

High-Performance Raman Distributed Temperature Sensing Powered by Deep Learning

Zhongshu Zhang, Hao Wu[✉], Can Zhao, and Ming Tang[✉], *Senior Member, IEEE, Member, OSA*

Abstract—Raman-based distributed temperature sensing (RRTS) can achieve temperature measurement of tens of kilometers and is widely used in temperature monitoring of crucial facilities. The accuracy of RRTS is highly determined by the signal-to-noise ratio (SNR) of the acquired spontaneous Raman scattering (SpRS) signals, which is about 60 dB weaker than the pump pulse. Therefore, the performance of single-mode fiber (SMF) based conventional RRTS is poor. To improve its performance, many methods have been proposed, including the use of special optical fibers, pulse coding techniques, and denoising algorithms. However, these methods have their limitations. Here, we propose and experimentally demonstrate a deep 1-D denoising convolutional neural network (1DDCNN) to enhance the performance of RRTS. A simplified RRTS model is built to train and optimize the 1DDCNN. To verify the performance of the 1DDCNN on actual data, we experimentally measure the SpRS data of a 10-km SMF with an average time of 1 s and a spatial resolution of 3 m. The well-trained 1DDCNN reduces the temperature uncertainty from 6.4 °C to 0.7 °C with high fidelity. As a comparison, the commonly used wavelet denoising algorithm can only reduce it to 3.7 °C. Besides, the 1DDCNN does not require manual parameter adjustment and is not affected by the sampling rate, giving it significant advantages in practical applications compared with conventional denoising algorithms. Moreover, we believe that the proposed 1DDCNN can be applied to other distributed optical fiber sensing systems by fine-tuning the network with appropriate training data.

Index Terms—Fiber optics sensors, raman scattering, distributed optical fiber sensors, convolutional neural networks.

I. INTRODUCTION

RAMAN-BASED distributed optical fiber temperature sensing (RRTS) has been investigated for decades due to its prominent advantages, such as fully distributed measurement over long distance, resistance to electromagnetic radiation, and real-time continuous measurement [1]. RRTS has been widely

Manuscript received July 28, 2020; revised October 4, 2020; accepted October 15, 2020. Date of publication October 19, 2020; date of current version January 15, 2021. This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1801002, in part by the National Natural Science Foundation of China under Grant 61722108 and Grant 61931010, and in part by the Innovation Fund of Wuhan National Laboratory for Optoelectronics(WNLO). (*Corresponding author: Hao Wu*)

The authors are with the Wuhan National Laboratory for Optoelectronics (WNLO) and National Engineering Laboratory for Next Generation Internet Access System, School of Optics and Electronic Information, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: 2350584655@qq.com; 2019507013@hust.edu.cn; zhao_can@hust.edu.cn; tangming@mail.hust.edu.cn).

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JLT.2020.3032150

used for temperature monitoring in power grids, oil and gas pipelines, and nuclear industry [2]–[5]. In an RRTS system, a pulsed light is injected into the sensing fiber and generates spontaneous Raman scattering (SpRS) whose intensity is sensitive to the fiber temperature. The temperature accuracy of RRTS is highly dependent on the signal-to-noise ratio (SNR) of the SpRS. However, the intensity of the SpRS is about 60 dB weaker than the pump pulse power. Besides, the signal decreases along with the fiber due to the transmission loss. As a result, the SNR is quite low for a typical RRTS, limiting its accuracy for long-distance temperature measurements.

To improve the performance of RRTS, many methods have been proposed to enhance the SNR. These methods can be broadly divided into two categories. One is by optimizing the implementation scheme of the system, such as optical pulse coding methods or by using special optical fibers [6], [7]. However, these methods also increase the complexity and cost of the system. The use of special optical fibers, such as multi-mode fiber and few-mode fiber, means that many existing single-mode fiber (SMF) resources for communication are wasted. The cost of laying special optical fibers often exceeds the sensors. The other way is to use digital signal processing algorithms to denoise the acquired data, such as short-time Fourier transform and wavelet denoising (WD) [8], [9]. These algorithms decompose the data on a specific basis to separate the noise in the data. However, the noise of RRTS is mainly white noise. The noise and signal cannot be completely separated on a single basis. The useful information in the signal is often lost while denoising. Moreover, there are many parameters of these algorithms that require manual adjustment, making it difficult to achieve an optimal result. And these parameters need to be readjusted for different systems and application scenarios. As a result, these algorithms are difficult to use in practical applications. Recently, two-dimensional (2D) signal processing methods have been proposed and turned out to be more efficient [10], [11]. However, RRTS data is natural one-dimensional (1D). To obtain 2D data, it is necessary to combine multiple consecutively acquired data, which significantly reduces the response speed of the system. Besides, the 2D denoising algorithms exploit the correlation between multiple measurements. However, for rapid temperature changes, the correlation between each measurement is limited, and 2D denoising methods will lose their advantages.

