Optica Applicata, Vol. LIV, No. 2, 2024
DOI: 10.37190/oa240208

Markov transition fields and auto-encoder-based preprocessing for event recognition of *Φ*-OTDR

Xin HU[1], Jingyi DAI[2], Ziyi WEI[2], Wei SHEN[1], Hao YU[1], Haiyang WU[1], Yingwen XU[1], Chengyong HU[2], Chuanlu DENG[2], Yi HUANG[2,*]

[1]China State Grid Taizhou Power Supply Company, Taizhou 225300, China

[2]Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai University, Shanghai 200444, China

*Corresponding author: huangyi1008@shu.edu.cn

To improve the model training efficiency and the classification performance of the phase-sensitive optical time-domain reflectometer (*Φ*-OTDR) in disturbance events recognition, a preprocessing method based on Markov transition fields (MTF) and auto-encoder (AE) is proposed. The phase time series, derived from demodulation of the original scattering signals, are converted into images by using the MTF method. Subsequently, an auto-encoder is introduced to perform a dimensionality reduction characterization of the MTF images, and the outputs of the encoder will be used as features for classification. The experimental results demonstrate that, compared with directly processing time series using 1-D CNN and classifying MTF images using CNN, the features obtained by the proposed method can accelerate the training process and improve the recognition performance of the classification model. The recognition accuracy for the four classes of events on the fence reaches 95.6%, representing a 12% increase.

Keywords: distributed optical fiber sensing, *Φ*-OTDR, disturbance recognition, Markov transition fields, auto-encoder.

1. Introduction

As a typical distributed optical fiber sensing technology, the phase-sensitive optical time-domain reflectometer (*Φ*-OTDR) is equipped with fundamental characteristics such as anti-electromagnetic interference, corrosion resistance, and long sensing distance [1-4]. Moreover, *Φ*-OTDR stands out in vibration sensing for its advantages such as accurate positioning capability, wide measuring range, and high sensitivity [5], which is widely used in electric power pipeline safety monitoring [6], traffic track monitoring [7], and security intrusion detection [8]. However, complex environments in practical applications, where plenty of interferences including system operation noise,

raining noise, and pedestrian footstep noise, lead to a high nuisance alarm rate (NAR) for *Φ*-OTDR [9].

In order to reduce NAR and improve the performance of monitoring target threat events, it is critical to investigate effective recognition methods for *Φ*-OTDR. In early researches, most efforts were devoted to extracting features manually from time series of signals, which were mainly phase temporal signals due to their linear correlation with external vibrations [10]. In 2017, XU *et al*. proposed to select short-time energy ratio, short-time level crossing rate (ST-LCR) and vibration duration as features, and the recognition rate of the four events reached more than 90% by using support vector machine (SVM) [11]. In 2019, WANG *et al*. extracted 14 features from vibration signals as input vectors for a random forest (RF) classifier, with an average recognition rate of 96.58% for four events [12]. To achieve effective classification results, these methods often require the artificial design of distinctive features tailored to specific event types or the construction of extensive feature sets, resulting in a high computational load and low feature utilization. In addition, it is difficult for researchers to discover effective features of complex events in practical applications, and thus the recognition results cannot meet expectations.

The introduction of deep learning methods has solved this problem. Compared to the manual feature extraction mentioned previously, end-to-end networks are better at capturing the complex relationships within the data. In 2018, XU *et al*. fed spectrograms of different vibration signals into a convolutional neural network (CNN) for feature extraction and classification, and the recognition rate was over 90% [13]. In 2019, WU *et al*. proved that one-dimensional (1-D) CNN outperforms CNN in both recognition accuracy and computational speed [14]. In 2023, LIU *et al*. proposed a deep learning network based on the deep belief network (DBN) and the gated recurrent unit (GRU) to effectively recognize five single events and four composite events with accuracies of 96.72% and 90.94%, respectively [15]. However, many of these deep learning algorithms are primarily derived from the speech and image processing fields, where the complex network structures present challenges in terms of model training costs.

