

Early Safety Warnings for Long-Distance Pipelines: A Distributed Optical Fiber Sensor Machine Learning Approach

Yiyuan Yang,¹ Yi Li,¹ Taojia Zhang,² Yan Zhou,² Haifeng Zhang^{3*}

¹ International Graduate School at Shenzhen, Tsinghua University, Shenzhen, China,

² PetroChina Pipeline Company, Langfang, China, ³ Research Institute of Tsinghua, Pearl River Delta, Guangzhou, China
yangyy19@mails.tsinghua.edu.cn, liyi@sz.tsinghua.edu.cn, {zhangtj01, kjzhouyan}@petrochina.com.cn, zhanghf@tsinghua-gd.org

Abstract

Automated pipeline safety early warning (PSEW) systems are designed to automatically identify and locate third-party damage events on oil and gas pipelines. They are intended to replace traditional, inefficient manual inspection methods. However, current PSEW methods cannot achieve universality for various complex environments because they are sensitive to the spatiotemporal stability of the signal obtained by its distributed sensors at various locations and times. Our research aimed to improve the accuracy of long-distance oil-gas PSEW systems through machine learning. In this paper, we propose a novel real-time action recognition method for long-distance PSEW systems based on a coherent Rayleigh scattering distributed optical fiber sensor. More specifically, we put forward two complementary feature calculation methods to describe signals and build a new action recognition deep learning network based on those features. Encouraging empirical results on the data collected at a real location confirm that the features can effectively describe signals in an environment with strong noise and weak signals, and the entire approach can identify and locate third-party damage events quickly under various hardware conditions with accuracies of 99.26% (500 Hz) and 97.20% (100 Hz). More generically, our method can be applied to other fields as well.

Introduction

Oil and gas pipelines are known as the backbone of global energy. Their small size and fast construction make them widely used in the field of energy transportation. Currently, the length of long-distance transportation pipelines worldwide exceeds 3.5 million km and is increasing by approximately 30,000 km per year¹. However, because buried pipes are not easy to supervise, and the environments along those pipelines are complicated, accidents are unfortunately not rare enough. If an accident occurs, it may cause oil and gas leakage or even an explosion, and cause great economic

losses, casualties, environmental pollution, and extremely negative publicity. In 2016, an excavator in Alabama, USA damaged a gas pipeline, causing a gas leak that triggered an explosion and forest fire, affecting the lives of millions of people and causing fluctuations in gas prices². In 2019, an oil pipeline explosion caused by man-made destruction in Hidalgo, Mexico resulted in 142 casualties³.

To ensure the safety of oil and gas transportation and supply, it is necessary to monitor the safety of oil and gas pipelines in real time using technical means and issue a warning before the damage happens. Nevertheless, current pipeline safety early warning (PSEW) systems are still based mainly on manual patrolling. With increasing pipeline mileages, these inefficient and high-cost monitoring approaches have become a major obstacle to the development of a modern, intelligent, and safe transportation pipeline system.

In recent years, the rapid development of sensors and artificial intelligence has provided an opportunity to promote the construction of intelligent pipelines. On one hand, fiber optic sensors are widely used in industrial settings such as transportation pipelines because of their long detection distances, good real-time performance, and weak radiation. On the other hand, big data and data storage technology can provide a large number of high-quality data resources for our algorithms. The goal is to provide datasets and deep learning (DL) to build a model with a powerful fitting ability for real-time early warning and spatiotemporal positioning.

Nevertheless, the following problems with PSEW systems remain: (1) The spatiotemporal features of sensor signals consistently change, making them difficult to process with a single algorithm. (2) Strong noise, weak signals, and signal fluctuations at the scene make an algorithm trained on an ideal condition difficult to fit. (3) A low-frequency signal can be processed at higher speed with a cheaper solution,

* Corresponding Author

Copyright © 2021, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹https://en.wikipedia.org/wiki/Pipeline_transport

²<https://www.nbcnews.com/news/us-news/1-dead-colonial-gas-pipeline-explosion-alabama-n676046>

³<https://www.nbcnews.com/news/world/20-dead-54-injured-ruptured-pipeline-explosion-outside-mexico-city-n960516>

but due to its transmission of less information per unit time, a higher requirement is imposed on the algorithm.

To address the problems currently faced by PSEW systems, we present a novel action recognition method based on a distributed fiber optic sensor network that jointly considers “temporal aggregation” and “spatial aggregation”. Our experiments show that the results of the proposed method markedly outperform those of other baselines.

In brief, the main contributions of this paper include the following:

- We present two complementary characteristic calculation methods based on the spatiotemporal information of distributed signals.
- We concatenate the features with a novel DL method for action recognition and spatiotemporal localization of damage events.
- On the basis of experiments in a real location, we show that the method we propose has better real-time environmental adaptability and model performance than other baselines.
- We prove that our approach enables a wider deployment and has greater adaptability and extendibility for various hardware setups than other baselines.

Background and Related Work

Oil and Gas PSEW Systems

Optical fiber is currently considered the best industrial signal carrier because it supports low-cost and long-distance laying. In particular, the coherent Rayleigh scattering distributed optical fiber sensor has higher sensitivity and greater detection range and needs only an ordinary communication optical cable (Bao and Chen 2012). This makes it more compliant with the international hardware requirements for building distributed long-distance transportation pipelines.

The optical-fiber-based oil-gas PSEW system shown in Figure 1 is an intelligent system that recognizes and locates dangerous behaviors, issues early warnings, conducts on-site inspections, and records data in real time. First, the signal transmitter sends a narrow pulse signal into the optical fiber through a coupler (①). When a dangerous event occurs (②), the signal carries its information back to the signal receiver (③). Then the early warning unit immediately processes and recognizes the signal (④). If the prediction is of a damage event, the system notifies the unmanned aerial vehicle (UAV) closest to the event to automatically gather the evidence and issue a warning immediately (⑤). Detailed information is sent to the central monitoring center for filing (⑥). Also, the system periodically uploads the information to the data center to optimize and update the model and to automatically download the latest model to the early warning unit (⑦).

