEMPORAL FEATURE EXTRACTION WITH 1D-CNNs

Various experimental results show that CNN can effectively extract structural features for the 1-D complicated speech or sensory signals [12]. In the application of DAS, considering that the information in spatial dimension is not as rich as that in temporal dimension, an array of 1D-CNNs is then adopted to extract the structural features of each signal at all the spatial nodes in the recognition area, while we do not use a direct two-dimensional CNN (2D-CNN). Besides, the 1D-CNN can extract the temporal feature brilliantly with fewer network parameters [12], which can improve the speed of model detection and prevent over fitting. Thus several identical 1D-CNNs are combined in parallel to form a 1D-CNN array, which is denoted as 1D-CNNs, in which each 1D-CNN is responsible for the temporal feature extraction for the signal at one spatial node.
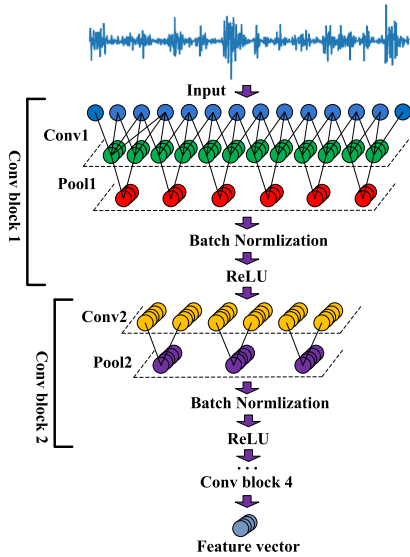
H. Wu et al.: Novel DAS Signal Recognition Method Based on Spatiotemporal Information Extraction With 1DCNNs-BiLSTM Network

IEEE Access

FIGURE 2. The detailed structure of the 1D-CNN network.

The detailed structure of each 1D-CNN in the model is demonstrated as in Fig. 2. It consists of four convolution blocks, and each convolution block is composed of a convolution layer, a pooling layer, and a ReLU activation function, which form a Convolution-Pooling-Relu structure. In the network, in order to alleviate the internal covariate shift phenomenon and increase the feature extracting ability, a Batch Normalization (BN) layer [36] is added following the output of each convolution block. Through the network, each column is processed and transformed into a deep feature vector, which is called the 1D-CNN feature in this paper. When all the columns of data are processed in the frame, the extracted 1D-CNN feature vectors are stitched into a feature sequence in the actual spatial order.

C. SPATIAL CONNECTION MINING WITH BiLSTM

In the second step, a bidirectional Long Short Term Memory (BiLSTM) network is designed further to extract the spatial connection among the deep structural feature vectors learned by CNN. LSTM is a deform of Recurrent Neural Network(RNN). It mainly solves the long memory problem of traditional RNN and remembers only useful information. Here we regard the spatial association of the signal nodes as a kind of sequence relationship, and use a bidirectional LSTM to extract the contextual connection of the spatial distribution rule. The typical structure of BiLSTM is shown as in Fig. 3. A one layer bidirectional LSTM is used, including a backward and a forward layer, to interpret the spatial relationship from left to right and from right to left respectively. In the BiLSTM network, each LSTM cell is designed for one spatial node. The structural details of each LSTM cell is also illustrated on the upper right corner in Fig. 3. And its calculation principle in each LSTM cell is shown in (2) to (7), and the involved symbols are described in Table 1.
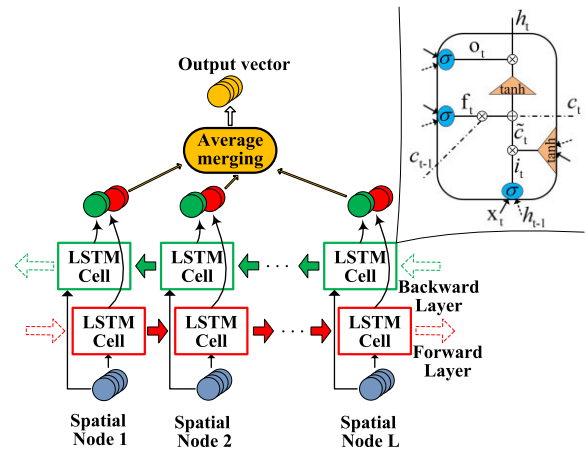
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

FIGURE 3. The internal structure of the BiLSTM network.

TABLE 1. Symbol description in the BiLSTM network.

Symbol	Description
i_t	Input Gate
f_t	Forget Gate
o_t	Output Gate
x_t	The input signal at the present cell
c_t	The memory state of the present cell
h_t	The output state of the present cell
$\tilde{c}_t$	The updated memory state of the present cell
σ	Mapping function of Sigmoid
tanh	Mapping function of Hyperbolic Tangent
W_i, W_f, W_o, W_c	The weight matrices of the input gate, forget gate, output gate, updated state.
b_i, b_f, b_o, b_c	The bias values of the input gate, forget gate, output gate, updated state.

$$\tilde{c}_t = \tanh(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{3}$$
$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{4}$$
$$c_t = f_t^* c_{t-1} + i_t^* \tilde{c}_t \tag{5}$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$
$$h_t = o_t^* \tanh(c_t) \tag{7}$$