Representation learning is a preprocessing method aimed at improving the characterization of data, so that it can be better understood and utilized by models. In this work, a preprocessing method based on Markov transition fields (MTF) and auto-encoder (AE) is proposed. The phase time series are converted into MTF images, which contain richer structural information and facilitate better capture of signal features [16]. However, converting the *Φ*-OTDR signals into MTF images requires the utilization of a more complex classification model, which may increase the computational cost. AE is applied to map MTF images to low-dimensional representations, known as encoding during the process, and outputs of the encoding are fed into the classification layer for event recognition. By utilizing the encoding and decoding process of the AE, key features can be extracted from the *Φ*-OTDR signals, and the computational cost of classification model can be reduced. In the experiment, the proposed method is employed

to identify four typical threat events on the optical fiber, which is fixed to the security fence. Compared with directly taking phase time series and MTF images as inputs to the deep learning networks, a higher recognition accuracy of up to 95.6% can be achieved with the proposed preprocessing method even if it is only in conjunction with a simply structured classification model, which is an improvement of more than 12% over the former two methods. At the same time, the training efficiency of the classification model is also significantly enhanced.

2. Principle

The processing flow of the proposed method is illustrated in Fig. 1, and the phase time series at each space node on the fiber is demodulated from the *Φ*-OTDR detected signals by using the in-phase/quadrature (I/Q) demodulation method. The preprocessing of the phase time series involves two main aspects, *i.e.*, converting the series into MTF images and encoding the images by AE. Unlike most methods that utilize phase time series for classification, data denoising is not involved in the proposed method, as the random noise present in the signal cannot alter the overall waveform trend and has no significant impact on the MTF images. Subsequently, the encoded output will be fed into the classification layer for training and testing to obtain the final event class.

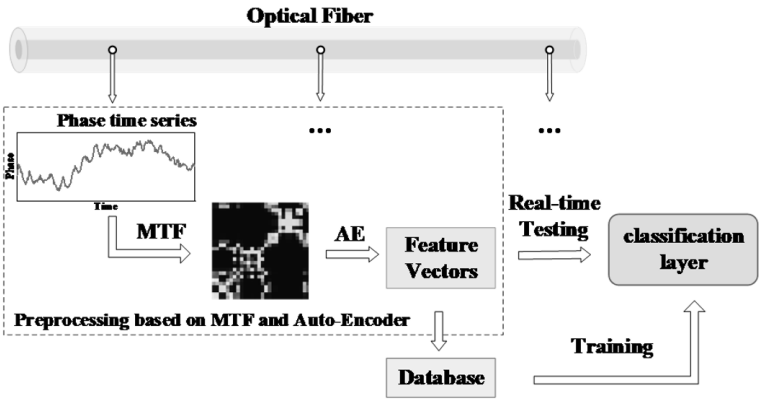

Fig. 1. The processing flow of the proposed method based on MTF and AE.

2.1. Markov transition field (MTF)

MTF is associated with Markov chains, which are represented by the probability distribution of state-to-state transitions within a system, and such the probability distribution is also known as Markov transition matrix [17]. A state transition diagram of a three-state Markov chain is shown in Fig. 2, where circles represent states, and edges with arrows represent transition probabilities between states.

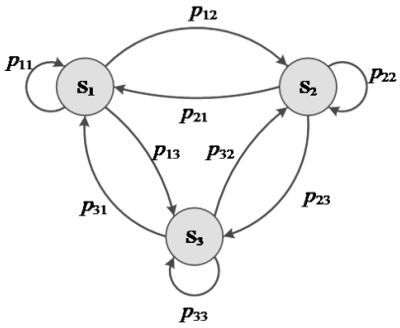

Fig. 2. The state transition diagram of a three-state Markov chain.

Then the Markov transition matrix corresponding to Fig. 2 can be expressed as,

$$\mathbf{P}_{33} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{bmatrix} \tag{1}$$

where p_{ij} is the transition probability from state i to state j, $0 \le p_{ij} \le 1$, and the sum of transition probabilities p_{ij} in each row is equal to 1. When a Markov chain has N states, the size of the transition matrix $\mathbf{P}_{NN}$ is $N \times N$ as follows,

$$\mathbf{P}_{NN} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix} \tag{2}$$

Vibration signals typically exhibit non-stationarity and randomness, which can be characterized by MTF. Given an obtained phase-time series $\mathbf{X} = \{x_1, x_2, \ldots, x_n\}$, the series X is firstly discretized by setting Q bins, with each series value x_i $(1 \le i \le n)$ corresponding to a unique bin q_j $(1 \le j \le Q)$. Then, the weight adjacency matrix $\mathbf{W}$ is constructed by calculating the transition probabilities among bins along the time axis [18], expressed as,

$$\mathbf{W} = \begin{bmatrix} w_{11} \left| p(x_i \in q_1 \middle| x_{i-1} \in q_1) \right. & \cdots & w_{1Q} \left| p(x_i \in q_1 \middle| x_{i-1} \in q_Q) \right. \\ \vdots & \ddots & \vdots \\ w_{Q1} \left| p(x_i \in q_Q \middle| x_{i-1} \in q_1) \right. & \cdots & w_{QQ} \left| p(x_i \in q_Q \middle| x_{i-1} \in q_Q) \right. \end{bmatrix} \tag{3}$$

where $\mathbf{W}$ is of size $Q \times Q$, and w_{ij} is the probability that the next point of a point in the bin q_i is in q_j. However, the insensitivity of $\mathbf{W}$ to temporal distribution leads to excessive information loss, so $\mathbf{W}$ is extended to MTF defined as follows [19],

