





# Federated learning-based wavelength demodulation system for multi-point distributed multi-peak FBG sensors

XUAN HOU,<sup>1</sup> SUFEN REN,<sup>1</sup>  KEBEI YU,<sup>1</sup> YULE HU,<sup>2</sup> HAoyang XU,<sup>1</sup> CHENYANG XUE,<sup>3,4</sup> SHENGCHAO CHEN,<sup>5,7</sup>  AND GUANJUN WANG<sup>6,8</sup>

<sup>1</sup>School of Information and Communication Engineering, Hainan University, Haikou 570228, China

<sup>2</sup>College of Mathematical Sciences, BoHai University, Jinzhou, Liaoning 121010, China

<sup>3</sup>Key Laboratory of Instrumentation Science and Dynamic Measurement of Ministry of Education, Taiyuan 030051, China

<sup>4</sup>School of Instrument and Electronics, North University of China, Taiyuan 030051, China

<sup>5</sup>Australian AI Institute, FEIT, University of Technology Sydney, NSW 2007, Australia

<sup>6</sup>School of Electronic Science and Engineering, Hainan University, Haikou 570228, China

<sup>7</sup>shengchao.chen.uts@gmail.com

<sup>8</sup>wangguanjun@hainanu.edu.cn

**Abstract:** Machine learning-based demodulation of multi-peak fiber Bragg grating (FBG) sensors has been extensively studied, demonstrating superior performance compared to conventional algorithms because it can neglect potential physical constraints. As the number of real-world monitoring points increases, the volume of fiber-optic sensing data grows exponentially. This necessitates aggregating data from various regions (e.g., via Wi-Fi), unlike traditional single-point monitoring, which challenges server storage capacity and communication efficiency. To address these issues, this paper proposes a federated learning (FL)-based framework for efficient wavelength demodulation of multi-peak FBGs in multipoint monitoring. Specifically, an arrayed waveguide grating (AWG) with multiplexing capability is employed at different monitoring points to convert spectral features into multi-channel transmission intensities, serving as training data for local models. Subsequently, the local model parameters, trained independently at each point, are uploaded to a central server to derive the optimal global model for demodulation across different monitoring points. The proposed demodulation framework is validated through stress demodulation experiments on multi-peak FBG sensors. Experimental results indicate strong multi-peak extraction performance with a demodulation error of  $\pm 0.64$  pm. Additionally, the method demonstrates excellent applicability for generating high-precision global demodulation models through multi-node cooperative work under various scenarios.

© 2024 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

## 1. Introduction

Fiber Bragg grating (FBG) has been widely researched for its multiplexing capability, small size, anti-electromagnetic interference, and long-time monitoring advantages [1–3]. In many applications, multiplexed FBG sensors are required for large-scale, high-precision multipoint measurements [4–7].

As a quasi-distributed sensor, FBG has the unique property of wavelength coding and wavelength multiplexing up to dozens of FBG gratings in a single fiber [8,9]. Common multi-peak FBG wavelength interrogation techniques are based on unbalanced Mach-Zehnder interferometry [10], Michelson interferometry [11], tunable Fabry-Perot filters, and ring-cavity fiber laser scanning [12], which require complex systems and expensive optics for their implementation [7,13]. However, the technical bottlenecks inherent in optical and electronic instruments make it

difficult to meet the needs of large-scale, high-precision, high-density, and high-speed sensing networks.

Demodulation algorithms for multi-peak FBG sensors have garnered significant attention. Yang et al. [14] used an improved Buneman frequency estimation algorithm to track FBG wavelengths, realizing an FBG wavelength demodulation scheme under low-resolution spectrometers or tunable lasers and reducing the cost of FBG interrogation systems. Liu et al. [15] proposed a multi-peak detection algorithm based on the Hilbert transform for solving the problem of multi-peak extraction of FBG reflectance spectra in wavelength-division multiplexing (WDM) systems, which improves the demodulation precision and speed of distributed FBG sensor networks. Theodosiou et al. [16] demonstrated a demodulation algorithm based on a combination of cross correlation and Hilbert transform, which converts the task of locating the wavelength maximum into a task of finding the over-zero point, enabling multi-peak demodulation of FBGs in multimode fibers. However, these methods focus on traditional demodulation, where measurements are obtained by parsing and extracting valid information, rely on the scientific knowledge and reasoning ability of the researcher [17], and require a lot of manual effort for data acquisition and information processing. They face challenges in meeting the demand for automated and cost-efficient mass surveillance.

In recent years, the emergence of machine learning (ML) has provided new ideas for the development of fiber optic sensing demodulation systems, and ML techniques have significantly improved the precision, efficiency, and flexibility of sensor signal processing, especially in dealing with complex and nonlinear problems, showing great potential [18–22]. Cao et al. [23] used a deep convolutional neural network (DCNN) to demodulate the spectra of an FBG sensor and achieved highly accurate temperature measurements with a mean-square error (MSE) of 0.108°C by extracting full-spectrum information. Zhao et al. [17] utilized the gramian angle field (GAF) technique to convert signals captured by FBG sensors into two-dimensional images and parsed these images by a convolutional neural network (CNN) to achieve highly accurate temperature demodulation. Sheng et al. [24] used a hybrid model incorporating a temporal convolutional network (TCN) and a light gradient boost machine (LightGBM) to solve the problem of FBG demodulation error drift of Fabry-Perot (F-P) filters in temperature-varying environments, and realized high-precision and fast FBG demodulation calibration. Ren et al. [25] proposed an improved multi-FBG filter demodulation method based on two cascaded artificial neural networks for improving the performance of FBG sensors in structural health monitoring, which achieves high-precision wavelength and strain measurements by adjusting the wavelength difference between matched FBGs and optimizing by the neural network algorithm. Currently, machine learning-driven demodulation systems effectively optimize the complexity of fiber optic sensor demodulation and achieve high precision measurements. However, in the sensing scenario of multiple monitoring points, the volume of fiber optic sensing data exhibits exponential growth, which is significantly different from the traditional single monitoring point method. Under these circumstances, it is necessary to effectively aggregate and analyze data from various monitoring areas (e.g., via Wi-Fi). In the engineering application of fiber optic sensing demodulation systems that incorporate machine learning technology, traditional centralized data analysis methods are encountering limitations [26,27]. The challenge mainly arises from the limitations in the storage capacity and communication efficiency of the underlying infrastructure, such as servers, which consequently hinder the speed and accuracy of data processing.

To address the deficiencies of the demodulation methods described above, this paper proposes multi-peak FBG demodulation method employing a federated learning (FL) algorithm. Specifically, the monitoring point utilizes the multiplexing capability of arrayed waveguide gratings (AWGs) to convert the response to external perturbations into transmitted intensities of different channels to train the neural network used for wavelength demodulation. The central server collaborates with the monitoring points to perform the optimization in multiple iterations

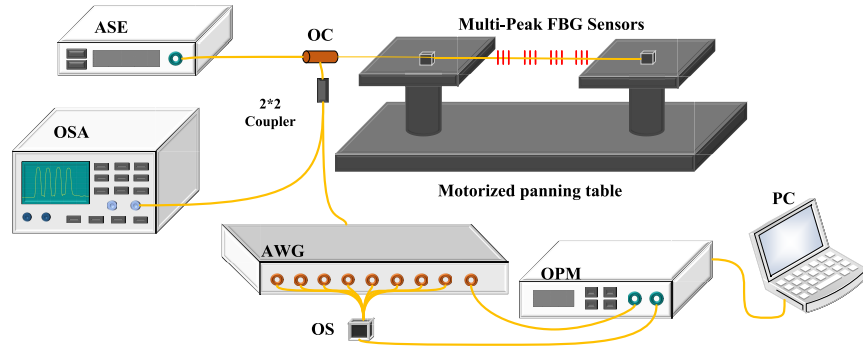
by means of parameter aggregation, which helps to establish a nonlinear relationship between the transmitted optical intensity and the multi-peak wavelengths in order to derive the global demodulation model. In the demodulation experiments performed on multi-peak FBG sensors in the 73 nm spectral range, the mean absolute error (MAE) of the global demodulation model is 1.79 pm, 1.51 pm, and 0.64 pm when three different numbers of monitoring points are set up. In addition, the inference time of the model for each data set is less than 0.02 ms. These results confirm that the demodulation method is not only accurate high, but also has a fast response time. It provides an effective technical way to realize intelligent large-scale structural health monitoring.