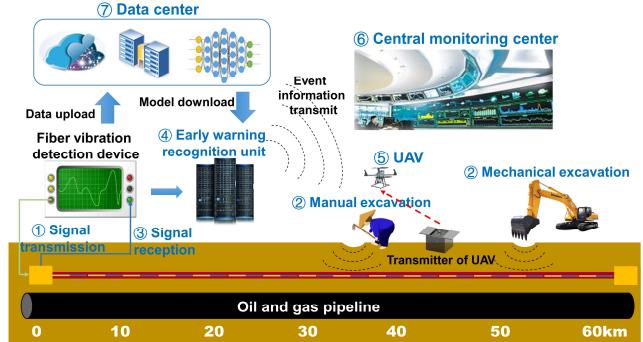


Figure 1. Structure of oil and gas PSEW system

Related Work and Solutions for PSEW

Research on PSEW algorithms is relatively recent, whereas many effective models dealing with PSEW have been extensively studied. They mainly take one of two directions:

Traditional methods. (Tanimola and Hill 2009) proposed method using distributed temperature sensing (DTS) and distributed acoustic sensing (DAS), and (Wu et al. 2017) used wavelet decomposition (WD) and wavelet packet decomposition (WPD) to extract features. Also, (Jiang et al. 2018; Zhang et al. 2018) applied frequency domain signal features such as mel-frequency cepstral coefficients for reference, and (Tabi Fouda et al. 2018) presented estimation methods for a frequency domain power spectrum.

Machine learning methods. (Kabir, Sadiq, and Tesfamariam 2016; Guo et al. 2018) used Bayesian network and its variant for PSEW, and (Sheng et al. 2019) updated the stochastic configuration network (SCN) proposed by (Wang and Li 2017) based on truncation singular value decomposition, calling it TSVD-SCN. (Wu et al. 2019) applied hidden Markov model (HMM) to extract the event areas and judge event categories. Also, (Yang et al. 2019) used a modified convolutional neural network (CNN), and (Kong et al. 2020) applied a probabilistic neural network to solve this problem.

However, most of them do not fuse the information of the spatial and temporal dimensions, which is the biggest difference between distributed and undistributed systems.

Action Recognition Methods

Action recognition is a subarea of computer vision that uses the spatiotemporal information for analysis. Common methods include CNNs, the optical flow method, recurrent neural networks (RNNs), and graph neural networks (Zhang et al. 2019; Hussein, Gavves, and Smeulders 2019). Similarly, distributed signals depend on both spatial and temporal dimensions, hence action recognition methods such as CNNs and RNNs can be used to deal with distributed signals.

As for CNNs, it can extract useful features in short fixed-length segments of the signal. Moreover, it is somewhat effective when the location of the segmental feature is not

highly relevant (LeCun, Bengio, and Hinton 2015). Specifically, the kernel in a convolution can be considered as a filter that removes outliers and acts as a feature extractor, which is defined as maximizing the response to a specific time series within the kernel time span (Zeng et al. 2014). Several CNN baselines (Krizhevsky, Sutskever, and Hinton 2012; Szegedy et al. 2015; Iandola et al. 2016; Szegedy et al. 2016; He et al. 2016) have been used in industry (Huang et al. 2018; Wang et al. 2019; Zhou et al. 2019; Hannun et al. 2019).

Also, RNNs and their variants, long short term memory (LSTM), bi-directional long short term memory (Bi-LSTM), and gated recurrent units (Graves and Schmidhuber 2005; Chung et al. 2014; Greff et al. 2016), through a recursive unit and a memory unit (Hochreiter and Schmidhuber 1997; Pascanu, Mikolov, and Bengio 2013), can be used for feature extraction of signals with high correlation, variable correlation lengths, and long-term dependence in the time domain. They are widely used in machine translation, speech recognition, and time series analysis.

Methodology

Signal Attenuation Compensation

When a signal is transmitted through an optical fiber, its energy decreases exponentially with the propagation distance, so a signal booster is used to compensate for signal attenuation every approximately 25 km for a real long-distance pipeline. To ensure the universality of the algorithm for signals at different distances, we compensated for the segmented attenuation of the signals between signal boosters based on least squared error, as shown in Figure 2.

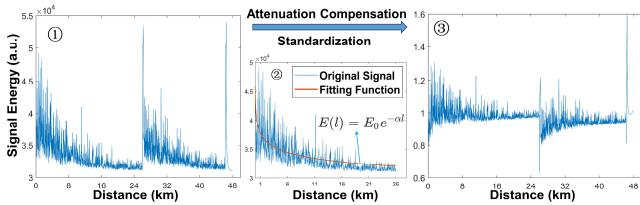


Figure 2. Segmental compensation of signal attenuation. ① Original signal. ② A least squared error and an exponential function fit of the signal between signal boosters (the signal attenuation compensation of the first segment in ① is shown). ③ The signal is spliced to create a complete signal.

Feature Generator

We propose two complementary features based on the spatiotemporal information: a peak feature, M_{peak} and an energy feature, M_{energy} . The peak feature describes the signal's high frequency and instantaneousness, and the energy feature describes the low frequency and continuous information. Algorithm I shows the calculation method in detail.