In this paper, we propose and experimentally demonstrate a deep 1D denoising convolutional neural network (1DDCNN) to enhance the performance of RRTS. The 1DDCNN is trained and optimized using synthetic RRTS signals. To verify the

performance of the trained network, experimentally acquired RDTs data are processed using the 1DDCNN and WD. With no deterioration in spatial resolution, the 1DDCNN shows better denoising performance than WD, resulting in a more significant reduction in temperature measurement uncertainty. And unlike WD, the 1DDCNN maintains the same excellent denoising effect at a low sampling rate. Besides, the end-to-end processing and the absence of manual parameter adjustment make the network of higher utility value. Moreover, the fully convolutional network structure ensures its computing speed for real-time processing.

II. PRINCIPLES AND METHODS

A. RDTs

When a short laser pulse couples into a sensing fiber, an inelastic interaction between the pulsed light and the optical phonon in the fiber generates SpRS. Since SpRS is generated due to the interaction of laser pulses and lattice vibration modes, a small portion is backscattered to Stokes light with a downshift frequency and anti-Stokes light with an upshift frequency. The intensity of SpRS is sensitive to the fiber temperature because the lattice vibration is affected by the ambient temperature [1]. Anti-Stokes light which is more sensitive to temperature is usually selected as the sensing signal. And Stokes light is used as a reference to eliminate instability caused by laser power fluctuations. The expression of the ratio of anti-Stokes light to Stokes light intensity $R(T_0)$ at temperature T_0 is [12]:

$$R(T_0) = \frac{P_{AS}(T_0)}{P_S(T_0)} = \frac{K_{AS}}{K_S} \left(\frac{v_{AS}}{v_S} \right)^4 \exp \left(-\frac{h\Delta v}{kT_0} \right) \exp [-(\alpha_{AS} - \alpha_S)L] \quad (1)$$

where P_{AS} (P_S) and v_{AS} (v_S) are the power and frequency of anti-Stokes (Stokes) light, respectively, Δv is the Raman frequency shift, α_{AS} (α_S) is the transmission loss factor, K_{AS} (K_S) is the scattering coefficient, h is the Planck constant, k is the Boltzmann constant, and L is the position of the scattering occurs. Comparing $R(T_0)$ and $R(T)$ measured at temperature T , we can get the expression of T :

$$\frac{1}{T} = \frac{1}{T_0} - \frac{k}{h\Delta v} \ln \frac{R(T)}{R(T_0)} \quad (2)$$

Therefore, the temperature accuracy of RDTs is related to the SNR of the acquired SpRS data. Denoising SpRS data before demodulation can effectively improve the temperature accuracy.

B. Wavelet Denoising

The wavelet threshold denoising algorithm is the most commonly used method for RDTs denoising. WD is based on the idea of wavelet analysis and reconstruction [9]. The process is shown in Fig. 1. Wavelet analysis uses the wavelet function to decompose the signal and obtain the corresponding wavelet coefficients. The low-frequency coefficients represent the approximate components of the signal and can be further decomposed. The high-frequency coefficients represent the detailed components of the signal. After wavelet decomposition of the

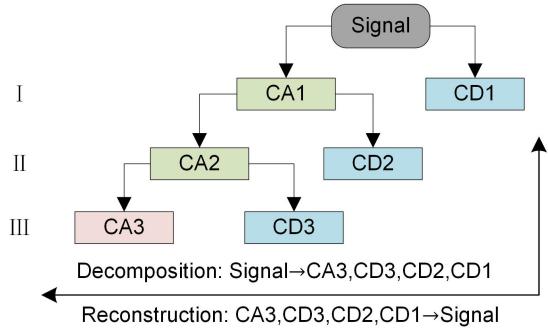


Fig. 1. Wavelet three-layer decomposition and reconstruction process. CA is low frequency information and approximate components, and CD is high frequency and detail components.

signal, we get a series of wavelet coefficients. Among them, the wavelet coefficient of the noise is smaller than that of the signal. By choosing a suitable threshold, wavelet coefficients larger than the threshold are considered to be generated by signals and should be retained; those smaller than the threshold are considered to be generated by noise and set to zero to achieve the purpose of denoising. Then the processed wavelet coefficients and wavelet function are used to reconstruct the signal. We choose the db3 wavelet function to decompose the RDTs signal in three layers, and the soft threshold method to process wavelet coefficients of each layer [13]. In this work, the thresholds are manually adjusted to reduce noise as much as possible without deteriorating spatial resolution.

