IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 29, 2019, accepted May 20, 2019, date of publication May 27, 2019, date of current version June 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2919138

Optimized Feedforward Neural Network Training for Efficient Brillouin Frequency Shift Retrieval in Fiber

YONGXIN LIANG[1], JIALIN JIANG[1], YONGXIANG CHEN[1], RICHENG ZHU[1], CHONGYU LU[1], AND ZINAN WANG[1,2], (Senior Member, IEEE)

[1]Key Laboratory of Optical Fiber Sensing and Communications, University of Electronic Science and Technology of China, Chengdu 611731, China
[2]Center for Information Geoscience, University of Electronic Science and Technology of China, Chengdu 611731, China

Corresponding author: Zinan Wang (znwang@uestc.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 41527805 and Grant 61731006, in part by the Sichuan Youth Science and Technology Foundation under Grant 2016JQ0034, in part by the Guofang Keji Chuangxin Tequ, and in part by the 111 project under Grant B14039.

ABSTRACT Artificial neural networks (ANNs) can be used to replace the traditional methods in various fields, making signal processing more efficient and meeting the real-time processing requirements of the Internet of Things (IoT). Recently, as a special type of ANN, the feedforward neural network (FNN) has been used to replace the time-consuming Lorentzian curve fitting (LCF) method in Brillouin optical time-domain analysis (BOTDA) system to retrieve the Brillouin frequency shift (BFS), which could be used as the indicator in temperature/strain sensing and so on. However, the FNN needs to be re-trained if the generalization ability is not satisfactory, or the frequency scanning step is changing in the experiment. This is a cumbersome and inefficient process. In this paper, the FNN only needs to be trained once with the proposed method, and 150.62 km BOTDA is built to verify the performance of the trained FNN. The simulation and experimental results show that the proposed method is promising in BOTDA because of its high computational efficiency and wide adaptability.

INDEX TERMS Feedforward neural networks, Brillouin optical time-domain analysis, Lorentzian curve fitting, sensor fusion, optical fiber sensors.

I. INTRODUCTION

The Internet of Things (IoT) is constantly progressing as multiple technologies are evolving, especially the sensing technologies [1]. As an important branch of sensors, optical fiber sensors, have been studied by many researchers due to their unique advantages [2]. Particularly, distributed optical fiber sensing (DOFS) systems are of great interests since they can turn fiber cables into massive sensor arrays [3]–[6]. Brillouin optical time-domain analysis (BOTDA) system is an important type of DOFS, which could achieve high precision, long distance, and fast scan-rate sensing [7]–[9].

The stimulated Brillouin scattering (SBS) effect is the basis of BOTDA [10], [11]. In BOTDA, usually pump pulse and continuous-wave (CW) probe light are counter-propagating inside the fiber to sample the Brillouin gain spectrum (BGS), then the Brillouin frequency shift (BFS) is retrieved for the purpose of temperature/strain sensing [12], [13]. Before the traditional Lorentzian curve fitting (LCF) method is used to find the BFS [14], the non-local means (NLM), wavelet denoising (WD) or block-matching and 3D filtering (BM3D) can be used to reduce the noise of the BGS in general [15]. All of those processes are time-consuming operations, especially for longer sensing distance and finer spatial resolution. Recently, denoising convolutional neural network (DnCNN) is used for BOTDA filtering [16], which can achieve real-time filtering as long as the DnCNN is trained properly. However, the training process of DnCNN is complicated.

On the other hand, artificial neural networks (ANNs) have been applied in BOTDA [17]–[19]. As an important type of ANN, the feedforward neural network (FNN) can be used to replace the traditional LCF method and improve the processing speed. Ideal BGSs are used for FNN training [17]. During the FNN testing phase, the noise in the measured BGS needs to be reduced as much as possible,

The associate editor coordinating the review of this manuscript and approving it for publication was Rui-Jun Yan.

which means that the filtering operation cannot be omitted. However, using traditional filtering methods will not meet the requirements of real-time analysis. Although using DnCNN can achieve fast filtering, the training process is complicated. Without filtering, it means that FNN trained by ideal BGSs may not has generalization ability for the noisy BGSs in the experiment. Regularization method is needed in order to enhance the generalization ability of FNN [20]–[22]. This is a time-consuming task.

Also, the frequency scanning step is one of the key parameters in BOTDA operation. It usually varies between 1 MHz and 10 MHz depending on the required accuracy and speed of the measurement. Previous work required training different FNNs to meet the needs of different frequency scanning steps [17], which is a tedious task.

In this paper, an optimized FNN training method is proposed for BOTDA. As a result, BFS retrieval from the BGS can be efficiently carried out with a wide range of BGS linewidths, signal-to-noise ratios (SNRs) and frequency scanning steps. The results show that the accuracy of FNN is similar to LCF, while FNN is much faster than LCF. The 23.95 km and 150.62 km BOTDA at the frequency scanning step of 1 MHz and 4 MHz are established respectively to experimentally verify the performance of FNN.

This paper is organized as follows. Section II describes the principles of LCF and FNN to get BFS from BGS. Section III describes the FNN training and test phase. Section IV explains how to enhance the adaptability of the trained FNN. Section V presents the experiments and analyses the results. Section VI gives the conclusion.

II. THE PRINCIPLES OF LCF AND FNN

The principles for BFS retrieval from the BGS using LCF and FNN are explained respectively in this section.

A. THE PRINCIPLE OF LCF

In BOTDA, the obtained BGSs conform to the shape of the Lorentzian curve. Equation (1) is the core of BGSs simulation [10].

$$g(v, z) = \frac{g_B(z)}{1 + [2(v - v_B)/\Delta v_B]^2}. \tag{1}$$

where Δv_B is the BGS linewidth. z is the position of the fiber. $g_B(z)$ is the gain coefficient. v is the scanning frequency. In the experiments, the scanning frequencies v are discrete values, and the frequency scanning step varies between 1 MHz to 10 MHz usually. $g(v, z)$ is the measured BGS. v_B is BFS which reflects information about temperature or strain. If the fiber is placed in a stable environment, the change of BFS is only caused by temperature theoretically [23]:

$$v_{B2} - v_{B1} = C_T(T_2 - T_1). \tag{2}$$

where C_T is the temperature coefficient. Fig. 1 shows the relationship between temperature and BFS. In order to obtain BFS, the least squares estimation (LSE) method is used in LCF.

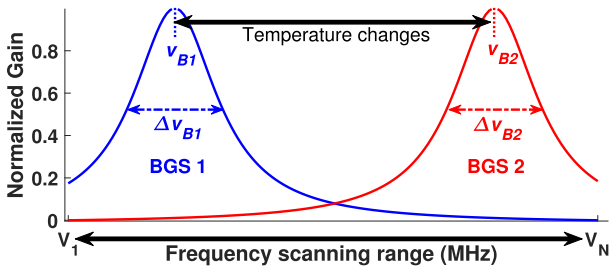

FIGURE 1. BFS changes with temperature.

