Measurement 228 (2024) 114391

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Application of machine learning in optical fiber sensors

Yifan Zhou [a], Ya-nan Zhang [a,b,c,*], Qi Yu [a], Lirong Ren [a], Qi Liu [a], Yong Zhao [a,b,c,*]

[a] College of Information Science and Engineering, Northeastern University, Shenyang 110819, China
[b] Hebei Key Laboratory of Micro-Nano Precision Optical Sensing and Measurement Technology, Qinhuangdao 066004, China
[c] State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China

ARTICLE INFO

Keywords:
Machine Learning (ML)
Optical fiber sensor
Distributed optical fiber sensor
Fiber Bragg grating (FBG) sensors
Spectrum demodulation

ABSTRACT

In recent years, with the increasing demand for intelligent society, intelligent photonics has developed rapidly. Machine learning (ML), as a subset of artificial intelligence (AI), has played an important role in the intelligent evolution of optical fiber sensors. Its impact extends beyond enhancing sensor performance by introducing innovative problem-solving approaches. Specifically, ML algorithms have become instrumental in signal demodulation and elevating the efficacy of discrete and distributed sensors, and have also greatly promoted the development of optical fiber speckle pattern processing. This paper presents the latest advancements in ML-based optical fiber sensors, outlines the problems faced by conventional demodulation methods and the common ML algorithms applied in optical fiber sensors, and emphasizes key applications. Additionally, this paper delves into the challenges and future development of this emerging research direction.

1. Introduction

Due to advancements in powerful computing tools, hardware, widespread use of cloud data, and the evolution of the Internet of Things, the accessibility of large datasets has significantly improved. This has promoted the emergence of many efficient machine learning (ML) algorithms. These algorithms, aided by continuous advancements in data analysis technology, have demonstrated effectiveness in training models and solving specific problems. If ML algorithms are complex and deep enough in terms of layers, then ML can become so-called deep learning (DL) [1]. Regardless of the name, ML seems to have a wide range of applications in modern society, because the achievements of artificial intelligence (AI) often exceed those achieved by manual control design.

In 2021, Nature Photonics further advanced its exploration of the hot topic of optics/photonics through a focused issue called "Machine learning of Light". As an important type of microphysics, optics seems to have similarities in complexity with some of the "black box" processes in ML. In the Q&A session, David Pille and Aydogan Ozcan discussed the two primary drivers behind the current application of ML in photonics [2]. One direction is to use AI methods on hardware to design optical structures and devices with specific performance. Another direction is to use optical systems as fully optical/hybrid statistical inference models to achieve AI calculations.

Optics and photonics applied to ML computing offer unique advantages, allowing for ultrafast calculations at extreme frame rates and low energy consumption [3–6]. For instance, Feldmann et al. [7] proposed a fully optical neural synaptic system. They use wavelength division multiplexing technology to implement a scalable circuit architecture for photonic neural networks (NNs) and use optical pulse signals to regulate the phase change materials that encapsulate neurons. This phase change material can regulate transmittance and reflectivity through a controllable laser. This system can perform both supervised and unsupervised learning. The work successfully demonstrates direct pattern recognition in the full optical system. This photonic NN has the potential to leverage the inherent high-speed and high bandwidth of optical systems, thereby enabling the direct processing of full optical communication and visual data. However, logic gates, computation, and NNs are all linear (passive) systems that cannot address nonlinear (active) issues. The lack of effective nonlinear optical processes as activation functions for many nodes poses challenges for all-optical implementation [8].

Compared to fully optical systems that integrate perception, storage, and computing, hybrid systems use optical devices as the front-end, combine with back-end electronic hardware systems, and optimize computing through software algorithms [5]. Optical devices can provide richer, more relevant, and more specific inputs for algorithms. This hybrid system will optimize algorithms to maximize trainable optical data utilization, significantly reducing the computational burden on

* Corresponding authors at: College of Information Science and Engineering, Northeastern University, Shenyang 110819, China.
E-mail addresses: zhangyanan@ise.neu.edu.cn (Y.-n. Zhang), zhaoyong@ise.neu.edu.cn (Y. Zhao).

https://doi.org/10.1016/j.measurement.2024.114391
Received 20 December 2023; Received in revised form 5 February 2024; Accepted 22 February 2024
Available online 23 February 2024

electronic hardware systems. In addition, how to design hybrid optical systems is very important. Statistical methods are usually used to monitor and predict the optical performance of the system, ensuring system robustness against issues like noise and unknown inputs. To ensure the performance of the optical system, ML can be used to set the threshold for errors generated in the optical system and optimize the design of the optical system [9]. Furthermore, use the laws of physics as prior knowledge of the ML model to connect the ideas of statistical learning with physical reality, enhancing the detection efficiency of hybrid optical systems. Hybrid measurement systems have broad prospects and application prospects, improving the speed and efficiency of hybrid optical systems, and are one of the influential exploration directions of ML in optics.

There are various forms of front-end data acquisition devices in hybrid optical systems, including microscopes, cameras, and optical fiber sensors. Optical fiber sensors, a type of optical sensor utilizing the principle of total light reflection, have numerous advantages such as high sensitivity, low loss, low cost, ease of operation, resistance to electromagnetic interference, and suitability for remote and online monitoring [10]. Fig. 1 shows the use of optical fiber sensors to collect signals, which can be analyzed by ML. Optical fiber sensors can be categorized into discrete/point and integrated distributed types based on the number of sensors. Discrete sensors can achieve high measurement accuracy through structural design, and can also realize multi-parameter sensing through surface materials design. Plug-and-play online sensing and monitoring have found widespread research on the applications of optical fiber sensors in wearable healthcare and industrial process monitoring [10]. The reason why distributed sensors are becoming increasingly popular is that they can continuously monitor objects along long-distance optical fibers, which makes them attractive in structural health monitoring of pipelines and bridges, production gas leakage monitoring, and multi-parameter detection of ocean [11]. Conventional demodulation methods make it difficult for optical fiber sensors to face problems such as nonlinear interference and limited measurement range. ML solves the inherent cross-sensitivity issue of some sensors and assists in improving their performance. Another industrial challenge that ML can help solve is the development of optical devices. Training models with “portability” from optical devices to other similar devices in production can help with large-scale manufacturing and assembly systems for optical products [12].

The existing work in ML and optical fiber sensors have focused on specific sensor types or applications [11,13–16], this article provides a broader perspective by covering various sensors and their applications in different fields. The next section will introduce the application and requirements of optical fiber sensors in hybrid optical systems. The third section introduces some ML algorithms, while the fourth section introduces the ML algorithms applied to optical fiber sensors. Finally, a conclusion is drawn, elaborating on the current problems and future development directions of ML in the application of optical fiber sensors.

2. Hybrid optical system based on optical fiber sensors

The article focuses on employing optical fiber sensors as the front-end components in a hybrid optical system. Optical fiber sensors have gained widespread use in monitoring systems due to their simultaneous transmission and sensing capabilities. With the improvement of fiber fabrication technology, the fiber transmission loss is reduced. Optical fiber-based transmission offers long-term reliability and low-cost advantages. When properly packaged, these sensors exhibit high durability and entail lower maintenance costs compared to electronic devices [17].

Optical fiber sensors are prepared in optical fibers and the most widely studied optical fiber sensors include optical fiber interferometer sensors, sensors, optical fiber Bragg grating (FBG) sensors, and optical fiber surface plasmon resonance (SPR) [18–20]. Optical fiber interferometer sensors have been widely studied because of their advantages of diverse structures and simple manufacturing. Because of the changes in the external environment, the effective refractive index during beam propagation changes. This results in a change in the optical path difference, which leads to a change in the interference spectrum. However, the sensitivity of general optical fiber interferometers is not high. Enhancing their sensitivity requires coupling the optical light propagating in the fiber into the surrounding medium through methods like bending, etching, or misalignment [10]. However, it makes the sensor structure more fragile and demands a more elaborate preparation process. The refractive index in the core of the FBG exhibits periodic or quasi-periodic changes along the axis, resulting in light of a specific wavelength (Bragg wavelength) being reflected during propagation. Any change in physical quantities in the sensing area will cause the changes in refractive index or the FBG period, which will cause a shift of Bragg’s reflection wavelength. However, the production of FBG requires expensive and complex operating systems, which makes sensing probes expensive to produce. Optical fiber SPR sensors are covered with a layer of precious metal film on the surface of the optical fiber, and the surface plasma wave is excited when the resonance condition is met. The

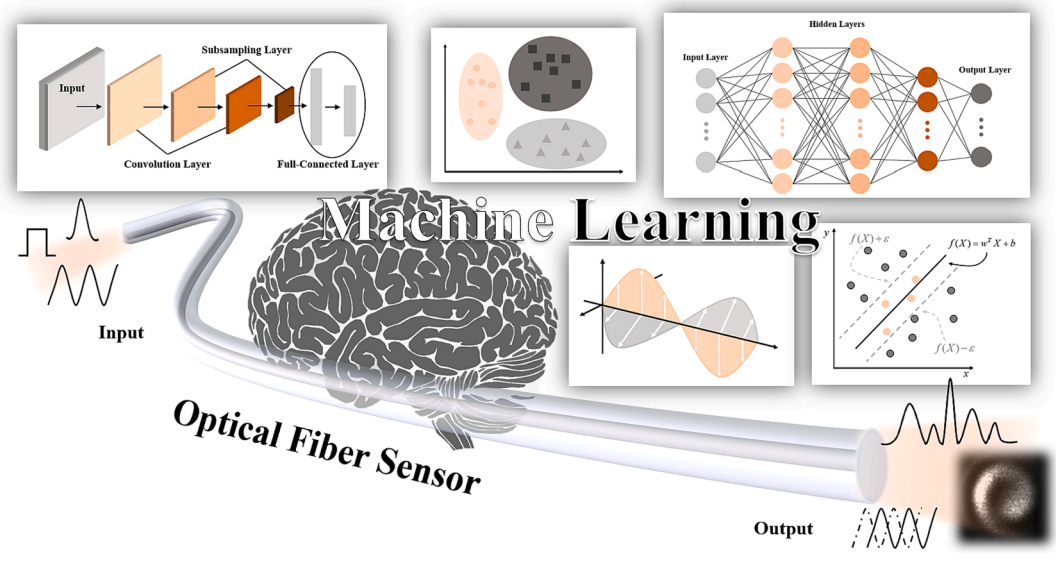

Fig. 1. Using optical fiber sensors to collect datasets, analyzed by ML algorithms.

conduction band electrons on the metal surface undergo resonance after being excited by an external electromagnetic field, usually with high sensitivity. However, the SPR sensor with an all-fiber structure has a wide resonant bandwidth, which is not conducive to high-precision demodulation.

In the common near-infrared and visible light bands, the interference spectrum exhibits a sinusoidal distribution in the frequency domain (wavelength). Strictly speaking, the interference spectrum is a sinusoidal distribution of wave numbers. The conventional demodulation of the spectrum means that utilizes the phase and amplitude information of light to separate and extract the desired frequencies from the spectrum. Conventional demodulation methods can be divided into dynamic demodulation and static demodulation. Dynamic demodulation typically involves more hardware in detection systems to extract the phase signal of the sensor and achieve signal demodulation [21]. The static demodulation method is based on the analysis of intensity changes or shifts of interference dip and has high resolution but slow demodulation speed. In the interference spectrum, there are multiple interference dips with different orders and similar shapes. The difficulty in extracting spectrum features hinders the demodulation of spectrums. In addition, due to the limitation of the free spectrum range, the spectrum is prone to an "interference dip mixing" phenomenon, which limits its application where high sensitivity and high measurement range are required. Moreover, in the demodulation of multi-parameter sensors, conventional methods are linear and ignore the nonlinear effects between measurement parameters, which limits the accuracy of demodulation. Demodulation of optical FBG sensors typically calculates the shift of Bragg wavelength. In multi-parameter measurement, a sensor array is formed by multiplexing the FBG sensors into one or more optical fibers. However, the conventional demodulation of the FBG array requires a fixed operating wavelength range for each sensor and the range cannot overlap [22]. Although the number of FBGs can be increased by using a wavelength division multiplexing system, it is also limited by the operating range and total bandwidth of the light source [23]. Using ML algorithms for spectrum demodulation allows automatic feature learning and extraction, enabling comprehensive spectrum analysis.

Distributed sensors are a network of sensors that utilizes the advantages of optical fiber sensors for long-distance transmission and enables multi-sensor multiplexing within the same system. They can offer high spatial resolution for temperature, strain, and acoustic detection [24–26]. However, in distributed transmission, the system is more susceptible to interference, leading to chaotic signals, particularly evident in long-distance transmission scenarios. Taking wearable sensor devices as an example, the human body signal patterns generated by different individuals are weak and only have slight differences between them. The longer transmission process brings more unknown interference and some subtle signals may be covered by noise. Successful identification of individuals not only requires high-quality data (such as high signal-to-noise ratio (SNR)) but also powerful data analysis algorithms. Moreover, in long-distance optical fiber transmission, the capacity of the optical transmission system is limited. Introducing compensation methods through conventional methods such as digital signal processing will increase hardware complexity. ML algorithms can achieve compensation for systems without increasing hardware complexity.

Since multimode fibers (MMFs) are highly scattering media, each mode that propagates independently in MMFs has a different speed. Because the light propagates along the fiber, resulting in a mixture of amplitude and phase, the output of the MMF at the other end appears as randomly arranged bright and black spots, known as speckle patterns. Due to the susceptibility of different modes of propagation in optical fibers, this speckle pattern is highly sensitive to external interferences such as vibration, temperature, or strain [27]. With the development of optical fiber networks, the use of optical fibers as a means of image propagation has gained more favor, but the accompanying noise issues and other impacts have become the main factors limiting the propagation of high-definition images. The application of ML in pattern recognition or pattern reconstruction provides a faster speed and high accuracy for analyzing large datasets [28–30]. Provide clearer analysis for classifying data and processing complex data, especially in the demodulation of distributed signals [21]. In addition, the ML algorithm is suitable for signals with large degrees of freedom and has been applied to quantum optics [29], optical communication [28], and super-resolution microscopy [30]. Additionally, in the application of optics, ML assists in analyzing complex multidimensional and nonlinear optical phenomena. Modeling using conventional methods is very challenging and sensitive to noise. Although we can eliminate some impacts through the design of spatial light paths, it will increase hardware complexity. By using the ML algorithm, it is easier and more consistent to analyze these data, ultimately achieving regulation of complex processes.

3. Examples of ML algorithms

As more and more complex process data is collected by discrete sensors or distributed sensors, the data preprocessing procedure requires more advanced analysis methods to extract useful information from them. ML, as an important branch of AI, can effectively handle data with complex characteristics like multidimensional and nonstationary, and provide powerful tools for meaningful interpretation of sensor data [8]. Many algorithms have been widely applied to various learning tasks in different applications, achieving many achievements [31–47]. This paper mainly introduces several algorithms that have been applied in optical fiber sensors, including linear regression [31], logistic regression (LR)[32], gaussian process regression (GPR) [33,34], decision tree (DT) [35,36], random forest (RF) [37,38], support vector machine (SVM) [39–42], K-means [43], and NN [44–47]. Fig. 2 provides schematic diagrams of some of these ML algorithms. In this paper, we classify them into three categories: regression, classification/clustering, and NNs. By introducing these algorithms, one can better understand the principles of ML and its applications in optical fiber sensors.

In different works [48–50], the above-mentioned methods can be summarized in different ways from different perspectives. For example, in reference [49]. The authors tend to divide the ML algorithms according to their training process and the authors divide these methods into unsupervised learning, semi-supervised learning, and self-supervised learning. While in reference [48]. the authors tend to divide the methods based on their implementation purposes, Thus the methods are divided into regression, classification, clustering, and so on. While in another reference [50], the authors tend to divide these methods in a more specific way. As for a researcher who wants to apply these methods to fulfill their actual needs as quickly as possible. What this method can do is the most important thing that they need to know. Thus, classifying these methods from the functionality perspective is more suitable for researchers, especially for those whose own areas of expertise are not related to the algorithm. Thus, in this work, we will summarize the above-mentioned methods from the functionality perspective of what they are used to do.

3.1. Regression

Regression methods are usually used to evaluate the relationship between the input and the output, and then quantitatively represent such a relationship with a regression model. Then, with this regression model, one can predict one or multiple variables with the corresponding input information. For example, the regression model can be implemented in the soft sensory task [51,52], process monitoring task [53,54], and content analysis task [55]. The key point of these tasks is to predict an important but hard to be observed variable based on the other observed variables. Here is a brief introduction to some typical regression methods.

We first introduce the linear regression method, it is a simple but effective regression method, where the linear means that the relationship between the input and the output is assumed to be linear. Usually,

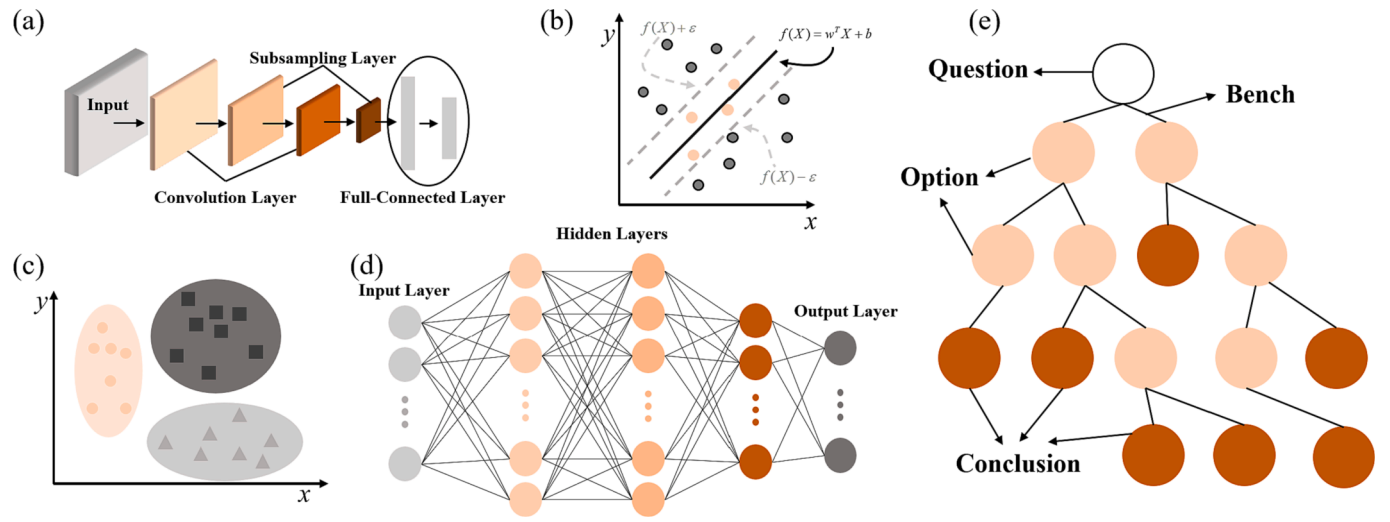

Fig. 2. ML algorithms (a) Convolutional neural network (CNN), (b) SVM, (c) K-means, (d) Artificial neural network (ANN), (e) DT.

due to the linear characteristics, the parameters of this model can be easily obtained. The model can be represented as a linear mapping function.

$$Y = WX + B \tag{1}$$

where X is the input, Y is the output. W and B denote the parameters of this model. Usually, the parameters can be obtained by optimizing the loss function with least squares estimation and its loss function is usually denoted as:

$$loss = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - (w_i x_i + b_i))^2 \tag{2}$$

where the subscript i denotes the i^{th} sample of the corresponding matrix, y_i is the target value of the i^{th} sample, and $\hat{y}_i$ is the predicted value of the i^{th} sample.

The second model we introduced is LR. Although LR is also known as a classification method, it can be used to establish a regression model. It can achieve binary or multi-classification tasks by mapping the output of the linear regression model to a probability value. The model can be also represented as a linear mapping function.

$$z = w^T x + b \tag{3}$$

where w and b are the parameters of this model. Before making a prediction, the LR uses logical functions (such as the sigmoid function) to convert the output into a probability value, where the sigmoid function represents as:

$$g(x) = \frac{1}{1 + e^{-z}} \tag{4}$$

The third model we introduced is GPR, which is a non-parametric model. This model uses kernel tricks to fit the relationship between the input and the output. GPR will use the entire sample information to perform predictions. Consequently, the GPR is not suitable for fitting the amount of data. The GPR model is commonly used to solve nonlinear, low-dimensional, and small sample problems. By adjusting and selecting appropriate kernel functions based on demand, the best prediction performance can be achieved. The brief model function of GPR is represented in Eq(5).

$$f(x) = GP(m(x), \sum(x)) \tag{5}$$

where GP denotes a Gaussian process, $m(x)$ and $\sum(x)$ denote the mean value and the covariance value of the process.