The remainder of this paper is organized as follows. Section 2 presents the multi-peak FBG sensor demodulation system and principles; Section 3 presents the federated learning-based multi-peak FBG sensors demodulation framework; Section 4 presents the experimental dataset and the model training setup; Section 5 presents the experimental results and discussion; Section 6 concludes this paper.

## 2. Multi-peak FBG demodulation

### 2.1. Demodulation system

The multi-channel strain sensing system we built is shown in Fig. 1. The multi-peak FBG sensors are mounted on the slide rails at both ends of the motorized translation table. The commands from a computer program control the movement of the slide rails, which in turn applies strain to the multi-peak FBG sensors. The light from the Amplified Spontaneous Emission (ASE) light source in the system first passes through an optical circulator (OC) and reaches the multi-peak FBG sensors. The sensors reflect light of a specific wavelength, which is evenly distributed to two different detection channels via a 2\*2 coupler (50:50 beam splitting ratio). These channels direct the light to the optical spectrum analyzer (OSA) and the AWG, respectively. The transmitted light from the arrayed waveguide grating (AWG) is passed through an 8-channel MEMS optical switch (OS) (MEMS-FSW8-SM-A) and output to an optical power meter (OPM) for measurement.



**Fig. 1.** The multi-channel strain sensing system comprises a broadband light source, a motorized panning table, an optical circulator, an AWG, an 8-channel MEMS Optical switch, an optical power meter, and a personal computer.

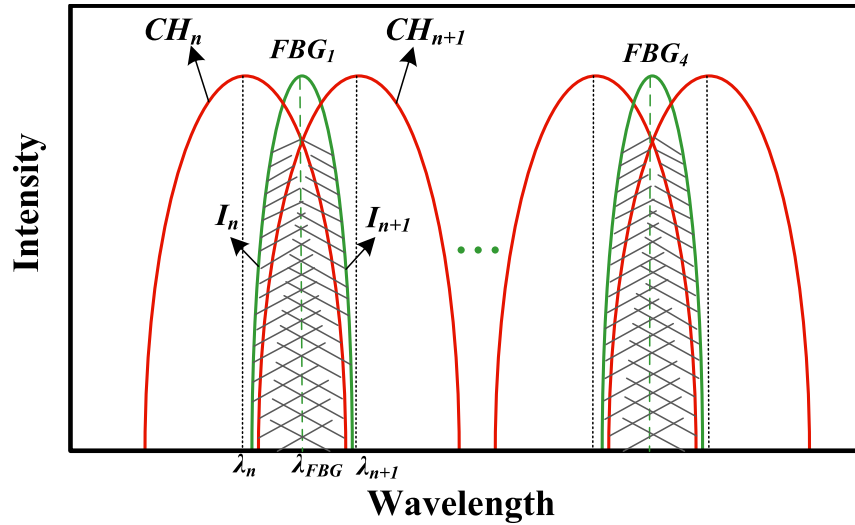
### 2.2. Demodulation principle

In this system, the spectral changes in FBG induced by strain are converted by an AWG into variations in transmitted intensity across multiple channels. Figure 2 shows the demodulation schematic of FBG sensor using AWG filters. Where  $\lambda_{FBG}$  represents the central wavelength of a single FBG.  $CH_n$  and  $CH_{n+1}$  represent the spectra of two neighboring channels in the AWG, with

corresponding center wavelengths of  $\lambda_n$  and  $\lambda_{n+1}$ , respectively. Two adjacent AWG channels work together to form a filter. After filtering through the adjacent two channels ( $CH_n$  and  $CH_{n+1}$ ) of AWG, the transmitted intensity of the reflected light is  $I_n$  and  $I_{n+1}$ , respectively. The shaded portion of the overlapping spectrum in Fig. 2 represents the transmission intensity. As the spectrum of the FBG sensor shifts, the size of the transmission strength changes. Therefore, we can capture the changes in the FBG spectrum due to external disturbances by observing the changes in the transmitted intensity of  $\lambda_{FBG}$ . This can be represented by an algorithm to calculate the central offset of the FBG sensor, as shown in Eq. (1).

$$\ln \left( \frac{I_{n+1}}{I_n} \right) = \frac{8(\ln 2)\Delta\lambda_c}{\Delta\lambda_{FBG}^2 + \Delta\lambda_n^2} \lambda_{FBG} - \frac{4(\ln 2)(\lambda_{n+1}^2 + \lambda_n^2)}{\Delta\lambda_{FBG}^2 + \Delta\lambda_n^2}, \quad (1)$$

where  $\Delta\lambda_c$  is the difference in center wavelength between the two channels,  $\Delta\lambda_{FBG}$  is the full width at half height (FWHM) of the FBG, and  $\Delta\lambda_n$  and  $\Delta\lambda_{n+1}$  are the FWHM of the  $n$ th and  $n+1$ th channels, respectively.



**Fig. 2.** Demodulation schematic of FBG sensors using AWG filters, where the transmission strength is represented by the overlap region of the FBG spectrum with the AWG channel.

### 3. Federated learning-based demodulation strategy

#### 3.1. Local model training

In this demodulation system, changes in external physical quantities are recognized by monitoring fluctuations in the transmitted intensity of the AWG channels. Specifically, the optical intensity values of the nine AWG channels captured by the OPM are introduced as input features for the neural network. Meanwhile, the sensor's peak wavelength data, measured by the OSA, was obtained as the target output value for the training process. These collected data were used to train the neural network model with the aim of constructing a model capable of describing the complex nonlinear relationship between transmitted intensity and peak wavelength, which can be represented by Eq. (2). In this way, the model is able to accurately infer the corresponding changes in physical quantities from the changes in transmission intensity, thus enabling highly



accurate demodulation of the multi-peak FBG sensors.

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4 = \text{Net}(I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8, I_9), \quad (2)$$

where Net is the neural network used to demodulate the multi-peak FBG sensors,  $\lambda_1$ - $\lambda_4$  are the four peak wavelengths of the sensor, and  $I_1$ - $I_9$  are the transmitted intensities of the nine AWG channels.

Before the dataset consisting of  $\lambda_1$ - $\lambda_4$  and  $I_1$ - $I_9$  is fed to the local model training, the data is scaled to between [0,1] using max-min normalization as shown in Eq. (3). This process enhances the model's generalization ability and improves prediction precision [28]. By ensuring that all input transmission intensity values are on a consistent scale, the training efficiency of the algorithm is optimized, enabling faster convergence towards the optimal solution on the error surface.

$$I'_{Chn} = \frac{I_{Chn} - \min(I_{Chn})}{\max(I_{Chn}) - \min(I_{Chn})}, \quad (3)$$

where  $I_{Chn}$  denotes the transmitted intensity detected by the  $n$ th AWG channel,  $I'_{Chn}$  denotes the Min-Max normalized intensity for the same channel.