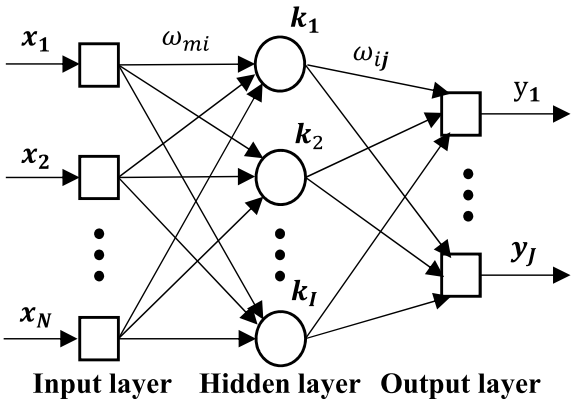

FIGURE 2. Typical FNN with three layers.

$$R^2 = \sum_{i=1}^{N} [g(v_i, z) - y_i]^2. \tag{3}$$

The purpose of (3) is to minimize the R^2 by modifying parameters during iterations. At the end of the iteration, y is the result of LCF.

B. THE PRINCIPLE OF FNN

A typical FNN is shown in Fig. 2. The number of the input layer, hidden layer, and output layer are N, I and J, respectively. The input and output of the hidden layer k_i are u_i and v_i, respectively. The weights from x_m to k_i and from k_i to y_j are ω_{mi} and ω_{ij}, respectively. The FNN accepts a vector of length N as the input data, and finally generates a vector of length J as the output data.

The input of the hidden layer k_i in the n^{th} iteration is:

$$u_i(n) = \sum_{m=1}^{N} \omega_{mi}(n) x_m(n). \tag{4}$$

The output of the hidden layer k_i is:

$$v_i(n) = f(u_i(n)). \tag{5}$$

where $f(\cdot)$ is the sigmoid function. The output of the network is:

$$y_j(n) = \sum_{i=1}^{I} \omega_{ij}(n) v_i(n). \tag{6}$$

$$Y(n) = [y_1(n), y_2(n), \ldots, y_J(n)]. \tag{7}$$

The expected output of the network is:

$$H(n) = [h_1(n), h_2(n), \ldots, h_J(n)]. \tag{8}$$

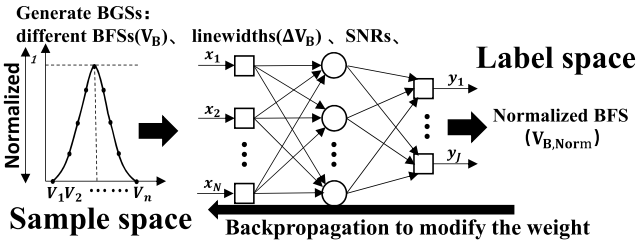

FIGURE 3. FNN training process.

The error signal in the n^{th} iteration is defined as:

$$e_j(n) = h_j(n) - y_j(n). \tag{9}$$

In order to minimize the errors, the basic method is to use the gradient descent algorithm for backpropagation (BP) calculations, which modifies the weights according to (10) and (11).

$$\Delta\omega_{ij}(n) = -\eta\frac{\partial e(n)}{\partial\omega_{ij}}. \tag{10}$$

$$\omega_{ij}(n+1) = \Delta\omega_{ij}(n) + \omega_{ij}(n). \tag{11}$$

where $e(n)$ is the cost function.

The input vector is the BGS, and the output vector is BFS of the BGS. It should be noted that both the input BGS and the output BFS need to be normalized. As the sensing distance increases, the gain of the BGS decreases, and the normalization eliminates the effects of different gains. At the same time, when the output BFS is normalized, the actual BFS of different parameters in the experiment can be conveniently calculated. The training process of FNN is shown in Fig. 3. In the sample space, the features of the BGS include BFS, BGS linewidth and SNR. It should be noted that the BFS changes with the temperature. The BGS linewidth will be broadened due to the shorter pulse width or the other nonlinear effects, and it can be narrowed by proper synthesis of loss/gain spectrum. The SNR of BGS is mainly determined by the average times in the experiments.

The SNR of the simulated BGS [24] is calculated according to (12). Due to the normalization, g_B is equal to 1.

$$SNR = {g_B}^2/{\sigma_n}^2 = 1/{\sigma_n}^2. \tag{12}$$

where ${\sigma_n}^2$ is the noise variance. The SNR of the measured BGS [24] is calculated by the (13) and (14).

$$SNR = {g_B}^2/E({N_r}^2). \tag{13}$$

$$N_r = y_F - y_M. \tag{14}$$

where y_M is the measured BGS, and y_F is the fitted curve. N_r is a vector of residual which can be calculated by subtracting the measured BGS from the fitted curve. E(Âů) is the arithmetic mean. g_B is the gain factor of the fitted curve.

This section provides the principles for BFS retrieval by these two methods. As a traditional method, LCF has a guaranteed accuracy but a time-consuming process, while FNN has the ability to calculate quickly but it requires verification of accuracy.

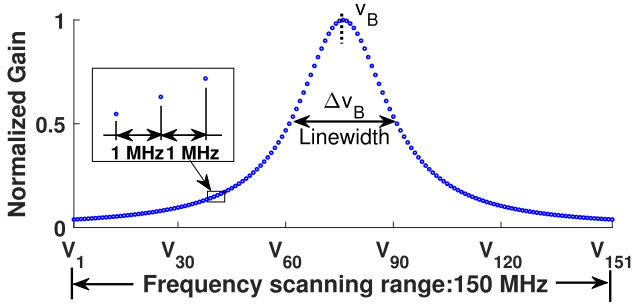

FIGURE 4. A Simulated BGS with the frequency scanning range of 150 MHz at a frequency scanning step of 1 MHz.

III. OPTIMIZED FNN TRAINING: SIMULATION RESULTS

Training an FNN with good generalization ability is not an easy task. The optimized training process includes two aspects. One is adding noise to ideal BGS in order to improve the generalization ability of FNN [25]. The other is using the symmetry of BGS in order to reduce the amount of training data to improve training efficiency. It will be described in this section.