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} w_{ij} \left| x_1 \in q_i, x_1 \in q_j \right. & w_{ij} \left| x_1 \in q_i, x_2 \in q_j \right. & \cdots & w_{ij} \left| x_1 \in q_i, x_n \in q_j \right. \\ w_{ij} \left| x_2 \in q_i, x_1 \in q_j \right. & w_{ij} \left| x_2 \in q_i, x_2 \in q_j \right. & \cdots & w_{ij} \left| x_2 \in q_i, x_n \in q_j \right. \\ \vdots & \vdots & \ddots & \vdots \\ w_{ij} \left| x_n \in q_i, x_1 \in q_j \right. & w_{ij} \left| x_n \in q_i, x_2 \in q_j \right. & \cdots & w_{ij} \left| x_n \in q_i, x_n \in q_j \right. \end{bmatrix} \tag{4}$$

where M_{ij} denotes the transition probability from bin q_i to bin q_j, q_i and q_j denote the corresponding bins of the series signals at time tamp i and j (namely x_i and x_j), respectively. By taking the temporal ordering into account, MTF achieves multi-span transition probability encoding [20], and M_{ij} can be viewed as pixels, allowing time series to be converted into images.

2.2. Auto-encoder (AE)

AE is an unsupervised learning algorithm that can automatically extract important features from unlabeled sample data [21], which includes two parts: the encoder and the decoder. The former maps the input data to a lower-dimensional representation, while the latter reconstructs the encoded result into the original input. In deep-structured networks, AEs are usually trained individually and then stacked together when training is complete. A standard AE has a relatively simple structure, consisting of an input layer, a hidden layer, and a reconstruction layer, as shown in Fig. 3.

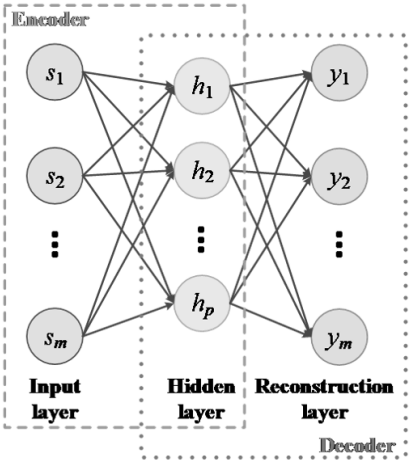

Fig. 3. The structure of the AE.

Given an unlabeled input sample $\mathbf{s} = [s_1, s_2, \ldots, s_m]^{\mathrm{T}}$, m is the number of nodes in the input layer. The encoder, which is composed of an input layer and a hidden layer, maps $\mathbf{s}$ to the hidden layer by a nonlinear transformation, and the obtained feature vector $\mathbf{h} = [h_1, h_2, \ldots, h_p]^{\mathrm{T}}$ can be expressed as,

$$\mathbf{h} = f_1(\mathbf{W}_1\mathbf{s} + \mathbf{b}_1) \tag{5}$$

where $\mathbf{W}_1$, $\mathbf{b}_1$, and $f_1(\cdot)$ are the weight, bias, and activation function of the encoder, respectively. Similarly, the decoder maps the feature vector $\mathbf{h}$ to the reconstruction layer via a nonlinear transformation, and the reconstructed signal $\mathbf{y} = [y_1, y_2, \ldots, y_m]^{\mathrm{T}}$ output by the decoder can be expressed as,

$$\mathbf{y} = f_2(\mathbf{W}_2\mathbf{s} + \mathbf{b}_2) \tag{6}$$

where $\mathbf{W}_2$, $\mathbf{b}_2$, and $f_2(\cdot)$ are the weight, bias, and activation function of the decoder, respectively. It is noteworthy that the input and reconstruction layers have the same number of nodes. Leaky rectified linear unit (ReLU) is employed as the activation function in this work, which is an improved ReLU activation function and is commonly used to address the issue of gradient vanishing and neuron death. The expression of Leaky ReLU is,

$$\text{Leaky ReLU}(x) = \begin{cases} x, & x > 0 \\ \alpha x, & x < 0 \end{cases} \tag{7}$$