The prediction result of GPR is illustrated in Eq(6), where the typical solution and the model structure can be found in reference [56].

$$\begin{aligned} m(x^*) &= K_{Xx^*}^T (K_{XX} + \delta^2 I)^{-1} Y \\ \sum(x^*) &= \delta^2 + k(x^*, x^*) - k_{Xx^*}^T (K_{XX} + \delta^2 I)^{-1} k_{Xx^*} \end{aligned} \tag{6}$$

where K_{Xx^*} denotes the kernel matrix with the element $k(X,x^*)$, and $k(X, x^*)$ denotes the kernel function (the kernel function is selected manually) with input pair X and x^*, δ^2 denotes the noise variance; Y and X denote the historical target value and input variables.

3.2. Classification/Clustering

Different from the regression, the classification and clustering methods tend to mark each sample with an integer so one can divide them into different parts based on these integers. Thus, from the functionality perspective, classification and clustering methods can be summarized in the same class. However, there are differences between them. Specifically, although they are usually used to do the same task, which is to separate the data into different parts, their model fitting procedures are different. Here is a brief illustration of the differences. In general, classification methods need the labeled data to develop a useful model in advance, and this is a supervised learning procedure. The clustering methods only need the modeling data that comes without corresponding labels or responses, and this is an unsupervised learning procedure. Although they have differences, considering that from the application perspective, they can do similar tasks, we summarized them together in this part. Here, we provide a detailed introduction to the two most representative methods: SVM (classification) and K-means (clustering).

The classification method SVM handles classification tasks by mapping data into a high-dimensional space and finding a hyperplane to maximize the space between different categories. As shown in Fig. 2 (b), the linear equation of the hyperplane is $w^T x + b = 0$, where w is the weight parameter and b is the bias, it determines the distance between the hyperplane and the data point. Therefore, the loss function of SVM is indicated as:

$$\begin{aligned} &\max_{w,b} \frac{a}{\|w\|} \\ &s.t. y_i(w^T x_i + b) \geqslant a, \forall i \end{aligned} \tag{7}$$

SVM has a strong ability to separate the data which is linearly inseparable. By selecting different kernel functions, they can transfer the linearly inseparable data into a high-dimensional nonlinear space, and

thus transfer the task into a nonlinear classification task. However, SVM also has disadvantages, which are the training cost as the data increases (proportional prediction time to the number of support vectors) and high computational complexity.

The clustering method K-means divides the given dataset into K clusters, the number of K is set manually, and continuously allocates sample points to each cluster. During the fitting process, the method will update the center of each cluster after each allocation is completed, and the center of each cluster is calculated based on the average distance contained in the cluster. The target function can be represented as:

$$T = \sum_{i=1}^{k} \sum_{x_i \in w_i} \|x_i - c_i\|^2 c_i = \frac{1}{n} \sum_{x_i \in w_i} x_i \tag{8}$$

where c_j is the center of the j^{th} cluster.

This algorithm attempts to find the K data points that minimize the distance T. The specific steps of K-means are as follows:

Input: Number of clusters K, and the dataset to be processed;

Output: The cluster results.

Step 1. Randomly select K samples as the center point of the initial cluster;

Step 2. Divide each sample into the closest clusters;

Step 3. Update the cluster center c_j for each cluster;

Step 4. Repeat Step 2. and Step 3. until the T converges or the cluster center c_j is stable.

Although this method is easy to implement, it also has disadvantages. In practice, it is necessary to manually set the K value. However, with the unreasonable K values, the clustering performance may be very poor. Specifically, the inappropriate selection of initial clustering center points may lead the model to converge to local minimum values. Besides, this method converges slowly while facing large-scale data, and it is sensitive to outliers and noisy data. Therefore, when applying this method, it is necessary to choose an appropriate clustering center and clean the outliers.

3.3. NNs

Here, leaving aside the functionality of methods, based on the concept of reference [45,57–60], different from the regression and classification/clustering methods, we have a general definition for the NNs. A network of neurons is the composition of nonlinear functions of two or more neurons, where a neuron is also a nonlinear, parameterized bounded function. Typically, ANN [58–61], recurrent neural network (RNN) [62], CNN [63–66], and backpropagation (BP) network [67,68] all belong to the NN model. In practice, we know that different NN models can do different tasks, like identifying multiple different images and then finding out what each of them is (classification/clustering) [69,70], predicting a following value of a time series (regression)[71]. What the NN model can do depends on the model designer (Most NN models can only do classification tasks, since this model needs to be trained with labeled data). Thus, we list and introduce this model solely here. In practice, the NN model usually has a flexible model structure, which means that the number of neurons and layers can be designed and changed according to the requirement. Consequently, as for a NN model, there are too many hyper-parameters that can be adjusted, which makes the design of a model over complicated. Typically, for the application of optical fiber sensors, there are three widely used models, which are ANN, RNN, and CNN. We will briefly introduce them here to illustrate the advantages and disadvantages of the NN model.

ANNs can also be referred to as multi-layer perceptron, as the input only undergoes unidirectional forward processing, hence they are also known as feedforward NNs. The ANN consists of the input layer, hidden layer, and output layer. The input layer receives the input and usually has one layer, the hidden layer consists of a single or multiple layers to process the input features according to the needs of the task, and the output layer is responsible for generating the result. Each layer has numerous neurons. Basically, each layer needs to train weights. The output of each neuron is the sum of the input value multiplied by the weight. Nonlinear activation functions can achieve nonlinearity between the input and output of neurons, enhancing the expression ability of ANNs. Moreover, in image classification issues, ANNs need to convert two-dimensional (2D) images into one-dimensional (1D) vectors. As the image size increases, the number of network parameters increases dramatically. This algorithm will also miss the spatial features of the image.

RNNs add loop connections on the hidden layer, which can capture sequential information in input data. However, conventional RNNs still face the problem of being unable to capture long-term dependencies and gradient explosions. More advanced models such as gated recurrent units (GRUs) [72,73] and long short-term memory (LSTM) networks [74,75] have replaced conventional RNNs. Typically, GRU controls the flow of information through learnable gates, better capturing dependencies with larger time intervals in time series. LSTM effectively filters redundant information using storage units, achieving more efficient information extraction and avoiding the problem of vanishing or exploding gradients in RNNs.

CNNs are the most common in image processing because they can effectively capture and recognize features in images. CNN transforms a sliding window on the input image and uses convolution operations to extract relevant features from the image, which is the function of convolutional layers. The pooling layer is used to reduce the size of the feature, reduce computational burden, and address overfitting issues caused by redundant features. Finally, after passing through a fully connected layer and unfolding into a 1D vector, the final recognition probability is obtained after another calculation. The final recognition probability assists in determining the final classification result in the classification task.

All the models tend to fit an accurate relationship between the input and the output, and thus, it is important to evaluate the accuracy of the fitting result. Besides, in practice, these evaluation indexes are usually set as the loss function of these models. Specifically, a suitable loss function can benefit the training process and thus one can obtain a model with good performance. Researchers choose different indicators according to different needs. The following is an introduction to several common evaluation indicators. R^2 is the correlation coefficient, which indicates the degree of agreement between experimental data and the fitting function. The closer the R^2 value is to 1, the higher the consistency, and the closer it is to 0, the lower the consistency. R^2 is represented as:

$$R^2 = 1 - \left[\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \right] \tag{9}$$

where $\bar{y}$ is the average value. R^2 is suitable for LR models, but its limitation is that it only considers the relationship between the dependent variable and the independent variable, without considering other influencing factors. Therefore, when choosing R^2, it is necessary to consider whether the model only explains a portion of the variation in the data.

The mean absolute error (MAE) is used to measure the average absolute error between predicted values and real values. A smaller MAE indicates a better model. MAE is represented as:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{10}$$

MAE is more sensitive to outliers as they can seriously affect the average error. However, the calculation of MAE is relatively simple and has good applicability for continuous variable data.

The mean square error (MSE) is expressed as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{11}$$

MSE is the average square of the prediction error, with a higher penalty for larger errors. The result of MSE differentiation is easier to calculate than MAE, so MSE is usually used as the loss function when training NNs and other models that require differentiation. MAE is commonly used for regression problems, while MSE can be used for both regression and classification problems.

Root mean squared error (RMSE) is the square root of MSE. RMSE is used to measure the deviation between the predicted value and the real value. RSME is represented as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{12}$$

RMSE imposes significant penalties for large errors. In RMSE, the error is squared before averaging, which means that RMSE assigns higher weights to larger errors. This indicates that RMSE is more useful when there are major errors that greatly affect the performance of the model.

Choosing appropriate indicators needs to be considered based on specific application scenarios and models. MAE, MSE, and RMSE are mainly used in regression problems to measure the error between predicted and actual values. The smaller the MSE, RMSE, and MAE, the better the performance of the regression model. R^2 is suitable for evaluating the degree to which the model interprets the data; MAE is suitable for continuous variable data; MSE is suitable for situations that are sensitive to errors; RMSE is suitable for evaluating the performance of regression models compared to other stochastic models.

4. Application of ML in optical fiber sensors

The steps of using ML for spectrum demodulation are shown in Fig. 3, including data acquisition, data preprocessing, feature extraction, model selection and training, model evaluation, and model application. Data acquisition is the collection of raw spectrums under different responses using a hybrid optical system. Data preprocessing is aimed at improving data quality, including steps such as noise elimination, filtering, and spectrum smoothing. Feature extraction often involves extracting the most relevant feature information to the target response from complex spectrums, with principal component analysis being the most commonly used. Model selection and training include selecting appropriate machine learning algorithms to train the extracted features and establishing predictive models. Model evaluation is the process of selecting appropriate evaluation indicators based on objectives to evaluate a model, and optimizing the model based on the evaluation results to improve predictive performance. The model application uses trained models for tasks such as predicting and classifying unknown spectrums. In spectrum demodulation, the training and evaluation of models are usually carried out separately, by dividing the data into training and testing sets. The training set is used to train the model, while the test set is used to evaluate the performance of the model. Compared to the million-level datasets obtained from industrial sites, the sample sizes involved in spectrum demodulation are relatively small. For small-scale data, the data is usually divided into 60 % −80 % of the training set, and the rest is used for the testing set. To better adjust the model, a validation set can also be added. The usual partitioning ratio is to use 70 % −80 % of the data for the training set, 10 % −15 % for the validation set, and the remaining portion for the testing set. This ratio can be adjusted based on the complexity of the specific task and the amount of available data.

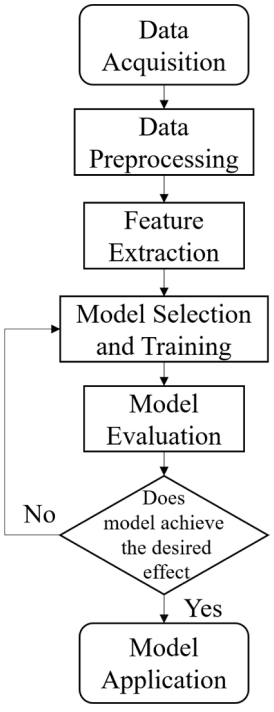

Fig. 3. Machine learning based spectrum demodulation flowchart.

The entire process of using ML algorithms for spectrum demodulation not only requires knowledge of spectroscopy, ML, and data processing but also requires expertise in experimental design and data analysis. In practical applications, data labeling and model overfitting issues also need to be considered to ensure the accuracy and reliability of the results.

4.1. Discrete/Point optical fiber sensors

Discrete/point sensors have a relatively clear signal compared to distributed sensors. However, taking interferometric sensors as an example, when their sensitivity is high or the measurement range is large, there will be an “interference dip mixing” phenomenon. Conventional demodulation methods cannot solve this problem, which limits the sensor’s application in situations where high sensitivity and large measurement range are required. In multi-parameter sensing, conventional linear demodulation methods cannot solve the impact of nonlinear effects between different measurement parameters. ML algorithms can not only conveniently solve those problems, but also effectively reduce noise in different applications. To facilitate an intuitive comprehension for readers, we have summarized the sensors mentioned in this section from the algorithm types, sensor types, measurement objects, sensitivity, and algorithm effects. The summary is presented in Table 1.

4.1.1. Demodulation performance

There are two conventional demodulation methods for discrete optical fiber sensors: static demodulation and dynamic demodulation. Static demodulation limits the speed of demodulation, while dynamic demodulation brings higher complexity hardware requirements, which is not conducive to the development trend of real-time, miniaturization and portability of sensing systems. With the assistance of ML, the signal demodulation process of sensors is not only more convenient, but also brings improvements in demodulation accuracy, measurement range, and anti-interference performance to the sensors.

To achieve a larger dynamic range, Yin et al. utilized nonlinear regression to demodulate the spectrum of the interferometric optical fiber sensors [76]. However, they are only focused on nonlinear regression in a polarization state, with each polarization state corresponding to different parameters. The increase in polarization state not only easily reduces measurement accuracy, but also requires the use of polarization maintaining fibers or the addition of polarizers to ensure the existence of only one polarization, which increases the system costs.

Table 1
Discrete/point sensors combined with different ML algorithms.

	Type of algorithm	Structure	Measurement	Sensitivity	Evaluation of model performance	Ref
Demodulation performance	ANN	SMF-NCF-HCF-SMF	Curvature (Regression)	16.34 dB/m^{-1}	Expand dynamic range from 0.5-1.5 $m^{-1}$to 0.5–3.87 m^{-1}	[77]
	ANN	SMF-HCF-SMF	Curvature (Regression)	–	Expand dynamic range to −13.67–13.67 m^{-1}	[78]
	GPR	FP	Strain (Regression)	288 kHz/με	Simplify the hardware design of the detection system	[79]
	LSTM	Michelson interferometer based on SMF	Glycerol concentration (Classification)	–	Accuracy of 97.5 %	[80]
	SSA-LSTM	SMF	Stress (Regression)	0.286 rad/MPa	Expand the measurement range (to 3492 με) and improve the accuracy	[81]
	GAF-LSTM	Side hole fiber	Pressure (Regression)	20.57 nm/MPa	R^2 was 0.9999908 and RMSE was 4.365 × 10^{-3} MPa	[83]
	K-NN	SMF	Liquid volume (Regression)	RMSE of 0.211 mL	Accuracy of 99.4 %	[84]
	SVM	ECF	Directional bending (Classification)	–	100 % identified accuracy and estimated the flexible rod trajectory	[85]
	SVM	FP (SMF-SMF-SMF)	Reflective index (Classification)	–	Large measurement range and anti-interference	[86]
	ANN	FP (UV glue, hollow silica tubes, and SMF)	Pressure (Regression)	–	Reconstruction error of 0.039 nm	[88]
	ELM	FBG	Strain (Regression)	–	Reduced the offline training time and enhanced the detection accuracy	[89]
	QPSO-KRELM	FBG	Strain (Regression)	–	Few sample points provide high precision	[91]
	CNN	TFBG	Reflective index (Regression)	MSE was 2.818 × 10^{-7} RIU	Test accuracy of 99.82 %	[92]
	GPR	FBG	Weight (Regression)	The average error in weight estimation of 2.7 g	improve the measurement speed and accuracy	[93]
	DT	FBG	Strain (Classification)	–	Accuracy of 94 %	[94]
	OS-ELM	FBG	Temperature (Regression)	The absolute error of the long/short-term stability of ~0.5 ℃	Improved the accuracy and generalization performance	[96]
	CNN	FBG	Temperature (Regression)	MSE was 0.1080 °C	Accuracy of 99.95 %	[97]
	DT	FBG	Temperature (Regression)	–	Accuracy of 99.83 %	[99]
	GPR	FBG	Temperature (Regression)	–	Lower error and higher detection accuracy	[100]
Multiparameter demodulation	MUSIC and NN	SMF-MMF-SMF	Strain and curvature (Regression)	–	R^2 values of 0.994	[101]
	K-NN	Four FBG	Magnetic field position and intensity (Regression)	–	Accuracy of magnetic field position classification was higher than 86 %, and the error of intensity estimation was less than 5 %	[102]
	Gradient boosting	FBG	Strain and temperature (Regression)	–	Accuracy of 90 %	[103]
	SVR	Tapered fiber half-covered with PDMS	Temperature, salinity, and pressure (Regression)	−2.312 nm/℃, 0.631 nm/‰, and 3.775 nm/MPa	Errors of TSP were 10.67 %, 5.25 % and 16.76 %, respectively	[104]
	NN	TFBG	Strain and temperature (Regression)	–	RMSEs were 2.31 ℃ and 21.84 με	[105]
	K-NN	U shaped fiber	Curvature and pressure (Regression)	–	Accuracy of 99.375 %	[107]
	SVR	PDMS covered SMF-MMF-SMF	Radial artery and brachial artery (Regression)	The average deviation was 0.06 mmHg and standard deviation was 1.54 mmHg	R^2 values of 0.98	[108]
	BP	Two FBG	Force and temperature (Regression)	9 pm/N and 9.32 pm/℃	Gesture recognition accuracy of 97.02 %	[109]
	CNN	FBG	Curvature and 3D-shape (Regression)	–	Accuracy of 91 %	[110]

Zhu et al. designed a highly sensitive curvature sensor and established a simple ANN to demodulate the spectrum [77]. They addressed the limited dynamic range problem in conventional phase demodulation methods by establishing a one-to-one correspondence between ML and curvature. The dynamic range of sensors in the range of 0.55 to 1.45 m^{-1} was extended to 0.55 to 3.87 m^{-1}. Based on this idea, a structure of fused capillary tubes between two single-mode fibers (SMF) was prepared, and ANN was used to achieve wide dynamic range directional bending sensing without sacrificing sensor performance [78]. Moreover, by reducing sampling points (as shown in Fig. 4 (a), (b)) and inputting different wavelengths spectrum (as shown in Fig. 4 (c), and (d)), the ANN model was used to analyze the sensor spectrum. The results show that ML-based analysis does not require careful selection of the bandwidth of the light source and the resolution of the spectrometer. Fig. 4 (e) shows the error histogram between the real curvature and predicted curvature obtained from the model, which indicates most of the data samples were predicted with an error of less than 0.07 m^{-1}. Zhu et al. utilized the overlap of vernier interferograms and micro delay vernier interferograms to obtain data through the self-vernier effect [79]. The application of GPR significantly reduces the frequency observation range and the sampling points number, which improves the measurement speed without affecting the sensor measurement accuracy. Fig. 4 (f) is an overview of the algorithm.

Unlike conventional RNNs, LSTM effectively filters redundant information using memory units, leading to more efficient information extraction. Xue et al. prepared a Michelson interferometer, and to improve demodulation accuracy and avoid the impact of tilt angle selection on a measurement range and demodulation accuracy [80], the LSTM algorithm was used for demodulation. Fig. 5 (a) shows the comparison of real data and predicted results, the insert is the RMSE (red) and loss function (blue) during the training process. Compared with conventional peak wavelength tracking, it was more stable and accurate, achieving a prediction accuracy of 97.5 %. Yu et al. designed a new strain demodulation model based on the LSTM algorithm and used the sparrow search algorithm (SSA) to optimize the LSTM model [81]. The addition of a sparrow search algorithm effectively reduces the feature extraction time of the network and reduces noise interference by extracting features more relevant to test parameters, preventing LSTM from falling into local optima. The optimized demodulation model can accurately illustrate the relationship between the strain applied to the sensor and the phase difference. The demodulation sensitivity has been increased from 0.257 rad/MPa to 0.286 rad/MPa. However, the sample size was relatively small, the amplitude of gestures changed significantly, and the signal features were pronounced. Filosadengren et al. proposed a wearable FBG device for predicting respiratory flow, using LSTM to construct a prediction model [82]. The RMSE of the test set was 10.99 L/min when only two FBGs were used. The work considers breathing situations under different postures but does not consider the additional pressure on FBG sensors during additional challenging activities such as exercise. More heterogeneous subjects and more respiratory behaviors should be considered to improve the performance of the sensing system. Mei et al. used an improved LSTM to expand the measurement range of optical fiber pressure sensors [83]. They used a low-cost spectrometer to record and construct 1D spectrum data using scaled spectrum intensity, the LSTM model predicts an R^2 of 0.9996148 and RMSE of 2.559 $\times$ 10^{-2} MPa. Fig. 5 (b) is the schematic structure of the LSTM memory block. Utilizing the Grauman angle field (GAF) enhancement of the LSTM algorithm to obtain 2D data, reveals more relevant information by establishing 2D data compared to the 1D spectrum. The improved model R^2 was 0.9999908 and RMSE was 4.365 $\times$ 10^{-3} MPa. Compared with the model trained on the 1D spectrum, the prediction error of this model had increased by nearly an order of magnitude, but the demodulation speed has slowed down. In practical applications, a balance needs to be made between accuracy and speed.