C. Convolutional Neural Networks Denoising

Neural networks are a type of artificial intelligence techniques, and CNNs are currently the most widely studied and applied neural networks. CNNs have shown better results than conventional denoising algorithms in the fields of image, audio, and video [14]. A CNN can be understood as a filter, and it can achieve arbitrary filtering effects through multiple layers of convolution and nonlinear operations. By training a CNN with a large amount of data, the CNN can gradually optimize its filtering parameters to achieve optimal statistical results on the training data. This training process corresponds to the process of parameter optimization in conventional denoising algorithms. Therefore, it can be assumed that the CNN learns a prior knowledge of the data through training. This statistical process is only for the training data, so the prior knowledge it obtains is more accurate. Therefore, CNN can provide better denoising result than conventional algorithms. And the CNN has the generalization capability to denoise data with the similar characteristics when the amount of training data is sufficient.

D. 1DDCNN

Here we use a relatively plain network structure to analyze and denoise the SpRS data. As shown in Fig. 2, the 1DDCNN consists of 1D convolution (Conv) layers [14], 1D batch normalization (BN) layers [15], and rectified linear units (ReLU) [16]. Convolution operation can extract different features of SpRS

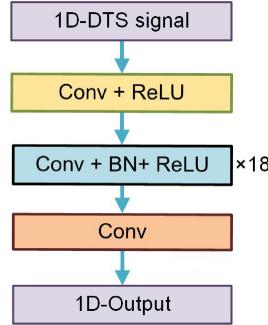


Fig. 2. Network structural diagram of the 1DDCNN.

data using different convolutional filters. The size of the convolution kernel is set to 3. And the channel number of the first and last convolutional layer is 1 corresponding to 1D SpRS data. The channel number of the rest convolutional layer is set to be 64 to extract more features. The size of the data is constant during the processing by padding zeros after each convolution operation. In shallow layers, the elementary features are obtained. And these features are combined in higher layers to learn characteristics of SpRS data. This is why CNN can restore a signal. BN is employed to normalize the data during training to help the network converge more quickly. By normalizing the layer's input without changing what the previous layer represents, BN can reduce the internal covariate shift problem in the training process [15]. As the most commonly used activation function, ReLU is used to increase the nonlinearity of the 1DDCNN. ReLU turns the result less than 0 into 0, and the data greater than 0 remains unchanged. So, it is not added in the last layer to avoid outputting only positive numbers. The entire network consists of 20 convolutional layers. So, the receptive field of 1DDCNN is 41, which means that each point of the output is related to 41 points of the input data.

E. Training Data and Training Process

For different RDTs systems and application scenarios, the data characteristics vary due to pulse width, scattering coefficient, transmission loss, and ambient temperature. Therefore, it is impossible to get enough training data through experiments. Here, we propose a simplified simulation method based on the characteristics of RDTs data. For a small section of fiber, where transmission loss can be ignored, we use randomly varying intensities to simulate signal changes due to temperature, loss, or reflection. Each intensity is consistent within a random number of sampling points to represent a relatively consistent temperature. Considering that the receptive field of the network is 41, we set a random range of 1 to 41 for the number of points with consistent intensity. And the synthetic data does not take into account the convolutional effects of the pump pulse so that it can satisfy an arbitrary spatial resolution. Then, Gaussian white noise with a standard deviation of 0.005 is added to the synthesized data. Since the random range of signal intensity is from 0 to 1, the maximum SNR of the synthetic data is 23 dB. We generate 4800 synthetic SpRS traces, of which 4000 traces are