The objective of AE training is to minimize the loss between the reconstructed signal $\mathbf{y}$ and the input signal $\mathbf{s}$ by optimizing the parameter set $\boldsymbol{\theta} = \{\mathbf{W}_1, \mathbf{b}_1, \mathbf{W}_2, \mathbf{b}_2\}$ [22], which is achieved by solving the following optimization problem,

$$\min_{\theta} L_{\mathrm{MSE}}(\mathbf{s}, \mathbf{y}) = \min_{\theta}\left(\frac{1}{2}\|\mathbf{s} - \mathbf{y}\|^2\right) \tag{8}$$

After the AE has been trained, the better the potential features it learns, the greater its ability to reconstruct the signal, which means that the reconstruction loss is lower. The dimension of the hidden layer in AE is often much lower than that of the input and reconstruction layers [23], which forces the model to learn the primary features of the data while disregarding noise or unimportant details. Eventually, the encoder's output $\mathbf{h}$ will be fed as features into the classifier for vibration event recognition.

3. Experimental results and discussion

3.1. Experimental setup

A heterodyne coherent Φ-OTDR system is used for data acquisition in the experiment, and its schematic diagram is shown in Fig. 4 (a). Continuous light with a central wavelength of 1550 nm is emitted from a narrow-linewidth laser (NLL) and divided into

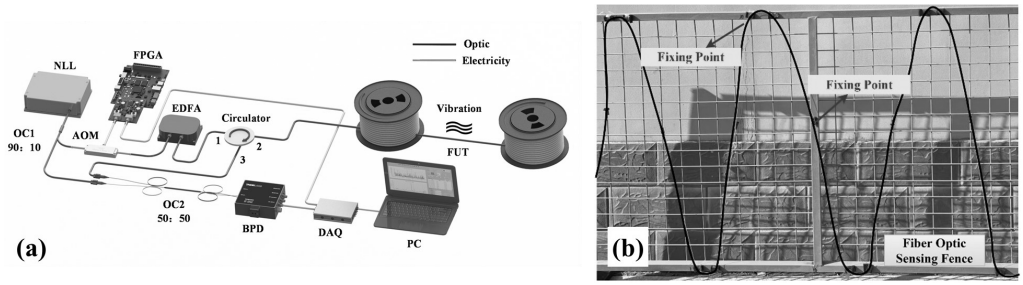

Fig. 4. (a) The heterodyne coherent *Φ*-OTDR system. (b) The S-shaped optical fiber on the security fence.

two parts at a 90:10 coupler (OC1). 90% of the light is used as the probe light for transmitting the sensing information and is modulated into optical pulses with 500 ns pulse width by an acousto-optic modulator (AOM) with 80 MHz frequency shift, which is driven by the pulse signal generated from the field programmable gate array (FPGA). Then, after being amplified by an erbium-doped fiber amplifier (EDFA), the optical pulses are launched into the sensing optical fiber via port 1 of the circulator. The backward Rayleigh scattering light returned from the optical fiber is collected by the circulator and is mixed with 10% light at a 50:50 coupler (OC2). The interference light is received by a balanced photodetector (BPD) and converted into electrical signals. Finally, the sampling is performed by a digital acquisition card (DAQ) with a sampling rate of 250 MS/s, which is controlled by synchronization signals output from the FPGA. A 14.8 km sensing optical fiber is connected to the Φ-OTDR system, and a section of the optical fiber is fixed in an S-shaped on the security fence to facilitate the application of disturbance, which is about 25 m long, as shown in Fig. 4(b).

3.2. Database preparation

According to the security fence scenarios constructed in the experiments, four noteworthy events are chosen as recognition targets in the work, including climbing, knocking, wind blowing and false disturbance. Tape was used to wrap the optical fiber on the fence. Because of the limitations of the field, the fiber is only taped on the fence at the test positions, where the experimenters would simulate various fence vibration events. All the event data were collected when actual events occurred on the fence or in its vicinity. The fence and optical fiber remain in the same arrangement and placement in all events. The climbing event records the situation where people are climbing the fence, resulting in compression of the optical fibers on the fence due to stepping or pressing, accompanied by strong shaking. The knocking event is generated by people striking the fence with tools such as plastic bottles or sticks. Compared to the climbing event, the knocking event exerts less strain on the optical fibers and has a limited impact range, shorter duration, and milder shaking. Wind-induced vibrations are also selected as a monitoring target, as wind blowing is an important factor leading to fence collapse. The wind blowing event was collected by placing the fence in an area with strong wind. Additionally, some events that would not pose a security threat to the fence

T a b l e. The database construction for vibration events.