The architecture of the established neural network is shown in Fig. 3, which consists of an output layer, three hidden layers, and an input layer, with full connectivity between each neural network layer. The input layer consists of nine independent neurons corresponding to the transmitted intensities of the nine channels of the AWG, and the four neurons of the output layer corresponding to the peak wavelengths of the multi-peak FBG sensors. The activation functions used between hidden layers are Leaky Relu [29] and Tanh [30], and the function expressions are Eq. (4) and Eq. (5), respectively.

$$\text{LeakyRelu}(x) = \begin{cases} x, & \text{if } x \geq 0 \\ \alpha x, & \text{if } x < 0 \end{cases}, \quad (4)$$

$$\text{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad (5)$$

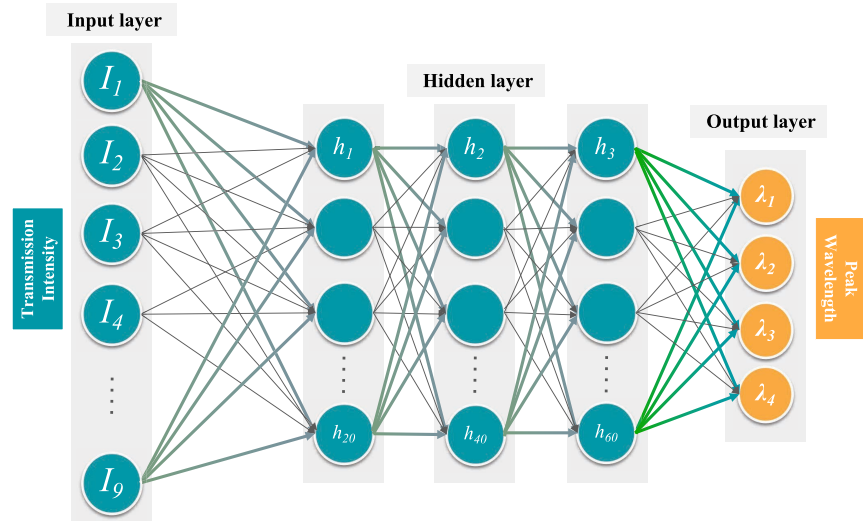
where  $x$  is the input to the neuron and  $\alpha$  is a very small negative number, usually 0.01.

### 3.2. Model parameter update-aggregation

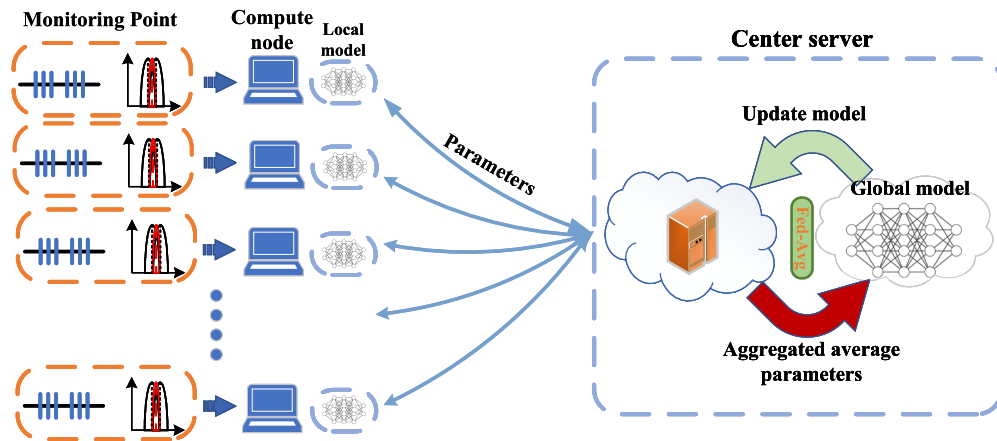
Figure 4 illustrates the construction process of the global model under the FL demodulation framework. In this framework, the system consists of two main parts: local devices and a central server. Specifically, the local device trains the local model using the data collected at the monitoring points, while the central server coordinates compute nodes. Model updating is achieved by passing neural network parameters between the local computing units and the central server, and the whole process does not involve the interaction of monitoring data.

Throughout constructing the global demodulation model, we employ parameter update-aggregation mechanisms. Local devices conduct model training and parameter updating using their respective data and transmit the updated model parameters to the central server. The central server aggregates these parameters to obtain the continuously optimized and improved global demodulation model. The specific parameter update and aggregation steps are as follows:

1. In the  $i$ th communication round, the central server sends the current global model parameters  $w_i$  to the local devices to initialize the local model parameters.
2. Local devices receive the parameters, assigning  $w_i$  to the local model parameters  $w_i^k$ , and then uses the data from the sub-data set to train the local model.



**Fig. 3.** The demodulation network model, consisting of 9 neurons in the input layer, 60 neurons in the hidden layer, and 4 neurons in the output layer.



**Fig. 4.** Federated learning based demodulation model training framework. The central server coordinates each compute node to train neural networks suitable for wavelength demodulation of multi-peak FBG sensors. Model updating is achieved by passing the neural network parameters between the local computing units and the central server

3. After the local model is trained, the local devices transmit the updated model parameters  $w_{i+1}^k$  obtained from the training updates to the central server.
4. In the  $i + 1$ th communication round, the central server aggregates the model parameters  $w_{i+1}^k$  transmitted from all local models, averaging them to update the global demodulation model parameters  $w_{i+1}$ .

Notably, federated learning never transmits raw data subsets between devices or to the central server. Training occurs exclusively on locally stored data, with only model parameters shared for aggregation. This approach safeguards data privacy and adheres to the fundamental principle of federated learning: raw data remains on local devices throughout the training process.

This collaborative approach enables the global model to leverage the data and knowledge contributed by each local device. The specific implementation of distributed federated learning is shown in Algorithm 1.

---

**Algorithm 1. Parameter Aggregation Updating Algorithm**

---

```

1: Initialize: Global Model Parameter:  $w$ 
2: Number of Local Devices:  $K$ , Mini-batch Size:  $B$ , Local epochs:  $E$ , Learning Rate:  $lr$ ,
   Local Dataset:  $D$ , Maximum Communication Rounds:  $T$ .
3: for communication round  $i \leftarrow 1$  to  $T$  do
4:   for compute code  $k \leftarrow 1$  to  $K$  do
5:      $w_{i+1}^k \leftarrow \text{LocalDeviceUpdate}(k, w_i, B, E, lr, D_k)$ 
6:      $w_{i+1} \leftarrow \sum_{k=1}^K \frac{n_k}{n} w_{i+1}^k$ 
7:   end for
8: end for
9: function LOCALDEVICEUPDATE( $k, w_i, B, E, lr, D_k$ )
10:  Input: Monitoring points:  $k$ , Global model parameter:  $w_i$ , Mini-batch size:  $B$ , Local
    epochs:  $E$ , Learning rate:  $lr$ , Local dataset:  $D_k$ 
11:  Output: Updated local model parameter:  $w_{i+1}^k$ 
12:  for local epoch  $j \leftarrow 1$  to  $E$  do
13:    for batch  $b \leftarrow 1$  to  $B$  do
14:       $w \leftarrow w - lr \cdot \partial f(w)$ 
15:    end for
16:  end for
17:  return  $w_{i+1}^k$ 
18: end function

```

---

The global and local models employ an identical neural network structure depicted in Fig. 3 to ensure compatibility during the parameter update-aggregation process. During the training of the global demodulation model, the objective optimization function of the model can be expressed as Eq. (6).

$$\min_{w,b} f(w,b) \quad \text{where} \quad f(w,b) = \frac{1}{n} \sum_{i=1}^n f_i(w,b), \quad (6)$$

where  $f(w)$  is the model training loss function, in this system the mean square error, and  $w$  is the model parameter. By iteratively optimizing the global model, it gradually incorporates and merges the knowledge from each local device, enabling it to adapt to the overall data distribution and characteristics and enhance the demodulation performance of the system as a whole.