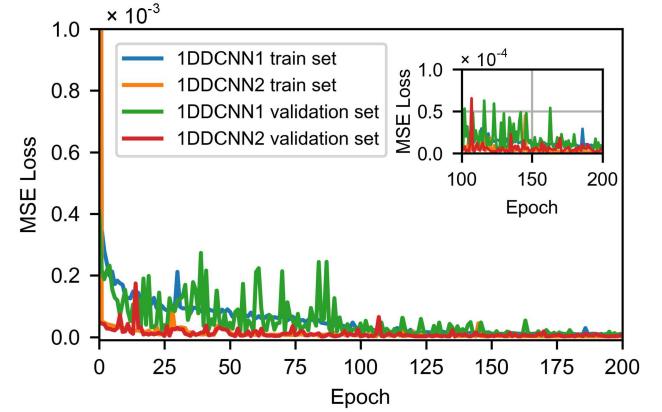


Fig. 3. MSE loss of each training epoch.

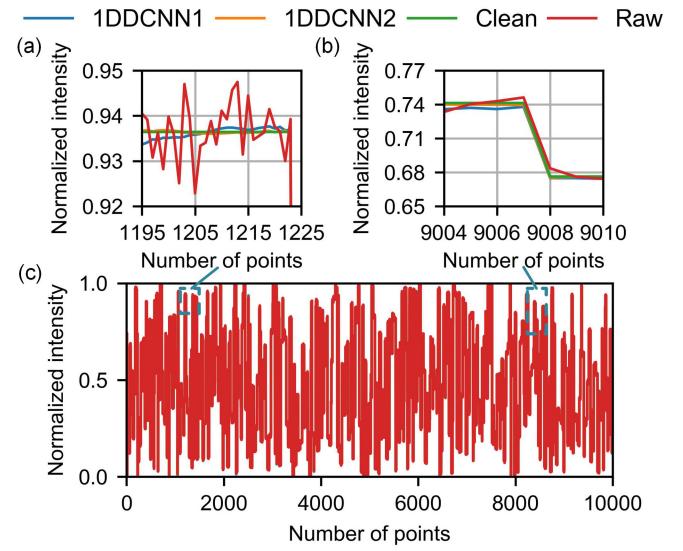


Fig. 4. Processed result of the networks on a validation curve. (a) Flat area; (b) mutation area; (c) entire validation curve.

used as the training set and the remaining 800 as the validation set. Each trace contains 10000 sampling points.

The neural network is trained for 200 epochs with a batch size of 16 and a learning rate of 0.001. In each epoch, the input data is first propagated forward, then the mean squared error (MSE) of the difference between the output and the target is calculated and propagated backward. The Adam optimization algorithm is used to update the network parameters [17]. It calculates the first-order moment estimation and second-order moment estimation of the gradient to design independent adaptive learning rates for different parameters. It takes about 58 minutes to complete the training process with Pytorch running on a PC with NVIDIA TITAN RTX GPU (24G).

The training target of a denoising neural network can be the clean trace or the corresponding noise. Here, we train two neural network models using clean trace (1DDCNN1) and noise (1DDCNN2) as the target output, respectively. As shown in Fig. 3, the loss of 1DDCNN2 decreases faster and smoother than 1DDCNN1. Fig. 4 shows the processed result of the trained

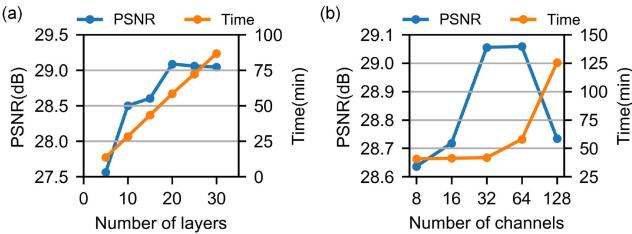


Fig. 5. PSNR of the validation data and training time with different number of convolution layers (a) and channels (b).

networks on a validation curve. Both networks have the denoising effect on the signal. The noise level of the validation curve processed using 1DDCNN1 drops to 0.0022 from 0.005. And the noise level of the validation curve using 1DDCNN2 is 0.0013. 1DDCNN2 works better using noise as the target because it is easier for the deep neural network to converge as the average value of noise is 0 [18]. Besides, using noise as the training target is consistent with the BN process, which can further improve the denoising effect. Therefore, we use the 1DDCNN2 to process the RDTs data.