In BOTDA, the measured BGS needs to be scanned over a wide frequency range, which is typically greater than 150 MHz, in order to correctly locate the BFS and achieve wide temperature monitoring range. As for the frequency scanning step, generally 1 MHz is chosen, which will change according to the actual needs in the experiment. Here, the frequency scanning range and frequency scanning step of the simulated BGSs are set to 150 MHz and 1 MHz, respectively. Therefore, each simulated BGS forms a vector of length 151 as a sample, and the label of each sample is its BFS, as shown in Fig. 4. In the simulation, the starting frequency is V_1 =10.7 GHz, and the ending frequency is V_{151} =10.85 GHz. If the BFSs of two Lorentzian curves satisfies the following symmetry condition:

$$v_{B_2} + v_{B_1} = (V_{151} + V_1) \equiv 2 \times V_{76}. \tag{15}$$

Bring (15) to (1), we can get the relationship between two Lorentzian curves that are symmetric with respect to the line where V_{76} is:

$$g_2(v, z) = g_1(-v + 2V_{76}, z). \tag{16}$$

It means that only half of the data is needed to represent the entire sample space.

A. FNN TRAINING PHASE

In [17], the features of the simulated BGSs only contain BFS and linewidth. However, the addition of random noise to training data can enhance the generalization ability of FNN. Therefore, in this work, the features of BGSs include BFS, linewidth, and SNR. As shown in Table 1, Datasets 1 and 2 are created for FNN training, and Dataset 3 is used as the validation-dataset in training phase. Dataset 4 is used as test-dataset after FNN training. For the samples in Dataset 1, the BFSs range from 10.7075 GHz to 10.7825 GHz at

TABLE 1. The simulated four datasets.

Dataset 1 for training

Features	Range	Step	Number of copies
Samples (Num: 76×51×3×20=232,560):			
v_B	10.7075 GHz - 10.7825 GHz	1 MHz	1
Δv_B	10 MHz - 60MHz	1 MHz	1
SNR	15 dB - 35 dB	10 dB	20
Labels (Num:76×3,060=232,560):			
$v_{B,Norm}$	5% - 55%	1/150	51×3×20 =3060

Dataset 2 for training

Features	Range	Step	Number of copies
Samples (Num: 76×51=3,876):			
v_B	10.7075 GHz - 10.7825 GHz	1 MHz	1
Δv_B	10 MHz - 60 MHz	1 MHz	1
SNR	/	/	/
Labels (Num: 76×51=3,876):			
$v_{B,Norm}$	5% - 55%	1/150	51

Dataset 3 for validation

Features	Range	Step	Number of copies
Samples (Num: 76×51×21=81,396):			
v_B	10.7075 GHz - 10.7825 GHz	1 MHz	1
Δv_B	10 MHz - 60 MHz	1 MHz	1
SNR	15 dB - 35 dB	1 dB	1
Labels (Num: 76×1071=81,396):			
$v_{B,Norm}$	5% - 55%	1/150	51×21 =1,071

Dataset 4 for test

Features	Range	Step	Number of copies
Samples (Num: 136×51×31×10=2,150,160):			
v_B	10.7075 GHz-10.8425 GHz	1 MHz	1
Δv_B	10 MHz - 60MHz	1 MHz	1
SNR	10 dB - 40 dB	1 dB	10
Labels (Num: 136×15,810=2,150,160):			
$v_{B,Norm}$	5% - 95%	1/150	51×31×10 =15,810

a step of 1 MHz. The linewidths range from 10 MHz to 60 MHz at a step of 1 MHz, and the SNRs range from 15 dB to 35 dB at a step of 10 dB. Meanwhile, since the added noise is random, 20 noisy copies of each sample are created based on their SNR. The number of samples in Dataset 1 is 76×51×3×20=232,560. As for the labels in Dataset 1, the BFSs can be normalized according to (17):

$$v_{B,Norm} = (v_B - V_1)/(V_{151} - V_1). \tag{17}$$

where V_1 and V_{151} are the starting frequency and the ending frequency of the BGS, respectively. $v_{B_{B,Norm}}$ is the normalized BFS, which is the label of each sample. After normalization, normalized BFSs range becomes 5% to 55%. There is no need to set the normalized BFSs range to 5% - 95%, because according to (15) and (16), the sample space can be fully expressed in half range of the data. Three samples and its labels are selected for observation from Dataset 1 and shown in Fig. 5.

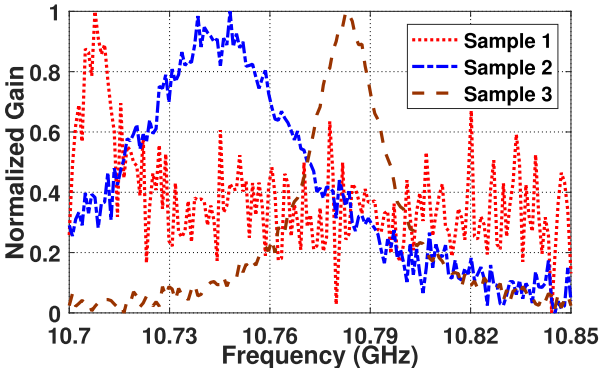

FIGURE 5. The simulated BGSs with different BFSs, linewidths and SNRs. (Sample 1: $v_B = 10.7075$ GHz $\Delta v_B = 10$ MHz, SNR=15 dB, Label 1: $v_{B,Norm} = 5\%$. Sample 2: $v_B = 10.7450$ GHz, $\Delta v_B = 60$ MHz, SNR=25 dB, Label 2: $v_{B,Norm} = 30\%$. Sample 3: $v_B = 10.7825$ GHz, $\Delta v_B = 35$ MHz, SNR=35 dB), Label 3: $v_{B,Norm} = 55\%$.

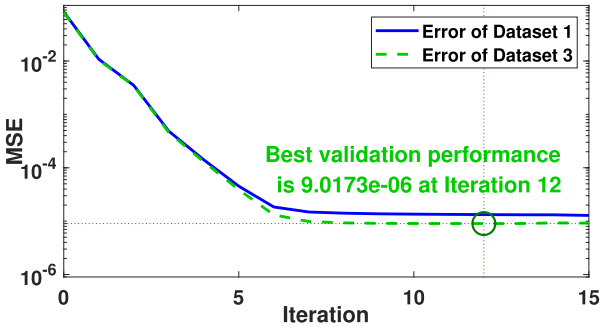

FIGURE 6. The mean squared error (MSE) of each iteration of the FNN trained by the noisy BGSs.

The dimensions of each sample and label are 151 and 1, respectively. According to the number of training data, and the number of input and output nodes, the layout of the hidden layers is set to 40-15 after several attempts. Therefore, the layout of FNNs is set to 151-40-15-1. Early stopping happens when validation performance has increased more than 3 times since the last time it decreased. Weights and biases are initialed by the Nguyen-Widrow algorithm. Using the Levenberg-Marquardt (LM) training algorithm to solve the problem of slow convergence of the gradient descent algorithm, it takes about 3 hours in 15 iterations for FNN training. The errors of the Dataset 1 and Dataset 3 have the same order of magnitude at Iteration 15, indicating the trained FNN can accurately retrieve the BFSs from the samples in Dataset 3, as shown in Fig. 6.