In the BiLSTM network, with the deep CNN feature at each spatial node as its input, the learning process of each LSTM cell is controlled by three gates: the input gate i_t, the forget gate f_t, and the output gate o_t. In (2) and (3), the input x_t and the output state of the previous cell h_{t-1} are transformed into the input information of the input gate and the memory state of the present cell respectively through the mapping functions of Sigmoid σ and Hyperbolic Tangent tanh. In (4),

IEEE *Access*

H. Wu *et al.*: Novel DAS Signal Recognition Method Based on Spatiotemporal Information Extraction With 1DCNNs-BiLSTM Network
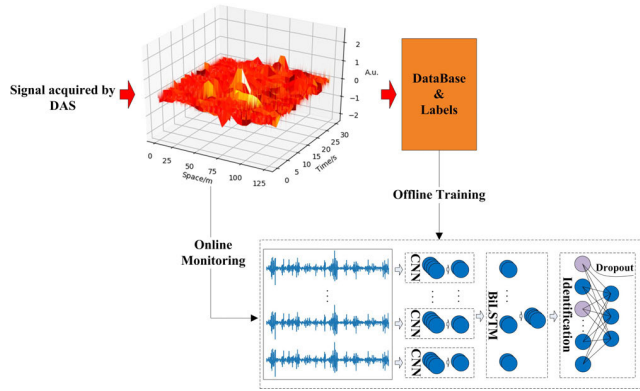
FIGURE 4. Flowchart of the proposed 1DCNNs-BiLSTM algorithm.

a forget gate f_t is formed through the sigmoid function with the input x_t and the output state of the previous cell h_{t-1}, which determines how much information of the previous cell is remained or forgotten in the present cell. In (5), the state of the present cell c_t is updated by adding the remainder information of the previous cell in the forget gate, and the input information in the input gate. In (6), the output gate at the present cell is obtained by fusing the output state of the previous point and the input with a sigmoid function. And in (7), the output state of the present cell is finally obtained by passing the updated memory state of the present cell through the output gate. And the output state of the cell is taken as the extracted feature in the BiLSTM network. In the forward direction, the 2nd cell contains the information of the 1st cell; and the 3rd cell contains the information of the 1st and the 2nd cell; and the last cell contains all the information of the previous cells. And in the backward LSTM, it is applied in the same way in an opposite direction. The two output feature in the bidirectional LSTM are contaminated together to combine a new feature vector. It contains the bidirectional information of all the cells in the network. Then an average mergence is carried out for the output of all the cells, and the merged feature vector finally contains all the spatiotemporal information and prepares for the following recognition.

D. IDENTIFICATION

In the last step, we use the fully connected (FC) layer followed by the 1DCNNs-BiLSTM network to identify the event type. And a dropout layer is adopted to follow the FC layer to avoid over-fitting phenomenon. At each iteration cycle of the training process, the dropout layer randomly discards a certain proportion of the cell units in the FC layer, which makes it to be equivalent to train a new network each time. Thus the model robustness can be improved and over-fitting can be prevented.

Finally, flowchart of the proposed 1DCNNs-BiLSTM algorithm is constructed as in Fig. 4, and its parameters in structure are detailed in Table 3 in Section IV. The event centered space-time samples acquired by DAS are prepared to construct a database of typical events. The 1DCNNs-BiLSTM network is trained with the database offline first;

(a)Background with no threats

(b) Traffic flow traveling

(c) Machine excavator operation

(d) Road breaker operation

(e) Manual digging

FIGURE 5. Field scenarios of five typical events.

when it achieves at its optimal state, the well-trained network is used for online monitoring and identification. Cross entropy is used as the loss function to train the whole 1DCNNs-BiLSTM network, which is calculated as in (8),

$$L = -\frac{1}{N}[y \ln a + (1 - y) \ln(1 - a)] \qquad (8)$$

In (8), N is the batch size or the number of samples in the data batch used for training, and y is the true label of the sample, and a is the predicted label of the sample. The obtained cross entropy L represents the difference of the true class and the predicted class. In the training process, the loss function value will decrease iteratively until the difference vanishes, or when the model converges.