The assistance of ML algorithms can help researchers achieve more accurate classification. Duque et al. studied liquid volume measurement of optical fiber sensors [84]. The accuracy of using k-nearest neighbor (k-NN) to determine liquid volume was 99.4 %. Manuel et al. collected a specific spectrum of multimode elliptical core fibers (ECFs) under different bending degrees and used SVM for identification [85]. Fig. 5 (c) is the schematic diagram of the entire experimental system and process of the algorithm. The motion trajectory of the flexible rod can be estimated through algorithms. Rodolfo et al. used SVM to eliminate the ambiguity in sample measurements between effective refractive indices

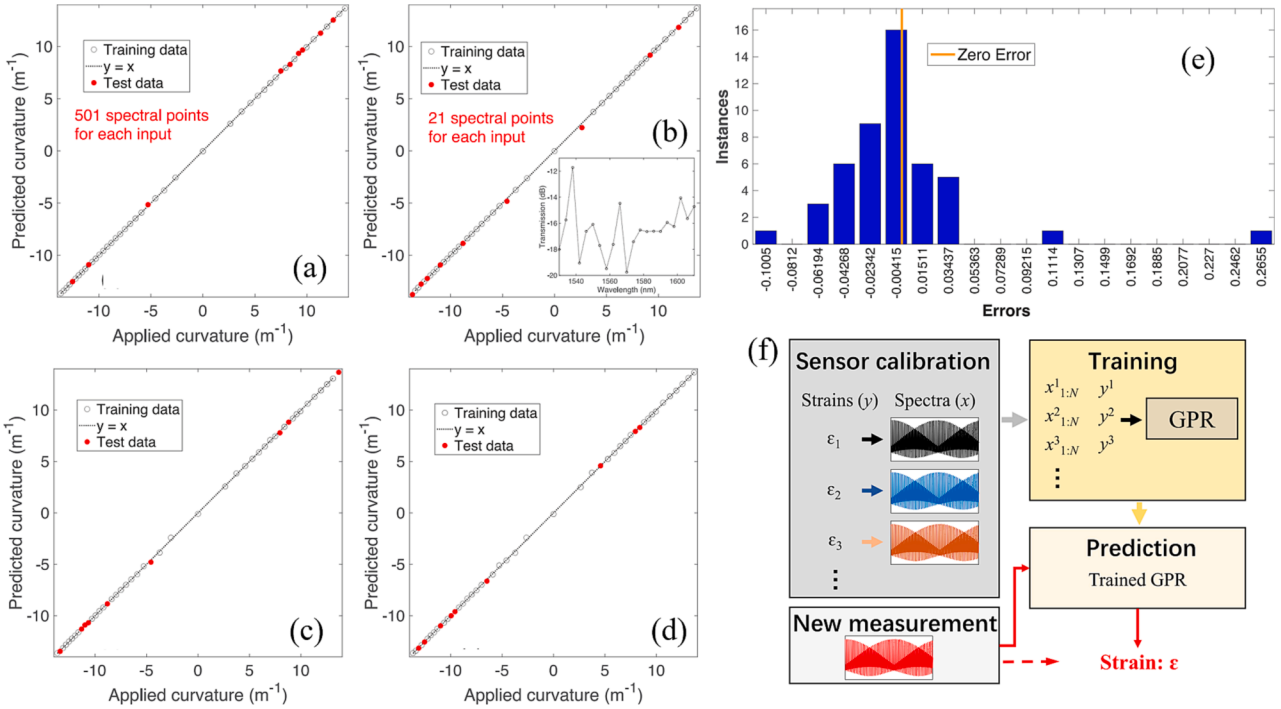

Fig. 4. (a) 501 points, (b) 21 points, (c) 1540–1560 nm, (d) 1590–1610 nm, (e) Error histogram [78], (f) Overview of sensor GPR demodulation methods [79].

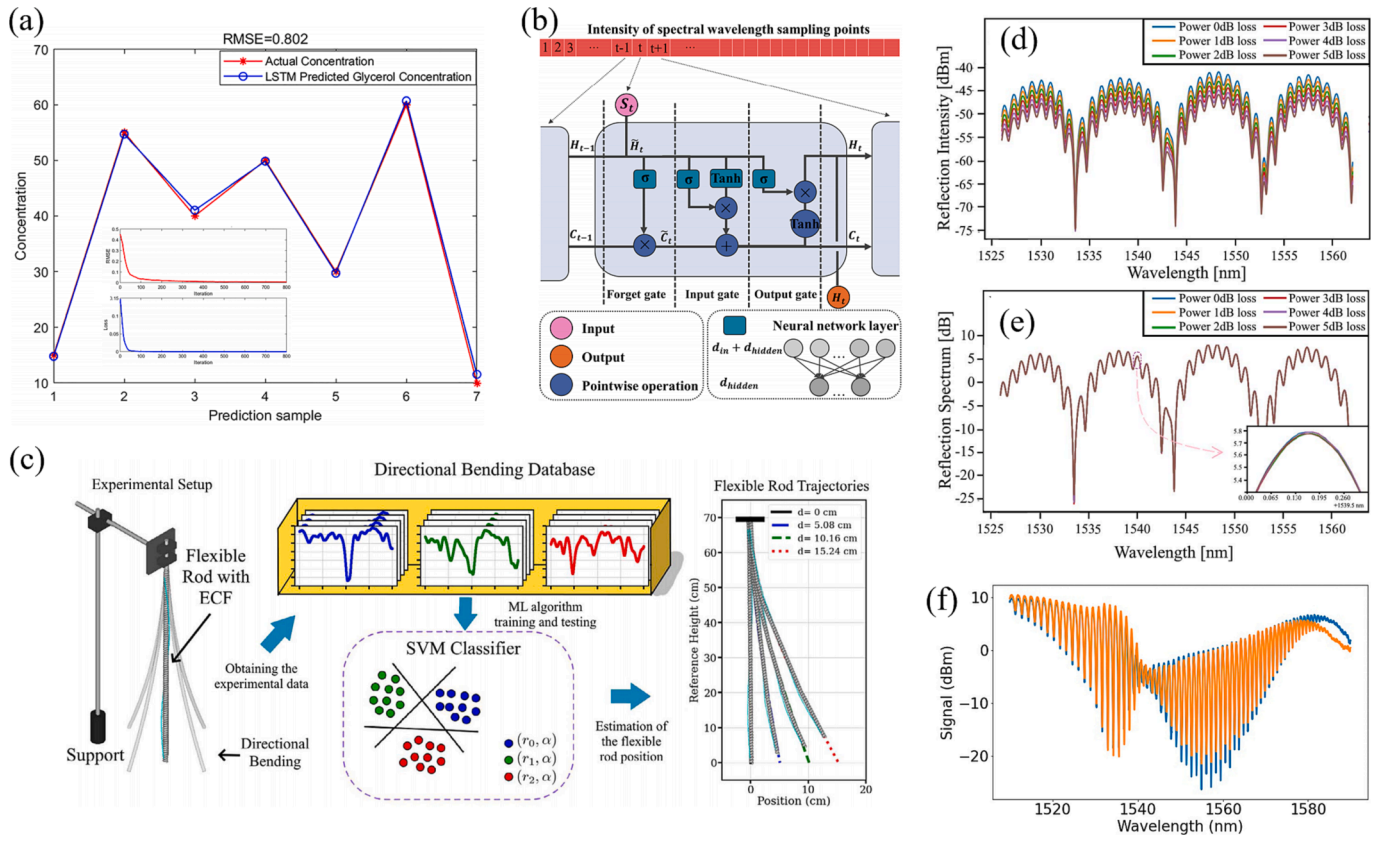

Fig. 5. (a) Comparison of actual and predicted data of different samples (insert is the RMSE and loss function) [80],(b) Schematic structure of LSTM memory block [83], (c) Schematic diagram of the experimental system and algorithm [85], (d) Detection signals under different light source intensities, (e) The detection signal after performing self-reference processing [86], (f) Transmission spectrum of different sensor [87].

above and below the fundamental mode of optical fiber sensors [86]. Moreover, they eliminated the influence of light source intensity on the spectrum (as shown in Fig. 5 (d) and (e)). While their approach showed promising results with 10 experimental samples, its limitation lay in focusing solely on discrete measurement objects. The transferability of ML models is also an important criterion for measuring the practicality of demodulation methods. Chubchev et al. utilized the ML algorithm to assist in SPR sensor data processing [87]. The model realized the transfer between different sensors produced with the same process parameters and achieved a resolution of 10^{-6} RIUs in the demodulation of sensors not involved in model training. However, its transferability was limited, it was only suitable for the spectrum characteristics between sensors (as shown in Fig. 5 (f)) were still similar. Ren et al. proposed a simple spectrum reconstruction method that utilizes an optical system to convert the spectrum into transmission intensity, which reduces requirements for optical instruments [88]. The model based on ANN learned the nonlinear relationship between intensity signals and spectrums, achieving spectrum reconstruction. This work supplements the entire spectrum by adding an unsupervised linear unit to the network, based on nonlinear relationships and masking labels, without the need for additional labeling or computational work. The experimental reconstruction error was 0.039 nm, and the reconstruction range was 73 nm.

The spectrum changes caused by the external environment in FBG sensors are mainly reflected in the shift of Bragg wavelength. The conventional demodulation methods calculate the shift of wavelength to determine the changes in the measured object. ML algorithms directly achieve one-to-one mapping between spectrum and measurement parameters, while reducing the requirements for instruments. Jiang et al. established a regression model and solved the detection problem of Bragg wavelength in FBG sensing networks using extreme learning machines (ELMs). Fig. 6 (a) shows the sensing principle diagram [89]. The Bragg wavelength of the sensor can be accurately detected even when the FBG spectrum overlaps entirely. ELM is an improved algorithm based on single hidden layer (SHL) feedforward networks. ELM produces better generalization performance by obtaining the minimum weight norm and runs fast. The training time of ELM was 56.14 times faster than that of the least squares support vector regression (LS-SVR). The inspection times for ELM and LS-SVR were 0.215 s and 0.578 s, respectively. Elliathy et al. used adaptive threshold algorithms to assist FBG based strain detection in low SNR scenarios [90]. When the ratio of the short-term window average energy to the long-term window average energy of the signal exceeds the set threshold, a peak was detected and interference was identified using LR. The adaptive threshold improves classification performance, with an accuracy of 99.17 % when the SNR was less than 0.5 dB. Xu et al. proposed a FBG demodulation method based on a quantum particle swarm optimization kernel regularized ELM (QPSO KRELM) [91]. As shown in Fig. 6 (d), (e), (f), and (g), the proposed algorithm provides higher demodulation accuracy compared to others. The use of the QPSO algorithm overcomes the problem of the difficult selection of KRELM hyperparameters and reduces the impact of artificial experience on model demodulation accuracy. Reducing the sampling point from 126 to 25 can also achieve high-precision demodulation. Cao et al. proposed a new tilted fiber Bragg grating (TFBG) sensor spectrum demodulating method using residual block CNN [92]. The model can be applied to the demodulation of under the same process parameters sensors with slight spectrum differences through transfer learning, and can still maintain a high MSE at low sampling rates. Pal et al. analyzed the performance of FBG strain sensors using linear regression, SVM, NN, and GPR, as shown in Fig. 6 (h), (i), (j), and (k) [93]. The prediction accuracy of these four algorithms was compared. Mapping between sensor Bragg wavelength changes and application

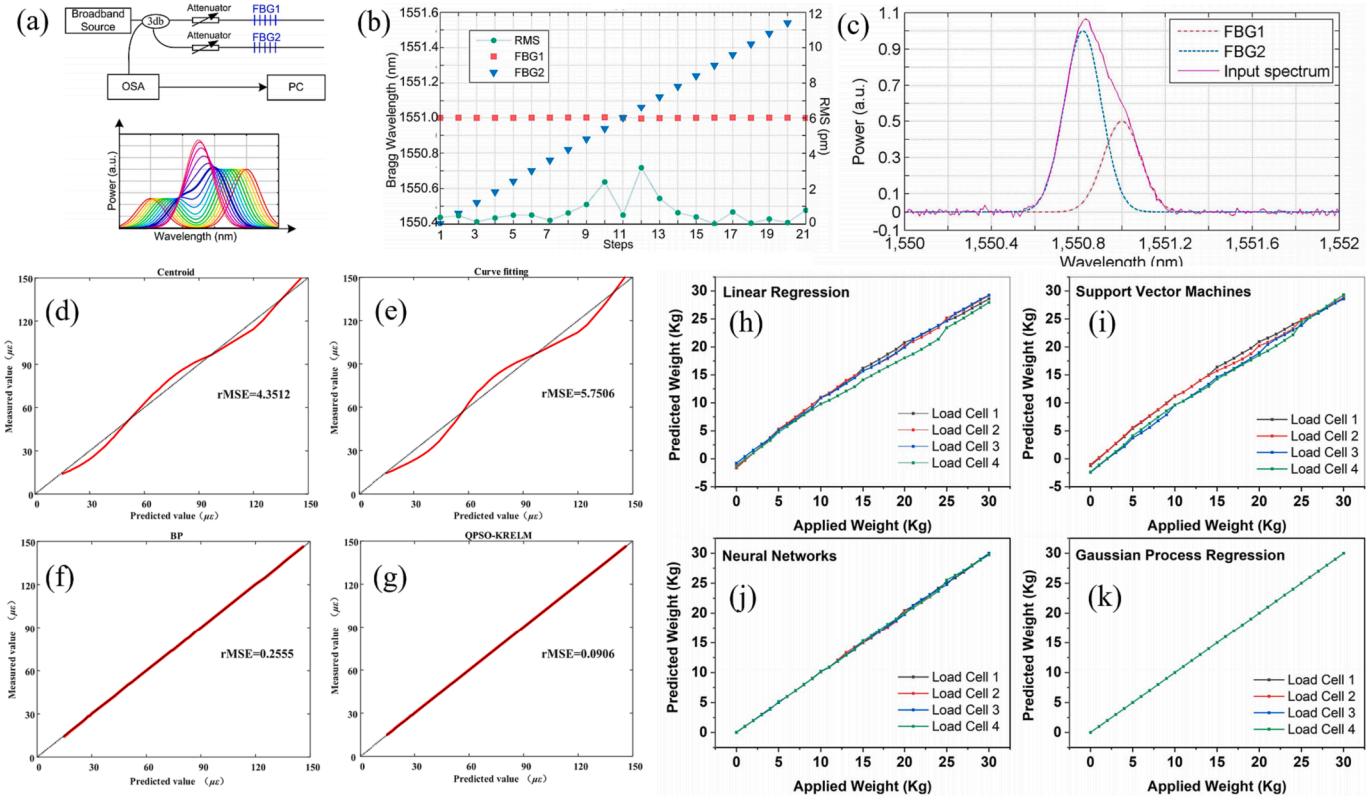

Fig. 6. (a) Schematic diagram system (up) and spectrum of the sensor (down) (b) different strains were applied to FBG2, the Bragg wavelengths shift of the two FBGs, (c) Input and output after applied the ELM [89], generalization performance comparison of the (d) centroid, (e) curve fit, (f) BP, and (g) QPSO-KRELM [91], comparison between the (h) linear regression, (i) SVM, (j) NN, and (k) GPR [93].

weights through ML to improve accuracy and measurement range. Pegorini et al. established a classification system for experimental chewing patterns in ruminants [94]. The FBG sensor measures the biomechanical strain during mandibular movement, and DT was responsible for classifying the relevant chewing patterns. In the experiment of identifying 5 chewing processes, the pattern classification algorithm achieved an accuracy of 94 %.

The calibration process of FBG sensors is crucial to optimizing sensor performance. There is a nonlinear relationship between Bragg wavelength and temperature, and the manufacturing process also affects the sensor's dependence on temperature [95]. Therefore, it is necessary to calibrate the fiber Bragg grating sensors. Shang et al. continuously updated the dynamic calibration model using online sequence limit learning machines, reducing computational complexity and improving prediction speed [96]. The absolute error of both long-term and short-term stability experiments was about 0.5 ℃. During the detection process, the model did not require further training but was constantly updated, reducing tedious calculations and improving prediction speed. Real-time samples can continuously help improve the model, solving the problem of insufficient training samples during static calibration. This method enhances the long-term stability of FBG sensors, ensuring model prediction accuracy and generalization ability. Cao used CNNs to demodulate the FBG spectrums, with a measurement accuracy of 99.95 % and the MSE of 0.1080 ℃ [97]. Ren et al. utilized a genetic algorithm (GA) optimized SVR to achieve high-precision rapid demodulation of large-scale temperature changes [98]. As shown in Fig. 7 (a), the training and testing stages of the GA-SVR model. Obtain the external environmental temperature and growth rate when the external environment reaches thermal equilibrium from the transient spectrum. The algorithm can predict external environmental temperature before the FBG reaches thermal equilibrium. The detection range of the sensor was expanded to 400–1000 °C, with an accuracy of 4.8 °C above 700 °C. The demodulation time has been reduced to about 15 s, which was only 3.14 % of the time for the sensor to reach thermal equilibrium. Fig. 7 (b) and (c) show the percentage error and RMSE of different external environmental temperatures, respectively, indicating that the demodulation

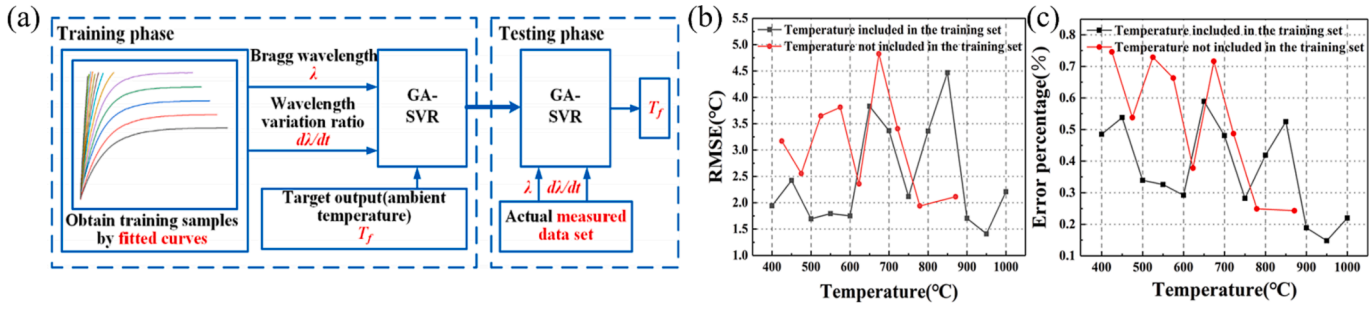

Fig. 7. (a) Training and testing stages of the model, (b) RMSE, and (c) Error percentage at different external temperatures [98].

accuracy of this method was close to the actual situation. Dhanalakshmi et al. used FBG to monitor the temperature of solar photovoltaic panels and compared the demodulation performance of linear regression, RF, and DT. The hyperparameter tuning assisted DT algorithm achieved the best results with an accuracy of 99.83 % [99]. Djurhus et al. proposed an alternative algorithm for temperature measurement signal processing based on FBG, which uses GPR to provide more accurate temperature calculations when there were fewer sampling points in the spectrum [100]. Due to the less time required for measurement, this algorithm helps to perform queries faster.

The conventional spectrum acquisition and processing of discrete sensors are independent, and the accurate and fast identification of multi-point discrete signals poses challenges to conventional methods. From the above works, compared with conventional demodulation methods, the ML algorithm not only improves the measurement range, and achieves higher demodulation accuracy without sacrificing sensor performance. The assistance of the ML algorithm eliminates the influence of noise, making prediction results more accurate than conventional demodulation methods. Even with fewer sampling points (less measurement time and lower instrument requirement), they can achieve reliable accuracy.

4.1.2. *Multiparameter demodulation*

Multi-parameter sensors have complex structures, which leads to greater noise impact on the spectrum, and conventional demodulation methods also need to consider the sensitivity matching between different parameters. These problems bring more difficulties to the multi-parameter demodulation of optical fiber sensors using conventional demodulation methods.

Ushakov et al. improved the processing efficiency of interference signals by using the MUSIC algorithm and used NNs to predict signal components from the covariance matrix distribution of the signal to improve the anti-interference performance of the original signal demodulation [101]. Compared with the fast Fourier transform processing method, it was more suitable for multi-parameter perception and improved tuning accuracy. However, the algorithm analyzed the linear response and did not consider the phenomenon of multi-parameter cross-nonlinear optics. Moreover, the model's accuracy can also be enhanced by prior knowledge judgment or other algorithm assistance. Leal et al. proposed a sensor based on ML assistance that utilizes an array of four FBGs to simultaneously measure magnetic field position and intensity [102]. By using the K-NN algorithm, the Bragg wavelength shift was calculated and the magnetic field position was classified. Estimated magnetic field intensity through signal changes in the FBG at the closest magnetic field position. Because the signal change of the nearest FBG was the most significant. The accuracy of the position classification was higher than 86 %, and the relative error of the magnetic field intensity prediction was less than 5 %. Sarkar et al. reduced the variance of the model by using bagging and RF averaging, enhanced it using AdaBoost and gradient tree enhancement algorithms, and made the final prediction by sequentially combining multiple weak prediction factors, reducing the algorithm's bias [103]. Fig. 8 (a) reveals the basic model of the boosting method. The algorithm uses the independent estimator to train the training spectrum dataset and combines the results of all estimators to make the final prediction. Fig. 8 (b) presents the comparison of the testing and training errors for the four types of methods on both experimental and simulation data. The experimental results have shown that the model is stable and well fitted, and the difference between the testing and training errors of the algorithm was less than 10 %. This achieved 90 % accuracy in predicting strain and temperature.