Algorithm I Matrix of Peak and Energy Features M_{peak} , M_{energy}

Input: Origin data X , Background noise data X_{base} .
Output: Matrix of Peak and Energy Features M_{peak} , M_{energy}
variable: Length of window and step N_{win} , N_{step} , Number of observation points L , Number of data in time dimension T , Number of observation points $N_{d-point}$ and windows N_{d-win} to be considered, Threshold α and β .

```

1: Attenuation compensation and standardization of  $X$  and  $X_{base}$ .
2: for each  $i = 1, \dots, L$  do
3:   for each  $j = 1, \dots, \frac{T}{N_{win}}$  do
4:      $F_{peak}[i, j] \leftarrow \text{Count peak}(X[i, j * N_{step}:j * N_{step} + N_{win}])$ 
5:     for each  $k = 1, \dots, N_{win} - 2$  do
6:        $F_{energy}[i, j] \leftarrow \frac{\sum(\text{data}[k+\alpha]-\text{data}[k])^2}{T * X_{base}}$ 
7:     end for
8:     Set  $F_{energy}[i, j] \leftarrow 1$  if  $F_{peak}[i, j] > \beta$ 
9:   end for
10: end for
11: for each  $m = \frac{N_{d-point}}{2}, \dots, L - \frac{N_{d-point}}{2}$  do
12:   for each  $n = 1, \dots, \frac{T}{N_{win} * N_{d-win}}$  do
13:      $M_{peak}, M_{energy} \leftarrow F_{peak}, F_{energy}[m - \frac{N_{d-point}}{2}; \frac{N_{d-point}}{2} + \frac{m+1}{2}, n * \frac{N_{d-win}}{2}:n * \frac{N_{d-win}}{2} + N_{d-win}]$ 
14:   end for
15: end for

```

Action Recognizer Based on Deep Learning

Because of the spatiotemporal information of a distributed signal, we considered using the action recognition method introduced in Section 2.3 to further analyze the two complementary features in Section 3.2 on spatiotemporal dimensions. Specifically, the signals in a time domain have a constant correlation length and no long-term dependency, so we used a 1D CNN to extract the spatial features. Temporal signals have long-term dependence and variable correlation length bidirectionally because of the signal scattering, reflection, or reverse propagation shown in Figure 1. Therefore, Bi-LSTM can obtain bidirectional and complex relations appropriately. We merged the model after the features were further extracted by the above networks and input the full connected network for specific event recognition and location, as shown in Figure 3.

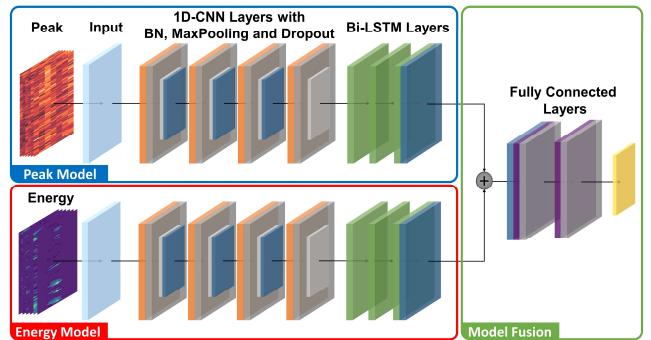


Figure 3. Action recognizer model. Orange is the Conv1D layer, dark grey is batch normalization, light grey is max-pooling, blue is dropout, green is the Bi-LSTM layer, and purple is the fully connected layer. The standardized peak and standardized energy features were expressed by heat maps and contour maps respectively.

Also, we applied batch normalization to keep data away from saturation zones and avoid distributed data bias, thereby reducing the training duration and improving the accuracy of the model (Ioffe and Szegedy 2015). Furthermore, we used max-pooling (Murray and Perronnin 2014) and dropout (Srivastava et al. 2014) strategies to simplify the model and alleviate overfitting.

Figure 4 shows the collaborative workflow of the feature generator and action recognizer. In the feature generator, we classified background noise preliminarily with the energy feature to ensure the efficiency of the algorithm. The remaining samples were then divided and input into the action recognizer for training, validation, and testing. From this, the best model was selected based on performance in the test set. Thereby, when new data came, the algorithm could recognize actions and spatiotemporal locations in real time.

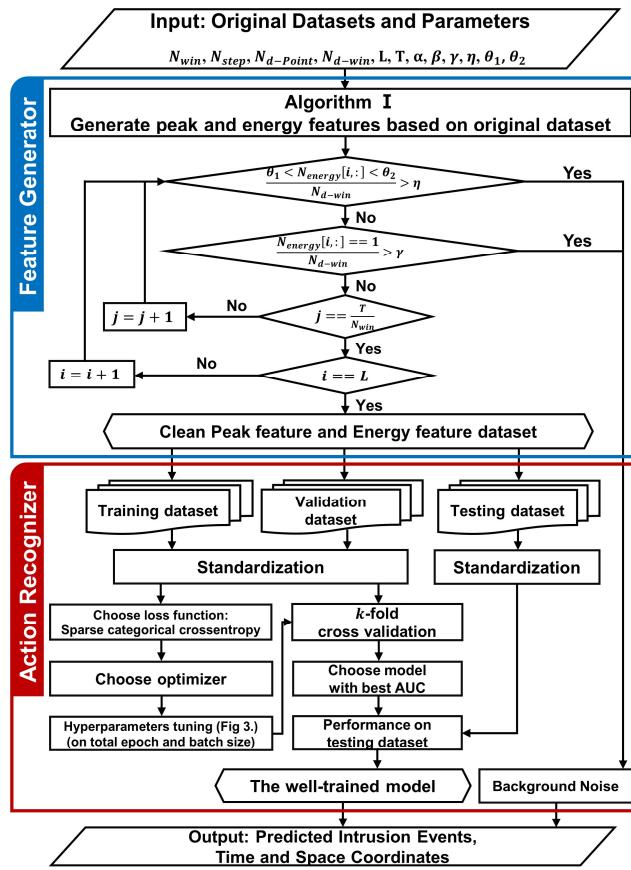


Figure 4. Collaborative workflow of the feature generator and action recognizer

Experiments and Results

This section introduces the datasets used for evaluating our work, as well as our preprocessing routines, experimental setting, final results, and the discussion on different algorithms and practical applications.