IV. EXPERIMENTS

A. DATA PREPARATION AND MODEL CONSTRUCTION

We applied the proposed algorithm into the field test for safety monitoring of underground communication cable in Wuhan, Hubei province in April 2019. The monitoring cable is about 40km long, and 0.8-1.5 meters deep buried in underground in urban. One spare core of the optical fiber communication cable is taken as the probe fiber. The temporal sampling rate of the system is 500Hz, and its spatial sampling rate is 5.16m. In the test, we use the DAS system as shown in Fig. 1 to collect five types of typical events, including background with no threats, traveling traffic flow, machine excavator operation, road breaker operation, and manual digging, which are labeled as type 1, 2, 3, 4, and 5 respectively. The spatiotemporal signal sample of each event is composed of 25 adjacent spatial points (~125m) in space, and the collection time is 30 seconds in time. Fig. 5 and Fig. 6 show the field scenarios and the corresponding samples of the five commonly interested events. As shown in Fig. 6, besides the
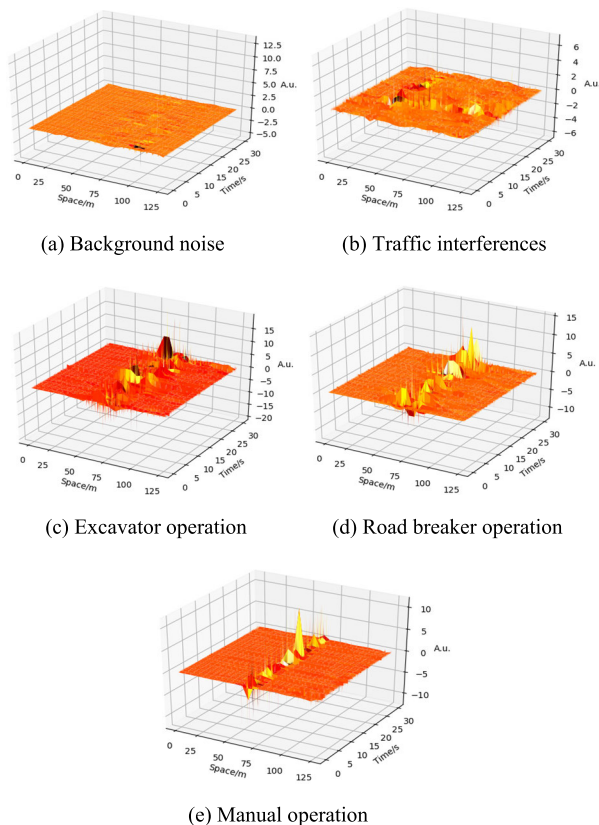
H. Wu *et al.*: Novel DAS Signal Recognition Method Based on Spatiotemporal Information Extraction With 1DCNNs-BiLSTM Network

IEEE Access

(a) Background noise

(b) Traffic interferences

(c) Excavator operation

(d) Road breaker operation

(e) Manual operation

FIGURE 6. Space-time samples of the five typical events in Fig. 5.

TABLE 2. Database preparation with real field data.

Event type	Training/Testing set	Label
Background noise	246/62	1
Traffic interferences	321/80	2
Excavator operation	209/52	3
Road breaker operation	213/53	4
Manual digging	206/52	5

difference in the fluctuation the temporal signal, the spatial distribution modes of different events are somewhat different. As we can see that the background with no threats has a wide mild noise distribution in the space-time matrix; the spatial distribution of the traffic interferences in main traffic arteries is disorderly, and the law of time evolution in long term is consistent with the regular urban population life; the signal samples of excavator digging and road breaker construction are similar from the perspective of space-time matrix, in both of which the signal fluctuation amplitude is relatively strong and the influenced spatial range is more concentrated than traffic interferences, but wider than most manual operation as in Fig. 6(e); and the influence range of the manual digging is the narrowest one, and its signal strength is weaker on the whole but occasionally strong locally. The field database prepared for training and testing is detailed as in Table 2.

TABLE 3. Structural parameters of the proposed CNN-BiLSTM network.

Layers	Kernel size/Stride/Padding	Input size
Conv1	1*25*32/1*1/0*12	25*15000*1
Pool1	1*25*32/1*25	25*15000*32
Conv2	1*25*64/1*1/0*12	25*600*32
Pool2	1*25*64/1*25	25*600*64
Conv3	1*5*128/1*1/0*2	25*23*64
Pool3	1*12*128/1*12	25*23*128
Conv4	1*5*256/1*1/0*2	25*2*128
Pool4	1*2*256/1*2	25*2*256
BiLSTM	hidden:256	25*256
FC	neurons:512	1*512
softmax	classifier	1*5

Structural parameters of the proposed CNN-BiLSTM model in the test are specified in Table 3. Four convolution layers are used in the CNN network, and a forward and a backward 25*256 LSTM are simultaneously used in the BiLSTM network. In the first two convolution layers, a 1*25 filter kernel is used for extracting the local detailed structure of more or less one second of the signal; while in the last two convolution layers, a 1*5 filter kernel is used for larger scale structure extraction in global, such as the signal trends. The filter size could be different for different signals with varying sampling rates, which has to be optimized through parameter adjustment. In each layer, the Kernel size/Stride/Padding and the Input size of the layer are also detailed respectively in Table 3. In the forward and backward 25*256 LSTMs, 25 LSTM cells stand for the spatial nodes in the sample, and 256 hidden states accepts the 1D-CNN feature of 256 channels. Finally, a 1*512 FC layer is used to stitch the output of the two LSTMs, and a softmax classifier of 1*5 is used to classify the five event targets. And the configuration of CNN and BiLSTM in Table 3 has been optimized. For example, one-layer BiLSTM is tested to have better performance than the two-layer LSTM structure. And convolution layer, convolution kernel size, pooling size and the activation function Relu in CNN network are all optimized through experiments. Table 3 actually contains the optimized configuration and parameters of the CNN-BiLSTM network for the DAS application in this paper.

B. FEATURE VISUALIZATION

As stated in the related work, 1-D CNN and 2-D CNN are the frequently used deep learning networks at present for the one-dimensional DAS signal recognition at each fiber location in [12]–[15], and 1-D CNN proves better than the others including the traditional machine learning methods with fixed hand crafted feature. In this paper, we mainly compare the proposed method with the networks which can consider both the spatio- and temporal- information but not the only temporal information. Then, we take the 1-D CNN as the 1st method to be compared, and it represents all the
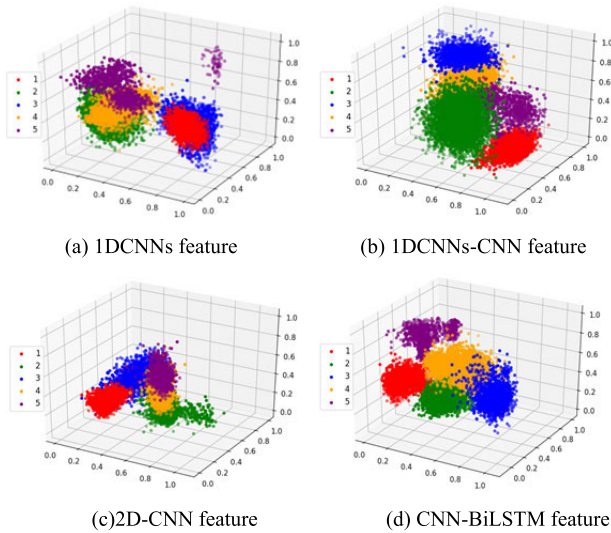
IEEE *Access*