Compared to conventional demodulation methods, ML algorithms can provide more accurate demodulation. Liu et al. developed an optical fiber sensor for the detection of the temperature, salinity, and pressure (TSP) of the ocean [104]. Due to the impact of cross-sensitivity on signal demodulation, a combination of SVR and sensitivity matrix method was used to reduce measurement errors in the multi-parameter demodulation process. The errors of TSP measurement were 10.67 %, 5.25 %, and 16.76 %, respectively. Fig. 9 (a), (b), and (c) show the comparison of predicted TSP with electronic device measurement results. Because the sensitivity between the three parameters was not on the same order of magnitude, there may be differences in errors under the same demodulation method. Kikuchi et al. used NNs and SVR to determine temperature and strain from the transmission spectrum [105]. Compared with the demodulation method of calculating the shift of interference dip, this algorithm had higher accuracy and was less than one-fifth of the RMSE value obtained by other methods (Fig. 9 (d) and (e)). Reyes used plastic optical fiber sensors to collect spectrum in gait classification experiments [106]. 14 different models analyzed the experimental results. Among them, the RF model with adjacent mean features had the best classification performance, with an average validation score of 90.84 % ± 2.46 %. The monitoring of curvature was crucial for geotechnical structures, bridge safety, and building safety. Huang et al. developed a flexible optical fiber curvature sensor that combines light-emitting diodes and photodetectors to form a measurement system [107]. Bending the optical fiber will change the spectrum intensity. Using supervised learning algorithms including K-NN, SVM, LR, and unsupervised learning algorithms K-means for classification, almost all algorithms have demodulation accuracy above 85 %. The classification results of the K-NN algorithm are shown in Fig. 9 (h), (i), and (j).

Human health monitoring is a dynamic process, and due to the complexity of the measurement environment and the diversity of feature parameters, signal demodulation of multi-parameter human health sensors poses challenges. Pang et al. proposed a dual channel inter-mode interferometer sensor encapsulated in polydimethylsiloxane (PDMS) for simultaneous detection of brachial and radial artery blood pressure

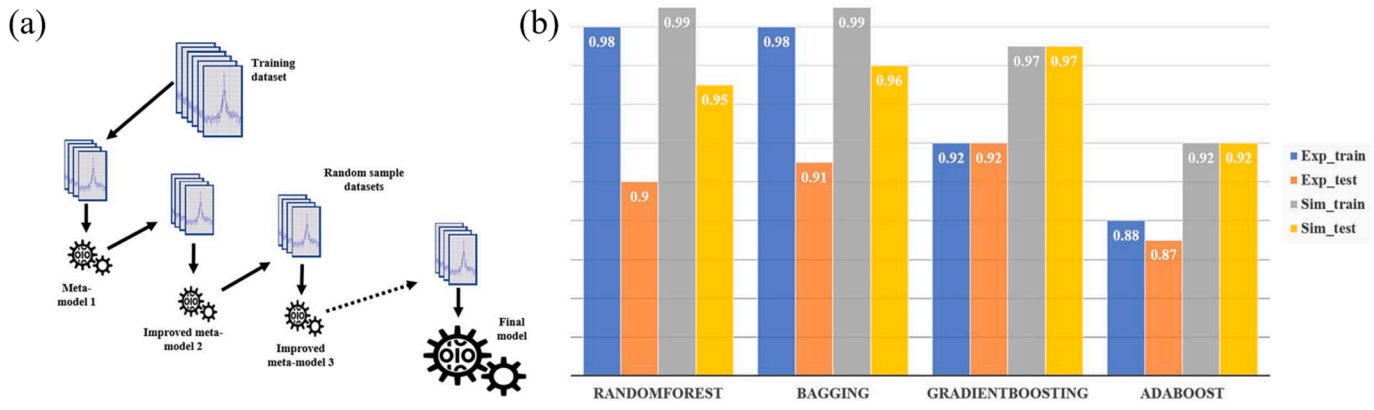

Fig. 8. (a) Overview of boosting method;(b) Comparison of training and testing errors for applied four algorithms on simulation and experimental datasets [103].

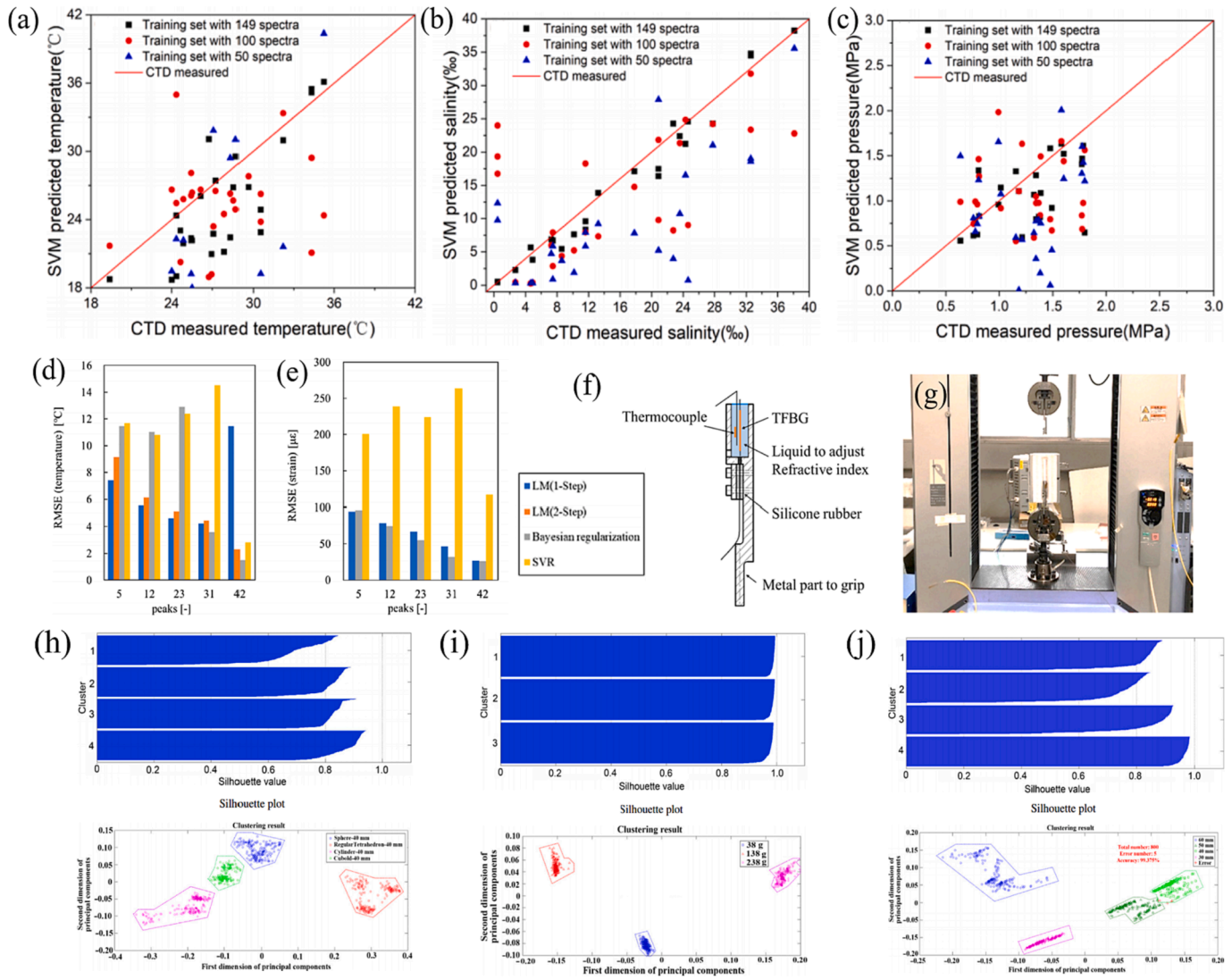

Fig. 9. Comparisons of SVM predicted (a) temperature, (b) salinity, and (c) pressure with electronic device measured results [104], (d)temperature and (e)strain RMSEs of different algorithms at different peak numbers, (f)cross-section of the lower jig, (g) photograph of a test rig [105], Clustering results in recognition for (h) object shape, (i) weight, and (j) size [107].

[108]. Continuous quasi-lateral monitoring of systolic and diastolic blood pressure through SVR achieved the effect of predicting the subject blood pressure. However, their implementation was to detect objects under static conditions without considering the unexpected situations in the dynamic process of human motion and actual usage environment. Li et al. developed a wearable flexible optical fiber sensor by weaving FBG sensors with yarn [109]. A gesture recognition program based on wearable sensors was developed using dual FBG signal acquisition and BP model, with an accuracy of 97.02 %. With the long-term monitoring of wearable sensors, with many available signals collected, the behavior patterns of the wearer can be determined through classification feedback, and timely intervention can be taken when assistance is needed. Gruionu et al. reconstructed the three-dimensional (3D) path of the FBG probe through CNNs to verify the feasibility of a FBG-based pulmonary airway navigation system for early cancer diagnosis [110]. The average prediction accuracy for identifying 10 lung airways was 91 %, with a standard deviation of 5 %. The entire instrument did not require radiation or electromagnetic navigation and meets non-invasive biopsy requirements, providing new feasible solutions for the health, low-cost, and miniaturization of medical equipment.

Compared with linear static demodulation or dynamic demodulation methods, the ML algorithm using nonlinear compensation provides higher demodulation accuracy without increasing hardware complexity, and improves the learning speed of multi-parameter demodulation, achieving one-to-one correspondence between input and output. The ML demodulation is not limited by the conventional demodulation method to match the multi-parameter sensitivity of the sensor. The accuracy of ML in signal classification has also been verified.

To better facilitate readers' understanding, we have summarized the algorithms appearing in Table 1 according to different abilities in Fig. 10, and it can be seen that NNs are the most frequently used model. The good adaptability, nonlinear mapping ability, self-learning, self-organization, and self-adaptability, make it more preferred in the spectrum demodulation of point optical fiber sensors.

4.2. *Speckle patterns processing*

Compared to electrical imaging, optical imaging transmits different modes through optical fibers and outputs speckle patterns with different characteristics. These patterns measure the amplitude and phase of each input mode corresponding to the output mode, constructing a complex matrix for demodulation [111,112]. However, this method introduces an external reference beam into the fiber output to generate an interferogram and extract complex light fields (amplitude and phase) from it.

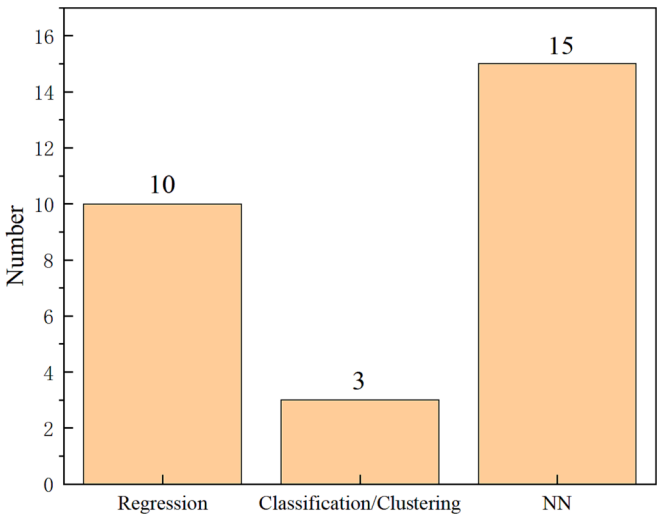

Fig. 10. Bar chart of algorithms summarized by function in Table 1.

Careful phase tracking is required to correct phase drift, which makes the implementation more complex. To address these challenges, a speckle pattern demodulation algorithm based on ML has been proposed, mainly focusing on deep CNNs. In contrast to conventional dense neural networks, CNN uses convolutional operations within the neural network layer instead of general matrix multiplication [113], Inspired by biological processes in visual perception. CNN has achieved great success in computer vision, making it natural to study the use of them for speckle pattern reconstruction.

4.2.1. *Interference identification*

The primary challenge faced by optical fibers to transmit the speckle pattern signals is the susceptibility to external interference, impacting signal authenticity and recognition at the output. Gu et al. etched a series of fluid channels on optical fibers to enhance the interaction between light and detection object [114]. They used a CNN model to classify the speckle patterns (Fig. 11 (a)). The demodulation accuracy reached 99.68 % within the refractive index range of 13326–1.3679. When external disturbances affect transmission, it is particularly important to figure out the bending degree and disturbance position of the optical fiber to guide how to solve the problem. Lu et al. introduced a plastic fiber bending sensor capable of detecting multi-point bending without additional hardware [115]. Using CNN to classify output speckle patterns under different bending states. When the bending range was 15°, 10°, and 5°, the recognition accuracy of the model was 99.2 %, 96.1 %, and 93.5 %, respectively. Fig. 11 (b) shows the speckle patterns and feature maps extracted by CNN. However, ensuring the stability of the entire experimental system remains crucial during experiments. The system's instability may lead to fluctuating speckle patterns, impacting the accuracy of speckle recognition. This underscores the algorithm's need to consider anti-interference measures for reliable performance. Cuevas et al. investigated the problem of using speckle patterns to locate the external disturbances of optical fiber sensors [116]. When using CNN to process collected speckle patterns, it was necessary to classify them between 3 different positions, with a classification accuracy of up to 99 %. When the number of positions increases, any minor modifications to environmental conditions or alignment of the optical fiber with the camera will result in significant misclassification. It had been observed that disturbance identification generated near the sensing element was more accurate, which may be due to higher light intensity and more

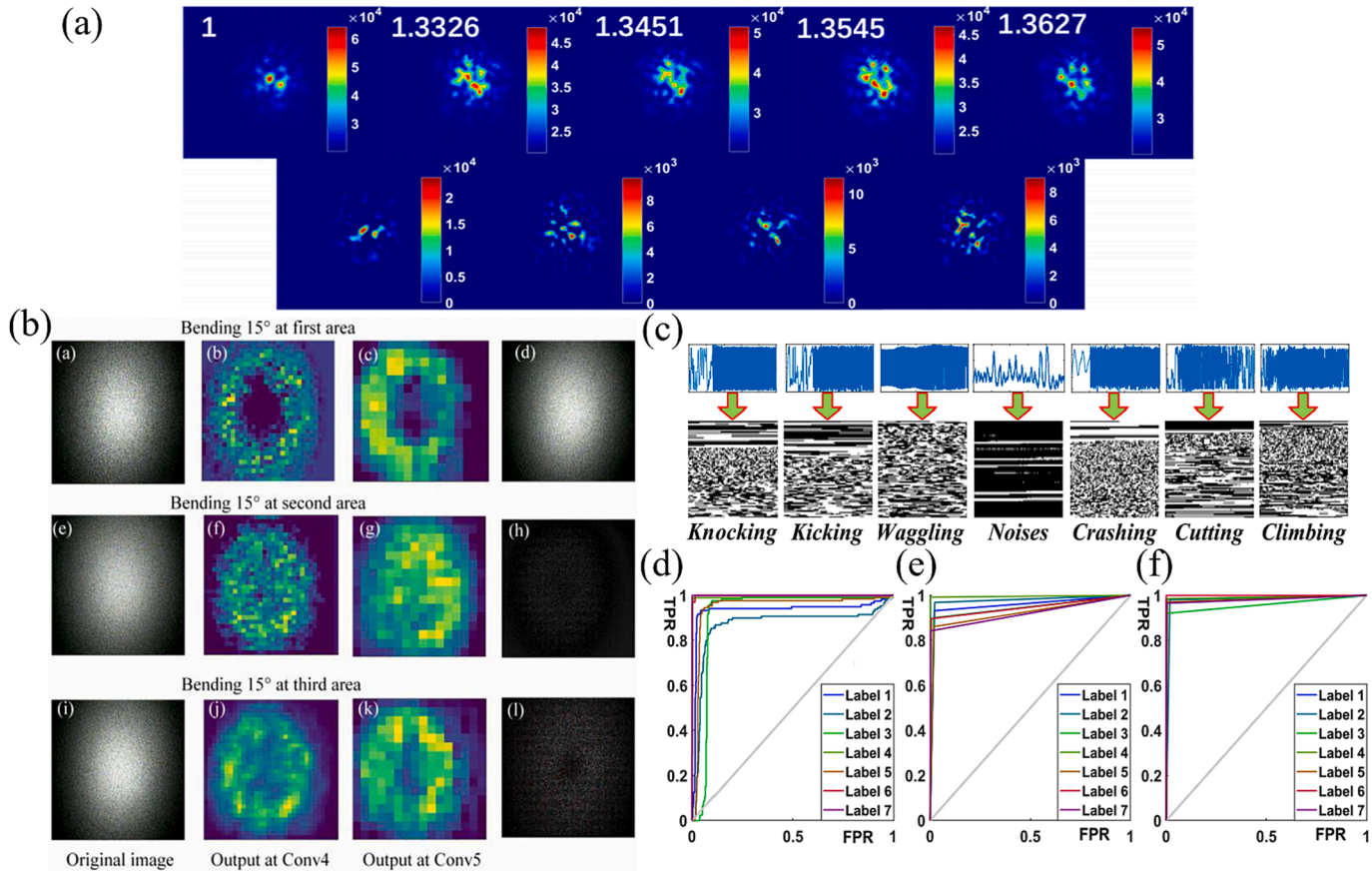

Fig. 11. (a) Speckle patterns correspond to different refractive indexes [114], (b)Visualization of speckle grams and CNN extracted feature maps [115], (c) Time domain waveforms of the seven events and their grayscale images, ROC curves of (d) handcrafted model, (e) 2D-CNN model, (f) 2DCNN-LSTM model [118].

pronounced characteristics compared to small signals that propagate further. Yang et al. used speckle patterns to characterize pattern interference changes caused by curvature in optical fibers and used CNNs to identify bending states and disturbance positions [117]. The trained model exhibited an accuracy of 99.13 %. The classification accuracy of the model was similar on both the training and testing sets, indicating that the model had satisfactory generalization ability. Sun et al. developed an improved algorithm based on a serial fusion feature extraction model [118]. This approach transformed time-domain signals into grayscale images and extracted convolutional features within each frame and temporal features between frames. Identification of 7 common sensing modes (as shown in Fig. 11 (c)) collected from vibration systems based on optical fiber. This algorithm can automatically extract effective features and perform precise recognition, as shown in Fig. 11 (d), (e), and (f). The 2D-CNN-LSTM model shows better performance. The average accuracy of this algorithm reached 96.16 %. The recognition response time of each sample was reduced to 11 ms.

Using ML algorithms for speckle patterns demodulation not only achieves anti-interference, and enhances sensor robustness, but also facilitates interference source identification, ensuring more precise recognition and resolution. Moreover, ML-based approaches significantly enhance response times compared to conventional methods.

4.2.2. Pattern reconstruction

In recent years, the development of endoscopes for image transmission utilizing lightweight and compact optical fibers has gained traction. The application of CNN in complex speckle pattern reconstruction caused by modal mixing and modal dispersion has long been confirmed. However, multimodal transmission in MMF may lead to stacking and dispersion, which does not affect individual points but rather the entire transmitted image [119]. CNN performs well in local features but performs poorly in global feature processing [120]. Generally speaking, the greater the difference in features between the test set and the training set, the worse the model performance will be. How to improve CNN or preprocess collected data to improve performance may become a solution.

Rahman et al. demonstrated that CNN can learn the nonlinear relationship between the amplitude (phase information loss) of speckle patterns and the phase or amplitude [121]. Compared with the speckle patterns recovered using a system characterized by complex measurement matrices, the image fidelity (correlation) was ~98 % and the image projection was ~94 % when reconstructing images in MMF. Zhu et al. utilized a SHL dense NN to accurately reconstruct speckle patterns collected within a week after the training set stopped [65]. The SHL network's efficacy proved comparable to CNN's performance across the entire cycle. The proposed SHL-CNN model achieved the same structural similarity index measurement (SSIM) in only 8 % of the conventional method's time (16 min). Fig. 12 (a), (b), and (c) show the SSIM, variance, and classification accuracy of SSIM, respectively. The image reconstruction fidelity of SHL-CNN was good, requiring less computational resources and shorter training time compared to conventional algorithms. Liu et al. introduced an all-fiber structure integrating fiber output pulse lasers, fiber probes, and side pump couplers [122]. This innovative setup leverages high-mode dispersion in MMF to convert 2D spatial information into 1D temporal pulse streams for high-speed all-fiber imaging. By training the CNN model, the speckle pattern was reconstructed from the time waveform recorded by the ultra-fast photodiode connected to the fiber output. Fig. 12 (d) shows an example image of letters and clothing, as well as the restored image. In Fig. 12 (f), the time signals collected at MMF lengths of 1000 m and 400 m had high frame rates. The experimental verification accounted for effects due to bending and temperature changes in the endoscope application. When the bending radius range was 22 cm to 28 cm, the average fidelity can be maintained at over 70 %. For temperature changes within 0.5 ℃, the image fidelity reaches 70 % and has high robustness.