Datasets

The datasets we used were collected at real pipeline sites that had several types of noise, strong noise, weak valuable signal information, and signal attenuation characteristics specific to long-distance pipelines.

Specifically, the data were gathered in 2016 at a China National Petroleum Corporation pipeline from 10 May to 2 June and from 19 November to 17 December. The total data size was approximately 494 GB. The test pipeline was approximately 48 km long with 2,400 observation points, each of which was 20 m apart, i.e., the spatial resolution was 20 m. The main signal categories included four types of events: background noise (no pipeline-damaging events), manual excavation (may have been oil theft by drilling), mechanical excavation (third-party construction damaging the pipeline), and vehicle driving (potential threat of heavy vehicles rolling over the pipeline). We labeled the precise categories and spatiotemporal coordinates, and the waveforms of the signals are shown in Figure 5.

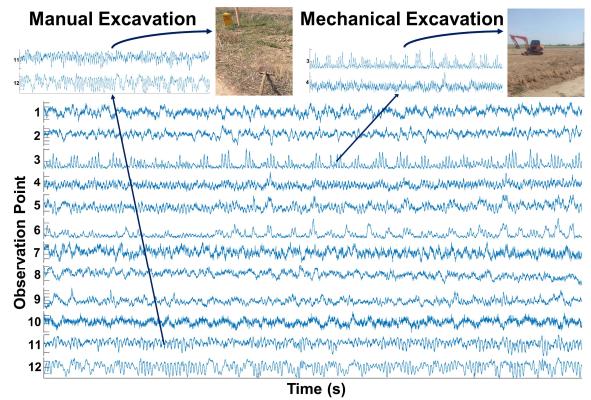


Figure 5. Waveforms of signals collected at a China National Petroleum Corporation pipeline

There were two main difficulties in collecting this type of dataset: (1) *Vast manpower consumption*. It required professionally skilled technicians to travel along the pipeline to simulate intrusions in poor and complex environments. Additionally, to acquire sufficient and more diverse data, intrusions had to be simulated at various locations in dozens of pipelines. (2) *Very time consuming*. To verify the signal drift and the robustness of the algorithm, professionals did long experiments from May to June and from November to December in real pipeline sites in harsh environments.

We extracted original signals according to their labels and classified them into four categories. Also, considering the applicability and cost with different devices, we did experiments at 100 Hz and 500 Hz. The lower the frequency used, the lower the equipment price, but the fewer data per unit time will be obtained. Therefore, higher requirements were put forward to the algorithm. It made engineering sense to test using signals at 500 Hz and 100 Hz separately.

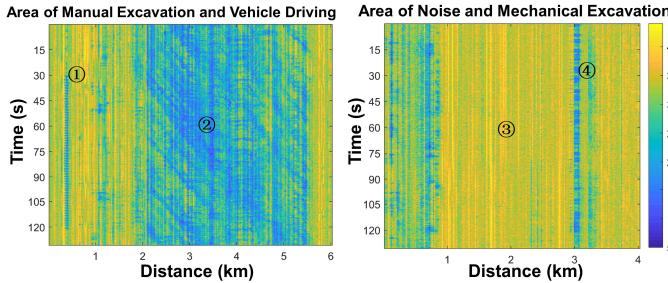


Figure 6. Heat maps of peak features at 500 Hz

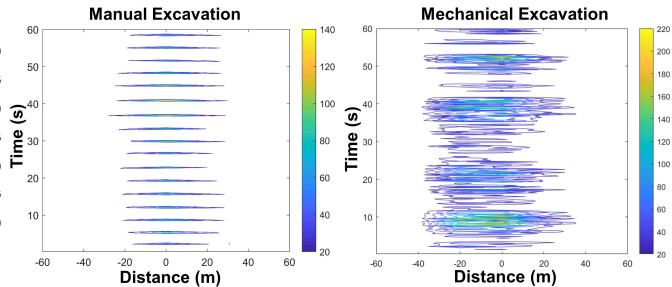


Figure 7. Contour maps of energy features at 500 Hz

Preprocessing and Feature Generation

We used the method in Section 3.1 to divide the pipe with 2,400 observation points into three segments, each of which was fitted using different and separate exponential functions. The signal average of each observation point was taken as the reference value of itself, and the piecewise compensation results were finally spliced into the complete results.

Then we calculated the two signal features described in Section 3.2. After repeated experiments, we found that the best number of temporal dimension windows N_{d-win} was 100 (i.e., 20 s), the best spatial dimension $N_{d-point}$ was 7 (i.e., 120 m of distance), the best window size N_{win} was 100, and the best movement step N_{step} was 100. Figure 6 shows that manual excavation (Figure 6①) had an obvious period and brief background noise between two excavation events. Vehicle driving (Figure 6②) created some slashes as a result of a road intersecting the pipeline in the corresponding area, and the dark slashes in Figure 6 represent cars passing on the road. The steeper the slash, the slower the vehicle, and the denser the slash, the greater the traffic flow. However, background noise (Figure 6③) appeared to be unchanged. Moreover, although mechanical excavation (Figure 6④) and manual excavation were periodic, there was no obvious brief background noise between two mechanical excavations, which was mainly due to its strong transient energy. Also, mechanical excavation had marked signal attenuations.