H. Wu *et al.*: Novel DAS Signal Recognition Method Based on Spatiotemporal Information Extraction With 1DCNNs-BiLSTM Network

(a) 1DCNNs feature   (b) 1DCNNs-CNN feature

(c)2D-CNN feature   (d) CNN-BiLSTM feature

FIGURE 7. Feature visualization for the four deep learning methods.

methods in which only temporal feature of the signal at each sampling node is considered here. And the other three are all methods of spatiotemporal information extraction and identification networks possibly used in DAS. The 1st method is denoted as 1DCNNs, in which a four layer 1D-CNN is specified as in Table 3, all the CNN feature vectors obtained in the spatial order are stitched together and input into a FC layer for classification directly; the 2nd one is denoted as 1DCNNs-CNN, in which the feature extraction process is the same with the 1st one, but in the classification step, another one layer 1D-CNN with kernels of 1*3 in 256 channels is used to mine the sequence relationship of the spatially stitched 1D-CNN features; the 3rd one is the 2D-CNN, in which the spatio-temporal sample is taken as a 2-D image, and a 2D-CNN is directly used to extract the contour pattern on the horizontal and vertical axes; and the last one is the proposed 1DCNNs-BiLSTM method in the paper, in which the four layer 1D-CNN is the same as that in the first two methods, and the obtained 1D-CNN feature at each node is taken as the input to each LSTM cell. And the spatial distribution characteristics among the signals is extracted in the BiLSTM network.

To observe the event distinguishability of the above four methods, the feature vectors learned by each of them from the database in Table 2 are visualized as in Fig. 7. The high dimensional feature obtained by the well-trained networks is reduced into a three-dimensional space through the Linear Discriminant Analysis(LDA) algorithm. From the four figures in Fig. 7, it shows the visualized feature of the five events can be basically distinguished for all of the above four methods, which explains all the four learning methods achieve at their optimal states in this test. And in more details, the spatio-temporal features extracted by 1DCNNs-CNN in Fig. 7(b) and the proposed CNN-BiLSTM in Fig. 7(d) has larger classification distances than the other two methods. At this stage,
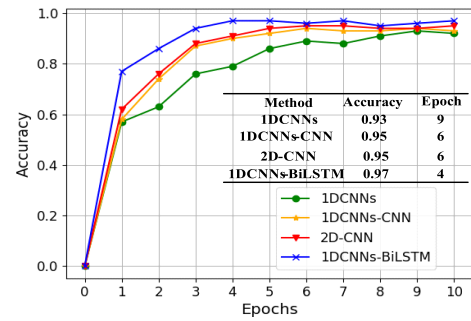

FIGURE 8. A comparison of the convergence process of the four methods.

(a) 1DCNNs   (b) 1DCNNs-CNN
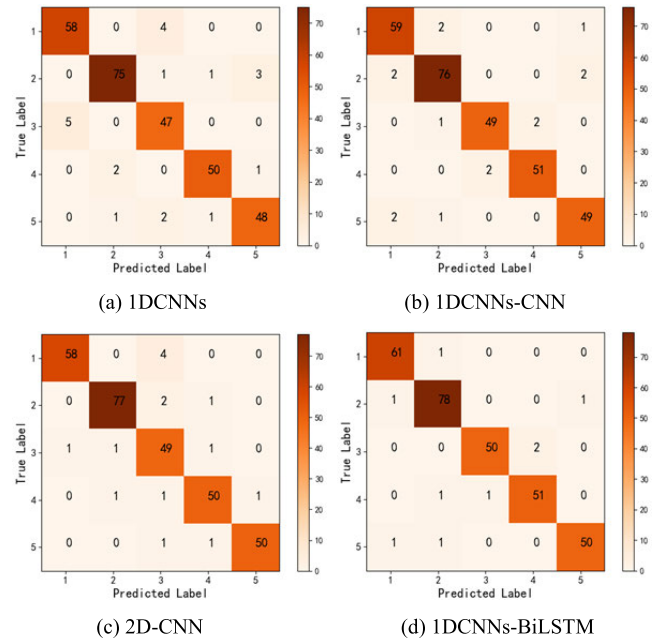
(c) 2D-CNN   (d) 1DCNNs-BiLSTM

FIGURE 9. Confusion matrices obtained by the four methods.

it is convenient to compare the classification performance further.

C. RECOGNITION PERFORMANCE
Then more detailed classification performance of the four methods are compared in this section. Firstly, in the training process, the converging processes of the four algorithms are compared as shown in Fig. 8. It shows the 1DCNNs has the slowest convergence speed of 9 epochs and the lowest accuracy of ∼93%, while the proposed 1DCNNs-BiLSTM has the fastest convergence speed of 4 epochs and the highest accuracy of ∼97% in this test; and the convergence speed and recognition accuracy of the 1DCNNs-CNN and the 2D-CNN are both in the middle.