For comparison, Dataset 2 without noisy samples is used as the training-dataset. Other training parameters are consistent. The errors of the Dataset 2 and 3 have different orders of magnitude at Iteration 7, indicating that the trained FNN cannot retrieve the BFSs from the samples in Dataset 3 accurately, as shown in Fig. 7. The error of the Dataset 2 is driven to a small value while the error of the Dataset 3 still at a high level, which means the overfitting occurs.

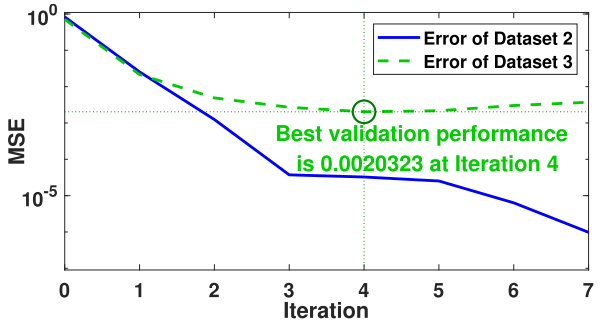

FIGURE 7. The MSE of each iteration of the FNN trained by the ideal BGSs.

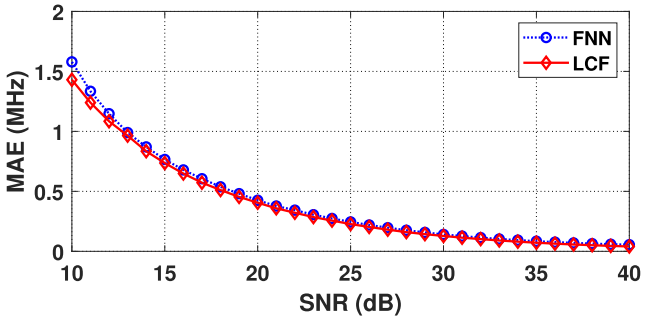

FIGURE 8. The MAE versus SNR.

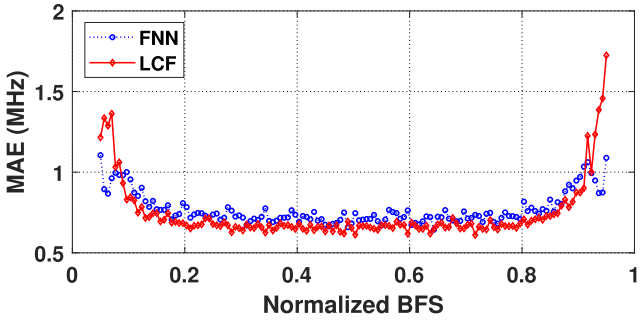

FIGURE 9. The MAE versus normalized BFS.

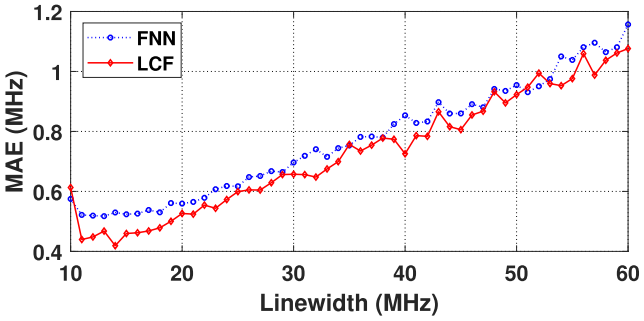

FIGURE 10. The MAE versus linewidth.

B. FNN TEST PHASE

Dataset 4 is used to demonstrate FNN and LCF ability to retrieve the BFS after FNN training. Since the labels in Dataset 4 range from 5% to 95%, the BFS can be retrieved by FNN using (18):

$$v_{B(F)} = \begin{cases} F(X) & \times F_{range} + V_1,\ x_{Pmax} < V_{76} \\ (1 - F(X')) & \times F_{range} + V_1,\ x_{Pmax} \geq V_{76}. \end{cases} \quad (18)$$

where $v_{B(F)}$ is BFS retrieved by FNN. X is a normalized BGS. X' is the X flipped in the up-down direction. $F(X)$ is the output of the trained FNN. F_{range} is the frequency scanning range and its value is equal to 150 MHz. x_{Pmax} indicates the location of the maximum power of the X. The mean absolute error (MAE) can be calculated from the retrieved BFS and the label of the samples according to (19).

$$\text{MAE} = E(|v_B - v_{B(F/L)}|). \quad (19)$$

where v_B is the BFS of the sample. $v_{B(F/L)}$ is the BFS predicted by FNN or LCF. The distribution of MAE calculated by FNN and LCF of the samples in Dataset 4 at different SNRs is shown in Fig. 8. In this range of SNR, the accuracy of the FNN can be compared to the LCF.

The error distribution for all 15 dB samples in Dataset 4 is analyzed under different BFSs and linewidths, as shown in Fig. 9 and 10. The errors of FNN are better than LCF in the edge regions. This makes sense for increasing the effective range of temperature recognition in a limited frequency scanning range. Moreover, for both FNN and LCF, the error increases as the BGS linewidth increases, which is the same as the theoretical conclusion in [26].

The trained FNN can be used to calculate BFS accurately from BGSs with different BFSs, linewidths and SNRs. Through this optimized training process, the trained FNN has generalization ability and greatly improves the training efficiency without additional regularization methods.

IV. THE FNN ADAPTABILITY: SIMULATION RESULTS

The FNN trained in the previous section will not be available if the frequency scanning step is not 1 MHz in the experiment, which means that samples at the corresponding frequency scanning step need to be simulated again and used as the training-dataset to train another FNN, as shown in Fig. 11(a). In long-distance BOTDA, it is more desirable to use a higher frequency scanning step, such as 4 MHz. This will improve the speed of measuring BGS at the expense of the accuracy of the temperature measurement. Correspondingly, a lower frequency scanning step, such as 1 MHz, is selected under high precision requirements.

However, the FNN training is a cumbersome process that involves generating simulated data, adjusting hidden layers and other parameters until the FNN has generalization ability and satisfactory error results. Therefore, it will be convenient if only one FNN needs to be trained, which can retrieve BFS from BGSs at different frequency scanning steps. In order to achieve this process, the linear interpolation method can be used to reconstruct the other frequency scanning step of the BGS to 1 MHz.