### 3.3. Global demodulation model distribution test

After completing the construction process of the global demodulation model described above, a global model  $F(w)$  applicable to the demodulation of multi-peak FBG sensors can be obtained.

The computational process of demodulation prediction can be expressed as Eq. (7).

$$(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4) \leftarrow F(w; I_1, I_2, \dots, I_9), \quad (7)$$

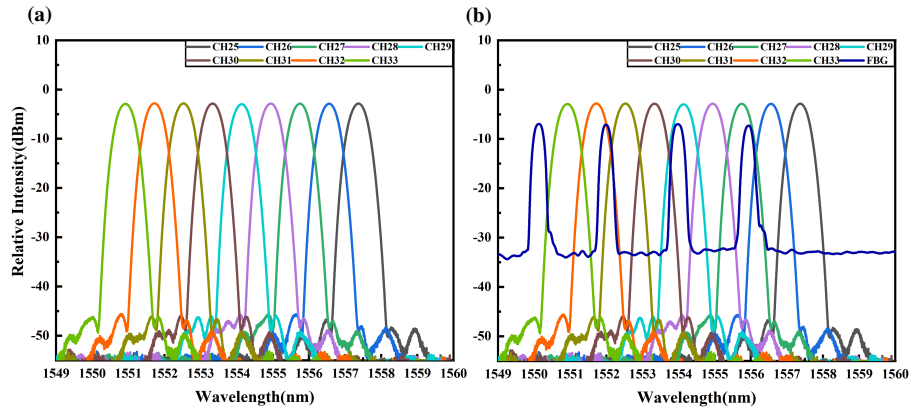
where  $D_k$  is the transmitted intensity dataset of the  $k$ th sub-data set.  $I_1, I_2, \dots, I_9$  denote the transmitted intensity.  $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4$  denote the predicted peak wavelengths.

## 4. Experiments

### 4.1. Multi-peak FBG sensors strain dataset

The experimental data were obtained using the experimental setup shown in Fig. 1, where the multi-peak FBG sensors used therein consist of four FBGs in series on a single optical fiber. The gratings used as Sen-FBG were based on SMF-28e with center wavelengths of 1550nm, 1552nm, 1554nm, and 1556nm, respectively, with a reflectivity of 90%, and the ambient temperature of the experiment was 26 °C.

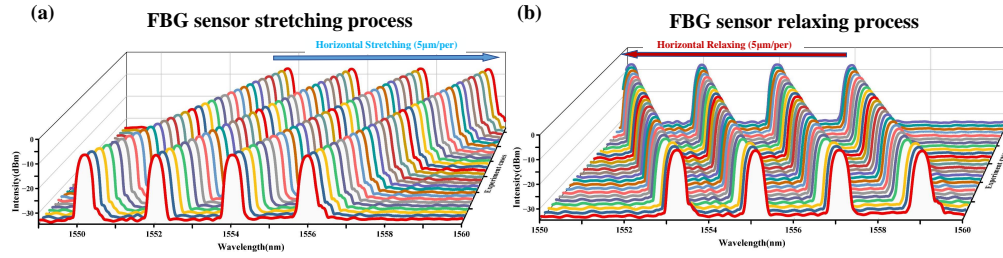
Figure 5(a) shows the reflectance spectra of the nine channels of the AWG used in the experiment, where channel  $CH_{31}$  is directly connected to  $CH_2$  of the optical power meter, and the rest of the selected channels are connected to  $CH_1$  of the optical power meter via optical switch. Figure 5(b) shows the mixing spectrum of the reflected light from the multi-peak FBG sensors with AWGs, with the four peak wavelengths of the sensor measured by OSA.



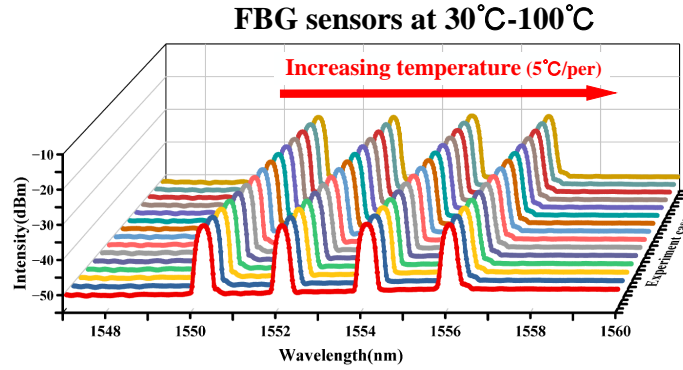
**Fig. 5.** (a) Reflectance spectra of the AWG channel used; (b) Combined spectra of the FBGs sensor and the AWG channel.

In the experimental operation, we achieved the stretching and relaxation control of the multi-peak FBG sensors by precisely adjusting the motorized panning table. Meanwhile, an optical power meter was used to collect the transmitted intensity of the AWG channel in real time, and an OSA was applied to obtain detailed spectral information. Figure 6(a) shows the variation of the interference spectrum of the sensor during stretching, where the interference fringes are gradually shifted to the right with stretching. Figure 6(b) shows the variation of the interference spectrum of the sensor during relaxation, where the interference fringes are gradually shifted to the left with stretching.

A key question during the application of FBG is the cross-talk of temperature sensitivity and strain sensitivity. We characterized the sensor's temperature response by placing a multi-peak FBG sensor in a tubular heating furnace. The temperature was incrementally increased by five degrees Celsius each time. Specifically, the temperature was raised from 30°C to 60°C in 5°C steps. At each step, the temperature was held for 10 minutes. Figure 7 illustrates the sensor spectra at different temperatures.

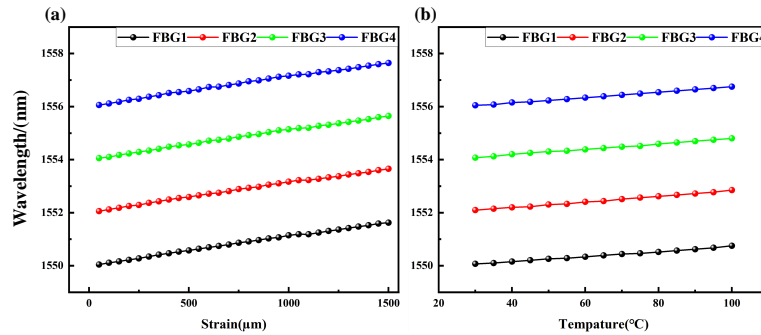


**Fig. 6.** (a) shows the schematic diagram of the spectral displacement of the sensor in the horizontal stretching state; (b) shows the schematic diagram of the spectral displacement in the horizontal relaxation state.