Event type	Training samples	Test samples	Total	Label
False disturbance	150	50	200	0
Climbing	291	97	388	1
Knocking	298	100	398	2
Wind blowing	298	99	397	3
Total	1037	346	1383	–

are considered false disturbance events, including system background noise, pedestrian passing noise, and rainfall noise. All these events are recognized as only single targets in this work, and the constructed event database is divided into training set and testing set in a 7.5:2.5 ratio, with details shown in the Table.

3.3. Signal preprocessing and analysis

After applying the I/Q demodulation method to demodulate the detected signal, the obtained phase signals are used for preprocessing. The measurement time for each phase signal sample is 1 s, which is longer than the periods of most vibrations. Due to the fact that the sampled signals are discrete, it is actually the phase time series that

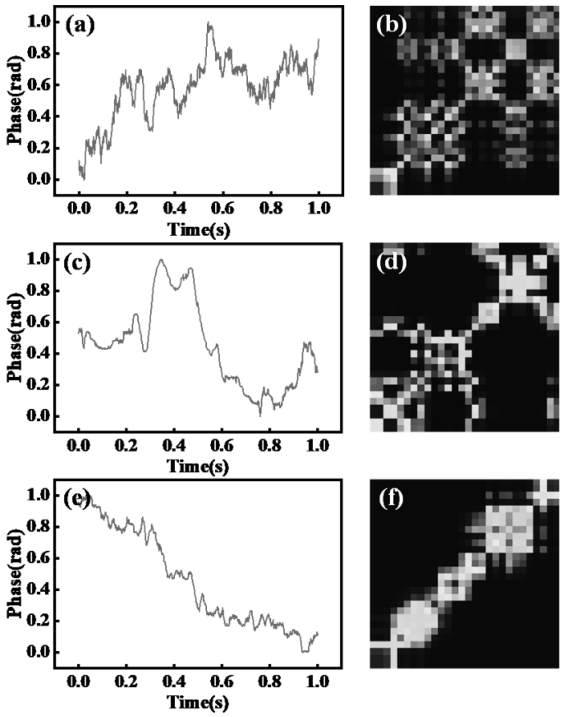

Fig. 5. Phase series and MTF images of the three events. (a), (b) Climbing. (c), (d) Knocking. (e), (f) Wind blowing.

are processed. Therefore, the length of each series corresponds to 1000 points, which is determined by the pulse emission frequency of the *Φ*-OTDR system, and it is set to 1 kHz in this work.

The phase series curves and their corresponding MTF images of the three threat events (climbing, knocking, and wind blowing) are shown in Fig. 5, where the resolutions of the MTF images are 369×369 to better present the information contained in the images. However, higher resolution will increase the time required for the conversion of series into images and model training, which significantly affects the recognition real-time performance, so the resolution of the image is reduced to 32×32 in the actual processing. It is obvious that the phase time series of different vibration types can also be quite different when converted into MTF images. The MTF images present large areas of similar color when the signal changes smoothly, while more segmented areas appear when the signal changes more rapidly.

Although converting series to MTF images can enrich the signal features, such dimensionality expansion will increase the complexity of the recognition model, resulting in model training speed reduction and performance degradation. Therefore, AE is employed to extract latent features from MTF images, which are challenging to be mined from the original sequence. The feature distributions obtained after encoding the MTF images of different event types by AE are shown in Fig. 6(a). To demonstrate the effectiveness of the MTF method being added, AE is also used to extract features from the original phase time series, the distribution of which is depicted in Fig. 6(b). Compared to the features extracted directly from the original phase time series by AE, the features obtained from MTF images exhibit higher distinctiveness, displaying four clearly defined clusters with well-defined boundaries in the figure.