The frequency scanning range should not be lower than 150 MHz because, in the previous section, the FNN is trained with the samples in the frequency scanning range of 150 MHz

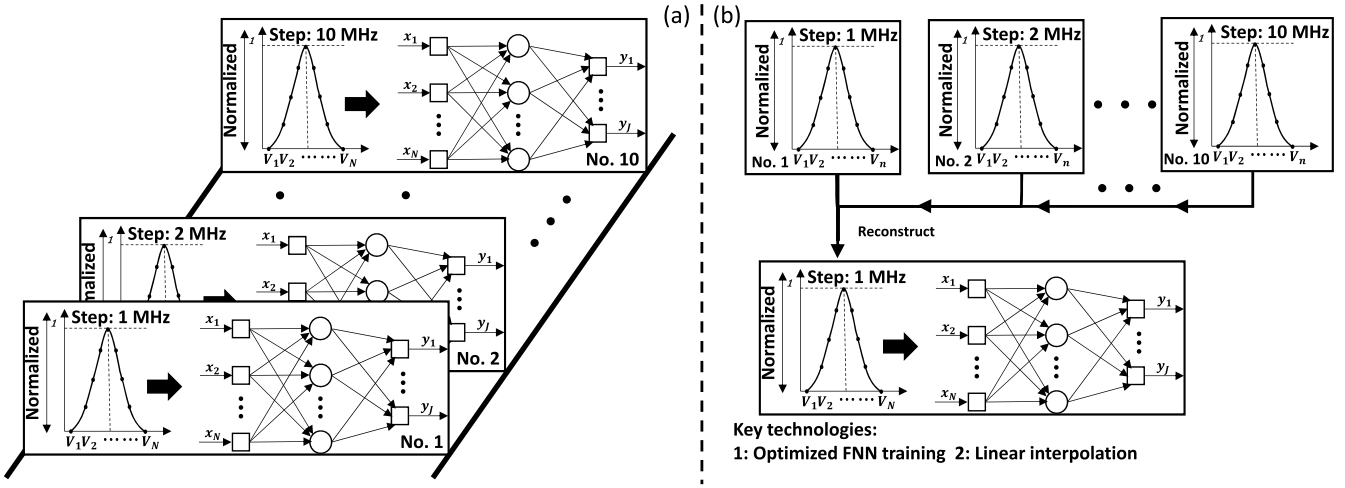

FIGURE 11. (a) Previous work needs to train separate FNNs at different frequency scanning steps; (b) combining the two technologies, only one FNN needs to be trained.

TABLE 2. Relationship between various parameters.

$Step$ / MHz	N	F_{range} / MHz	Layouts of separate FNNs
1	151	150	151-40-15-1
2	76	150	76-30-10-1
3	51	150	51-25-10-1
4	39	152	39-20-10-1
5	31	150	31-15-10-1
6	26	150	26-10-5-1
7	23	154	23-10-5-1
8	20	152	20-10-5-1
9	19	162	19-10-5-1
10	16	150	16-10-5-1

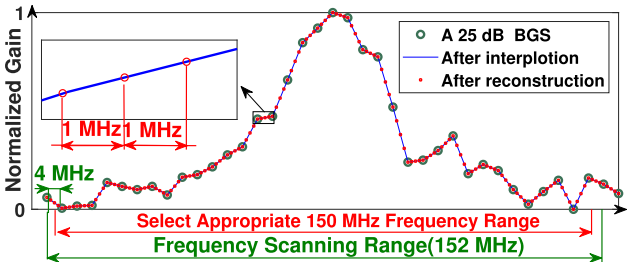

FIGURE 12. The process of reconstruction for a simulated 25 dB BGS at a step of 4 MHz.

at a step of 1 MHz. This can be expressed in (20) and itemized in Table 2.

$$F_{range} = (N - 1) \times Step \geq 150MHz. \tag{20}$$

where F_{range} is the frequency scanning range. N is the number of points of the BGS. $Step$ is the frequency scanning step.

As shown in Fig. 12, if the frequency scanning step is 4 MHz, the frequency scanning range should at least 152 MHz. Only in this way, after using linear interpolation, a BGS with a frequency scanning range of 150 MHz at a step of 1 MHz can be reconstructed. In short, discrete data are serialized by linear interpolation. Then reconstruct a vector of length 151 from the continuous function, and the BFS can be retrieved according to (18). After these processes, only need to train one FNN to handle BGSs with different frequency scanning steps, as shown in Fig. 11(b).

The BFS is retrieved by three methods from the 10 dB, 15 dB, 25 dB and 40 dB BGSs with different BFSs (10.7075 GHz - 10.8425 GHz) and linewidths (10 MHz - 60 MHz) at different frequency scanning steps (1 MHz - 10 MHz), and the error results are shown in Fig. 13(a). Method 1 is using the proposed method shown in Fig. 11(b). In order to verify the performance, Method 2 is using the separately trained FNNs shown in Fig. 11(a). The layouts of separately trained FNNs are shown in Table 2. Others training parameters are consistent. Method 3 is using LCF. Three methods are defined as FNN, FNNs and LCF in Fig. 13(a), respectively. The errors of FNN and FNNs are basically the same for the high SNR signals. For the BGS with 10 dB, the performance of the FNN is worse than that of the separate FNNs with the increase of the frequency scanning step. However, due to the large error in this case, the sensing information is lost. In conclusion, the sensing performance of Method 1 is similar to that of Method 2 for the effective sensing signals. Three different interpolation methods are compared. i.e.: 1. Linear interpolation: linear; 2. Piecewise Cubic Hermite Interpolating Polynomial (PCHIP): pchip; 3. Spline interpolation: spline. Their continuities are zero-order (C^0), first-order (C^1) and second-order (C^2), respectively. The results are shown in Fig. 13(b). The high-efficiency linear interpolation method has better performance in this situation.

In order to compare efficiency, the BFS is retrieved by FNN and LCF from 100,000 sets of the BGS at each frequency scanning step, as shown in Table 3. The consumption time for LCF is the Lorentzian curve fitting time and BFS retrieval time, while for FNN is interpolation time, reconstruction time and BFS retrieval time.

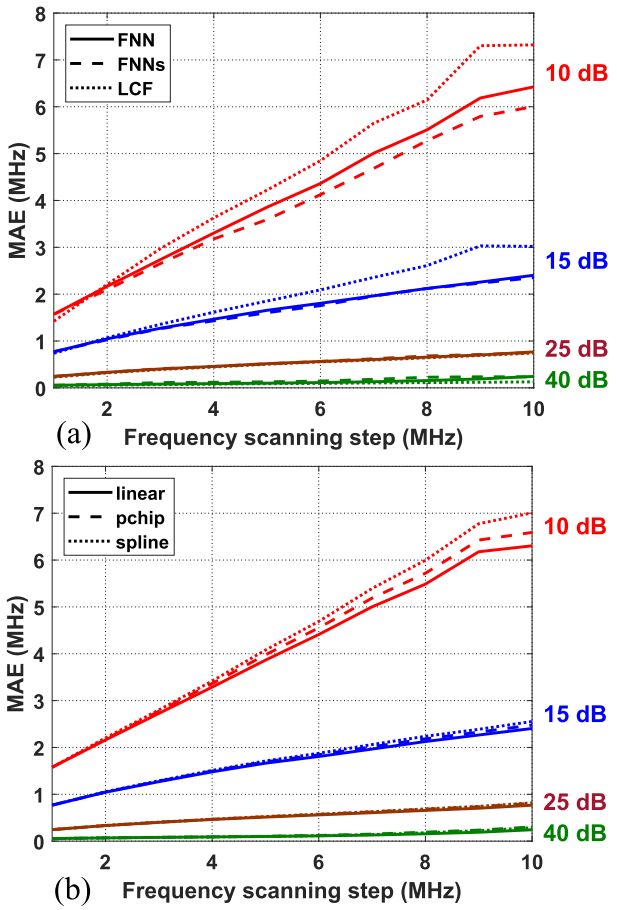

FIGURE 13. The MAE versus frequency scanning step. (a) Three methods for BFS retrieval; (b) three different interpolations for Method 1.