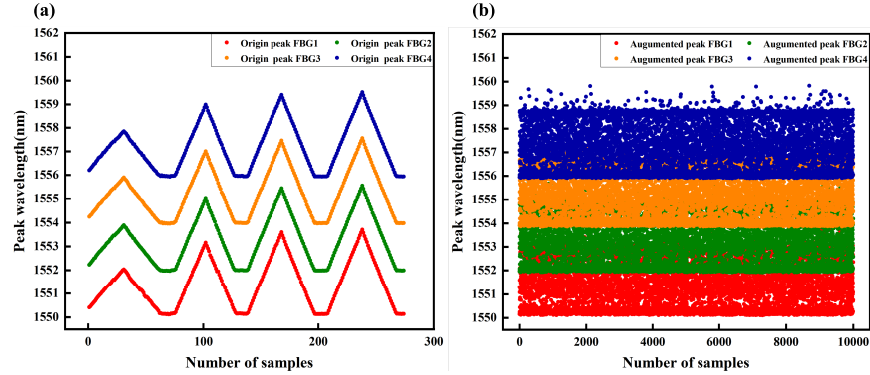
**Fig. 7.** Spectral shift of FBG sensors during the process of increased temperatures.

Figure 8 illustrates the sensor peak's wavelength shift during both stretching and heating. The sensor's four peak wavelengths shift with increasing temperature, exhibiting temperature sensitivities of 10.3 pm/°C, 10.7 pm/°C, 10.4 pm/°C, and 10 pm/°C, respectively. The strain sensitivities of the four peaks are 1.082 pm/μ $\epsilon$ , 1.101 pm/μ $\epsilon$ , 1.096 pm/μ $\epsilon$ , and 1.096 pm/μ $\epsilon$ , respectively. Thus, the temperature-to-strain crosstalk can be calculated as 9.519 μ $\epsilon$ /°C, 9.718 μ $\epsilon$ /°C, 9.489 μ $\epsilon$ /°C, and 9.124 μ $\epsilon$ /°C, respectively. To minimize temperature crosstalk and ensure reliable data for neural network testing, strain measurements were conducted at a constant ambient temperature.



**Fig. 8.** Wavelength shift of sensor peaks during stretching (a) and heating (b).

We employed data augmentation techniques [20,31] to extend the original real dataset with the aim of evaluating the demodulation efficacy of federated learning algorithms when dealing with different numbers of monitoring points. Figure 9(a) illustrates the raw dataset collected during the experiment, in which the peak wavelength distribution of the multi-peak FBG sensor is presented in the form of different colored dot groups, each representing a specific peak wavelength. Figure 9(b) shows the distribution of peak wavelengths of the multi-peak FBG sensor after data augmentation. The expanded dataset not only significantly increased in quantity, but also increased the diversity of samples while maintaining data quality, providing a more solid foundation for in-depth analysis and evaluation of algorithms.



**Fig. 9.** (a) and (b) are the peak wavelength distribution of the original data set and the peak wavelength distribution of the augmented data set, respectively.

#### 4.2. FL model training parameter settings

In this study, we have endeavored to construct an distributed training environment by tuning the parameter  $K$ . Specifically, the experiments have been designed to incrementally increase the value of  $K$ , thereby concurrently scaling up the volume of data engaged in training and the number of computational nodes. This approach was employed to evaluate the performance of the federated learning demodulation model under varying scales of monitoring points.

Specifically, the strain dataset is assigned to each computational node in order to train the local model independently on each node. We set the number of nodes  $K$  involved in training, and performed parallel training of the global demodulation model for three scenarios with 5, 10, and 20 nodes, each of which was assigned 500 pieces of monitoring data. In the experimental setup, the dataset is categorized into training, validation and testing sets with the proportions of 75%, 15%, and 10%, respectively. The number of iterations for global communication was set to 256 rounds. In the training phase, we chose the Adam optimizer [32] for parameter optimization. To prevent model overfitting, we introduce an early stopping strategy during training.

Both the local and global models use the neural network structure shown in Fig. 3. In the setting of local model training, the initial learning rate was adjusted to 0.001, the training period was fixed to 50 epochs, and the batch size was chosen to be 64 samples.

#### 4.3. Evaluation indicators

Mean square error (MSE), root mean square error (RMSE), coefficient of determination ( $R^2$ ), and mean absolute error (MAE) as indicators of model evaluation. They can build as follows:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\lambda_i - \hat{\lambda}_i)^2, \quad (8)$$



$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\lambda_i - \hat{\lambda}_i)^2}, \quad (9)$$

$$R^2 = 1 - \frac{\sum_i (\hat{\lambda}_i - \lambda_i)^2}{\sum_i (\bar{\lambda}_i - \lambda_i)^2}, \quad (10)$$

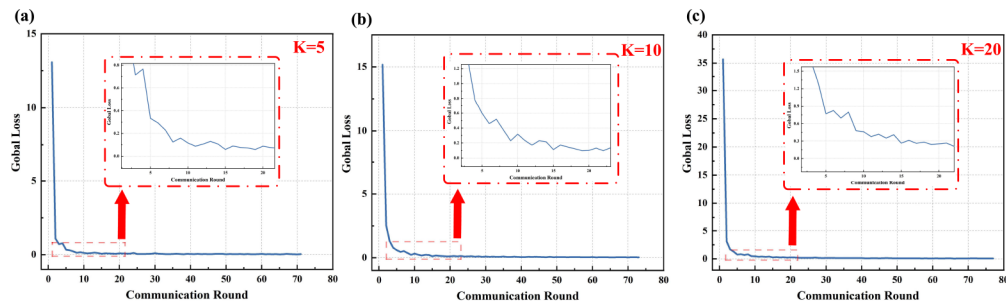
$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |(\lambda_i - \hat{\lambda}_i)|, \quad (11)$$

where  $n$  is the total number of samples tested,  $\lambda$  and  $\hat{\lambda}$  are the peak wavelength of the FBG and the predicted peak wavelength of the network model output, respectively,  $\bar{\lambda}$  is the actual average peak wavelength. MSE and RMSE assess the accuracy of the model in predicting wavelength peaks. MSE calculates the squared error between predicted and real values, making it ideal for precise peak detection. RMSE, expressed in the same units as the prediction and real values, provides a more intuitive measure of error magnitude.  $R^2$  evaluates the ability of the model to capture complex nonlinear relationships between transmitted light intensity and wavelength.  $R^2$  close to 1 indicates strong model fit and precise characterization of this physical correlation. MAE, as a complementary metric, measures prediction robustness, ensuring reliable performance across varying conditions.

## 5. Experimental results and discussion

### 5.1. Training of FL-DNN model and results

First, we monitor the training of the global demodulation model by analyzing the trend of the evolution of the global loss function during the communication process. Figures 10(a), 10(b) and 10(c) depict the evolution of the global loss during training. Observations show that the global loss decreases with training and converges after about 20 rounds of communication. The convergence trend of the model is not hindered as the amount of data increases and the computational size increases. This reflects the efficiency and stability of the distributed federated strategy. Although the increase in data volume makes the model training more challenging, the federated learning strategy effectively integrates the sensor demodulation information and constructs an accurate nonlinear model.



**Fig. 10.** Loss variation during global model training for different number of computing nodes (from left to right, K=5, 10, 20).