By increasing the number of convolution layers or convolutional channels, the number of parameters in a neural network can be increased, which generally corresponds to higher performance. However, the more parameters, the greater the amount of calculation. Too many parameters may cause difficulty in convergence during training, affecting the training time and performance. The numbers of layers and channels of 1DDCNN are chosen based on its performance on the validation data. Fig. 5(a) shows the PSNR of the validation data and training time when the number of layers is 5, 10, 15, 20, 25, and 30. The training time increases as the number of layers increases, but the PSNR does not increase all the time. When the number of layers is 20, the network achieves the best denoising effect. And as shown in Fig. 5(b), when the number of channels is 64, the PSNR of the validation data is the highest. When the number of channels is less than 64, the training speed is mainly determined by the speed of reading and writing data, so the difference in their training time is small. To achieve the best denoising effect, we set the number of network layers to 20 and the number of channels to 64.

III. EXPERIMENTAL RESULT

To verify the performance of the 1DDCNN on actual data, an RDTs system is set up, as shown in Fig. 6. The system consists of a pulse laser, an erbium-doped fiber amplifier (EDFA), an optical circulator, a Raman wavelength division multiplexer (WDM), two avalanche photodiodes (APDs), a data acquisition (DAQ) card, and fiber under test (FUT). The sensing fiber is a SMF which is about 10-km long. The fiber end is immersed in a 60 °C water bath. The pulse laser operates at a wavelength of about 1550 nm and a pulse width of 30 ns. The pulse light is amplified by the EDFA and then launches into the FUT through the optical circulator. The backscattered SpRS lights enter the WDM through the optical circulator to separate the anti-Stokes light and Stokes light. These optical signals are converted into

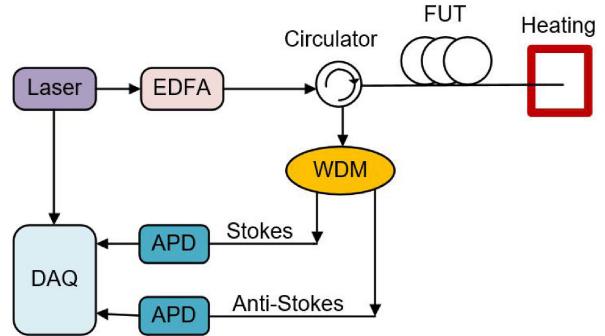


Fig. 6. Schematic diagram of the RDTs.

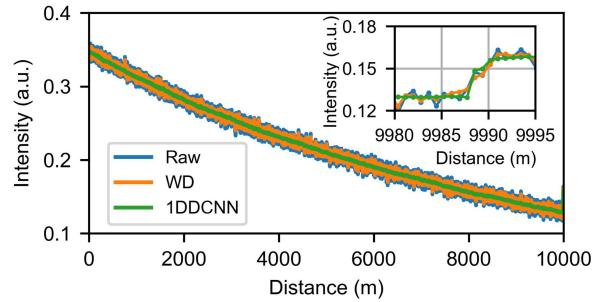


Fig. 7. Anti-Stokes signal processed results when the fiber end is heated to 60 °C.

electrical signals by the APDs and then obtained using a DAQ card at a sampling rate of 125 MSa/s. The DAQ card processes an average of 10000 times in 1 s.

It takes 1DDCNN about 0.2 s to process the acquired data, which means a real time processing compared with the sampling time. The processed anti-Stokes light curve is shown in Fig. 7. The use of both WD and 1DDCNN effectively reduces the noise of the signal. The result processed by the 1DDCNN is not distorted, and the fluctuation is smaller than the result using WD. Spatial resolution of RDTs is defined as the fiber length corresponding to the rising edge of the signal where the temperature changes. The rising edges of the processed data are steep, indicating no deterioration in spatial resolution.

The temperature is demodulated using the processed Stokes and anti-Stokes curves according to (2). As shown in Fig. 8, the temperature profile using signals processed by the 1DDCNN has a smaller fluctuation than using WD. And there is no degradation around the temperature change area. The demodulated temperature of heated fiber section matches the thermometer.

To evaluate the performance of the 1DDCNN more clearly, the absolute difference between two consecutive measurements is calculated. Then the difference is fitted by a quadratic function, which represents the temperature uncertainty, as shown in Fig. 9. Due to transmission loss, the temperature uncertainties increase with fiber length. For the raw data, the temperature uncertainty at the fiber end is 6.4 °C. After 1DDCNN and WD processing, the temperature uncertainty is reduced to 0.7 °C and 3.7 °C, respectively. The 1DDCNN provides better noise removal than WD and a greater improvement in temperature measurement accuracy.