TABLE 3. Comparison of efficiency between FNN and LCF for 100,000 BGS samples.

Step / MHz	FNN	LCF	
		Single thread	16 threads
1	1.20 s	17.88 min	155.91 s
2	1.26 s	30.50 min	173.81 s
3	1.19 s	19.39 min	112.94 s
4	1.20 s	18.86 min	187.48 s
5	1.19 s	15.17 min	142.65 s
6	1.18 s	16.62 min	187.55 s
7	1.20 s	17.88 min	343.23 s
8	1.20 s	19.43 min	146.70 s
9	1.19 s	16.12 min	170.80 s
10	1.18 s	16.52 min	91.68 s

The interpolation method enables BGSs at different frequency scanning steps to be input into a trained FNN and then obtain the BFS accurately. The high efficiency and wide adaptability of FNN will promote the development of BOTDA. Although linear interpolation is not smooth at the nodes, from another perspective, it can be considered as noise-induced. This disadvantage can be overcome well because the FNN is trained by Dataset 1 with noise samples. Therefore, by combining the two methods of optimized FNN training and linear interpolation, the more effective application of FNN in BOTDA is realized.

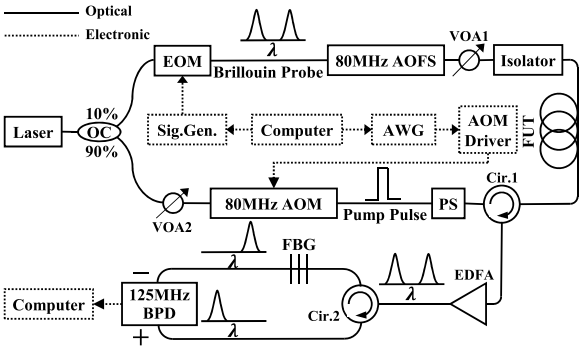

FIGURE 14. Experimental setup of 23.95 km BOTDA. AWG: Arbitrary waveform generator; OC: optical coupler.

V. EXPERIMENTAL DEMONSTRATION

In order to analyze the ability of the FNN for calculating the BFS from the BGSs in the experiment, the 23.95 km BOTDA with 200 MHz frequency scanning range at a frequency scanning step of 1 MHz and 150.62 km BOTDA with 156 MHz frequency scanning range at a frequency scanning step of 4 MHz are established, respectively.

A. EXPERIMENTAL SETUP OF THE 23.95 km BOTDA

The experimental setup of the 23.95 km BOTDA is shown in Fig. 14. This is a dual-sideband BOTDA experimental system.

The 10% branch of the laser output is modulated by a Mach-Zehnder electro-optic modulator (EOM). With a tunable microwave generator, the carrier of the probe signal is suppressed and two sidebands are generated. Through an 80 MHz acoustic-optic frequency shifter (AOFS), a variable optical attenuator (VOA1) and an isolator, the probe light is routed to the fiber. The 90% branch of the laser output is modulated by an 80 MHz acousto-optic modulator (AOM) to generate the pump pulse. The symmetrical dual-sideband of the probe wave relative to the pump pulse is formed since the frequency shift of AOFS is the same as the AOM. Then after a polarization scrambler (PS), the influence of polarization fading is reduced. The pump pulse and probe light are counter-propagating inside the fiber under test (FUT) to experience SBS. After that, the probe light is routed to the low-noise erbium-doped fiber amplifier (EDFA) for amplifying. The fiber Bragg grating (FBG) is used to separate the dual-band of the probe light. The 125 MHz balanced photodetector (BPD) has two input ports. One receives the Stokes lightwave, and the other receives the anti-Stokes lightwave. The photocurrent is sampled by a 100 MS/s A/D converter. The BGS is measured after averaging 300 times in the frequency scanning range of 200 MHz, with the frequency scanning step of 1 MHz. The spatial resolution is 16 m.

The heating location is around 23.7 km. The SNR is about 30 dB calculated from (13). As can be seen from Fig. 15, the position of the BFS changes due to heating.

The input of the FNN is the BGS in the frequency scanning range of 150 MHz at a frequency scanning step of 1 MHz.

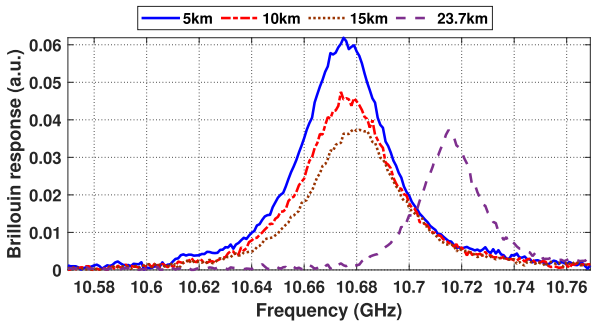

FIGURE 15. BGSs at different locations of 23.95 km BOTDA.

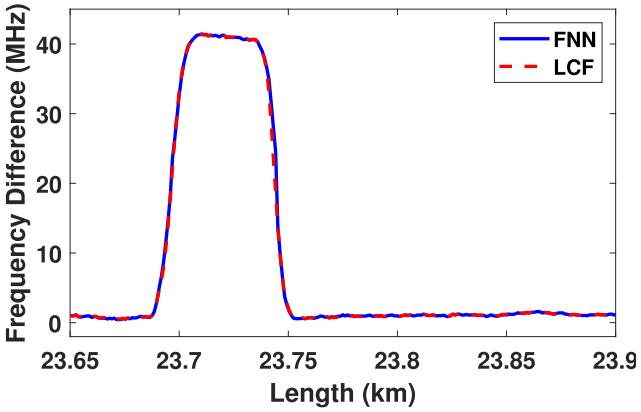

FIGURE 16. The frequency difference around the heating location in the 23.95 km BOTDA.

However, the frequency scanning range of the measured BGS is 200 MHz at a frequency scanning step of 1 MHz. Therefore, the BGS with appropriate 150 MHz frequency scanning range is selected according to the position of the power peak and input to the FNN to retrieve the BFS. Also, the BFS is calculated by LCF for comparison. Subtracting BFS before and after heating to obtain the frequency difference, which reflects the change of temperature, as shown in Fig. 16.