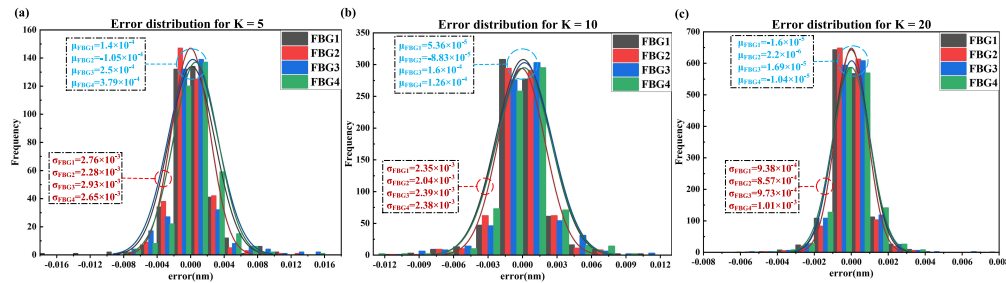
Table 1 provides a comprehensive overview of the test metric values for global demodulation models under different data size configurations, setting the stage for understanding the subsequent error distribution and statistical analysis. As the number of computational nodes increases from 5 to 20, alongside the expansion of data size, there is a noticeable enhancement in the model's performance. The value of MSE decreases significantly from  $4.58 \times 10^{-6}$  to  $1.12 \times 10^{-6}$ , indicating

a substantial reduction in the average squared difference between the predicted and actual values. Similarly, the MAE, which measures the average magnitude of the errors, drops from  $1.79 \times 10^{-3}$  to  $0.64 \times 10^{-3}$ , suggesting that the model's predictions are becoming more precise. The RMSE also exhibits a descending trend, further corroborating the improvement in prediction accuracy. Notably, the value of  $R^2$  remains consistently high at 0.99 across all configurations, implying an extremely close fit of the model to the data. These preliminary test metric results foreshadow the positive trends observed in the detailed error analysis and distribution figures, highlighting the model's robustness when more monitoring points are involved.

**Table 1. Test metric values for global demodulation models with different data size configurations.**

	$K = 5$	$K = 10$	$K = 20$
MSE ( $\times 10^{-6}$ )	4.58	3.45	1.12
MAE ( $\times 10^{-3}$ )	1.79	1.51	0.64
RMSE ( $\times 10^{-3}$ )	2.14	1.86	1.05
$R^2 (\times 10^{-2})$	99.99	99.99	99.99

Figure 11 illustrates the refinement in error distribution as the federated FL-DNN model scales up in terms of dataset size and number of computational nodes. Initially, with 2500 data points and 5 participants, the standard deviations for the FBG peaks are relatively higher, indicating broader error dispersion. Notably, as the dataset grows to 5000 data points with 10 participants, there is a slight yet consistent decrease in standard deviations ( $\sigma_{FBG}$ ), suggesting an improvement in demodulation precision. This trend continues positively with the largest dataset of 10000 data points and 20 participants, where standard deviations reach their lowest, notably narrowing the error distribution and showcasing the model's enhanced accuracy. The means ( $\mu_{FBG}$ ) also demonstrate a stable trend towards zero, reinforcing the model's reliability across various scales. These specific numerical values indicate that the FL-DNN model exhibits improved precision in demodulation as the dataset and participants increase, demonstrating the model's scalability and robustness in handling larger datasets.



**Fig. 11.** Distribution of peak sensor prediction errors for different amounts of data and number of computing nodes.

The error analysis presented in Table 2 complements the findings from Fig. 11 by providing a quantitative assessment of the demodulation performance for FBG peaks with varying numbers of monitoring points involved. As the dataset size increases from 2500 to 10000 and the number of monitoring points participating in the FL-DNN model grows from 5 to 20, a consistent trend of error reduction is observed. For instance, the average error for FBG1 decreases from 1.78 pm with 5 participants to 0.64 pm with 20 participants, indicating a significant improvement in demodulation accuracy. This trend is consistent across all FBG peaks, with the average errors generally decreasing as more nodes contribute to the model. The maximum and minimum errors

also show a decline, suggesting a reduction in both the upper bounds and the variability of errors. This aligns with the observations from Fig. 10, which indicated a more concentrated error distribution and higher precision with increased data and computational resources. Collectively, the data from both the table and figure underscore the scalable and robust nature of the FL-DNN model in demodulating FBG sensors, as it adapts effectively to larger datasets and distributed computing environments.

**Table 2. Absolute error analysis of FBG peaks with different number of computing nodes involved (unit: pm).**

Participants	Peak	Max	Min	Average
K=5	FBG1	16.3	0.014	1.78
	FBG2	12.7	0.012	1.61
	FBG3	15.1	0.005	1.91
	FBG4	15.9	0.007	1.85
K=10	FBG1	13.2	0.003	1.56
	FBG2	12.8	0.014	1.39
	FBG3	11.8	0.001	1.53
	FBG4	12.9	0.002	1.58
K=20	FBG1	7.8	0.0005	0.64
	FBG2	8.1	0.0002	0.58
	FBG3	7.2	0.0001	0.65
	FBG4	6.9	0.001	0.7

The data in Table 3 shows the time required for the training and inference processes in the parallel configuration of the FL algorithm. From the results, it can be seen that the training time of the global demodulation model remains relatively stable and does not change much as the  $K$  value increases. Demonstrates high efficiency in parallel processing of monitoring data. In addition, the wavelength interrogation speed of the model is maintained within 0.02 seconds.

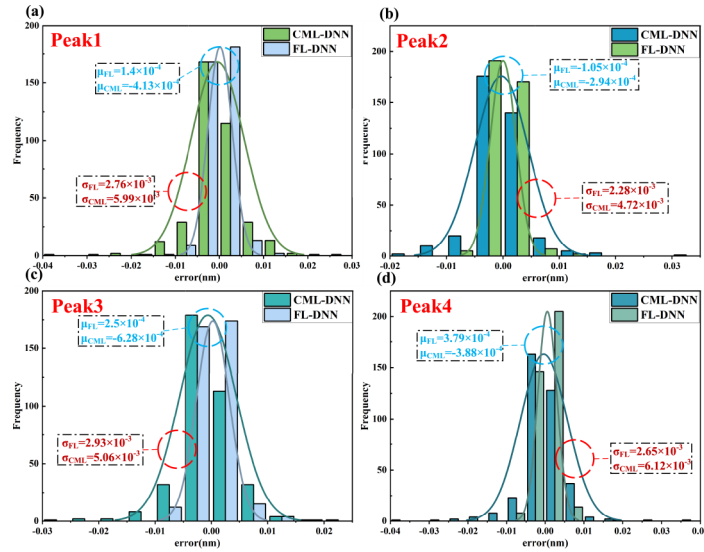
**Table 3. Training and testing time for parallel state computation.**

Participants	Training time(min)	Testing time(ms)
5	2.3	0.02
10	2.2	0.018
20	2.2	0.016

## 5.2. Centralized and FL performances with deep neural network (DNN) model

In this section, the non-federated (CML) performance of the DNN model is compared with the DNN performance under the distributed federated learning framework. Where the same settings are used for the network structure and training parameters of the model.