In the testing process, the confusion matrices of the four methods are all obtained as in Fig. 9 by using the testing set in Table 2. And from Fig. 9, the performance indices of Precision, Recall, F1-score and Accuracy of each event for all of the four methods are respectively computed and compared as in Table 4. In these four methods, the proposed 1DCNNs-BiLSTM has the best recognition performance indices for

H. Wu *et al.*: Novel DAS Signal Recognition Method Based on Spatiotemporal Information Extraction With 1DCNNs-BiLSTM Network

IEEE *Access*

TABLE 4. A comparison of recognition performance for the four methods.

Model	Label	Precision	Recall	F1-score
1DCNNs Accuracy=0.929	1	0.921	0.935	0.923
	2	0.962	0.938	0.950
	3	0.870	0.904	0.887
	4	0.962	0.943	0.952
	5	0.923	0.923	0.923
1DCNNs-CNN Accuracy =0.949	1	0.937	0.952	0.944
	2	0.950	0.950	0.950
	3	0.961	0.942	0.951
	4	0.962	0.962	0.962
	5	0.942	0.942	0.942
2D-CNN Accuracy =0.949	1	0.983	0.935	0.958
	2	0.974	0.963	0.968
	3	0.859	0.942	0.898
	4	0.943	0.943	0.943
	5	0.980	0.962	0.971
1DCNNs-BiLSTM Accuracy =0.970	1	0.968	0.984	0.980
	2	0.963	0.975	0.969
	3	0.980	0.962	0.971
	4	0.962	0.962	0.962
	5	0.980	0.962	0.971

FIGURE 10. The ten-fold cross validation of the four methods.

(a) test time for one space-time sample   (b) test time for the whole line

FIGURE 11. Test time comparison of the four methods Notes: the test is on a typical processor (GPU:GTX1080Ti, CPU: Intel i7 dual core, Memory: 32G).

each types of events, and the best average F1-score in general. 1DCNNs-CNN is in the second place, which is better than the 2D-CNN method. And the 1DCNNs is the worst, in which only the temporal structural features are used.

Firstly, it preliminarily shows that the spatiotemporal features have more distinguishable information than the single temporal structural features. Secondly, as stated above, the detailed temporal features and the spatial distribution have obvious different contributions in the DAS signal recognition. The spatial information mainly reflects the difference of the distribution trend, while the temporal information contains much more detailed structural difference for different types of operation processes. The spatial information is actually not as rich as the temporal information. Thus they cannot be treated in the same way. That's why the 2-D CNN behaves worse when it is compared with the proposed 1DCNNs-BiLSTM and the 1DCNNs-CNN networks. As we all know that, CNN has unique advantages in extracting detailed feature, and the LSTM has advantages in mining the relationship in sequences. And in this paper, we use CNN to extract the detailed temporal structure difference and utilize LSTM to dig out the spatial relationship of different signal nodes, and treat the two dimensional information in a different way. And the test results on the field data proves that this is reasonable. The proposed 1DCNNs-BiLSTM finally behaves better than the 1DCNNs-CNN networks when considering this difference in this application.

On the other hand, recognition of the excavator operation is the most challenging because the action is sometimes strong and sometimes weak, and its signal and the affection area change randomly. That's why the accuracy of label 3 data is significantly lower with 1DCNNs based on only the temporal feature, and with 2D-CNN based on the two dimensional contour pattern.
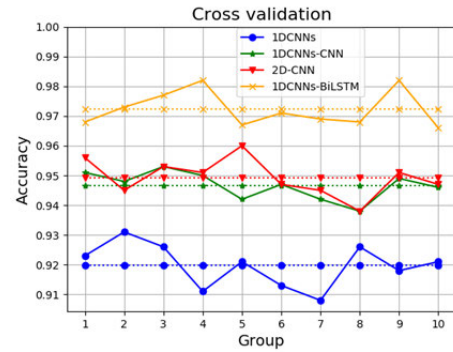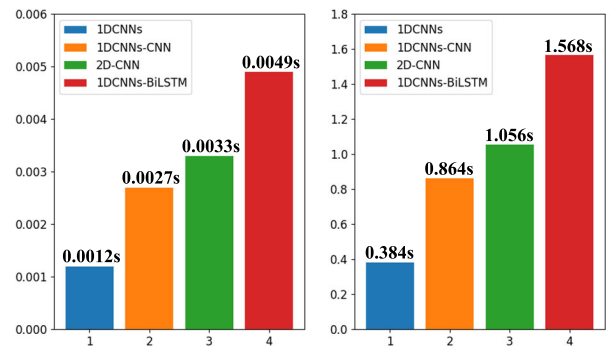
Furthermore, a ten-fold cross validation is also carried out on the database in Table 2 to verify the stability of the four models and the results are comparatively included in Fig. 10. The four lines with different marks stand for the obtained recognition accuracy for the above four methods in 10 random tests, and the four dotted lines represent the average accuracy for each method. In the ten-fold cross validation, the dataset in Table 2 is used and the ratio of training and testing sets is set as 8:2 in each test. The results in Fig. 10 show that the proposed 1DCNNs-BiLSTM always has a better recognition behavior and its average accuracy in the ten tests can be achieved at 97.2%, which is significantly higher than those of the other three methods; the 1DCNNs-CNN is in the second place, which behaves slightly better than, but more or less the same with the 2D-CNN, and the average accuracy of these two method are 95% and 94.2% respectively; and the 1DCNNs still performs the worst in all of the four methods, and its average accuracy can only be achieved at 92.5% in the ten tests. It shows that the proposed 1DCNNs-BiLSTM network behaves steadily the best in this field data test for DAS, which generally reveals that the 1DCNNs-BiLSTM has the best learning performance for the spatiotemporal information extraction in this application.