Figure 7 shows contour maps of the energy features. Similar to the peak features, the energy characteristics were periodic. Specifically, the manual excavation period was approximately 3.5 to 5 s, whereas that of the mechanical excavation was at least 10 s. As for the maximum values of the energy features, those of the manual excavation were much lower than those of the mechanical excavation. In brief, we could distinguish various categories from the periods and maximum values of the energy features.

Also, we applied a matrix operation from NumPy instead of using the loop in Algorithm I. According to the results, it took only 13.23 s to compute the features of a 48-km pipe with 4-min, 500 Hz of data using an Intel Core i7-8700 CPU at 3.2 GHz, a GTX1080ti GPU, and 32 GB of RAM. The corresponding 100 Hz data took 3.028 s.

	Training	Validation	Testing
Noise	8,500 / 8,500	1,500 / 1,500	540 / 540
Manual	12,954 / 9,375	2,286 / 1,655	1,740 / 1,030
Mechanical	7,097 / 6,502	1,253 / 1,148	850 / 450
Vehicle	10,378 / 8,202	1,832 / 1,448	1,710 / 650
Total event	38,929 / 32,879	6,871 / 5,751	4,840 / 2,670

Table 1. Descriptions of the training, validation, and testing datasets (100 Hz / 500 Hz data)

Action Recognition

Considering the industrial deployment, we used TensorFlow 2.0.0 and Keras 2.3.1 to build an action recognition model under Python 3.7.3. We applied the features obtained in Section 4.2 as input and the network of Figure 3 for training. To verify the performance and adaptability of the model with different time data (i.e., the signal drift problem), we divided all data into training and validation sets comprising the data from May and June and the testing set data from November and December. The datasets are described in Table 1.

Specifically, we used the sparse categorical cross-entropy loss function, which is specifically designed for multiclassification problems with label encoding for integers. Also, we applied the RMSprop optimizer to learn the parameters. RMSprop is commonly used in RNNs to control the acquisition of historical information through attenuation coefficients, thus speeding up gradient drops and model convergence (Hinton, Srivastava, and Swersky 2012). Moreover, we set the initial learning rate to 0.001, the batch size to 256 with 70 epochs. The dropout rate was the default of 0.15. The model's input was a two-feature matrix of $N \times 100 \times 7$, where N is the total amount of data, N_{d-win} is 100, and $N_{d-point}$ is 7. The output was 4, which corresponded to the predicted probability of 4 events.

Results and Discussion

In this section, we evaluate the model with datasets at 100 Hz and 500 Hz. First, we considered the effects of various optimizers on the model with the same parameters and data because the use of the optimizer is an argument crucial

for the training of a machine learning model. A wisely chosen optimizer is supposed to decrease the gradients of the loss function as quickly and as much as possible. In this case, we used six common optimizers for evaluation: root mean square propagation (RMSprop), adaptive moment estimation, stochastic gradient descent, adaptive gradient algorithm (AdaGrad), an extension of AdaGrad (AdaDelta), and Nesterov-accelerated adaptive moment estimation (Nadam). Figure 8 shows the training performance of 70 epochs on each optimizer with their own optimal parameters. The loss results in Figure 8 show that regardless of whether the data was at 100 Hz or 500 Hz, the Nadam, AdaDelta, and RMSprop optimizers had the highest convergence speed and lowest overall losses among the six common optimizers. Moreover, RMSprop clearly evaluated the test set more accurately and trained the model more quickly than the other optimizers. Specifically, we used data at 500 Hz (100 Hz) to train for 70 epochs. The duration and accuracy on the test set were as follows: AdaDelta, 813.8 s (830.9 s) and 98.327% (95.016%). Nadam, 926.6 s (832.1 s) and 98.328% (95.639%). RMSprop, 775.8 s (701.3 s) and 99.257% (97.196%). Therefore, we chose RMSprop because it could not only reduce the training time but also ensure the most accurate model.

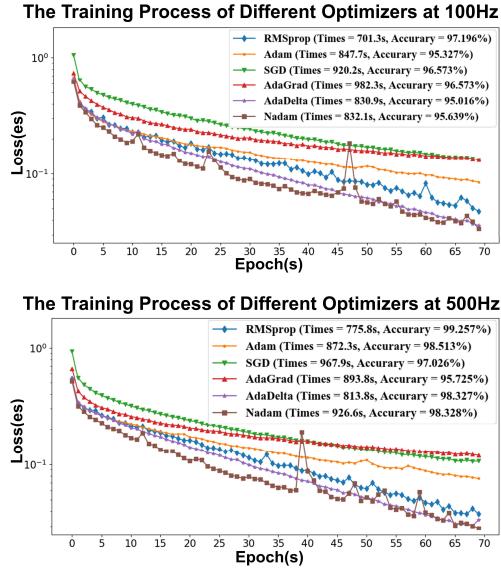


Figure 8. Comparison of various optimizers in the training process with data at 100 Hz and 500 Hz

Next, to further analyze how the number of Bi-LSTM layers impacted our approach, we tested the model performances that corresponded to one to five Bi-LSTM layers, as shown in Figure 9. For either 100 Hz or 500 Hz data, all model performances were optimal with three Bi-LSTM layers. In detail, for 500 Hz (100 Hz) data, the accuracy of the model corresponding to the three layers was 99.26% (97.20%), the macro-average was 99.52% (97.86%), and the

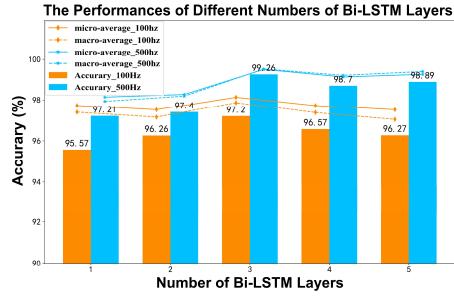


Figure 9. Comparison of performances under various numbers of Bi-LSTM layers with data at 100 Hz and 500 Hz

micro-average was 99.50% (98.13%) under the same conditions. Therefore, we believed it was best to use three Bi-LSTM layers to analyze the event features in the time domain. Increasing or decreasing the number of layers would result in inadequate or overfitted feature extraction.