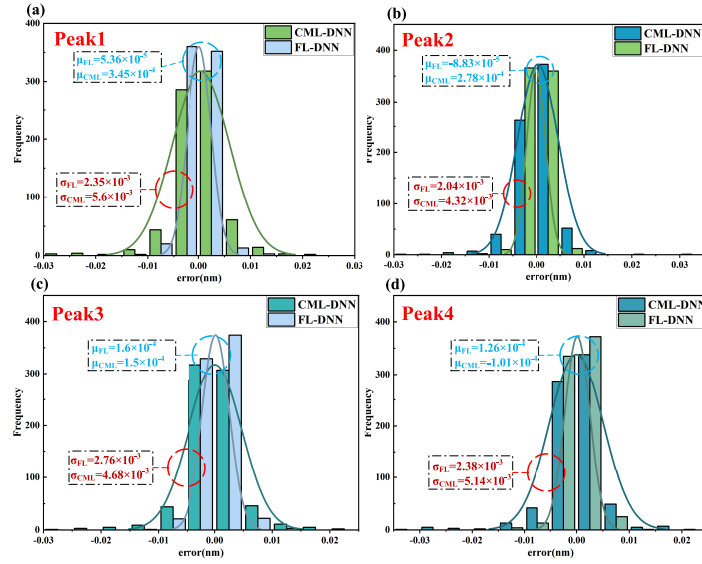
Figure 12 presents a comparative analysis of the error distribution in the demodulation of FBG strain datasets, with a size of 2500, using a CML-DNN and a FL-DNN demodulation model with 5 local devices. The data provided reveals that across all peak values (Peak1 to Peak4), the FL-DNN model exhibits a smaller standard deviation ( $\sigma$ ), indicating a more concentrated error distribution and higher demodulation accuracy. Specifically, the mean ( $\mu$ ) of the FL-DNN model is closer to zero, suggesting that its error center is more aligned with the ideal state. Furthermore, the consistently lower standard deviation of the FL-DNN model compared to the CML-DNN model implies less variability in the demodulation process, thereby demonstrating greater stability and reliability.



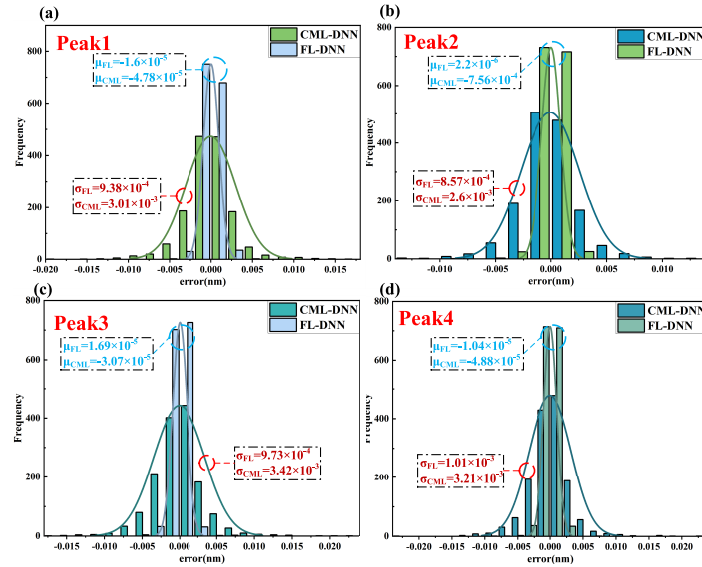
**Fig. 12.** Performance evaluation of CML-DNN and FL-DNN models: A comparative analysis of prediction errors with a 2500-entry dataset.

Figure 13 shows a comparative analysis of the error distribution for demodulation of the FBG strain dataset using the CML-DNN and FL-DNN demodulation models, where the FL-DNN dataset is of size 5000 and 10 nodes are involved in training. Extending the previous analysis, it is observed that the FL-DNN model continues to demonstrate a smaller standard deviation ( $\sigma$ ) across all peak values, indicating a more focused error distribution and higher precision in demodulation. The mean ( $\mu$ ) values for the FL-DNN model are closer to zero compared to the CML-DNN model, suggesting that the FL-DNN model maintains a more accurate error center. The lower standard deviation of the FL-DNN model across all peaks implies a more consistent performance with less variability, even with an increased dataset size and a higher number of local devices. This suggests that as the complexity of the dataset grows, the FL-DNN model retains its advantage over the CML-DNN model in terms of accuracy and robustness in demodulation performance.

Figure 14 presents a comparative analysis of the error distribution for demodulation of the FBG strain dataset using the CML-DNN and FL-DNN demodulation models, where the FL-DNN dataset is of size 10000 and 20 nodes are involved in training. The FL-DNN model, continues to exhibit a lower standard deviation ( $\sigma$ ) across all peak values. This is evidenced by the values such as  $9.38 \times 10^{-4}$ ,  $8.57 \times 10^{-4}$ ,  $9.73 \times 10^{-4}$ , and  $1.01 \times 10^{-3}$  for FL-DNN, which are markedly lower than the CML-DNN model's standard deviations of  $3.01 \times 10^{-3}$ ,  $2.6 \times 10^{-3}$ ,  $3.42 \times 10^{-3}$ , and  $3.21 \times 10^{-3}$ , respectively. The mean ( $\mu$ ) values for FL-DNN, being closer to zero, indicate a more accurate central tendency of the error distribution. The normal curves representing the error distributions further illustrate the comparative performance of the two models. The narrower spread of the FL-DNN model's normal curve suggests a more precise demodulation process, with errors tightly clustered around the mean. This is in contrast to the broader spread of the CML-DNN model's curve, indicating a higher likelihood of larger errors. The consistency of these results across different dataset sizes and configurations of the number of nodes in the FL-DNN model underscores its effectiveness in minimizing demodulation error and maintaining a stable performance.



**Fig. 13.** Performance evaluation of CML-DNN and FL-DNN models: A comparative analysis of prediction errors with a 5000-entry dataset.



**Fig. 14.** Performance evaluation of CML-DNN and FL-DNN models: A comparative analysis of prediction errors with a 10000-entry dataset.

## 6. Discussion

The following justifies our proposed approach by presenting the rationale for applying Federated Learning (FL) to distributed FBG demodulation, which encompasses three key aspects:

1. **Relevance of Federated Learning** Although large-scale quasi-distributed FBG networks are not yet common, the trend towards extensive sensor deployments in IoT and smart infrastructure anticipates increased demand for such systems. FL is well-suited for these applications by ensuring data privacy, reducing transmission costs, and enhancing security, which centralized models cannot efficiently provide.
2. **Advantages Over Centralized AI Methods** While some machine learning models achieve high demodulation performance, they typically require centralized data processing. This approach can lead to privacy issues and higher data transfer costs, especially as network size grows. FL mitigates these issues by keeping data localized, thus offering a scalable and secure alternative without compromising performance.
3. **Empirical Support for FL's Effectiveness** Our experiments demonstrate that FL not only maintains low RMSE but also improves model generalization and robustness in heterogeneous data environments common in multi-peak FBG applications. These benefits highlight FL's superiority in scenarios where data distribution varies, which is less effectively handled by traditional models.

These findings demonstrate that FL outperforms centralized AI methods in our demodulation scenario. Moreover, FL's effective balance between demodulation performance and key factors such as data privacy and transmission cost further validates the soundness and effectiveness of our approach.

## 7. Conclusion

This study proposes what we believe to be a novel distributed framework for multi-peak FBG sensor demodulation under multi-point monitoring. The framework employs a data-distributed training strategy that allows model training parameters to be shared across monitoring sites without direct data interaction, effectively overcoming the limitations of low-resource devices in engineering applications. In addition, a global demodulation model with strong generalization is constructed by employing a lightweight neural network across multiple monitoring sites and aggregating models trained on small local datasets through parameter aggregation updates. Experimental results show that the method achieves efficient demodulation of multi-peak FBG sensors with a demodulation linearity of about 0.99 at different scales and a maximum error of only 8.1 p.m. This provides a new and efficient solution for the demodulation of multi-peak FBG sensors under multi-point monitoring.