The temperature coefficient is 1.3 MHz/°C, and the applied temperature difference is 31.9°C. The differences in temperature measured by LCF and FNN are 31.60°C and 31.64°C, respectively. The measurement uncertainties calculated by LCF and FNN are ±0.25°C and ±0.26°C, respectively. The MAE of the frequency difference of BFS calculated by FNN and LCF is analyzed every 1 km according to (21).

$$\mathrm{MAE} = E(|FD_L - FD_F|). \tag{21}$$

where FD_L and FD_F is the frequency difference calculated by LCF and FNN, respectively. The error is shown in Fig. 17. Compared to LCF, it indicates that FNN can calculate BFS from the measured BGS accurately.

B. EXPERIMENTAL SETUP OF THE 150.62 km BOTDA

To further analyze the performance of the FNN, an advanced BOTDA of 150.62 km is built. The spatial resolution is 9 m, which means more than sixteen thousand sensing units are fused along the fiber to sense the change of temperature.

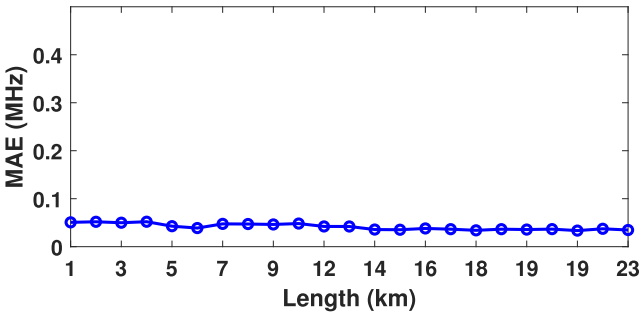

FIGURE 17. The MAE (every 1km) of the calculated BFS between FNN and LCF for the 23.95 km BOTDA.

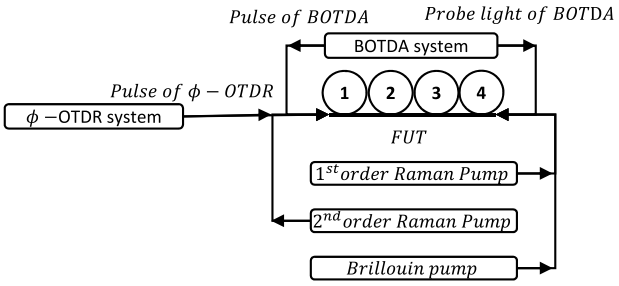

FIGURE 18. Experimental setup of 150.62 km BOTDA.

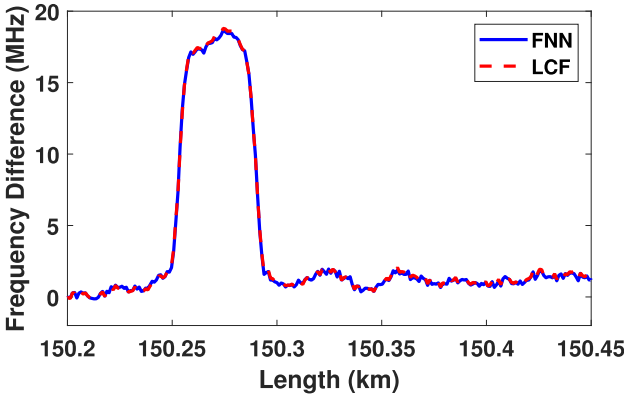

FIGURE 19. The frequency difference around the heating location in the 150.62 km BOTDA.

A variety of technologies, such as hybrid distributed amplification, frequency division multiplexing (FDM), wavelength division multiplexing (WDM), time division multiplexing (TDM), and 127-bit Simplex coding are combined and used in the experiment [7]. The experimental setup is shown in Fig. 18.

Since Simplex coding is used in 150.62 km BOTDA, measuring BGS is time-consuming especially in a wider frequency scanning range and lower frequency scanning step. Therefore, 156 MHz frequency scanning range at a step of 4 MHz is set to sample the BGS in 150.62 km BOTDA after averaging 16 times. The SNR of the BGS of the four-segment fiber is analyzed. The first three segments have the same SNR, about 30 dB, and the last segment is about 19 dB.

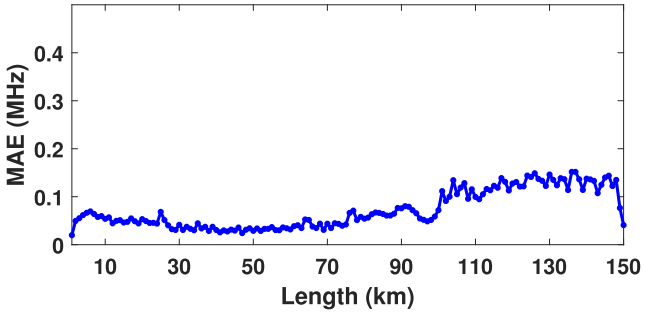

FIGURE 20. The MAE (every 1km) of the calculated BFS between FNN and LCF in the 150.62 km BOTDA.

Since the frequency scanning step of the measured BGS is 4 MHz, linear interpolation is required to reconstruct the BGS with appropriate 150 MHz frequency scanning range before it is input to the FNN. The frequency difference of BFS calculated by FNN is compared with LCF, as shown in Fig. 19. The temperature coefficient is 1.0 MHz/°C, and the applied temperature difference is 18.2°C. The differences in temperature measured by LCF and FNN are 18.67°C and 18.50 °C, respectively. The measurement uncertainties calculated by LCF and FNN are ±0.82°C and ±0.75°C, respectively. The measurement uncertainties are higher than the 23.95 km BOTDA at a frequency scanning step of 1 MHz.

Similarly, the MAE of the frequency difference is analyzed every 1 km according to (21), as shown in Fig. 20. It is acceptable that the MAE is less than 0.2 MHz compared with the measurement uncertainty.

By constructing and analyzing BOTDA experiments, it confirms that FNN can retrieve BFS accurately under different experimental parameters.

VI. CONCLUSION

Using the proposed FNN training method, BFS can be retrieved from the BGSs under various experimental parameters. Compared with LCF, FNN trained with the proposed method is much more efficient, without sacrificing the measurement uncertainty. This work proves the value of FNN in DOFS system and promotes the development of DOFS for real-world applications.

REFERENCES

[1] E. Hodo, X. Bellekens, A. Hamilton, P.-L. Dubouilh, E. Iorkyase, C. Tachtatzis, and R. Atkinson, "Threat analysis of IoT networks using artificial neural network intrusion detection system," in *Proc. Int. Symp. Netw., Comput. Commun. (ISNCC)*, Hammamet, Tunisia, May 2016, pp. 1–6.