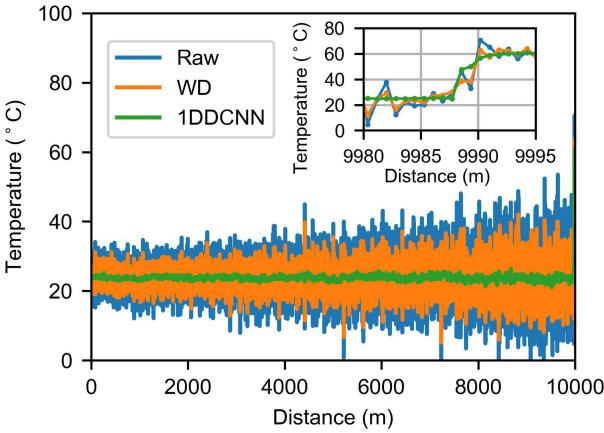


Fig. 8. Demodulated temperature profiles as a function of fiber length.

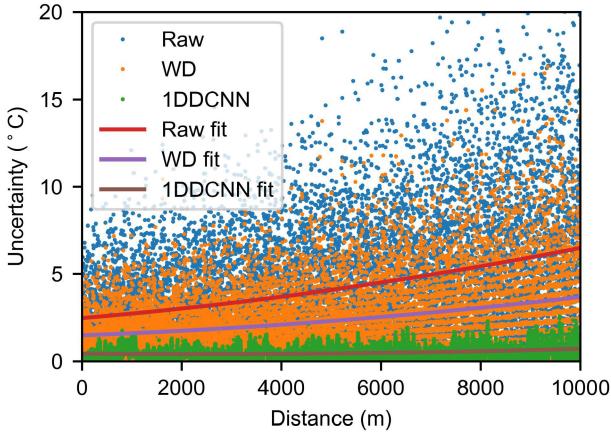


Fig. 9. Temperature uncertainties as a function of fiber length.

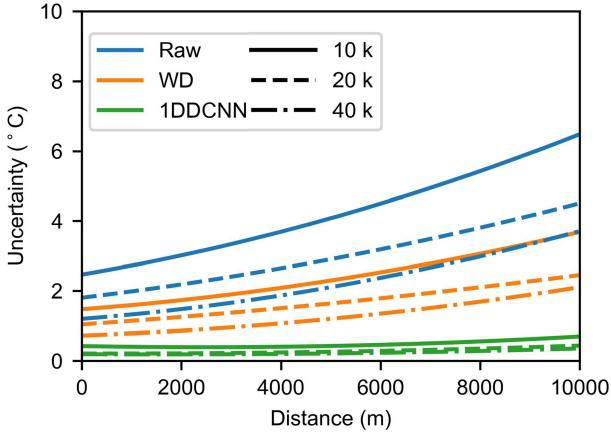


Fig. 10. Temperature uncertainties as a function of fiber length at different average times.

To demonstrate the universality of the 1DDCNN, we experimentally acquire RDTs data with different average times, heating temperatures, pulse widths and sampling rates. The SpRS signals with an average of 20 k and 30 k times are acquired and processed. Fig. 10 shows the calculated temperature uncertainties at different average times. The 1DDCNN shows obvious

TABLE I
TEMPERATURE UNCERTAINTIES AT THE FUT END WITH
DIFFERENT AVERAGE TIMES

Average times	RAW	WD	1DDCNN
10 k	6.4 °C	3.7 °C	0.7 °C
20 k	4.5 °C	2.5 °C	0.5 °C
40 k	3.7 °C	2.1 °C	0.4 °C