D. RECOGNITION SPEED DISCUSSION
Considering the practical application requirements, the computation speeds of the four methods are also compared as

shown in Fig. 11. For each test sample, the average test time of the four methods is included in Fig. 11(a); and for the whole line of 40km, the overall time of the four methods is compared in Fig. 11 (b). It demonstrates that: the computation speed of the 1DCNNs is the highest because it has only CNN learning network in it; and that of 1DCNNs-BiLSTM proposed in this paper is the slowest even though it has the best recognition performance, and the test time is about four times of the 1DCNNs; the 1DCNN-CNN is the second fastest method and the 2D-CNN is the third one in the computation speed. It shows BiLSTM takes a little more time because it has a more complicated network. Thus in this paper, a basic one-layer BiLSTM but not a two-layer structure as in [34] is used to improve the computation efficiency, which also avoids possible over fitting in the two-layer BiLSTM and improves the recognition performance of the whole model. Besides, the average merging operation for the hidden 256 states is also designed in the BiLSTM to improve the computation speed. In fact, the input sample of the CNN-BiLSTM network is a space-time matrix with a certain width but not a single temporal signal. In this field application, the monitoring cable is about 40km long and the spatial resolution of DAS is 5m, which means there are 8000 spatial points in total in the whole line. And in this application, every 25 spatial points are taken as a sample. Then there are 320 space-time samples to be identified at most in total. The recognition time for a line of 40km, takes about 1.59s for the proposed method, which is much less than the signal collection time of 30 seconds for the whole line in this case. In this way, the processing speed of the proposed method in this paper can still catch up the data collection rate well and it can be used on line in this application. Actually in other applications, it needs to weigh the tradeoff of the recognition rate and computation speed to choose proper algorithm. For example, the 1DCNN-CNN behaves better than the other two methods of 1DCNNs and 2-D CNN, but it has better computation efficiency.

V. CONCLUSION

In this paper, a novel 1DCNNs-BiLSTM based deep-learning model is proposed to automatically and accurately extract both the temporal structure feature and the spatial association characteristics for identifying the distributed DAS sensing signals. It helps to improve the recognition rate further compared with the frequently used models based on a single temporal feature extraction with 1D-CNN and the simultaneous space-time feature learning with 1DCNN-CNN and 2D-CNN. Moreover, the real-time computation efficiency of the proposed method is also discussed and compared with others, and it can be utilized on line. From the field test results, it shows the proposed 1DCNNs-BiLSTM is a promising method in the DAS signal recognition in practical complicated environments. It not only digs out the spatiotemporal information of the DAS signals in a deeper level, but also treats it in the two dimensions differently according to the different contribution of them.