By reconstructing optical fiber transient signals into images and distinguishing them through ML algorithms, various feature information in the signals can be identified. Reconstructing transient data under continuous sampling into multiple images can achieve analysis and classification of dynamic processes. Naku et al. used the optical fiber sensor to observe the process of liquid evaporation, immersing the sensor into the liquid and pulling it out immediately [123]. As droplets evaporate, the length of the liquid chamber decreases. The entire process of change can be described by observing the changes in the interference spectrum. Transient response was converted into an image by continuous wavelet transform, and CNN was used for classification. The classification of different liquids was achieved through the evaporation process, with an accuracy rate of over 98 %. There were relatively few detected 7 types of objects, and there were significant differences in the sample evaporation process. It was still worth studying how to identify the evaporation process under lower concentration differences of water/alcohol mixtures. Based on their earlier work, Naku proposed an algorithm for identifying volatile organic liquids [124]. The temporal transient response of 11 different liquids' evaporating droplets was monitored. The diversity of information collected using three sensors provides more features for ML-based methods, which perform better in classification accuracy.

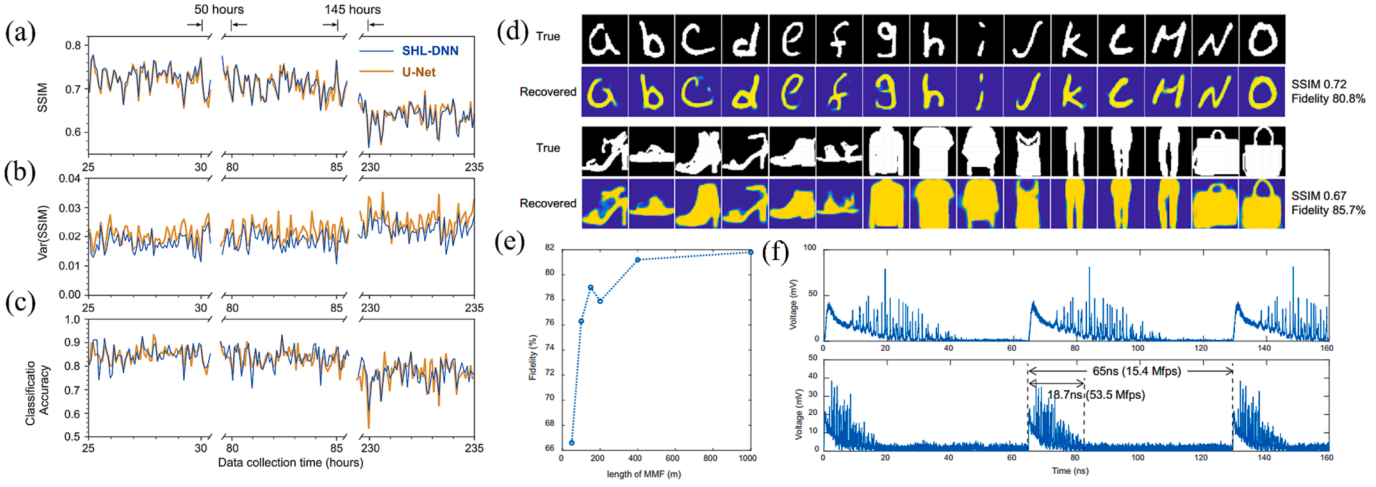

Fig. 12. (a) SSIM. (b) Variance of SSIM, (c) Classification accuracy, (d) Examples and their recovered [65], (e) Average fidelities with different fiber lengths. (f) Time signals with different fiber lengths (1000 m (up) and 400 m (down)) [122].

ML algorithms not only ensure that complex speckle patterns transmitted using optical fiber sensors have good fidelity after reconstruction but also enable rapid model training while ensuring model portability. The multi-objects processing and dynamic process recognition also show good performance.

4.2.3. 3D

The transmission of optical fibers in multimodal transmission allows the capture of 2D image information from detected objects. However, these optical fiber sensors exist within 3D space, offering the potential to acquire 3D information as well. Wang et al. used K-NN to identify the 3D geometric shape information of MMF output speckle data, achieving nearly 100 % classification accuracy [125]. Despite the absence of considerations for temperature and vibration changes, they attained a spatial resolution of 0.3 mm and a 3D resolution of 1°. Similarly, Razmyar et al. investigated the deflection direction of the MMF tip by scrutinizing speckle pattern shapes and structures [126]. Their study explored diverse strategies for integrating information from sensor arrays within the CNN model. Most estimated values differ by ± 5 degrees from the real values, and the accuracy of the model was not affected by temperature changes of 2–3 ℃. As mentioned in the [122], due to the increased transmission rate of lasers, there is the possibility of extending this principle to the reconstruction of 3D information using captured signals. This process of identifying 3D information concealed within 2D data can be termed data mining, representing a domain where the potential of ML is significantly harnessed.

4.3. Distributed optical sensors

The optical fiber can not only be used to prepare the discrete/point sensors, because the surrounding environment can affect the change of transmission state in the fiber, a long-distance optical fiber can provide thousands of “virtual sensors” distributed along the fiber. Distributed optical fiber sensors mainly collect and analyze scattered signals in optical fibers. There are three main scattering processes in optical fibers: Brillouin scattering, Rayleigh scattering, and Raman scattering. Their relationship is shown in Fig. 13 (b). As shown in Fig. 13 (a), changes in vibration and other factors from the surrounding environment can alter the scattering state of nearby optical fibers, leading to changes in fiber strain and refractive index. These scattered signals are weak, and in actual environments, longer fibers are more affected by different interference sources. The problem with distributed acoustic sensors is signal clutter. How to extract effective information from a complex spectrum is the most important situation faced by ML. To provide readers with a more accessible grasp of the sensors and algorithms discussed in this section, we have summarized some references based on algorithm types, sensor types, measurement objects, sensitivity, and algorithm implementation effects, and presented these details in Table 2.

4.3.1. Optical Time-Domain reflectometry (OTDR)

OTDR has the advantages of simple structure, high spatial resolution, and high location accuracy. But the SNR of signals is usually low because it is inevitably affected by background noise. In pursuit of enhancing denoising capabilities, Kong et al. developed a phase unwrapping algorithm based on quasi-Graman matrices and deep convolutional neural networks (DCNNs), which had the advantages of strong robustness, and low SNR, but multiple parameters [127]. Subsequently, an extended algorithm based on length and memory networks was proposed [128], which was one order of magnitude lower than [127]. The algorithm had two steps which are shown in Fig. 14 (a). The first step was to encode the 1D phase into a 2D array based on the quasi-Graman matrix. The second step used DCNN for phase unwrapping, with an SNR less than 4 dB. The RMSEs of the 5 algorithms at different SNR levels are shown in Fig. 14 (b). Saleh et al. improved the SNR by combining wavelet and normalized differentiation methods and used SVM to classify four types of human activities. The classification accuracy was greater than 95 % [129]. Pan et al. introduced lightweight NNs for distributed optical fiber vibration pattern recognition [130]. By extracting time–frequency sequence correlations from signals and spectrograms, and achieving better performance with fewer additional parameters and calculations. This study includes 8 different scenarios with a classification accuracy of 96.02 %. This processing method of time–frequency domain features can be flexibly integrated with different model architectures to improve classification performance and have universality. Huang et al. used the extremum gradient boosting (xgboost) algorithm to identify feature vectors of different vibration events [131]. Fig. 14 (c) shows the structure of the xgboost algorithm. This algorithm was based on a DT. Compared with SVM and k-NN, the proposed method had better learning ability and noise resistance. It performs better in dealing with classification and regression issues. The average recognition accuracy of the xgboost was 93.3 % for 8 event types. Fig. 14 (d) shows the classification performance of 5 algorithms with different sub-band intervals. Liu et al. reconstructed the distributed optical fiber system signal using an improved wavelet denoising algorithm and a dual threshold algorithm tailored to the characteristics of vehicle vibration signals [132]. By extracting features, they estimated the number of vehicles and the

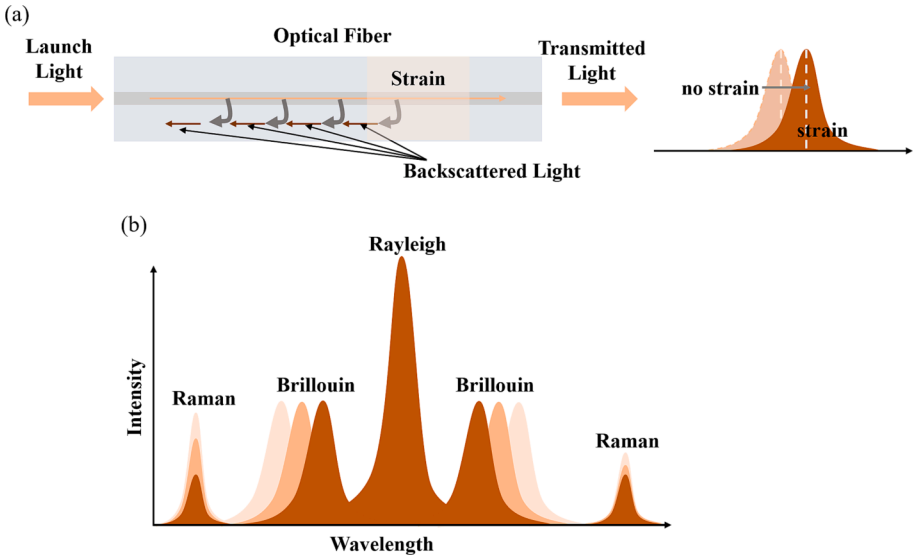

Fig. 13. (a) The Brillouin wavelength of the optical fiber is shifted due to external strain, (b) Common scattering phenomena.

Table 2
Distributed sensors combined with different ML algorithms.

	Type of algorithm	Measurement	Evaluation of model performance	Ref
OTDR	DCNN	predict the true phase	Work stably at negative SNR	[127]
	LSTM-DCNN	predict the true phase	The parameter amounts 1 order smaller than the DCNN	[128]
	SVM	Classification (3 Acoustic disturbance events)	Classification accuracy higher than 95 %	[129]
	CNN	Classification (8 Acoustic disturbance events)	Classification accuracy of 96.02 % and achieved better performance with a lower number of extra parameters and computations	[130]
	xgboosting	Classification (8 acoustic disturbance events)	Average recognition rate of 93.3 %	[131]
	SVM	Classification and vehicle speed	Detection accuracy was higher than 80 %, the vehicle classification accuracy was higher than 70 % and the speed estimation error was less than 5 %	[132]
	ELM	Classification (5 acoustic disturbance events)	Average identification rate of 95 %, shorter identification time (<0.1 s)	[133]
	VMD + AP	Classification (3 acoustic disturbance events)	Accuracy of ~97 %	[134]
	1-D RL-CNN	Classification (6 acoustic disturbance events)	Accuracy of 91.19 %	[136]
	CNN	Classification (5 acoustic disturbance events)	Accuracy of 91.6 %	[137]
	IP-CNN	Classification (9 acoustic disturbance events)	Accuracy of 88.2 %	[138]
	kS-DCGAN	Classification (2 acoustic disturbance events)	Average accuracy of 97 %	[139]
	K-means	Classification (8 acoustic disturbance events)	High spatial resolution and classification accuracy of 77.65 %.	[141]
	PCA + K-means	Classification (4 acoustic disturbance events)	The cluster number was 4, the accuracy was 87.1 %; when the cluster number was 3, the accuracy was 89.4 %	[142]
BOTDA/ BOTDR	ANN BOTDR	Acoustic	Acquisition time was 5.5 % of the normal sampling, temperature measurement uncertainty was 0.2 ℃, and 3 m spatial resolution	[144]
	K-ELM BOTDA	Temperature	Better measurement accuracy of 0.3 °C, processing speed over 120 times	[145]
	CNN BOTDA	Improve SNR	SNR improvement of 13.43 dB	[146]
	GLM BOTDA	Temperature	Improving the signal processing time (655 s to 1 s) and temperature prediction accuracy was 1.32℃	[147]
	SVM BOTDA	Temperature	Extraction accuracy of temperature information was improved about 4 °C	[148]
FBGs	BP	Strain	The best interrogation precision was ± 4 pm	[151]
	DCNN	Temperature	MAE of the predicted reflectance peak position was 7.8 pm	[154]
	CNN	–	RMSE of the predicted reflectance peak position was lower than 0.05 pm	[150]
	PSO-SVM	Strain	Test accuracy of 100 %	[155]
	GRU	Strain	Distinguishing overlapped spectrum with only 3 % intensity difference of FBG peaks	[153]
	SVM	Acoustic	Fault identification rate of 99 %	[152]
	DT	Temperature	Average accuracy of 89.54 %	[156]
	ANN	Impact	Average errors were 0.98 cm	[158]

vehicle's speeds. Then the SVM algorithm was used to classify the vehicles. The experiment results showed that the accuracy of determining whether it is a vehicle was greater than 80 %, and the estimation error of vehicle speed was less than 5 %. On this basis, vehicle classification has also been implemented.

In addition to denoising through algorithms, signal optimization can also be achieved through feature selection. Jia et al. used a recognition method combining ELM and Fisher scoring feature selection to reduce the interference alarm rate of OTDR [133]. Fig. 15 (a) shows the implementation process of OTDR identification based on the proposed model. By selecting 25 features, the average recognition rate of 5 interference events exceeded 95 %, accomplished within a recognition time of less than 0.1 s and a false alarm rate of 4.67 %. Fig. 15 (b) shows the effect of adjusting the activation function and neurons on the model performance. Fig. 15 (c) shows the recognition time of hardlim as an activation function in different numbers of neurons. Xu et al. enhanced the vibration signal through variational mode decomposition (VMD) and fused different statistical features to obtain a selection factor (SF) [134]. Using SF as the standard for selecting different event signals. Then, the affinity propagation (AP) clustering method was used to identify the disturbance event types. The flowchart of the proposed method and cluster results are shown in Fig. 15 (e) and (d), respectively. Nevertheless, Zhao et al. proposed an innovative method for classifying vibration events and measuring vibration frequencies in OTDR distributed vibration sensors, incorporating image processing techniques [135]. The normalized time series of signals are converted into Markov transition field images, and then the images are classified using CNN. The accuracy of the model can reach 97 %.

ML can accurately classify events by extracting different features between samples. However, most methods focus on the recognition of closed datasets. When there are unknown events, the model may make misclassification, thereby reducing recognition accuracy. Zhou et al. used OpenMax's 1D residual learning CNN (1D RL-CNN) to achieve an open set classification accuracy of 91.19 %, which was 18.47 % higher than the 1D CNN under SoftMax [136]. The utilization of OpenMax enabled system managers to receive timely reminders, prompting human intervention to recalibrate the model when necessary. Rizzo et al. took inspiration from faster region-based CNNs (R-CNN) and proposed the first 1D-NN for OTDR trace detection [137]. Their test results revealed further success in identifying unknown events compared to OpenMax's ability to report unknown events. By conducting a preliminary analysis of traces, unknown events were detected to identify candidate regions, indicating rare and unforeseeable situations that need to be reported.

In some applications where data samples are scarce, data samples need to be expanded. Wu et al. used CNN to learn features from the spectrum and introduced a method to enhance data by incorporating intensity and phase parameters [138], which expanded the dataset and achieved a classification accuracy of 88.2 %. Fig. 16 (a) shows the training flow of the proposed model. Fig. 16 (b)-(f) show the

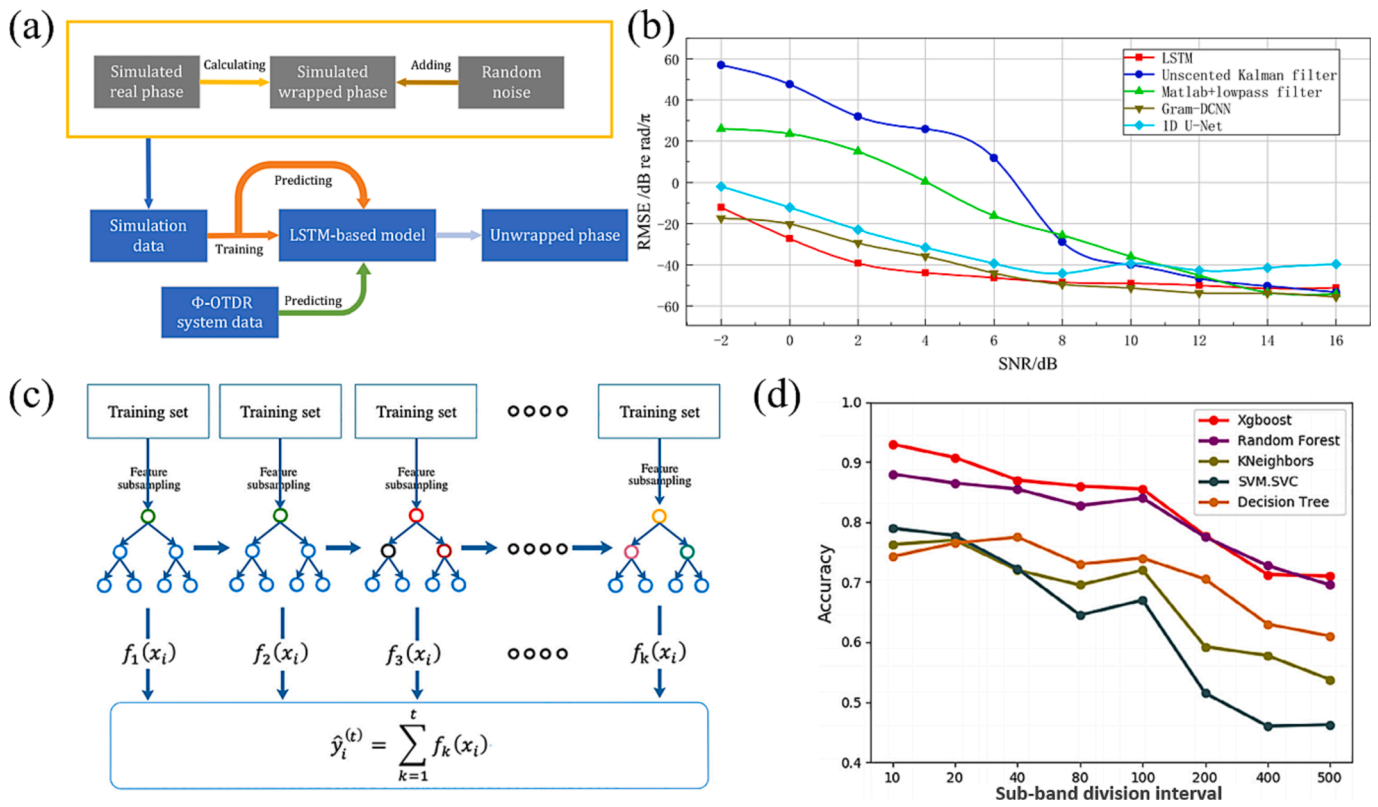

Fig. 14. (a) The proposed algorithm flows, (b) The anti-noise performance of the five algorithms [128], (c) The basic structure of the xgboost algorithm, (d) The classification performance of various algorithms [131].