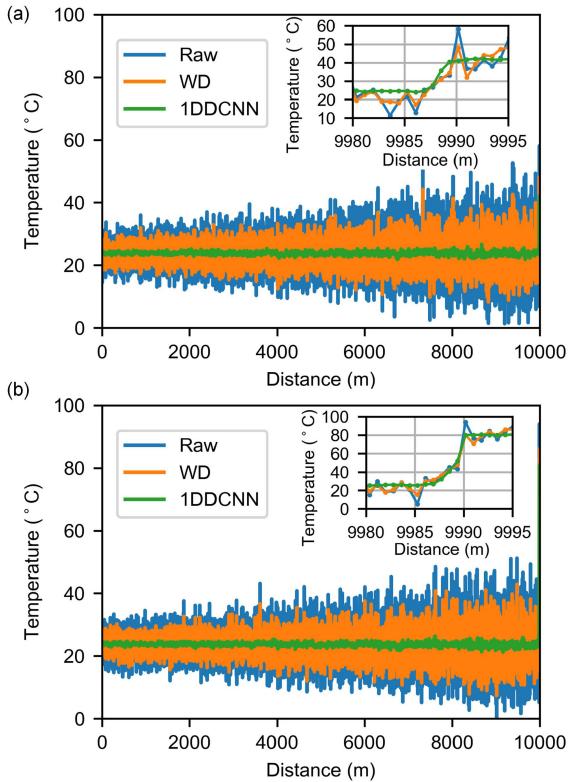


Fig. 11. Demodulated temperature profiles as a function of fiber length when the heating temperature is 40 °C (a) and 80 °C (b).

advantages at various average times, as shown in Table I. This means that 1DDCNN is valid for RDTs data with a relatively high SNR.

In order to study the effect of temperature change on the denoising result, the SpRS signals with heating temperatures of 40 and 60 °C are measured. The demodulated temperature profiles are shown in Fig. 11. For different heating temperatures, the 1DDCNN can maintain the denoising performance and obtain an accurate temperature change. In addition, the rising edge of the temperature change section is very steep, which means that the spatial resolution has not deteriorated.

The pulse width of RDTs systems varies depending on different applications. Fig. 12 shows the temperature demodulation results when the pulse width is 60 ns. The 1DDCNN shows the same denoising performance for data with a pulse width of 30 ns. And the temperature profile has not deteriorated. This is because the synthetic training data does not specify the pulse width, which means that it contains RDTs data with various

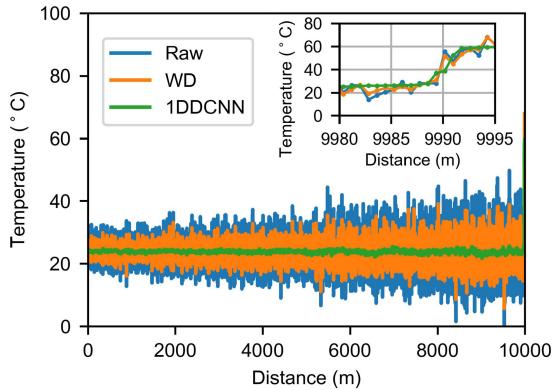


Fig. 12. Demodulated temperature profiles as a function of fiber length when the pulse width is 60 ns.

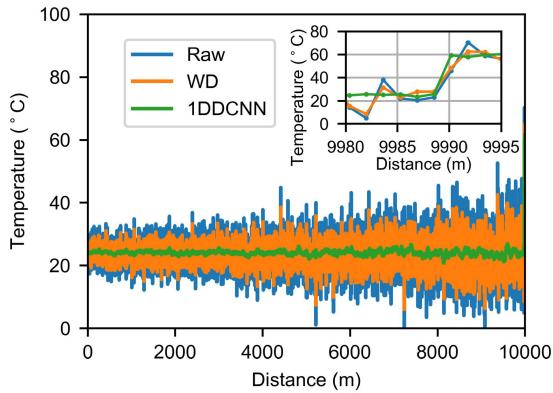


Fig. 13. Demodulated temperatures as a function of fiber length when the sampling rate is 62.5 MSa/s.

pulse widths. As a result, the 1DDCNN can adapt to RDTs data of any pulse width.

For conventional denoising algorithms, the denoising effect is influenced by the sampling rate, and a higher sampling rate is required to achieve a better denoising effect [19]. So, we reduce the sampling rate by half to 62.5 MSa/s to investigate its influence on the 1DDCNN. Due to the reduction of valid data, it is necessary to manually readjust the WD threshold to ensure that the spatial resolution does not deteriorate. The demodulated temperature curves are shown in Fig. 13. The temperature uncertainty at the fiber end of the data processed by WD increases to 4 °C, while the temperature uncertainty is still 0.7 °C when using 1DDCNN. The denoising ability of WD decreases with the sampling rate, resulting in a larger uncertainty. However, the performance of the 1DDCNN is not affected by the sampling rate, which means that using the proposed method can reduce the performance requirements of the sampling device.