[2] D. A. Krohn, T. W. MacDougall, and A. Mendez, *Fiber Optic Sensors: Fundamentals and Applications*, 4th ed. Bellingham, WA, USA: SPIE, 2014.

[3] P. Ferdinand, "The evolution of optical fiber sensors technologies during the 35 last years and their applications in structure health monitoring," in *Proc. EWSHM-7th Eur. Workshop Struct. Health Monitor.*, Jul. 2014, pp. 1–17. [Online]. Available: https://hal.inria.fr/hal-01021251

[4] Z. Wang, B. Zhang, J. Xiong, Y. Fu, S. Lin, J. Jiang, Y. Chen, Y. Wu, Q. Meng, and Y. Rao, "Distributed acoustic sensing based on pulse-coding phase-sensitive OTDR," *IEEE Internet Things J.*, to be published. doi: 10.1109/JIOT.2018.2869474.

[5] Z. Wang, L. Zhang, S. Wang, N. Xue, F. Peng, M. Fan, W. Sun, X. Qian, J. Rao, and Y. Rao, "Coherent Φ-OTDR based on I/Q demodulation and homodyne detection," *Opt. Express*, vol. 24, no. 2, pp. 853–858, Jan. 2016.

[6] S. Lin, Z. Wang, J. Xiong, Y. Fu, J. Jiang, Y. Wu, Y. Chen, C. Lu, and Y. Rao, "Rayleigh fading suppression in one-dimensional optical scatters," *IEEE Access*, vol. 7, pp. 17125–17132, 2019. doi: 10.1109/ACCESS.2019.2895126.

[7] Y. Fu, Z. Wang, R. Zhu, N. Xue, J. Jiang, C. Lu, B. Zhang, L. Yang, D. Atubga, and Y. Rao, "Ultra-long-distance hybrid BOTDA/Φ-OTDR," *Sensors*, vol. 18, no. 4, p. 976, Mar. 2018.

[8] D. Zhou, Y. Dong, B. Wang, C. Pang, D. Ba, H. Zhang, Z. Lu, H. Li, and X. Bao, "Single-shot BOTDA based on an optical chirp chain probe wave for distributed ultrafast measurement," *Light Sci. Appl.*, vol. 7, no. 1, Jul. 2018, Art. no. 32.

[9] M. Niklès and F. Ravet, "Distributed fibre sensors: Depth and sensitivity," *Nature Photon.*, vol. 4, no. 7, pp. 431–432, Jul. 2010.

[10] A. Motil, A. Bergman, and M. Tur, "[INVITED] state of the art of Brillouin fiber-optic distributed sensing," *Opt. Laser Technol.*, vol. 78, pp. 81–103, Apr. 2016.

[11] T. Horiguchi and M. Tateda, "BOTDA-nondestructive measurement of single-mode optical fiber attenuation characteristics using Brillouin interaction: Theory," *J. Lightw. Technol.*, vol. 7, no. 8, pp. 1170–1176, Aug. 1989.

[12] T. Kurashima, T. Horiguchi, and M. Tateda, "Distributed-temperature sensing using stimulated Brillouin scattering in optical silica fibers," *Opt. Lett.*, vol. 15, no. 18, pp. 1038–1040, Sep. 1990.

[13] T. Horiguchi, T. Kurashima, and M. Tateda, "A technique to measure distributed strain in optical fibers," *IEEE Photon. Technol. Lett.*, vol. 2, no. 5, pp. 352–354, May 1990.

[14] S. M. Haneef, Z. Yang, L. Thévenaz, D. Venkitesh, and B. Srinivasan, "Performance analysis of frequency shift estimation techniques in Brillouin distributed fiber sensors," *Opt. Express*, vol. 26, no. 11, pp. 14661–14677, May 2018.

[15] 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, Mar. 2018.

[16] H. Wu, Y. Wan, M. Tang, Y. Chen, C. Zhao, R. Liao, Y. Chang, S. Fu, P. P. Shum, and D. Liu, "Real-time denoising of Brillouin optical time domain analyzer with high data fidelity using convolutional neural networks," *J. Lightw. Technol.*, vol. 37, no. 11, pp. 2648–2653, Jun. 2019.

[17] A. K. Azad, L. Wang, N. Guo, H.-Y. Tam, and C. Lu, "Signal processing using artificial neural network for BOTDA sensor system," *Opt. Express*, vol. 24, no. 6, pp. 6769–6782, 2016.

[18] A. K. Azad, L. Wang, N. Guo, C. Lu, and H. Y. Tam, "Temperature profile extraction using artificial neural network in BOTDA sensor system," in *Proc. Opto-Electron. Commun. Conf. (OECC)*, Jun./Jul. 2015, pp. 1–3. doi: 10.1109/OECC.2015.7340143.

[19] B. Wang, N. Guo, F. Khan, A. K. Azad, L. Wang, C. Yu, and C. Lu, "Extraction of temperature distribution using deep neural networks for BOTDA sensing system," in *Proc. Conf. Lasers Electro-Opt. Pacific Rim (CLEO-PR)*, Jul./Aug. 2017, pp. 1–4. doi: 10.1109/CLEOPR.2017.8118961.

[20] A. P. Piotrowski and J. J. Napiorkowski, "A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling," *J. Hydrol.*, vol. 476, pp. 97–111, Jan. 2013.

[21] N. Murata, S. Yoshizawa, and S. Amari, "Network information criterion-determining the number of hidden units for an artificial neural network model," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 865–872, Nov. 1994.

[22] R. Gencay and M. Qi, "Pricing and hedging derivative securities with neural networks: Bayesian regularization, early stopping, and bagging," *IEEE Trans. Neural Netw.*, vol. 12, no. 4, pp. 726–734, Jul. 2001.

[23] C. A. Galindez-Jamioy and J. M. López-Higuera, "Brillouin distributed fiber sensors: An overview and applications," *J. Sensors*, vol. 2012, Aug. 2012, Art. no. 204121. doi: 10.1155/2012/204121.

[24] M. A. Farahani, E. Castillo-Guerra, and B. G. Colpitts, "A detailed evaluation of the correlation-based method used for estimation of the Brillouin frequency shift in BOTDA sensors," *IEEE Sensors J.*, vol. 13, no. 12, pp. 4589–4598, Dec. 2013.

[25] C. M. Bishop, "Training with noise is equivalent to Tikhonov regularization," *Neural Comput.*, vol. 7, no. 1, pp. 108–116, 1995.

[26] J. Dhliwayo, D. J. Webb, and C. N. Pannell, "Statistical analysis of temperature measurement errors in a Brillouin scattering-based distributed temperature sensor," *Proc. SPIE*, vol. 2838, pp. 276–286, Nov. 1996. doi: 10.1117/12.259808.