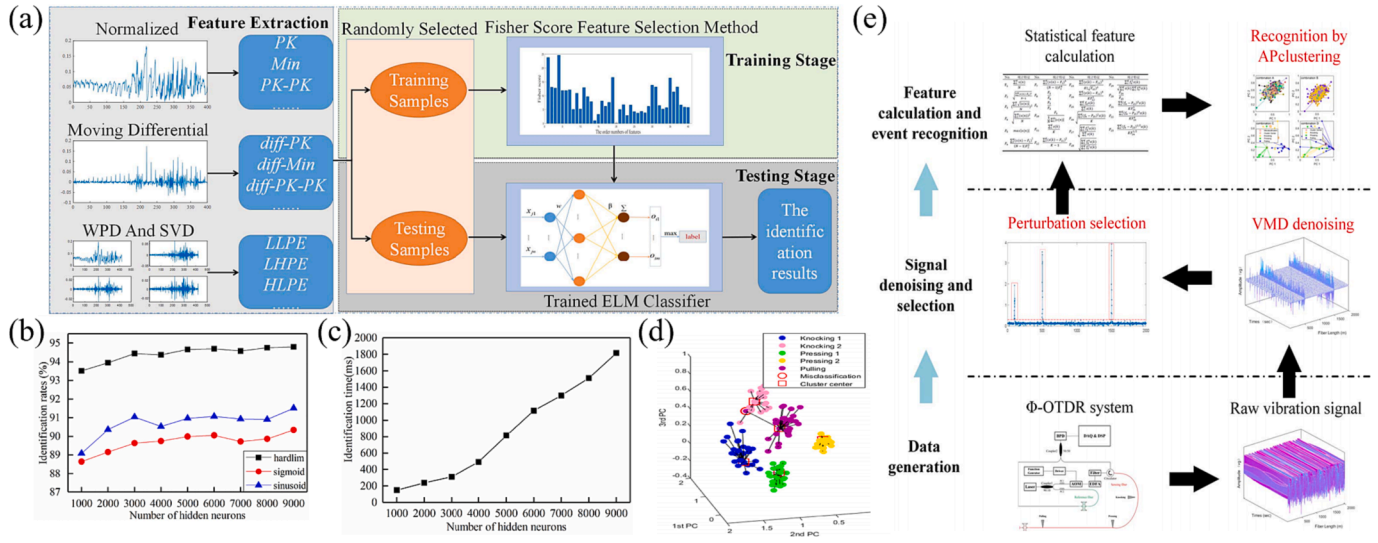

Fig. 15. (a) The process of the proposed method, (b) Model performance comparison of different activation functions and neurons, (c) The identification time with different numbers of neurons [133], (d) Flowchart of the proposed method, (e) 3D-Scatter plots of clustering results [134].

classification results of different enhancement techniques, clearly indicating that the (f) method mentioned in the work has a more pronounced classification effect. Shang et al. used the K-means clustering-synthetic minority oversampling technique deep convolutional generative adversarial network (kS-DCGAN) data augmentation method to increase the quantity and quality of small sample event data and meet the needs of classification network training [139]. The number of small sample events was increased after augmentation, and the accuracy of the classification task reached 97.22 %.

At present, the main work of distributed optical fiber signal classification is mainly based on supervised learning, which requires manual annotation and collection of many samples. For complex environments and some special applications, a large amount of labeled data may be difficult to obtain. Due to the lack of labeled data, supervised learning cannot be achieved. Therefore, the research on unsupervised learning is very meaningful because it can reveal the performance level of ML algorithms without too much prior knowledge [140]. Peng et al. explored the identification of distributed optical fiber sensing signals with

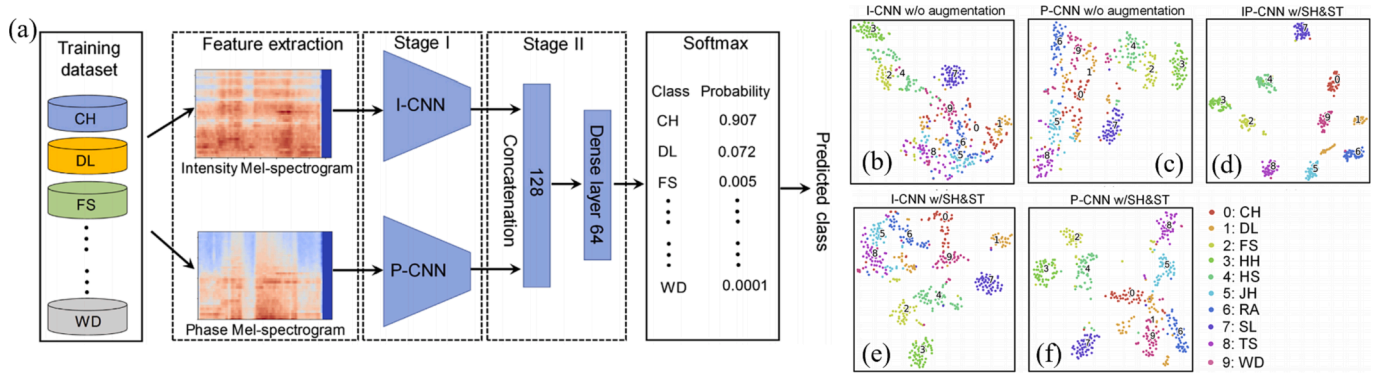

Fig. 16. (a) Training flow of the proposed algorithm, (b)-(f)Feature embedding from different augmentation [138].

enhanced Rayleigh backscattering caused by human motion vibration [141]. The femtosecond induced Rayleigh scattering centers improve the SNR of the signal. Using unsupervised and supervised ML algorithms to identify specific sound times, supervised learning achieved a recognition rate of 76.25 % using DCNN. The unsupervised K-means clustering algorithm achieved a recognition rate of over 77.65 %. However, these learning algorithms have been preprocessed by fast Fourier transform, and combined with some pre-classification. Zhang et al. proposed an unsupervised learning method for classifying OTDR vibration events [142]. In data processing, the principal component analysis dimensionality reduction method was used, while the k-means algorithm was used to cluster different vibration events. In the identification of 4 vibration events, when the clustering number was 4, the accuracy was 87.1 %. ML can play a more powerful role in processing large amounts of data without the need for preprocessing.

ML ensures good demodulation accuracy at different SNR levels and achieves better performance through prior knowledge, avoiding false positives. Accurate classification in extreme situations can be achieved with unsupervised learning methods. ML algorithms not only reduce the false alarm rate of classification but also provide solutions for classification needs in special situations such as demodulation with few samples and unlabeled data.

4.3.2. Brillouin optical time domain analysis (BOTDA)/Brillouin optical time domain reflectometer (BOTDR)

BOTDR detects backscatter signals from released Brillouin scattering. Conversely, BOTDA requires pulse light and continuous pump light to be incident from both ends of the fiber. When the frequency difference between the two lights is equal to the Brillouin frequency shift in the fiber, stimulated Brillouin method effects will occur, leading to energy transfer between beams. Fig. 17 (a) and (b) show the working principles of BOTDR and BOTDA, respectively.

The conventional Lorentz curve fitting (LCF) method requires converting the Brillouin frequency shift into temperature changes for analysis, but ML algorithms do not require this complex process and the input spectrum directly outputs temperature. Wang et al. used the ELM to train BOTDR data [143]. Compared with the LCF, ELM networks also have higher accuracy and better measurement error tolerance under large-frequency scanning steps. In addition, Zhang et al. proposed a temperature measurement method based on sparse frequency sampling of the Brillouin scattering spectrum (BSS) [144]. The NN was used to utilize the symmetry of BSS without increasing hardware complexity to restore BSS under sparse sampling. Fig. 18 (a) is a schematic diagram of the entire process. Significantly shortened the acquisition time of Brillouin optical time-domain sensors based on scanning frequency. Compared with normal sampling time, the data acquisition time was reduced by 5.5 %, and the temperature measurement error was 0.2 ℃. Zhang et al. proposed a raw spectrum partitioning method based on kernel ELM (K-ELM) [145]. The performance of this algorithm was studied through simulation and experiments under different SNRs, pump pulse widths, and frequency scanning steps. Compared with the LCF, K-ELM balanced the prediction accuracy and processing speed of temperature measurement, with an accuracy of 0.3 °C and a processing speed increase of more than 120 times. For different data sizes, the algorithm architecture was unchanged. Wu et al. applied CNN to denoise data images from the BOTDA across different sampling rates [146]. Collecting data from the BOTDA system at sampling rates of 500 MSA/s, 250 MSA/s, and 125 MSA/s, the trained CNN denoiser improved its SNR by 13.43 dB, 13.57 dB, and 12.9 dB, respectively. Improved the speed of denoising algorithms and spatial resolution at low sampling rates. Nordin et al. compared and analyzed 5 algorithms (generalized linear model (GLM), DL, RF, Gradient Enhancement Tree, and SVM) to improve the signal processing time and prediction accuracy of BOTDA [147]. Fig. 18 (b) and (c) show the comparison of temperature distribution and heating position error along the fiber after the GLM method and LCF demodulation, respectively. Research has found that the GLM method was significantly superior to the LCF method. GLM was 655 times faster than LCF. Although RF and SVM require a relatively long time, their prediction accuracy was high. The minimum RMSE of RF was about 0.48 ℃, and the SVM was about 0.69 ℃. If higher sampling accuracy is required, more time may be required. Zhu et al. investigated three optimization algorithms: PSO, GA, and firefly algorithm (FA) to optimize SVM parameters [148]. They evaluated the performance of the optimized SVM in simulations and experiments for temperature extraction under various Brillouin gain spectrum acquisition conditions. When the SNR was as low as 2.5 dB, the accuracy of temperature extraction was improved by about 4 ℃ compared to conventional SVM. When the SNR was greater than and less than 10 dB, PSO-SVM and GA-

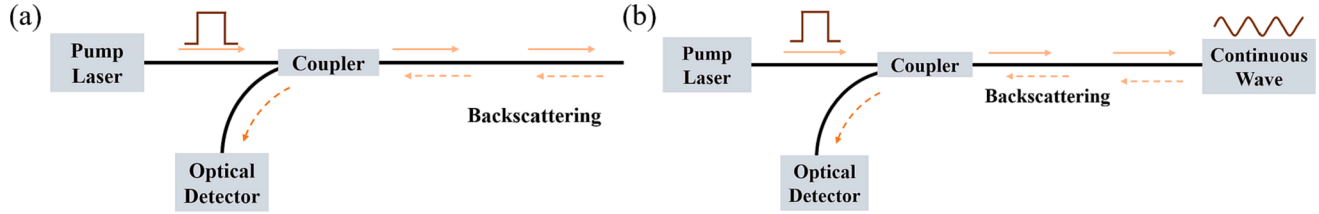

Fig. 17. Schematic diagram of (a) BOTDR and (b) BOTDA.

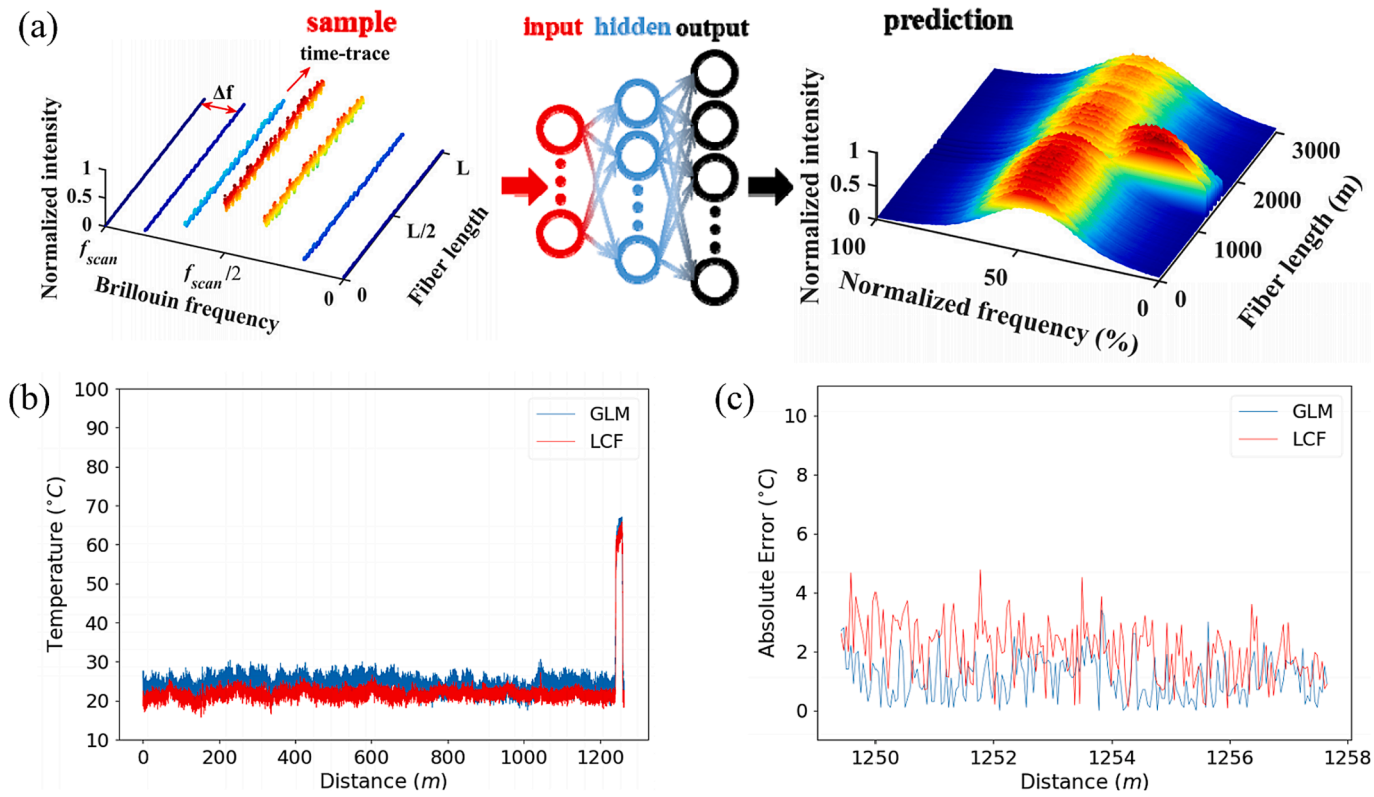

Fig. 18. (a) Low-resolution BSS with sparse frequency sampling is restored to a high-resolution BSS [144], (b) Temperature distribution along the fiber, and (c) Absolute error of heating position [147].

SVM had the best performance, respectively. The FA-SVM algorithm had the fastest processing speed. Although the optimized SVM algorithm requires more processing time compared to conventional SVM, the processing speed was more than 20 times faster than curve fitting methods, and the temperature accuracy was higher.

ML algorithms prove effective in recovering spectrum under sparse sampling, ensuring demodulation accuracy while substantially reducing distributed optical fiber sensor sampling times, and establishing a direct correspondence between spectrum and output without necessitating Brillouin frequency shift.

4.3.3. FBGs

The sensor array formed by carving multiple series connected FBGs onto an optical fiber can be called a quasi-distributed optical fiber sensor, where each FBG functions as an individual sensing unit. In addition to the significant differences in sensor demodulation technology, the biggest difference from the above-mentioned distributed systems lies in the existence of gaps and blind spots between sensors.

To solve those problems, Chen et al. introduced a wavelength detection method for FBG sensor networks based on LS-SVR [149]. Fig. 19 (a) illustrates how the model directly uses the spectrum as input to predict the Bragg wavelength as output. When the spectrum of any two FBGs in the FBG array overlap, the Bragg wavelength can also be determined separately, as shown in Fig. 19 (b). When the SNR was 20 dB, even in noisy environments (20–100 dB), the spectrum before and after using the LS-SVR model can be accurately determined. The LS-SVR

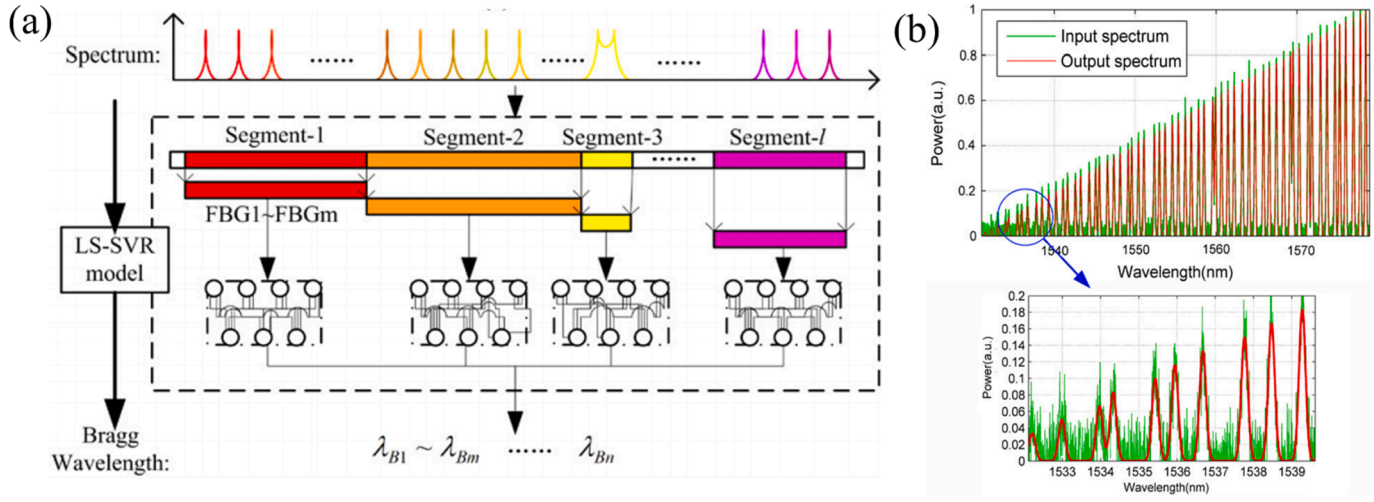

Fig. 19. (a) Schematic diagram of the LS-SVR model, (b) Spectrum before and after the ML model [149].

algorithm makes up for the shortcomings of SVM, such as too many support vectors and complex modeling. Moreover, it eliminates the need for insensitive parameters during the training process which improves the model convergence efficiency. Li et al. proposed an extended CNN model for the demodulation of FBG signals, with a RMSE of < 0.05 pm in peak wavelength measurement and the demodulation time of 15 ms [150]. The demodulation model can achieve a RMSE of < 0.47 pm even when the SNR was as low as 15 dB. Ren et al. proposed an efficient FBG array demodulation system based on NNs [151]. A low-cost data augmentation strategy has been introduced to enable NNs to achieve excellent performance when processing small-scale data. The demodulation model established the nonlinear relationship between emission intensity and actual wavelength, enabling precise peak wavelength determination in response to stress changes. But the method's applicability might be limited as the number of FBGs increases due to the interplay between FBG sensitivity and excitation band length.

A single FBG can be considered as a loss device in the FBG array. The loss of the single FBG and the multiple reflections between the FBGs can interfere with the reflection signal of the FBG array, limiting the number of integrated FBGs. With the increase of optical fiber network users and the continuous expansion of the scale of optical fiber networks, effective monitoring of optical link failures has become more important. OTDR detectors cannot simultaneously recognize many reflected signals, therefore they are not suitable for fault characterization in point-to-point networks. Usman et al. identified fiber defects in optical networks by deploying a distributed FBG network and evaluated the data by using the SVM algorithm [152], achieving a defect evaluation accuracy of 99 %. Fig. 20 (a) is the schematic diagram of the SVM-based system. However, the limitation of the study lies in its experimental setup, which included only four FBGs distributed across different links. The effect of cascading FBGs in series, where downstream gratings can be influenced by upstream grating transmission, leading to spectrum distortion, was not considered. This spectrum shadow effect can significantly impact demodulation. Manie et al. utilized intensity wavelength division multiplexing (IWDM) technology to increase the total number of FBG sensors multiplexed in sensor networks and utilized the GRU to solve the problem of large measurement errors caused by spectrum overlap and crosstalk in IWDM [153]. Fig. 20 (b) is the schematic diagram of the GRU unit. The strain sensing signal measurement performance of the GRU algorithm was verified under the conditions of 10 % (as shown in Fig. 20 (c)) and 3 % (as shown in Fig. 20 (d)) intensity differences between two overlapped FBG spectrums. Their proposed method demonstrated increased reliability and accuracy of the signal, concurrently expanding the number of FBG sensors in the system. The study conducted by Kokhanovskiydenrgen et al. applied CNNs to demodulate the problem of spectrum distortion caused by mutual interference in densely arranged FBG arrays [154]. The average absolute error of this work was 7.8 pm, which was higher than the 5 pm of the hardware multiplexing interrogator. The experimental content was to predict the Bragg wavelength by arranging dense sensors around sparsely arranged sensors, which increases the amount of FBG used but does not fully utilize each one's capabilities.

Due to the differences in data features, the most effective ML algorithms may not be the same for different problems. Shao et al. established four intelligent detection models: CNN, PSO-SVM, DT, and RF, to analyze road leakage data collected by FBG sensor arrays [155]. Fig. 21 (a) displays the accuracy of the training set and the test set of these 4 models. Although several models showed excellent detection accuracy, DT and RF models demonstrated overfitting, as their training accuracy was notably higher than the testing accuracy. Fig. 21 (b) shows the different error rates of 4 models. PSO-SVM stood out as the best-performing model in detecting various error types, achieving a remarkable testing accuracy of 100 %. Nascimento et al. proposed using a FBG array to estimate fluid level, using DT to classification [156]. Fig. 21 (c) presents the diagram of the analysis with the proposed FBG array. The classification accuracy was 89.54 %. Furthermore, the researchers incorporated a combination of techniques, including weighted linear regression, SVR, and kernel selection minimum cost SVR (SVRmin), at various stages of the classification process. This combination allowed them to achieve the lowest RMSE in estimating liquid levels (as shown in Fig. 21 (d)). Liu et al. introduced a sparse Bayesian learning mechanism based on data uncertainty of FBG sensors, which can describe the probability characteristics of uncertainty in a Bayesian framework [157]. Taking the time series of noise strain sensors of a certain highway bridge and railway bridge as an example, regression analysis and uncertainty analysis were conducted. The results have shown that the proposed sparse Bayesian learning mechanism was more effective than SVMs and least squares methods.