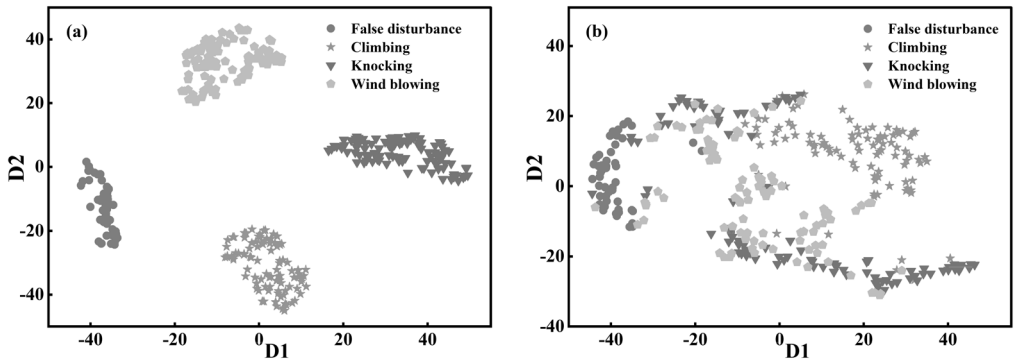

Fig. 6. Feature distributions distribution of the encoder output in AE. (a) Using MTF images as input. (b) Using the original phase series as input.

3.4. Recognition and results

To better validate the features that obtained by the proposed method can improve classification performance, a simple classification model is constructed for event recognition, which is a neural network consisting of two fully connected layers. The number

of input neurons in the model is 32, corresponding to the output feature dimension of the encoder in the AE. The number of output neurons is equal to the number of categories in the classification task, and it is 4 in this work. Softmax activation function is used in the output layer, namely the second fully connected layer, to transform the network's output into probability distributions for each category. Briefly, the model maps combinations of input features onto these categories to achieve classification. In addition, 1-D CNN has been reported previously to exhibit superior performance in *Φ*-OTDR recognition [14], so the original phase time series are fed into the simple classification model mentioned above and a 1-D CNN as a comparison here. All related computational processing is implemented by Python, using a laptop with Nvidia GPU (RTX 3050Ti). The training loss curves and testing confusion matrices of the classification models are shown in Fig. 7 and Fig. 8, respectively.

From the comparison of the training loss curves, it is obvious that the constructed simple classification model, which is used in conjunction with the proposed method, has the highest convergence speed. Whereas the training loss with the original phase

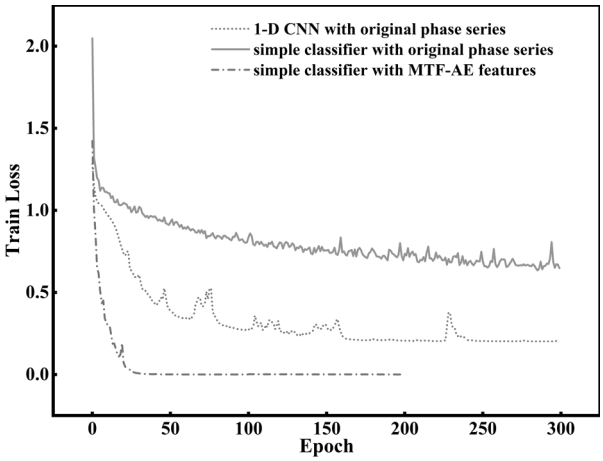

Fig. 7. Training loss curves.

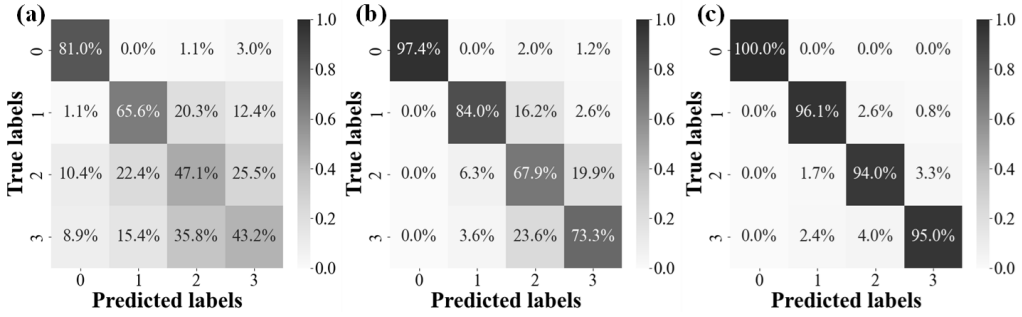

Fig. 8. Testing confusion matrices. (a) The simple classifier with original phase time series. (b) 1-D CNN with original phase time series. (c) The simple classifier with the proposed MTF-AE features.

series can only converge to about 0.6, much higher than that of the other two methods, indicating that the data is not well-understood and learned by the model. This explanation can be further supported by the confusion matrices in Fig. 8. The classification success rate for the three threat events is lower as shown in Fig. 8(b), while they are better distinguished in Fig. 8(a). The proposed method also achieves the highest average recognition accuracy up to 95.6%.