REFERENCES

[1] J. C. Juarez, E. W. Maier, K. Nam, and H. F. Taylor, "Distributed fiber-optic intrusion sensor system," *J. Lightw. Technol.*, vol. 23, no. 6, pp. 2081–2087, Jun. 2005.
[2] A. Masoudi, M. Belal, and T. P. Newson, "A distributed optical fibre dynamic strain sensor based on phase-OTDR," *Meas. Sci. Technol.*, vol. 24, no. 8, Aug. 2013, Art. no. 085204.
[3] L. Zhang, S. Wang, N. Xue, F. Peng, M. Fan, W. Sun, J. Qian, J. Rao, and Y. Rao, "Coherent Φ-OTDR based on I/Q demodulation and homodyne detection," *Opt. Express*, vol. 24, no. 2, pp. 853–858, 2016.
[4] Y. Dong, X. Chen, E. Liu, C. Fu, H. Zhang, and Z. Lu, "Quantitative measurement of dynamic nanostrain based on a phase-sensitive optical time domain reflectometer," *Appl. Opt.*, vol. 55, no. 28, p. 7810, 2016.
[5] P. Jousset, T. Reinsch, T. Ryberg, H. Blanck, A. Clarke, R. Aghayev, G. P. Hersir, J. Henninges, M. Weber, and C. M. Krawczyk, "Dynamic strain determination using fibre-optic cables allows imaging of seismological and structural features," *Nature Commun.*, vol. 9, no. 1, Dec. 2018, doi: 10.1038/s41467-018-04860-y.
[6] H. Wu, Y. Qian, W. Zhang, and C. Tang, "Feature extraction and identification in distributed optical-fiber vibration sensing system for oil pipeline safety monitoring," *Photonic Sensors*, vol. 7, no. 4, pp. 305–310, Dec. 2017.
[7] D. Tan, X. Tian, W. Sun, Y. Zhou, L. Liu, Y. Ma, J. Meng, and H. Zhang, "An oil and gas pipeline pre-warning system based on Φ-OTDR," in *Proc. Int. Conf. Opt. Fibre Sensors*, Jun. 2014, Art. no. 91578W, doi: 10.1117/12.2054698.
[8] H. Wu, X. Li, Z. Peng, and Y. Rao, "A novel intrusion signal processing method for phase-sensitive optical time-domain reflectometry (Φ-OTDR)," in *Proc. Int. Conf. Opt. Fibre Sensors*, Jun. 2014, Art. no. 91575O, doi: 10.1117/12.2058503.
[9] A. Lv and J. Li, "On-line monitoring system of 35 kV 3-core submarine power cable based on φ-OTDR," *Sens. Actuators A, Phys.*, vol. 273, pp. 134–139, Apr. 2018.
[10] J. Tejedor, J. Macias-Guarasa, H. Martins, D. Piote, J. Pastor-Graells, S. Martin-Lopez, P. Corredera, and M. Gonzalez-Herraez, "A novel fiber optic based surveillance system for prevention of pipeline integrity threats," *Sensors*, vol. 17, no. 2, p. 355, Feb. 2017.
[11] Q. Sun, H. Feng, X. Yan, and Z. Zeng, "Recognition of a phase-sensitivity OTDR sensing system based on morphologic feature extraction," *Sensors*, vol. 15, no. 7, pp. 15179–15197, 2015.
[12] H. Wu, J. Chen, X. Liu, Y. Xiao, M. Wang, Y. Zheng, and Y. Rao, "One-dimensional CNN-based intelligent recognition of vibrations in pipeline monitoring with DAS," *J. Lightw. Technol.*, vol. 37, no. 17, pp. 4359–4366, Sep. 1, 2019.
[13] M. Aktas, T. Akgun, M. U. Demircin, and D. Buyukaydin, "Deep learning based multi-threat classification for phase-OTDR fiber optic distributed acoustic sensing applications," *Proc. SPIE*, vol. 10208, Apr. 2017, Art. no. 102080G.
[14] H. Li, Z. Zhang, F. Jiang, and X. Zhang, "An event recognition method for fiber distributed acoustic sensing systems based on the combination of MFCC and CNN," *Proc. SPIE*, vol. 10618, Jan. 2017, Art. no. 1061804.
[15] C. Xu, J. Guan, M. Bao, J. Lu, and W. Ye, "Pattern recognition based on time-frequency analysis and convolutional neural networks for vibrational events in φ-OTDR," *Opt. Eng.*, vol. 57, no. 1, p. 1, Jan. 2018.
[16] Y. Lu, T. Zhu, L. Chen, and X. Bao, "Distributed vibration sensor based on coherent detection of phase-OTDR," *J. Lightw. Technol.*, vol. 28, no. ,22, pp. 3243–3249, Nov. 15, 2010.
[17] T. Zhu, X. Xiao, Q. He, and D. Diao, "Enhancement of SNR and spatial resolution in φ-OTDR system by using two-dimensional edge detection method," *J. Lightw. Technol.*, vol. 31, no. 17, pp. 2851–2856, Sep. 1, 2013.
[18] Z. Qin, L. Chen, and X. Bao, "Wavelet denoising method for improving detection performance of distributed vibration sensor," *IEEE Photon. Technol. Lett.*, vol. 24, no. 5, pp. 542–544, Apr. 1, 2012.
[19] X. Hui, S. Zheng, J. Zhou, H. Chi, X. Jin, and X. Zhang, "Hilbert–Huang transform time-frequency analysis in φ-OTDR distributed sensor," *IEEE Photon. Technol. Lett.*, vol. 26, no. 23, pp. 2403–2406, Dec. 1, 2014.
[20] H. Wu, J. Wang, X. Wu, and Y. Wu, "Real intrusion detection for distributed fiber fence in practical strong fluctuated noisy backgrounds," *Sensor Lett.*, vol. 10, no. 7, pp. 1557–1561, Jul. 2012.
[21] H. Zhu, C. Pan, and X. Sun, "Vibration waveform reproduction and location of OTDR based distributed optical-fiber vibration sensing system," *Proc. SPIE*, vol. 9062, Jan. 2014, Art. no. 89931P.

H. Wu *et al.*: Novel DAS Signal Recognition Method Based on Spatiotemporal Information Extraction With 1DCNNs-BiLSTM Network

IEEE *Access*

[22] N. Fang, L. Wang, D. Jia, C. Shan, and Z. Huang, "Walking intrusion signal recognition method for fiber fence system," in *Proc. Asia Commun. Photon. Conf. Exhibit.*, Nov. 2009, pp. 1–6.

[23] W. Zhaoyong, P. Zhengqing, Y. Qing, C. Haiwen, Q. Ronghui, and F. Zujie, "Fast pattern recognition based on frequency spectrum analysis used for intrusion alarming in optical fiber fence," *Chin. J. Lasers*, vol. 42, no. 4, 2015, Art. no. 0405010.

[24] H. F. Martins, D. Piote, J. Tejedor, J. Macias-Guarasa, J. Pastor-Graells, S. Martin-Lopez, P. Corredera, F. De Smet, W. Postvoll, C. H. Ahlen, and M. Gonzalez-Herraez, "Early detection of pipeline integrity threats using a smart fiber optic surveillance system: The PIT-STOP project," *Proc. SPIE*, vol. 9634, Sep. 2015, Art. no. 96347X.