IV. CONCLUSION

In this paper, a 1DDCNN is proposed for the performance enhancement of RDTs. And a simplified RDTs model is built to train and optimize the 1DDCNN. The temperature uncertainty of a 10-km SMF with 3 m spatial resolution and 1 s sampling time is reduced from 6.4 °C to 0.7 °C. Compared with conventional denoising algorithms, 1DDCNN provides a better denoising effect

resulting in higher temperature accuracy. Moreover, 1DDCNN is more flexible and reliable in practical applications because it does not require manual parameter adjustment. And its performance is not affected by the changes of average times, temperature, pulse width and sampling rate. The proposed 1DDCNN can be applied to any RDTs system and effectively improve its performance with no hardware modification. And we believe that the proposed method can be applied to other distributed sensing systems, such as optical time-domain reflectometer, distributed acoustic sensing, and distributed Brillouin sensing, by using suitable training data to carry out transfer learning on the 1DDCNN.

REFERENCES

- [1] J. P. Dakin, D. J. Pratt, G. W. Bibby, and J. N. Ross, "Distributed optical fibre raman temperature sensor using a semiconductor light source and detector," *Electron. Lett.*, vol. 21, no. 13, pp. 560–569, 1985.
- [2] A. Datta, H. Mamidala, D. Venkitesh, and B. Srinivasan, "Reference-free real-time power line monitoring using distributed anti-stokes raman thermometry for smart power grids," *IEEE Sens. J.*, vol. 20, no. 13, pp. 7044–7052, Jul. 2020.
- [3] T. Yamate, G. Fujisawa, and T. Ikegami, "Optical sensors for the exploration of oil and gas," *J. Lightw. Technol.*, vol. 35, no. 16, pp. 3538–3545, Aug. 2017.
- [4] D. Inaudi and B. Glisic, "Long-range pipeline monitoring by distributed fiber optic sensing," *J. Press. Vessel Technol.*, vol. 132, no. 1, 2010, Art. no. 011701.
- [5] A. F. Fernandez, P. Rodeghiero, B. Brichard, F. Berghmans, and A. P. Leach, "Radiation-tolerant Raman distributed temperature monitoring system for large nuclear infrastructures," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pp. 2689–2694, Dec. 2005.
- [6] F. Baronti *et al.*, "SNR enhancement of Raman-based long-range distributed temperature sensors using cyclic simplex codes," *Electron. Lett.*, vol. 46, no. 17, pp. 1221–1223, 2010.
- [7] M. Wang, H. Wu, M. Tang, Z. Zhao, and D. Liu, "Few-mode fiber based raman distributed temperature sensing," *Opt. Express*, vol. 25, no. 5, pp. 4907–4916, 2017.
- [8] M. K. Saxena, S. D. V. S. J. Raju, R. Arya, S. V. G. Ravindranath, S. Kher, and S. M. Oak, "Optical fiber distributed temperature sensor using short term fourier transform based simplified signal processing of raman signals," *Measurement*, vol. 47, no. 47, pp. 345–355, 2014.
- [9] M. K. Saxena *et al.*, "Raman optical fiber distributed temperature sensor using wavelet transform based simplified signal processing of raman backscattered signals," *Opt. Laser Technol.*, vol. 65, pp. 14–24, 2015.
- [10] M. A. Soto, J. A. Ramirez, and L. Thevenaz, "Intensifying the response of distributed optical fibre sensors using 2D and 3D image restoration," *Nat. Commun.*, vol. 7, no. 1, pp. 10870–10870, 2016.
- [11] H. Wu, C. Zhao, R. Liao, Y. Chang, and M. Tang, "Performance enhancement of ROTDR using deep convolutional neural networks," in *Proc. 26th Int. Conf. Opt. Fiber Sensors*, Sep. 2018, Paper TuE16.
- [12] J. Li *et al.*, "High-accuracy distributed temperature measurement using difference sensitive-temperature compensation for Raman-based optical fiber sensing," *Opt. Express*, vol. 27, no. 25, pp. 36183–36196, 2019.
- [13] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [14] H. Habibi Aghdam and E. J. Heravi, *Guide to Convolutional Neural Networks*. Berlin, Germany: Springer, 2017.
- [15] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inform. Process. Syst.*, Dec. 2012, pp. 1097–1105.
- [17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, 2015.
- [18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
- [19] H. Wu, L. Wang, Z. Zhao, N. Guo, C. Shu, and C. Lu, "Brillouin optical time domain analyzer sensors assisted by advanced image denoising techniques," *Opt. Express*, vol. 26, no. 5, pp. 5126–5139, 2018.