The impact between spacecraft and orbital debris or stars in space helps detect spacecraft faults and analyze and repair them. Jin et al. proposed a method based on the FBG array, which combines ANN and Markov distance for impact detection and localization of

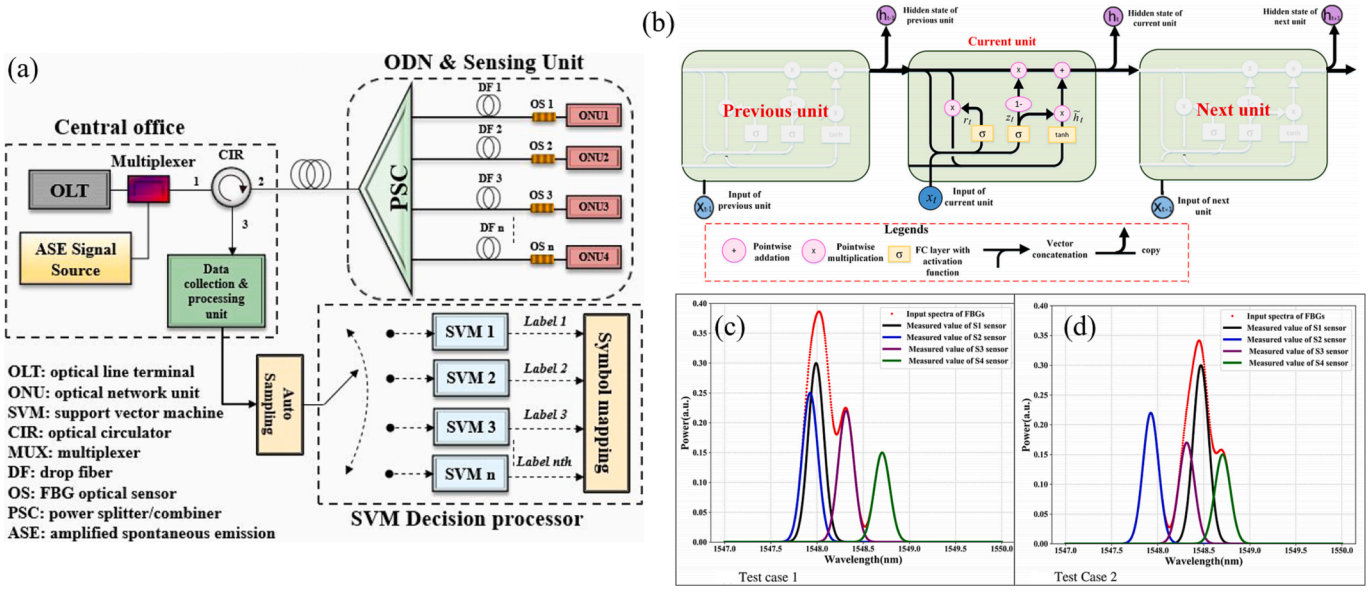

Fig. 20. (a) The schematic diagram of the monitoring system [152], (b) The structure of the GRU unit, The signal using GRU when the intensity difference between FBGs is (c) 10% and (d) 3% [153].

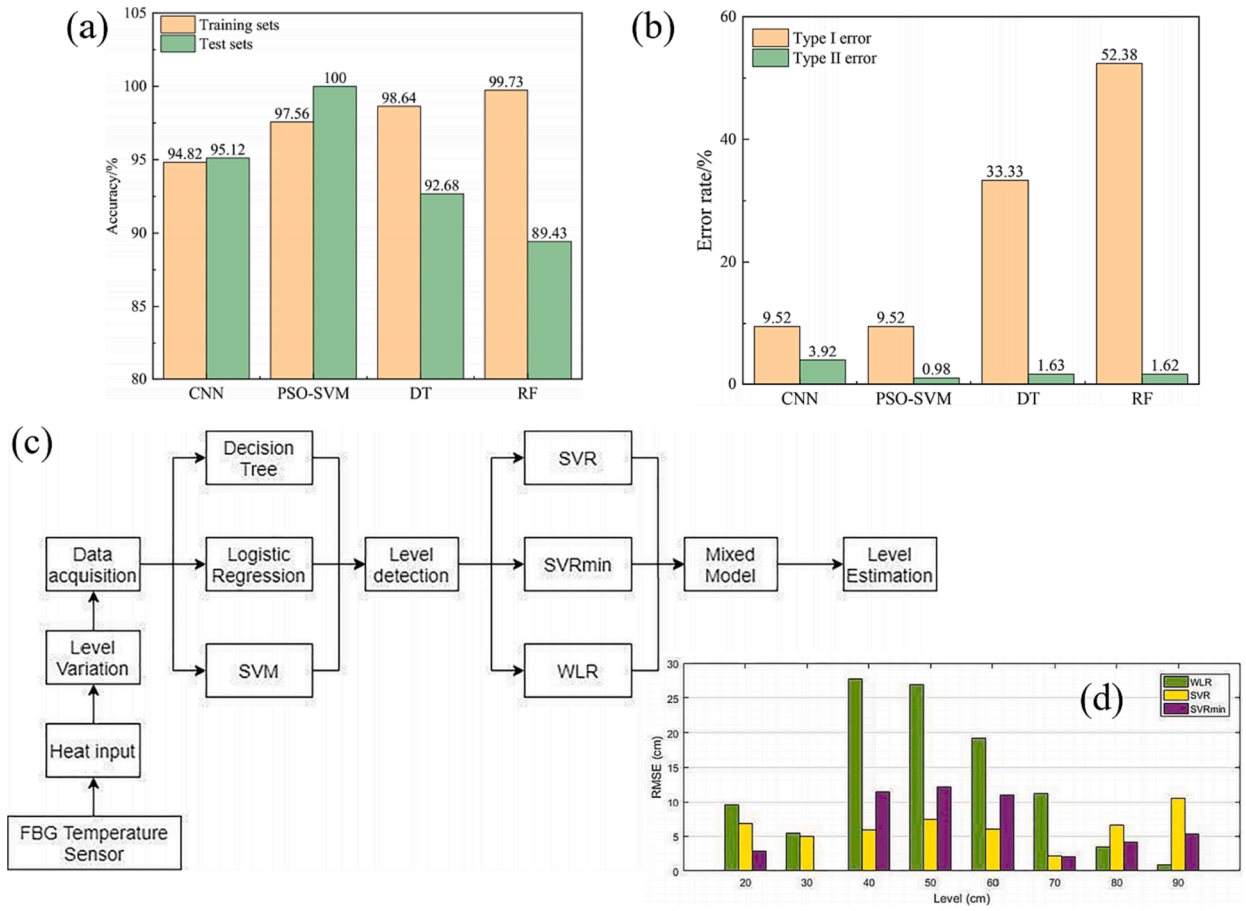

Fig. 21. (a) The training and test accuracy of 4 models, (b) The error rates of 4 models [155], (c) Block diagram of the method, (d) RMSE for the different studied levels of 3 models [156].

micrometeoroids and orbital debris [158]. Using Markov distance and collision point coordinates as inputs and outputs of the ANN, respectively. In 360 impact experiments, the average positioning error of this method is 0.89 cm. Compared to using the Markov distance discriminant analysis method and ANN alone, the error was reduced by 39 % and 31 %, respectively.

Unlike OTDR systems, if the central wavelength remains separated, the distributed FBG array can achieve mixed measurements of multiple parameters. However, the conventional demodulation method also requires bandwidth modulation and wavelength division multiplexing, and there is also an impact between sensor sensitivity and measurement range. The demodulation method of ML not only reduces the requirement for data SNR, but also directly establishes the complex relationship between actual wavelengths and measurement parameters. The demodulation method is not limited by the separation of the central wavelength, increasing the number of sensors in the sensor array, and expanding the sensing range.

To better facilitate readers' understanding, we have summarized the algorithms appearing in Table 2 according to different abilities in Fig. 22, and it can be seen that NNs are the most frequently used. Compared to the point optical fiber sensors' spectrum, the spectrum characteristics of distributed optical fiber sensors are more chaotic and prone to noise interference. NNs have strong adaptability and learning ability, which can automatically adjust the weights based on the input spectrum, thereby achieving accurate demodulation of the spectrum. Secondly, NNs can process large-scale distributed optical fiber sensor spectrums through parallel processing mechanisms, improving demodulation efficiency. Therefore, NNs play an important role in the spectrum demodulation.

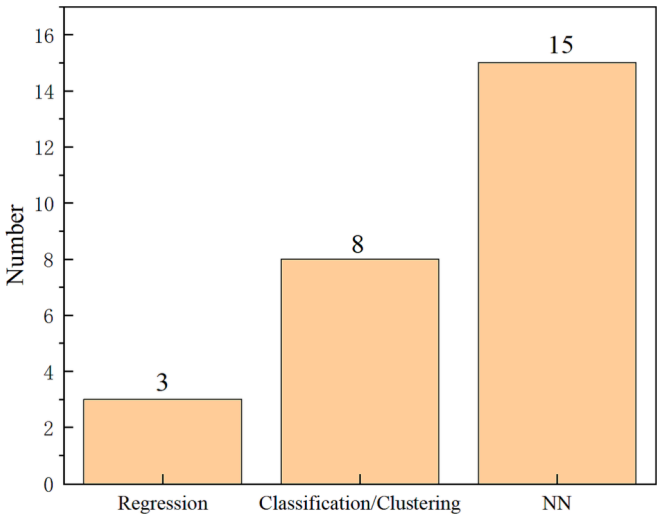

Fig. 22. Bar chart of algorithms summarized by function in Table 2.

4.4. Others

In addition to helping with signal demodulation in the optical fiber sensors mentioned earlier, ML can also help optimize the design of optical fiber sensors and accelerate design time. Lee et al. used NNs to analyze the reflectance of different sicknesses of SiO_2 films grown on Si substrates, guiding the design and production of optical devices. ANN

returned the optical refractive index, absorption coefficient, and film thickness related to multiple output frequencies, providing accurate production information for special requirements [159]. Sridevi et al. used NNs to guide the design of photonic crystal fibers (PCFs), and solved the problem of long calculation time and repeated design parameters in existing numerical simulation software based on finite element methods [160]. A PCF temperature sensor optical parameter prediction method based on acoustic emission and NN was proposed, which utilizes an automatic encoder network to increase data and provide enough data for network training. Fig. 23 (a) shows the cross-section of the PCF, indicating that many parameters need to be optimized in the design. Compared to the running time of finite element simulation, this model had a higher R^2 and lower computational time. In addition, compared to the NN model without self-coding, this model had a low MSE. Nithish et al. proposed a rectangular PCF sensor for detecting reproductive hormones in female blood samples in the terahertz band [161]. The optimal structure's identification was achieved by simulating various parameters of PCF to attain high performance. However, as shown in Fig. 23 (b), many parameters need to be optimized in the design, and simulation requires a certain amount of time. Therefore, local weighted linear regression was used for prediction and hyperparameter optimization. Dey et al. demonstrated a rapid training feedforward algorithm for studying the key characteristics of different FBGs and predicting their output spectrum [162]. This model efficiently identified nonlinearity and complexity within the output spectrum of FBGs. The model can accurately and effectively predict the effective refractive index, bandwidth, reflectance and wavelength of different types of FBG, and generate an accurate reflection spectrum to help and guide the production of FBG under different needs. Dwivedi et al. established a GPR model to adjust the design parameters of optical fiber plasma sensors, assisting researchers in rapid sensor design for different demands [163]. Kim et al. utilized NNs and Bayesian regularization algorithms to optimize the design of ultra-compact SPR optical fiber sensors using multi-guided waves [164]. For different applications, data has different characteristics, so the most suitable algorithm is different. However, it is undeniable that the use of ML algorithms is more efficient and quicker than conventional design processes.

5. Conclusion and future perspectives

This paper reviews the application of the ML algorithm in optical fiber sensors. In recent years, the advent of ML has greatly impacted optical development. Compared with static demodulation algorithms and complex spatial optical paths, the ML algorithm offers better signal processing mechanisms for optics. ML algorithms enable automated processing, analysis, and recognition of spectrums, thereby enhancing the efficiency and accuracy of spectrum analysis. Using ML algorithms to classify and identify the spectrum of different substances, enabling qualitative and quantitative analysis of substances. By learning from known corresponding spectrums and building models, ML algorithms can predict and analyze unknown spectrums. ML algorithms can extract and recognize features from the spectrum, analyze the content and proportion of various components, and reveal the intrinsic nature of substances. With the continuous development of spectroscopy and ML algorithms, the application prospects of ML in spectroscopy will become increasingly widespread. This advancement will contribute to more efficient and accurate analytical methods in chemical analysis, environmental monitoring, food safety, and other fields, promoting progress in scientific research and expanding application fields.

While the capability of ML algorithms to achieve demodulation without additional hardware circuits represents a significant advancement, the reliance on large amounts of labeled data for training poses challenges. Acquiring extensive labeled data for spectrum demodulation is particularly difficult, as it necessitates professional expertise and entails high measurement costs. To minimize data complexity and ensure accurate identification and classification of spectrums, application features often need expert annotation. The lack of labeled data limits the generalization ability of models as they cannot learn accurate patterns from limited data. Spectrum can be affected by various noises such as instrument noise, environmental interference, drift, etc., which can adversely affect the performance and robustness of ML algorithms. Dealing with data quality and noise requires appropriate preprocessing and noise reduction methods to improve data reliability and accuracy. ML algorithms can optimize spectrum, improving spectrum resolution and SNR. Meanwhile, because the ML algorithm relies on the distribution of training data, it is often difficult to accurately detect and handle outliers. Outliers are common in the spectrum and can have a significant impact on the analysis results. In such cases, prior knowledge is also required for the preprocessing spectrum. The DL algorithm may offer a solution by autonomously learning advanced features from the data. This eliminates the need for domain specific knowledge and core feature extraction. The pursuit of self-learning models is an important direction for the application of ML in optical fiber sensors. ML has been widely used in the signal processing of optical fiber sensors, particularly in predicting measurement parameters. Due to the cross-sensitivity problem in optical fiber sensors, conventional demodulation methods for the output (such as temperature) are prone to errors if other parameters (such as strain or humidity) change. ML demodulation avoids this problem and can achieve accurate demodulation. ML algorithms can also solve the nonlinear relationship between measurement parameters and spectrum by selecting nonlinear models such as SVMs and NNs.

Spectrum demodulation requires extracting useful features from the

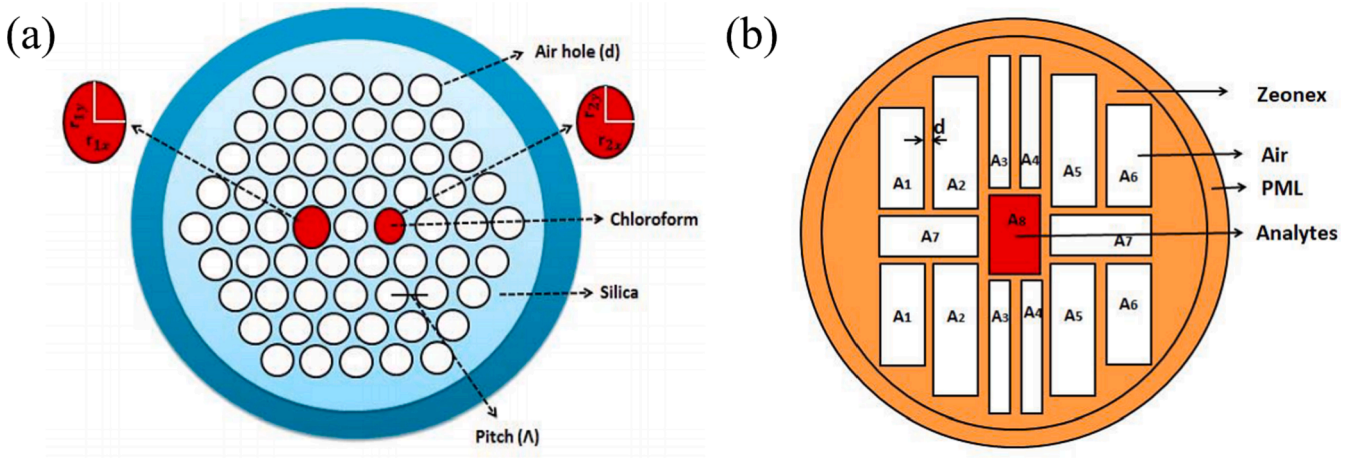

Fig. 23. (a) Cross-sectional view of dual-core PCF [160],(b) Cross-section view of the designed PCF [161].

original spectrum. However, feature extraction from the spectrum is a challenge, as the spectrum is typically high-dimensional and contains a large amount of feature information. Effectively processing and utilizing this information is one of the challenges that ML faces in spectroscopy. The large amount of labeled data also brings huge computational requirements, requiring powerful computing resources, including graphic processing units—similar to challenges faced by ML in other research fields. Feature selection and dimensionality reduction methods can help reduce the number of features while retaining important information and increasing the efficiency of data computation.

Choosing an appropriate ML model is also an important issue in spectrum demodulation, as different models come with distinct advantages and disadvantages. and the selection of models usually needs to be adjusted according to specific demodulation tasks. Selecting the appropriate model correctly requires a comprehensive consideration of factors such as data size, feature representation, and computational resources based on the specific spectrum demodulation tasks. Most ML algorithms can be applied to optical fiber sensors, whether they are demodulated in the time domain or frequency domain. Special ML algorithms can be applied according to special requirements, such as spectrum recovery under sparse sampling, to improve data acquisition speed. The application of ML algorithms in high-precision recovery of low-resolution signals and prediction of monitoring process trends makes the performance of optical fiber sensors more attractive. However, challenges in large-scale production of optical instruments impact model selection, training, maintenance, and standardization, underscoring the need for robust and adaptable ML models. This is not only an algorithmic problem, but also an industrial challenge. Due to the possible differences in spectrum under different conditions, the performance of a model on one dataset often cannot be well generalized to other datasets. To improve the generalization ability of the model, it is necessary to use more complex models (such as transfer learning) and use more data to train the model. Real-time signal processing and analysis capabilities are crucial for practical monitoring models. Most model training discussed in the paper is offline, while online monitoring encounters unexpected states, raising questions about model updating and maintenance. Although the decision process of many ML models is difficult to explain, this lack of interpretability may limit the reliability of the model. Despite the challenges, compared with conventional demodulation methods, the ML demodulation method offers better performance and remains an exciting prospect. Researchers are studying more interpretable ML algorithms. With the improved explanation of the ML model, the demodulation of the spectrum will become more persuasive.

In this paper, the limitations of different optical fiber sensors in conventional demodulation methods are introduced, and then examples of ML algorithms are introduced. Then, we emphasized the achievements of ML in optical fiber sensors and explained the challenges faced in further advancing this direction in the future. In general, the application of ML in spectrum demodulation still faces some challenges and shortcomings. But, by optimizing existing ML algorithms, and strengthening cross research in the fields of spectrum demodulation and ML, these problems are gradually being overcome. The development of better data acquisition methods, feature extraction techniques, model selection strategies, and interpretive ML algorithms will help improve the practicality and reliability of ML in spectrum demodulation. We believe that shortly, the advantages brought by the synergy between hybrid optical systems and AI technology can be better utilized, thereby better constructing a green intelligent world.

CRediT authorship contribution statement

Yifan Zhou: Writing – original draft, Resources, Investigation. **Yanan Zhang:** Project administration, Funding acquisition, Conceptualization. **Qi Yu:** Writing – review & editing, Data curation. **Lirong Ren:** Writing – review & editing, Resources. **Qi Liu:** Writing – review & editing, Data curation. **Yong Zhao:** Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under grants 62075036, the Liaoning Revitalization Talents Program under grants XLYC2203125, the State Key Laboratory of Synthetical Automation for Process Industries under grant 2018ZCX31, and the Natural Science Foundation of Hebei Province under grant F2020501040.