[25] B. Wang, S. Pi, Q. Sun, and B. Jia, "Improved wavelet packet classification algorithm for vibrational intrusions in distributed fiber-optic monitoring systems," *Opt. Eng.*, vol. 54, no. 5, May 2015, Art. no. 055104.

[26] Q. Tian, C. Zhao, Y. Zhang, and H. Qu, "Intrusion signal recognition in OFPS under multi-level wavelet decomposition based on RVFL neural network," *Optik*, vol. 146, pp. 38–50, Oct. 2017.

[27] A. K. Fedorov, M. N. Anufriev, A. A. Zhirnov, K. V. Stepanov, E. T. Nesterov, D. E. Namiot, V. E. Karasik, and A. B. Pnev, "Note: Gaussian mixture model for event recognition in optical time-domain reflectometry based sensing systems," *Rev. Sci. Instrum.*, vol. 87, no. 3, Mar. 2016, Art. no. 036107.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, vol. 1, 2012, pp. 1097–1105.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 1, 2016, pp. 770–778.

[30] S. Hershey, S. Chaudhuri, D. P. W. Ellis, J. F. Gemmeke, A. Jansen, R. C. Moore, M. Plakal, D. Platt, R. A. Saurous, B. Seybold, M. Slaney, R. J. Weiss, and K. Wilson, "CNN architectures for large-scale audio classification," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 131–135. [Online]. Available: https://ieeexplore.ieee.org/document/7952132/

[31] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018.

[32] J. Tejedor, H. F. Martins, D. Piote, J. Macias-Guarasa, J. Pastor-Graells, S. Martin-Lopez, P. C. Guillen, F. De Smet, W. Postvoll, and M. Gonzalez-Herraez, "Toward prevention of pipeline integrity threats using a smart fiber-optic surveillance system," *J. Lightw. Technol.*, vol. 34, no. 19, pp. 4445–4453, Oct. 1, 2016.

[33] H. Wu, X. Liu, Y. Xiao, and Y. Rao, "A dynamic time sequence recognition and knowledge mining method based on the hidden Markov models (HMMs) for pipeline safety monitoring with Φ-OTDR," *J. Lightw. Technol.*, vol. 37, no. 19, pp. 4991–5000, Oct. 1, 2019.

[34] A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, and S. W. Baik, "Action recognition in video sequences using deep bi-directional LSTM with CNN features," *IEEE Access*, vol. 6, pp. 1155–1166, 2018.

[35] Z. Duan, Y. Yang, K. Zhang, Y. Ni, and S. Bajgain, "Improved deep hybrid networks for urban traffic flow prediction using trajectory data," *IEEE Access*, vol. 6, pp. 31820–31827, 2018.

[36] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: http://arxiv.org/abs/1502.03167

MINGRU YANG was born in Anhui, China, in 1996. He received the B.Eng. degree in communication engineering from the Anhui University of Electronical and Information Engineering, in 2017. He is currently pursuing the master's degree in optical engineering with the University of Electronic Science and Technology of China. His research interests include distributed optical fiber sensing technology and data mining.

SIQI YANG was born in Hunan, China, in 1997. She received the B.Eng. degree in communication engineering from Central South University, in 2018. She is currently pursuing the master's degree in optical engineering with the University of Electronic Science and Technology of China. Her research interests include distributed optical fiber sensing technology, signal processing, and data mining.

HAO LU was born in Shanxi, China, in 1996. He received the B.Eng. degree in communication engineering from the University of Electronic Science and Technology of China, in 2018. He is currently pursuing the master's degree in Optical Engineering with the University of Electronic Science and Technology of China. His research interests include distributed optical fiber sensing technology and data analysis.

CHAOQUN WANG was born in Shandong, China, in 1994. He received the B.Eng. degree in communication engineering from the Jinan University of Communication Engineering, in 2016. He is currently pursuing the master's degree in electronic and communication engineering with the University of Electronic Science and Technology of China. His research interests include distributed optical fiber sensing technology and data mining.

HUIJUAN WU (Member, IEEE) was born in Henan, China, in 1980. She received the B.Eng. and Ph.D. degrees in opto-electronic engineering from Chongqing University, Chongqing, China, in 2003 and 2009, respectively. From 2007 to 2008, she studied at the Department of Engineering Science, Oxford University as a Joint Training Doctoral Student. She joined the University of Electronic Science and Technology of China, in 2009, where she is currently an Associate Professor. Her research interests include sensing signal processing and optical fiber sensing systems. She is a member of OSA.

YUNJIANG RAO (Fellow, IEEE) was born in Yunnan, China, in 1962. He received the M.Eng. and Ph.D. degrees in optoelectronic instruments from Chongqing University, Chongqing, China, in 1986 and 1990, respectively. He joined the Optoelectronics Division, Electric and Electronic Engineering Department, Strathclyde University, U.K., as a Postdoctoral Research Fellow, from 1991 to 1992, and the Applied Optics Group of Physics Department, Kent University, U.K., as a Research Fellow and then a Senior Research Fellow, from 1992 to 1999. From 1999 to 2004, he was a Chang-Jiang Chair Professor under the Program of Chang-Jiang Scholar Professorship at Chongqing University. He is currently the Director of Key Laboratory of Optical Fiber Sensing and Communications, Ministry of Education, University of Electronic Science and Technology of China. His research interests include fiber-optic sensors and applications, random fiber lasers, and all-fiber devices. He is a Fellow of OSA and SPIE.

• • •