Furthermore, we used other models to make comparisons under the same conditions. We repeated each experiment 10 times and reported the average results in Table 2. Note that the fusion model performed better than baselines that merely used a separate feature. Also, in optical fiber the signal is affected by scattering, reflection, and reverse propagation, which leads to long-term dependence and variable, bidirectional correlation length. Thus, Bi-LSTM was better at identifying intrusions than LSTM and CNN in temporal signals of optical fiber.

	CNN	CNN+LSTM	CNN+Bi-LSTM		
	Fusion	Fusion	Peak	Energy	Fusion
Background noise					
Pre(%)	100.0 / 100.0	100.0 / 100.0	99.86 / 99.26	98.18 / 98.13	100.0 / 100.0
Rec(%)	100.0 / 100.0	100.0 / 100.0	99.86 / 99.16	99.23 / 98.13	100.0 / 100.0
F1(%)	100.0 / 100.0	100.0 / 100.0	99.86 / 99.21	99.08 / 98.13	100.0 / 100.0
AUC	1.00 / 1.00	1.00 / 1.00	0.999 / 0.993	0.998 / 0.991	1.00 / 1.00
Manual excavation					
Pre (%)	100.0 / 100.0	100.0 / 100.0	100.0 / 100.0	98.84 / 94.39	100.0 / 100.0
Rec (%)	91.38 / 98.06	98.83 / 100.0	94.25 / 88.35	97.70 / 98.06	98.85 / 100.0
F1 (%)	95.50 / 99.02	99.40 / 100.0	97.04 / 93.81	98.27 / 96.19	99.42 / 100.0
AUC	0.957 / 0.990	0.994 / 1.00	0.971 / 0.942	0.986 / 0.977	0.994 / 1.00
Mechanical excavation					
Pre(%)	83.51 / 75.19	95.51 / 81.82	82.28 / 76.60	90.59 / 70.31	96.55 / 83.33
Rec(%)	95.29 / 100.0	100.0 / 100.0	76.47 / 80.01	90.59 / 100.0	98.82 / 100.0
F1(%)	89.01 / 85.71	97.70 / 90.00	79.27 / 78.26	90.59 / 82.57	97.67 / 90.91
AUC	0.959 / 0.973	0.996 / 0.982	0.867 / 0.880	0.944 / 0.966	0.991 / 0.984
Vehicle driving					
Pre(%)	97.70 / 100.0	100.0 / 100.0	89.30 / 86.67	95.32 / 100.0	99.42 / 100.0
Rec(%)	99.42 / 80.05	98.83 / 84.62	97.66 / 100.0	95.32 / 64.62	99.42 / 86.15
F1(%)	98.55 / 88.89	99.41 / 91.67	93.30 / 92.86	95.32 / 78.50	99.42 / 92.56
AUC	0.992 / 0.900	0.994 / 0.923	0.961 / 0.981	0.966 / 0.823	0.996 / 0.931
Total					
Acc(%)	96.28 / 95.33	99.22 / 96.89	93.68 / 93.44	96.28 / 92.05	99.26 / 97.20

Table 2. Performance comparison of various models on 500 Hz/100 Hz test sets

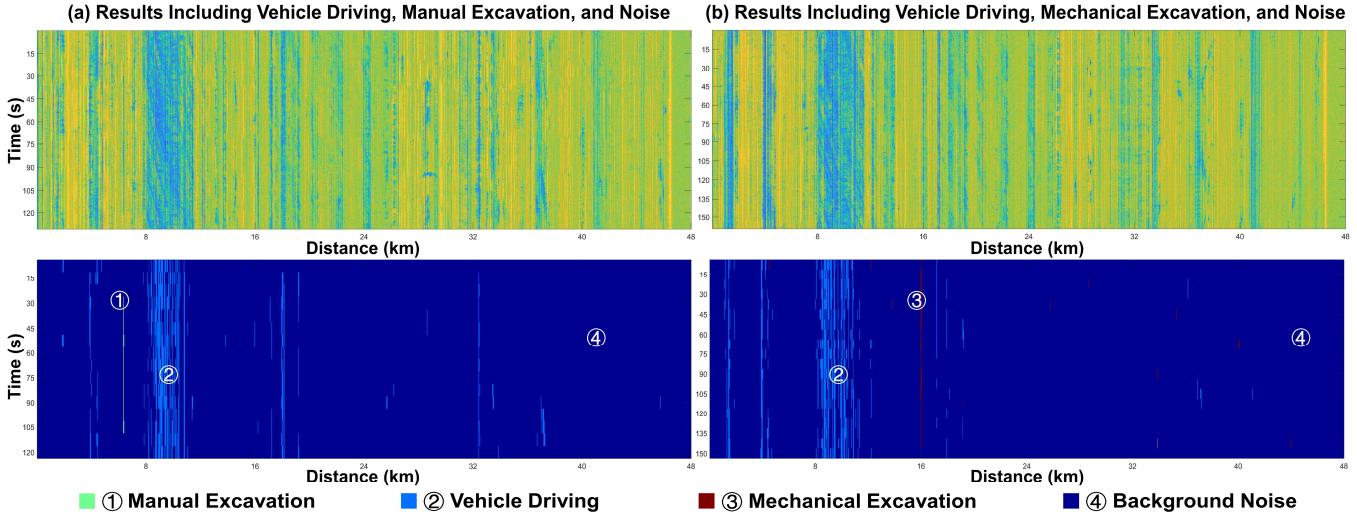


Figure 10. Heat maps comparing peak features and identification results from a real 48-km pipeline. The upper part of each graph shows the peak features of intrusion events, the lower half shows the identification results of intrusion events. (a) We can see that manual excavation events appeared at approximately 6 km and lasted nearly 90 s. (b) We can see that mechanical excavation events appeared at 16 km and lasted about 140 s. There were also continuous vehicle driving events at 8 to 10 km.