MTF images are fed into a CNN for recognition to demonstrate the advantage of adding AE for data dimensionality reduction in preprocessing. The time required for the CNN trained to converge using all MTF images in the training set is about 81 s, whereas that of the proposed method is only about 36 s, which includes the training time of the AE. Besides, recognizing a sample using CNN and using the proposed method takes 1.12 and 0.62 ms, respectively, implying that the computational time added by AE is much smaller than that reduced by model recognition. As can be seen in Fig. 9, the proposed method not only raises the classification model efficiency, but also further optimizes the classification results, leading to a 12% increase in recognition accuracy.

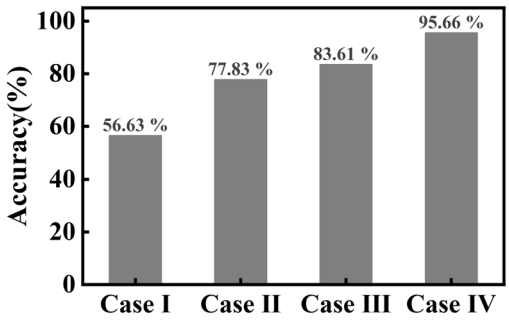

Fig. 9. Comparison of recognition results, where Case I: simple classifier with original phase series, Case II: 1-D CNN with original phase series, Case III: CNN with MTF images, Case IV: simple classifier with the proposed MTF-AE features.

4. Conclusions

In this work, MTF and AE are introduced to preprocess the phase signal for Φ-OTDR recognition. The phase time series without being denoised are converted into MTF images, and then the AE is applied to learn and extract latent features from these MTF images. Finally, the downscaled features obtained by the encoder are sent to the classification model for the recognition of vibration event types. The experimental results show that, compared to the 1-D CNN with original signals and the CNN with MTF images, only a straightforward model with the features obtained by the proposed method can achieve a higher recognition accuracy of 95.6%. Moreover, the proposed method can improve the training speed of the model, which means that the model can be trained in a shorter amount of time, and is beneficial for reducing the wastage of computational resources and time in practical applications of Φ-OTDR.

Funding
This work was funded by the Science and Technology Project of State Grid Jiangsu Electric Power Co., Ltd. (J2022101).

References

[1] HE Q., ZHU T., XIAO X.H., ZHANG B.M., DIAO D.M., BAO X.Y., *All fiber distributed vibration sensing using modulated time-difference pulses*, IEEE Photonics Technology Letters **25**(20), 2013: 1955-1957. https://doi.org/10.1109/LPT.2013.2276124

[2] PENG F., WU H., JIA X.-H., RAO Y.-J., WANG Z.-N., PENG Z.-P., *Ultra-long high-sensitivity Φ-OTDR for high spatial resolution intrusion detection of pipelines*, Optics Express **22**(11), 2014: 13804-13810. https://doi.org/10.1364/OE.22.013804

[3] BARRIAS A., RODRIGUEZ G., CASAS J.R., VILLALBA S., *Application of distributed optical fiber sensors for the health monitoring of two real structures in Barcelona*, Structure and Infrastructure Engineering **14**(7), 2018: 967-985. https://doi.org/10.1080/15732479.2018.1438479

[4] SHANG Y., YANG Y., WANG C., LIU X., WANG C., PENG G., *Optical fiber distributed acoustic sensing based on the self-interference of Rayleigh backscattering*, Measurement **79**, 2016: 222-227. https://doi.org/10.1016/j.measurement.2015.09.042

[5] HE X., XIE S., LIU F., CAO S., GU L., ZHENG X., ZHANG M., *Multi-event waveform-retrieved distributed optical fiber acoustic sensor using dual-pulse heterodyne phase-sensitive OTDR*, Optics Letters **42**(3), 2017: 442-445. https://doi.org/10.1364/OL.42.000442

[6] ZHANG X., WU J., SHAN Y., LIU Y., WANG F., ZHANG Y., *Online monitoring of power transmission lines in smart grid based on distributed optical fiber sensing technology*, Optoelectronic Technology **37**(04), 2017: 221-229.