**Funding.** Hainan Province Science and Technology Special Fund (ZDYF2022SHFZ304, ZDYF2023GXJS013); Scientific Research Starting Foundation of Hainan University (KYQD(ZR)1882); Major Science and Technology Program of Haikou City (2021-002); Natural Science Foundation of Hainan Province (2019CXTD400, 617079, 620RC554); National Natural Science Foundation of China (61762033, 61865005, 62175054).

**Disclosures.** The authors declare no conflict of interest.

**Data availability.** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## References

1. U. Jayawickrema, H. Herath, N. Hettiarachchi, *et al.*, "Fibre-optic sensor and deep learning-based structural health monitoring systems for civil structures: A review," *Measurement* **199**, 111543 (2022).
2. A. Kersey, M. Davis, H. Patrick, *et al.*, "Fiber grating sensors," *J. Lightwave Technol.* **15**(8), 1442–1463 (1997).
3. L. S. M. Alwis, K. Bremer, and B. Roth, "Fiber optic sensors embedded in textile-reinforced concrete for smart structural health monitoring: A review," *Sensors* **21**(15), 4948 (2021).



4. C. Li, J. Tang, C. Cheng, *et al.*, "Fbg arrays for quasi-distributed sensing: A review," *Photonic Sens.* **11**(1), 91–108 (2021).
5. Y.-W. Ren, Q. Yuan, J. Chai, *et al.*, "Study on the clay weakening characteristics in deep unconsolidated layer using the multi-point monitoring system of fbg sensor arrays," *Opt. Fiber Technol.* **61**, 102432 (2021).
6. M. Mieloszyk and W. Ostachowicz, "An application of structural health monitoring system based on fbg sensors to offshore wind turbine support structure model," *Mar. Struct.* **51**, 65–86 (2017).
7. F. O. Barino and A. B. dos Santos, "Lpg interrogator based on fbg array and artificial neural network," *IEEE Sens. J.* **20**(23), 14187–14194 (2020).
8. E. Li, J. Xi, Y. Yu, *et al.*, "Multichannel fbg sensing system using a dense wavelength division demultiplexing module," in *Advanced Sensor Systems and Applications II*, vol. 5634 (SPIE, 2005), pp. 211–218.
9. S. Weng, P. Yuan, W. Zhuang, *et al.*, "Soi-based multi-channel awg with fiber Bragg grating sensing interrogation system," in *Photonics*, vol. 8 (MDPI, 2021), p. 214.
10. M. D. Todd, G. A. Johnson, and B. L. Althouse, "A novel Bragg grating sensor interrogation system utilizing a scanning filter, a mach-zehnder interferometer and a 3×3 coupler," *Meas. Sci. Technol.* **12**(7), 771–777 (2001).
11. K. T. O'Mahoney, R. P. O'Byrne, S. V. Sergeyev, *et al.*, "Short-scan fiber interferometer for high-resolution Bragg grating array interrogation," *IEEE Sens. J.* **9**(10), 1277–1281 (2009).
12. A. Ray, A. Bandyopadhyay, S. De, *et al.*, "A simple scanning semiconductor diode laser source and its application in wavelength modulation spectroscopy around 825nm," *Opt. Laser Technol.* **39**(2), 359–367 (2007).
13. S. Chen, F. Yao, S. Ren, *et al.*, "Cost-effective improvement of the performance of awg-based fbg wavelength interrogation via a cascaded neural network," *Opt. Express* **30**(5), 7647–7663 (2022).
14. Y. Yang, J. Wu, M. Wang, *et al.*, "Fast demodulation of fiber Bragg grating wavelength from low-resolution spectral measurements using buneman frequency estimation," *J. Lightwave Technol.* **38**(18), 5142–5148 (2020).
15. F. Liu, X. Tong, C. Zhang, *et al.*, "Multi-peak detection algorithm based on the hilbert transform for optical fbg sensing," *Opt. Fiber Technol.* **45**, 47–52 (2018).
16. A. Theodosiou, M. Komodromos, and K. Kalli, "Accurate and fast demodulation algorithm for multipeak fbg reflection spectra using a combination of cross correlation and hilbert transformation," *J. Lightwave Technol.* **35**(18), 3956–3962 (2017).
17. B. Zhao, W. Li, L. Xia, *et al.*, "A cnn-based fbg demodulation method adopting the gaf-assisted ascending dimension of complicated signal," *Opt. Commun.* **499**, 127296 (2021).
18. Y. Zhou, Y. nan Zhang, Q. Yu, *et al.*, "Application of machine learning in optical fiber sensors," *Measurement* **228**, 114391 (2024).
19. T. Yamaguchi, H. Kawashima, H. Matsuda, *et al.*, "Experimental demonstration of peak wavelength measurement of multiplexing fiber Bragg gratings using convolutional neural network," *IEEE Sens. J.* **23**(9), 9343–9352 (2023).
20. H. Xu, S. Chen, S. Ren, *et al.*, "Dual-parameter demodulation of fbg-fpi cascade sensors via sparse samples: A deep learning-based perspective," *IEEE Sensors Journal* (2023).
21. S. Ren, S. Chen, J. Yang, *et al.*, "High-efficiency fbg array sensor interrogation system via a neural network working with sparse data," *Opt. Express* **31**(5), 8937–8952 (2023).
22. S. Ren, S. Chen, H. Xu, *et al.*, "Unsupervised representation learning-based spectrum reconstruction for demodulation of fabry-perot interferometer sensor," *IEEE Sensors Journal* (2023).
23. Z. Cao, S. Zhang, T. Xia, *et al.*, "Spectral demodulation of fiber Bragg grating sensor based on deep convolutional neural networks," *J. Lightwave Technol.* **40**(13), 4429–4435 (2022).
24. W. Sheng, X. Yin, J. Wen, *et al.*, "Accurate and fast calibration for fbg demodulation based on deep learning and ensemble learning," *Opt. Laser Technol.* **172**, 110476 (2024).
25. N. Ren, Y. Yu, X. Jiang, *et al.*, "Improved multi-grating filtering demodulation method based on cascading neural networks for fiber Bragg grating sensor," *J. Lightwave Technol.* **37**(9), 2147–2154 (2019).
26. S. Hu, X. Chen, W. Ni, *et al.*, "Distributed machine learning for wireless communication networks: Techniques, architectures, and applications," *IEEE Commun. Surv. Tutorials* **23**(3), 1458–1493 (2021).
27. D. Vimalajeewa, C. Kulatunga, D. P. Berry, *et al.*, "A service-based joint model used for distributed learning: Application for smart agriculture," *IEEE Trans. Emerg. Topics Comput.* **10**(2), 838–854 (2021).
28. S. Patro and K. K. Sahu, "Normalization: A preprocessing stage," *arXiv*, (2015).
29. J. Xu, Z. Li, B. Du, *et al.*, "Reluplex made more practical: Leaky relu," in *2020 IEEE Symposium on Computers and communications (ISCC)*, (IEEE, 2020), pp. 1–7.
30. E. Fan, "Extended tanh-function method and its applications to nonlinear equations," *Phys. Lett. A* **277**(4-5), 212–218 (2000).
31. S.-W. Huang, C.-T. Lin, S.-P. Chen, *et al.*, "Auggan: Cross domain adaptation with gan-based data augmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, (2018), pp. 718–731.
32. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv*, (2014).