Also, because our aim was to solve practical engineering problems, the real-time performance and size of the model were important, so we tested a 4-min data from a 48-km pipe and repeated the experiment 10 times, and the average results are shown in Table 3. The Algorithms could identify and locate events accurately in an extremely short time. Specifically, it took 17.22 s for full identification of 500 Hz data and 6.597 s for 100 Hz, which exceeded the real-time performance requirements of the real scene. Also, the model size was only 18.7 MB, which could be deployed in most embedded systems and was highly adaptable to hardware.

Dataset	Feature time	Model time	Model size
500 Hz	13.23 s	3.715 s	18.7 MB
100 Hz	3.028 s	3.489 s	

Table 3. The model size and time cost of a 4-min data from a 48-km pipe

Finally, Figure 10 shows the identification results from a real 48-km pipe in the form of a heat map: the upper part of each graph is the peak feature, the lower half is the identification results of intrusion events. Figure 10a shows that our method was almost 100% accurate in the recognition of manual excavation and space-time positioning, and the false alarm rate was almost 0%. Similarly, as Figure 10b shows, the model can accurately locate the mechanical excavation in the time-space domain, but in some space-time points there were some misidentifications with a rate of approximately 2.56%. The samples were in an almost completely discrete state, and most of the action time was within 20 s. Naturally, by constraining the minimum time of predicted intrusion behavior with a threshold in our algorithm, we

could filter out most of the error recognition results. It was also valid to exclude them by the real-time evidences and videos acquired by unmanned aerial vehicle (UAV) (Figure 1⑤) and by the decision of the central monitoring center (Figure 1⑥).

Conclusions

In this paper, we propose a novel action recognition technology based on a distributed optical fiber sensor network to monitor the safety conditions of long-distance oil and gas pipelines in real time. According to the experimental results of the dataset we collected from a real scene, the two features proposed in this paper could express the essential information of various events under complex environmental conditions. The described algorithm can accurately identify and promptly locate third-party intrusion events in various environments. Furthermore, our model fully meets the actual industrial requirements of real-time, easy deployment, adaptability to various types of hardware, and the ability to be used in other industrial deployment applications.

As for the future, we are interested in exploring the application of distributed signal early warning systems in other areas, such as early warnings of undersea and land earthquakes, traffic flow statistics for urban road networks, and illegal cross-border behavior monitoring.

Acknowledgments

This research is supported by the pipeline branch of the China National Petroleum Corporation under grant number GDGS-2020-JS-24.

Ethics Statement

In this paper, we propose a novel early safety warning method for long-distance pipelines, and have been applied in the real scene. In terms of the impact on ethics and society, on one hand, our approach uses artificial intelligence (AI) to help us automatically monitor the safety of energy pipelines, and replaces some traditional manual inspection methods. Moreover, it can improve supervision quality, which contributes to energy supply and has a positive impact on society. On the other hand, we cannot completely rely on the whole systems we propose to carry out fully automated monitoring. If we meet some unforeseen circumstances unfortunately that make it unable to work, the system may make wrong judgments to the safety of the pipeline, which might lead to the emergence of energy security problems and even have quite bad impacts on society. So at present, although the system we presented has been applied on the real site, we still need a few professionals to supervise and control the system to ensure the reliability of our solution and maximize the contribution of the system for society.