[7] CATALANO E., COSCETTA A., CERRI E., CENNAMO N., ZENI L., MINARDO A., *Automatic traffic monitoring by ϕ-OTDR data and Hough transform in a real-field environment*, Applied Optics **60**(13), 2021: 3579-3584. https://doi.org/10.1364/AO.422385

[8] JUAREZ J.C., TAYLOR H.F., *Field test of a distributed fiber-optic intrusion sensor system for long perimeters*, Applied Optics **46**(11), 2007: 1968-1971. https://doi.org/10.1364/AO.46.001968

[9] KANDAMALI D.F., CAO X., TIAN M., JIN Z., DONG H., YU K., *Machine learning methods for identification and classification of events in ϕ-OTDR systems: A review*, Applied Optics **61**(11), 2022: 2975-2997. https://doi.org/10.1364/AO.444811

[10] RAO Y., WANG Z., WU H., RAN Z., HAN B., *Recent advances in phase-sensitive optical time domain reflectometry (Φ-OTDR)*, Photonic Sensors **11**, 2021: 1-30. https://doi.org/10.1007/s13320-021-0619-4

[11] XU C., GUAN J., BAO M., LU J., YE W., *Pattern recognition based on enhanced multifeature parameters for vibration events in φ-OTDR distributed optical fiber sensing system*, Microwave and Optical Technology Letters **59**(12), 2017: 3134-3141. https://doi.org/10.1002/mop.30886

[12] WANG X., LIU Y., LIANG S., ZHANG W., LOU S., *Event identification based on random forest classifier for Φ-OTDR fiber-optic distributed disturbance sensor*, Infrared Physics & Technology **97**, 2019: 319-325. https://doi.org/10.1016/j.infrared.2019.01.003

[13] XU C., GUAN J., BAO M., LU J., YE W., *Pattern recognition based on time-frequency analysis and convolutional neural networks for vibrational events in φ-OTDR*, Optical Engineering **57**(1), 2018: 016103. https://doi.org/10.1117/1.OE.57.1.016103

[14] WU H., CHEN J., LIU X., XIAO Y., WANG M., ZHENG Y., RAO Y., *One-dimensional CNN-based intelligent recognition of vibrations in pipeline monitoring with DAS*, Journal of Lightwave Technology **37**(17), 2019: 4359-4366. https://doi.org/10.1109/JLT.2019.2923839

[15] LIU M., WANG X., LIANG S., SHENG X., LOU S., *Single and composite disturbance event recognition based on the DBN-GRU network in φ-OTDR*, Applied Optics **62**(1), 2023: 133-141. https://doi.org/10.1364/AO.477642

[16] Zhao X., Sun H., Lin B., Zhao H., Niu Y., Zhong X., Wang Y., Zhao Y., Meng F., Ding J., Zhang X., Dong L., Liang S., *Markov transition fields and deep learning-based event-classification and vibration-frequency measurement for φ-OTDR*, IEEE Sensors Journal **22**(4), 2022: 3348-3357. https://doi.org/10.1109/JSEN.2021.3137006

[17] Jiang J.-R., Yen C.-T., *Product quality prediction for wire electrical discharge machining with Markov transition fields and convolutional long short-term memory neural networks*, Applied Sciences **11**(13), 2021: 5922. https://doi.org/10.3390/app11135922

[18] Gao Z., Yang H., Gong Y., Mu Y., *Fault diagnosis of shunt capacitor based on Markov transfer field transformation of vibration signal*, Acta Metrologica Sinica **44**(9), 2023: 1339-1346.

[19] Wang Z., Oates T., *Imaging time-series to improve classification and imputation*, Proceedings of the 24th International Conference on Artificial Intelligence, Buenos Aires, Argentina, 2015: 3939-3945.

[20] Markovin P.A., Trepakov V.A., Guzhva M.E., Kvyatkovskii O.E., Razdobarin A.G., Itoh M., *A crystal optical study of short range polar order in the ferroelectric phase: Doped incipient ferroelectrics*, Ferroelectrics **538**(1), 2019: 35-44. https://doi.org/10.1080/00150193.2019.1569983

[21] Rumelhart D.E., Hinton G.E., Williams R.J., *Learning representations by back-propagating errors*, Nature **323**(6088), 1986: 533-536. https://doi.org/10.1038/323533a0

[22] Baldi P., *Autoencoders, unsupervised learning, and deep architectures*, [In] *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, Proceedings of Machine Learning Research, Vol. 27, 2012: 37-49.

[23] Theis L., Shi W., Cunningham A., Huszár F., *Lossy image compression with compressive autoencoders*, Proceedings of International Conference on Learning Representations, Palais des Congrès Neptune, Toulon, France, 2017: 1-19.

Received December 3, 2023
in revised form February 8, 2024