References

- Bao, X., and Chen, L. 2012. Recent Progress in Distributed Fiber Optic Sensors. *Sensors* 12.7: 8601-8639.
- Chung, J.; Gulcehre, C.; Cho, K. H.; Bengio, Y. 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv preprint arXiv: 1412.3555*.
- Graves, A., and Schmidhuber, J. 2005. Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures. *Neural networks* 18.5-6: 602-610.
- Greff, K.; Srivastava, R. K.; Koutník, J.; Steunebrink, B. R.; Schmidhuber, J. 2016. LSTM: A Search Space Odyssey. *IEEE transactions on neural networks and learning systems* 28.10: 2222-2232.
- Guo, X.; Zhang, L.; Liang, W.; Haugen, S. 2018. Risk identification of third-party damage on oil and gas pipelines through the Bayesian network. *Journal of Loss Prevention in the Process Industries* 54: 163-178.
- Hannun, A. Y.; Rajpurkar, P.; Haghpanahi, M.; Tison, G. H.; Bourn, C.; Turakhia, M. P.; Andrew, Y. N. 2019. Cardiologist-level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms Using a Deep Neural Network. *Nature medicine* 25.1: 65.
- He, K.; Zhang, X.; Ren, S.; Sun, J. 2016. Identity Mappings in Deep Residual Networks. In *European conference on computer vision*. Springer, Cham, pp. 630-645.
- Hinton, G. E.; Srivastava, N.; and Swersky, K. 2012. Lecture 6d: A Separate, Adaptive Learning Rate for Each Connection. *Slides of lecture neural networks for machine learning 5*.
- Hochreiter, S., and Schmidhuber, J. 1997. Long Short-term Memory. *Neural computation* 9.8: 1735-1780.
- Huang, S.; Tang, J.; Dai, J.; W, Y. 2018. Signal Status Recognition based on 1DCNN and its Feature Extraction Mechanism Analysis. *Sensors* 19.9.
- Hussein, N.; Gavves, E.; and Smeulders, A. W. M. 2019. Timeception for Complex Action Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 254-263.
- Iandola, F. N.; Han, S.; Moskewicz, M. W.; Ashraf, K.; Dally, W. J.; Keutzer, K. 2016. SqueezeNet: AlexNet-level Accuracy with 50x Fewer Parameters and < 0.5 MB Model Size. *arXiv preprint arXiv:1602.07360*.
- Ioffe, S., and Szegedy, C. 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv preprint arXiv:1502.03167*.
- Jiang, F.; Li, H.; Zhang, Z.; Zhang X. 2018. An event recognition method for fiber distributed acoustic sensing systems based on the combination of MFCC and CNN. In *2017 International Conference on Optical Instruments and Technology: Advanced Optical Sensors and Applications*. International Society for Optics and Photonics, 10618:1061804.
- Kabir, G.; Sadiq, R.; and Tesfamariam, S. 2016. A fuzzy Bayesian belief network for safety assessment of oil and gas pipelines. *Structure and Infrastructure Engineering* 12.8: 874-889.
- Kong, W.; Yu, J.; Yang, J.; Tian, T. 2020. Model building and simulation for intelligent early warning of long-distance oil & gas storage and transportation pipelines based on the probabilistic neural network. *IOP Conference Series: Earth and Environmental Science*. IOP Publishing, 546(2): 022009.
- Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in neural information processing systems*. pp. 1097-1105.
- Lecun, Y.; Bengio, Y.; and Hinton, G. E. 2015. Deep Learning. *Nature*, 521(7553): 436-444.
- Murray, N., and Perronnin, F. 2014. Generalized Max Pooling. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 2473-2480.
- Pascanu, R.; Mikolov, T.; and Bengio, Y. 2013, February. On the Difficulty of Training Recurrent Neural Networks. In *International conference on machine learning*. pp. 1310-1318.
- Peng, Z.; Jian, J.; Wen, H.; Gribok, A.; Wang, M.; Liu, H.; Huang, S.; Mao, Z.; Chen, K. 2020. Distributed fiber sensor and machine learning data analytics for pipeline protection against extrinsic intrusions and intrinsic corruptions. *Optics Express* 28.19: 27277-27292.
- Sheng, Z.; Zeng, Z.; Qu, H.; Zhang, Y. 2019. Optical Fiber Intrusion Signal Recognition Method based on TSVD-SCN. *Optical Fiber Technology*, 48, 270-277.
- Srivastava, N.; Hinton, G. E.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. 2014. Dropout: a Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15.1, 1929-1958.
- Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. 2015. Going Deeper with Convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1-9.
- Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. A. 2016. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *arXiv preprint arXiv:1602.07261*.
- Tabi Fouad, B. M.; Han, D.; An, B.; Lu, X.; Tian, Q. 2018. Events Detection and Recognition by the Fiber Vibration System based on Power Spectrum Estimation. *Advances in Mechanical Engineering*, 10.11, 1687814018808679.

- Tanimola, F., and Hill, D. 2009. Distributed fibre optic sensors for pipeline protection. *Journal of Natural Gas Science and Engineering* 1.4-5: 134-143.
- Wang, D., and Li, M. 2017. Stochastic Configuration Networks: Fundamentals and Algorithms. *IEEE transactions on cybernetics*, 47.10, 3466-3479.
- Wang, H.; Liu, Z.; Peng, D.; Qin, Y. 2019. Understanding and Learning Discriminant Features based on Multiattention 1DCNN for Wheelset Bearing Fault Diagnosis. *IEEE Transactions on Industrial Informatics*, 16.9, 5735-5745.
- Wu, H.; Qian, Y.; Zhang, W.; Tang, C. 2017. Feature extraction and identification in distributed optical-fiber vibration sensing system for oil pipeline safety monitoring. *Photonic Sensors* 7.4: 305-310.
- Wu, H.; Liu, X.; Xiao, Y.; Rao, Y. 2019. A dynamic time sequence recognition and knowledge mining method based on the hidden Markov models (HMMS) for pipeline safety monitoring with φ-OTDR. *Journal of Lightwave Technology* 37.19: 4991-5000.
- Yang, Y.; Li, J.; Tian, M.; Zhou, Y.; Dong, L.; He, J. X. 2019. Signal Analysis of Distributed Optic-Fiber Sensing Used for Oil and Gas Pipeline Monitoring. In *Proceedings of the 2019 International Symposium on Signal Processing Systems*. pp. 21-25.
- Zeng, M.; Nguyen, L. T.; Yu, B.; Mengshoel, O. J.; Zhu, J.; Wu, P.; Zhang, J. 2014, November. Convolutional Neural Networks for Human Activity Recognition Using Mobile Sensors. In *6th International Conference on Mobile Computing, Applications and Services*. pp. 197-205.
- Zhang, H.; Zhang, Y.; Zhong, B.; Lei, Q.; Yang, L.; Du, J.; Chen, D. 2019. A Comprehensive Survey of Vision-Based Human Action Recognition Methods. *Sensors*, 19.5:1005.
- Zhang, Y.; Zhao, L.; Tian, Q.; Fan, J. 2018. Optical Fiber Intrusion Signal Recognition Based on Improved Mel Frequency Cepstrum Coefficient. In *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics*. IEEE, pp. 1-9.
- Zhou, Y.; Guo, Q.; Sun, H.; Yu, Z.; Wu, J.; Hao, L. 2019. A Novel Data-driven Approach for Transient Stability Prediction of Power Systems Considering the Operational Variability. *International Journal of Electrical Power & Energy Systems*, 107, 379-394.