References

[1] W. Ma, Z. Liu, Z.A. Kudyshev, A. Boltasseva, W. Cai, Y. Liu, Deep learning for the design of photonic structures, Nat. Photonics 15 (2020) 77–90.
[2] D. Pile, Machine learning with light, Nat. Photonics 15 (2021) 68–69.
[3] G. Genty, L. Salmela, J.M. Dudley, D. Brunner, A. Kokhanovskiy, S. Kobtsev, et al., Machine learning and applications in ultrafast photonics, Nat. Photonics 15 (2020) 91–101.
[4] J. Liu, Q. Wu, X. Sui, Q. Chen, G. Gu, L. Wang, et al., Research progress in optical neural networks: theory, applications and developments, PhotoniX 2 (2021) 1–39.
[5] B.J. Shastri, A.N. Tait, T. Ferreira de Lima, W.H.P. Pernice, H. Bhaskaran, C. D. Wright, et al., Photonics for artificial intelligence and neuromorphic computing, Nat. Photonics 15 (2021) 102–114.
[6] Z. Wang, Y. Xiao, K. Liao, T. Li, H. Song, H. Chen, et al., Metasurface on integrated photonic platform: from mode converters to machine learning, Nanophotonics 11 (2022) 3531–3546.
[7] J. Feldmann, N. Youngblood, C.D. Wright, H. Bhaskaran, W.H.P. Pernice, All-optical spiking neurosynaptic networks with self-learning capabilities, Nature 569 (2019) 208–214.
[8] A. Argyris, Photonic neuromorphic technologies in optical communications, Nanophotonics 11 (2022) 897–916.
[9] D. Azzimonti, C. Rottondi, M. Tornatore, Reducing probes for quality of transmission estimation in optical networks with active learning, Journal of Optical Communications and Networking 12 (2019) A38–A48.
[10] Y. Zhou, Y.-N. Zhang, B. Han, L. Cheng, D. Li, W. Zheng, et al., Biochemical sensor based on functional material assisted optical fiber surface plasmon resonance: a review, Measurement 207 (2023) 112353.
[11] C. Karapanagiotis, K. Krebber, Machine learning approaches in brillouin distributed fiber optic sensors, Sensors (basel) 23 (2023) 6187.
[12] K. Venkatesan, A. Chandrasekar, P.G.V. Ramesh, On-demand DWDM design using machine learning, Soft. Comput. 26 (2022) 6577–6589.
[13] U.M.N. Jayawickrema, H.M.C.M. Herath, N.K. Hettiarachchi, H. P. Sooriyaarachchi, J.A. Epaarachchi, Fibre-optic sensor and deep learning-based structural health monitoring systems for civil structures: a review, Measurement 199 (2022) 111543.
[14] M. Johnson Singh, S. Choudhary, W.-B. Chen, P.-C. Wu, M. Kumar Goyal, A. Rajput, et al., Applications of fibre bragg grating sensors for monitoring geotechnical structures: a comprehensive review, Measurement 218 (2023) 113171.
[15] A.G. Leal-Junior, A. Frizera, C. Marques, M.J. Pontes, Optical fiber specklegram sensors for mechanical measurements: a review, IEEE Sens. J. 20 (2020) 569–576.
[16] J. Li, Y. Wang, P. Wang, Q. Bai, Y. Gao, H. Zhang, et al., Pattern recognition for distributed optical fiber vibration sensing: a review, IEEE Sens. J. 21 (2021) 11983–11998.
[17] A.D. Gomes, H. Bartelt, O. Frazão, Optical Vernier effect: recent advances and developments, Laser Photonics Rev. 15 (2021) 2000588.
[18] Z. Ran, X. He, Y. Rao, D. Sun, X. Qin, D. Zeng, et al., Fiber-optic microstructure sensors: a review, Photonic Sensors 11 (2021) 227–261.
[19] Y. Xu, P. Bai, X. Zhou, Y. Akimov, C.E. Png, L.K. Ang, et al., Optical refractive index sensors with plasmonic and photonic structures: promising and inconvenient truth, Adv. Opt. Mater. 7 (2019) 1801433.
[20] Y. Zhao, X.-G. Li, X. Zhou, Y.-N. Zhang, Review on the graphene based optical fiber chemical and biological sensors, Sens. Actuators B 231 (2016) 324–340.
[21] F. Cunzheng, L. Hao, H. Tao, Z. Shixiong, Y. Baoqiang, Y. Zhijun, et al., Large dynamic range optical fiber distributed acoustic sensing (DAS) with differential-unwrapping-integral algorithm, J. Lightwave Technol. 39 (2021) 7274–7280.

[22] X. Wen, H. Shuai, F. Zhu, D. Zhang, Fiber bragg gratings sensing network with a bus chain typology structure, Opt. Eng. 55 (2016) 066102.
[23] C.-H. Yeh, Y.-H. Zhuang, N. Tsai, C.-W. Chow, Capacity and capability enhancements of FBG sensor system by utilizing intensity and WDM detection technique, Smart Mater. Struct. 26 (2017) 035026.
[24] X. Gui, Z. Li, X. Fu, H. Guo, Y. Wang, C. Wang, et al., Distributed optical fiber sensing and applications based on large-scale fiber bragg grating Array: review, J. Lightwave Technol. 41 (2023) 4187–4200.
[25] P. Lu, N. Lalam, M. Badar, B. Liu, B.T. Chorpening, M.P. Buric, et al., Distributed optical fiber sensing: review and perspective, Appl. Phys. Rev. 6 (2019) 041302.
[26] M.F. Bado, J.R. Casas, A review of recent distributed optical fiber sensors applications for civil engineering structural health monitoring, Sensors (basel) 21 (2021) 1818.
[27] C. Zuo, J. Qian, S. Feng, W. Yin, Y. Li, P. Fan, et al., Deep learning in optical metrology: a review, Light Sci Appl 11 (2022) 39.
[28] A. Lugnan, A. Katumba, F. Laporte, M. Freiberger, S. Sackesyn, C. Ma, et al., Photonic neuromorphic information processing and reservoir computing, APL Photonics 5 (2020) 020901.
[29] A.M. Palmieri, E. Kovlakov, F. Bianchi, D. Yudin, S. Straupe, J.D. Biamonte, et al., Experimental neural network enhanced quantum tomography, npj Quantum Information, 6(2020) 20.
[30] A. Durand, T. Wiesner, M.A. Gardner, L.E. Robitaille, A. Bilodeau, C. Gagne, et al., A machine learning approach for online automated optimization of super-resolution optical microscopy, Nat Commun 9 (2018) 5247.
[31] L. Ho, C. Ane, A linear-time algorithm for gaussian and non-gaussian trait evolution models, Syst Biol 63 (2014) 397–408.
[32] S. Nusinovici, Y.C. Tham, M.Y. Chak Yan, D.S. Wei Ting, J. Li, C. Sabanayagam, et al., Logistic regression was as good as machine learning for predicting major chronic diseases, J. Clin. Epidemiol. 122 (2020) 56–69.
[33] H. Jing, C. Zhao, Adjustable piecewise regression strategy based wind turbine power forecasting for probabilistic condition monitoring, Sustainable Energy Technol. Assess. 52 (2022) 102013.
[34] H. Jing, C. Zhao, F. Gao, Non-stationary data reorganization for weighted wind turbine icing monitoring with gaussian mixture model, Comput. Chem. Eng. 147 (2021) 107241.
[35] B. Kaminski, M. Jakubczyk, P. Szufel, A framework for sensitivity analysis of decision trees, cent, Eur J Oper Res 26 (2018) 135–159.
[36] S.B. Kotsiantis, Decision trees: a recent overview, Artif. Intell. Rev. 39 (2011) 261–283.
[37] Y. Chen, W. Zheng, W. Li, Y. Huang, Large group activity security risk assessment and risk early warning based on random forest algorithm, Pattern Recogn. Lett. 144 (2021) 1–5.
[38] A. Hajjem, F. Bellavance, D. Larocque, Mixed-effects random forest for clustered data, J. Stat. Comput. Simul. 84 (2012) 1313–1328.
[39] Z. Tan, G. De, M. Li, H. Lin, S. Yang, L. Huang, et al., Combined electricity-heat-cooling-gas load forecasting model for integrated energy system based on multi-task learning and least square support vector machine, J. Clean. Prod. 248 (2020) 119252.
[40] S. Shamshirband, S. Hashemi, H. Salimi, S. Samadianfard, E. Asadi, S. Shadkani, et al., Predicting standardized streamflow index for hydrological drought using machine learning models, Eng. Appl. Comput. Fluid Mech. 14 (2020) 339–350.
[41] P. Rebentrost, M. Mohseni, S. Lloyd, Quantum support vector machine for big data classification, Phys Rev Lett 113 (2014) 130503.
[42] M. Tanveer, T. Rajani, R. Rastogi, Y.H. Shao, M.A. Ganaie, Comprehensive review on twin support vector machines, Ann. Oper. Res. (2022) 1–46.
[43] A.M. Ikotun, A.E. Ezugwu, L. Abualigah, B. Abuhaija, J. Heming, K-means clustering algorithms: a comprehensive review, variants analysis, and advances in the era of big data, Inf. Sci. 622 (2023) 178–210.
[44] A.A. Movassagh, J.A. Alzubi, M. Gheisari, M. Rahimi, S. Mohan, A.A. Abbasi, et al., Artificial neural networks training algorithm integrating invasive weed optimization with differential evolutionary model, journal of ambient intelligence and humanized, Computing 14 (2021) 6017–6025.
[45] H. Zhang, Z. Wang, D. Liu, A comprehensive review of stability analysis of continuous-time recurrent neural networks, IEEE Trans. Neural Networks Learn. Syst. 25 (2014) 1229–1262.
[46] X. Ma, J. Mo, J. Zhang, J. Huang, Optical fiber vibration signal recognition based on the fusion of multi-scale features, Sensors (basel) 22 (2022) 6012.
[47] R. Rakkiyappan, J. Cao, G. Velmurugan, Existence and uniform stability analysis of fractional-order complex-valued neural networks with time delays, IEEE Trans Neural Netw Learn Syst 26 (2015) 84–97.
[48] Y. Wang, M.L. Adam, Y. Zhao, W. Zheng, L. Gao, Z. Yin, et al., Machine learning-enhanced flexible mechanical sensing, Nanomicro Lett 15 (2023) 55.
[49] L. Schmarje, M. Santarossa, S.-M. Schroder, R. Koch, A survey on semi-, self- and unsupervised learning for image classification, IEEE, Access 9 (2021) 82146–82168.
[50] Z. Tian, J. Li, L. Liu, H. Wu, X. Hu, M. Xie, et al., Machine learning-assisted self-powered intelligent sensing systems based on triboelectricity, Nano Energy 113 (2023) 108559.
[51] S. Haustein, A.F. Jensen, Factors of electric vehicle adoption: a comparison of conventional and electric car users based on an extended theory of planned behavior, Int. J. Sustain. Transp. 12 (2018) 484–496.
[52] G. Knell, M.C. Robertson, E.E. Dooley, K. Burford, K.S. Mendez, Health behavior changes during COVID-19 pandemic and subsequent "stay-at-home" orders, Int J Environ Res Public Health 17 (2020) 6268.
[53] T. Benkedjouh, K. Medjaher, N. Zerhouni, S. Rechak, Health assessment and life prediction of cutting tools based on support vector regression, J. Intell. Manuf. 26 (2013) 213–223.
[54] A. Stetco, F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, et al., Machine learning methods for wind turbine condition monitoring: a review, Renew. Energy 133 (2019) 620–635.
[55] N. Fumo, M.A. Rafe Biswas, Regression analysis for prediction of residential energy consumption, Renewable Sustainable Energy Rev. 47 (2015) 332–343.
[56] C.K.I. Williams, C.E. Rasmussen, Gaussian processes for machine learning, MIT press, Cambridge, MA, 2006.
[57] S. Gao, Y. Huang, S. Zhang, J. Han, G. Wang, M. Zhang, et al., Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation, J. Hydrol. 589 (2020) 125188.
[58] R.M.A. Ikram, A.A. Ewees, K.S. Parmar, Z.M. Yaseen, S. Shahid, O. Kisi, The viability of extended marine predators algorithm-based artificial neural networks for streamflow prediction, Appl. Soft Comput. 131 (2022) 109739.
[59] E. Momeni, D. Jahed Armaghani, M. Hajihassani, M.F. Mohd Amin, Prediction of uniaxial compressive strength of rock samples using hybrid particle swarm optimization-based artificial neural networks, Measurement 60 (2015) 50–63.
[60] E. Momeni, R. Nazir, D. Jahed Armaghani, H. Maizir, Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN, Measurement 57 (2014) 122–131.
[61] D. Wang, Z.J. Shen, X. Yin, S. Tang, X. Liu, C. Zhang, et al., Model predictive control using artificial neural network for power converters, IEEE Trans. Ind. Electron. 69 (2022) 3689–3699.
[62] M.T. Hoang, B. Yuen, X. Dong, T. Lu, R. Westendorp, K. Reddy, Recurrent neural networks for accurate RSSI indoor localization, IEEE Internet Things J. 6 (2019) 10639–10651.
[63] X. Yang, Y. Ye, X. Li, R.Y.K. Lau, X. Zhang, X. Huang, Hyperspectral image classification with deep learning models, IEEE Trans. Geosci. Remote Sens. 56 (2018) 5408–5423.
[64] Y.-L. Chang, T.-H. Tan, W.-H. Lee, L. Chang, Y.-N. Chen, K.-C. Fan, et al., Consolidated convolutional neural network for hyperspectral image classification, Remote Sens. (Basel) 14 (2022) 1571.
[65] C. Zhu, E.A. Chan, Y. Wang, W. Peng, R. Guo, B. Zhang, et al., Image reconstruction through a multimode fiber with a simple neural network architecture, Sci Rep 11 (2021) 896.
[66] Z. Li, F. Liu, W. Yang, S. Peng, J. Zhou, A survey of convolutional neural networks: analysis, applications, and prospects, IEEE Trans. Neural Netw. Learn Syst. 33 (2022) 6999–7019.
[67] Y. Deng, X. Zhou, J. Shen, G. Xiao, H. Hong, H. Lin, et al., New methods based on back propagation (BP) and radial basis function (RBF) artificial neural networks (ANNs) for predicting the occurrence of haloketones in tap water, Sci Total Environ 772 (2021) 145534.
[68] C. Ren, N. An, J. Wang, L. Li, B. Hu, D. Shang, Optimal parameters selection for BP neural network based on particle swarm optimization: a case study of wind speed forecasting, Knowl.-Based Syst. 56 (2014) 226–239.
[69] S. Xiong, B. Li, S. Zhu, DCGNN: a single-stage 3D object detection network based on density clustering and graph neural network, Complex Intell. Sys. 9 (2022) 3399–3408.
[70] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, Y. Miao, Review of image classification algorithms based on convolutional neural networks, Remote Sens. (Basel) 13 (2021) 4712.
[71] J. Wen, J. Yang, B. Jiang, H. Song, H. Wang, Big data driven marine environment information forecasting: a time series prediction network, IEEE Trans. Fuzzy Syst. 29 (2021) 4–18.
[72] Y.-X. Wang, Z. Chen, W. Zhang, Lithium-ion battery state-of-charge estimation for small target sample sets using the improved GRU-based transfer learning, Energy 244 (2022) 123178.
[73] W. Zhang, H. Li, L. Tang, X. Gu, L. Wang, L. Wang, Displacement prediction of jiuxianping landslide using gated recurrent unit (GRU) networks, Acta Geotech. 17 (2022) 1367–1382.
[74] E. Chemali, P.J. Kollmeyer, M. Preindl, R. Ahmed, A. Emadi, Long short-term memory networks for accurate state-of-charge estimation of li-ion batteries, IEEE Trans. Ind. Electron. 65 (2018) 6730–6739.
[75] H. Palangi, L. Deng, Y. Shen, J. Gao, X. He, J. Chen, et al., Deep sentence embedding using long short-term memory networks: analysis and application to information retrieval, IEEE/ACM Trans. Audio Speech Lang. Process. 24 (2016) 694–707.
[76] R. Yin, L. Cao, Q. Huang, H. Yang, W. Ji, L. Lu, et al., Nonlinear regression: a possible solution to larger dynamic range for some spectrum-based optical sensors, Measurement 199 (2022) 111506.
[77] C. Zhu, Y. Zhuang, J. Huang, Machine learning assisted high-sensitivity and large-dynamic-range curvature sensor based on no-Core fiber and hollow-Core fiber, J. Lightwave Technol. 40 (2022) 5762–5767.
[78] C. Zhu, J. Huang, Machine learning boosts performance of optical fiber sensors: a case study for vector bending sensing, Opt Express 30 (2022) 24553–24564.
[79] C. Zhu, J. Huang, Self-Vernier effect-assisted optical fiber sensor based on microwave photonics and its machine learning analysis, J. Lightwave Technol. 41 (2023) 1890–1898.
[80] J. Xue, L. Zhao, T. Wang, L. Zhao, F. Cui, A novel compact fiber optic concentration sensing system based on machine learning demodulation, IEEE Photonics J. 15 (2023) 1–6.
[81] D. Yu, X. Qiao, X. Wang, Light intensity optimization of optical fiber stress sensor based on SSA-LSTM model, Front. Energy Res. 10 (2022) 972437.

[82] M. Filosa, L. Massari, D. Ferraro, G. D'Alesio, J. D'Abbraccio, A. Aliperta, et al., A meta-learning algorithm for respiratory flow prediction from FBG-based wearables in unrestrained conditions, Artif Intell Med 130 (2022) 102328.
[83] Y. Mei, S. Zhang, Z. Cao, T. Xia, X. Yi, Z. Liu, Deep learning assisted pressure sensing based on sagnac interferometry realized by side-hole fiber, J. Lightwave Technol. 41 (2023) 784–793.
[84] W.S. Duque, C.A. Rodriguez Diaz, A.G. Leal-Junior, A. Frizera, Fiber-Optic Hydrophone Based on Michelson's Interferometer with Active Stabilization for Liquid Volume Measurement, Sensors (Basel) 22 (2022) 4404.
[85] R. Martínez-Manuel, J. Esquivel-Hernández, L.M. Valentín-Coronado, S. LaRochelle, Fiber-optic sensor for directional bending monitoring based on a pattern recognition algorithm, IEEE Sens. J. 23 (2023) 4819–4824.
[86] R. Martinez-Manuel, L.M. Valentin-Coronado, J. Esquivel-Hernandez, K.-J.-J. Monga, S. LaRochelle, Machine learning implementation for unambiguous refractive index measurement using a self-referenced fiber refractometer, IEEE Sens. J. 22 (2022) 14134–14141.
[87] E. Chubchev, K. Tomyshev, I. Nechepurenko, A. Dorofeenko, O. Butov, Machine learning approach to data processing of TFBG-assisted SPR sensors, J. Lightwave Technol. 40 (2022) 3046–3054.
[88] S. Ren, S. Chen, H. Xu, X. Hou, Q. Yang, G. Wang, et al., Unsupervised representation learning-based Spectrum reconstruction for demodulation of fabry-perot interferometer sensor, IEEE Sens. J. 23 (2023) 13810–13816.
[89] J. Hao, C. Jing, L. Tundong, Wavelength detection in spectrally overlapped FBG sensor network using extreme learning machine, IEEE Photon. Technol. Lett. 26 (2014) 2031–2034.
[90] A. Elleathy, F. Alhumaidan, M. Alqahtani, A.S. Almaiman, A.M. Ragheb, A. B. Ibrahim, et al., Strain FBG-based sensor for detecting fence intruders using machine learning and adaptive thresholding, Sensors (basel) 23 (2023) 5015.
[91] X. Xu, Y. Wang, D. Zhu, J. Shi, Accurate strain extraction via kernel extreme learning machine for fiber bragg grating sensor, IEEE Sens. J. 22 (2022) 7792–7797.
[92] Z. Cao, T. Xia, S. Zhang, S. Zhang, Y. Mei, Z. Liu, et al., Improved spectral interrogation of tilted fiber bragg grating refractometer using residual convolutional neural networks, J. Lightwave Technol. 40 (2022) 7403–7411.
[93] D. Pal, A. Kumar, A. Gautam, J. Thangaraj, FBG based optical weight measurement system and its performance enhancement using machine learning, IEEE Sens. J. 22 (2022) 4113–4121.
[94] V. Pegorini, L.Z. Karam, C.S. Pitta, R. Cardoso, J.C. da Silva, H.J. Kalinowski, et al., In vivo pattern classification of ingestive behavior in ruminants using FBG sensors and machine learning, Sensors (basel) 15 (2015) 28456–28471.
[95] M.B. Reid, M. Ozcan, Temperature dependence of fiber optic bragg gratings at low temperatures, Opt. Eng. 37 (1) (1998) 237–240.
[96] Q. Shang, W. Qin, Fiber bragg grating dynamic calibration based on online sequential extreme learning machine, Sensors (basel) 20 (2020) 1840.
[97] Z. Cao, S. Zhang, T. Xia, Z. Liu, Z. Li, Spectral demodulation of fiber bragg grating sensor based on deep convolutional neural networks, J. Lightwave Technol. 40 (2022) 4429–4435.
[98] Y. Li, T. Chen, J. Si, R. Lv, X. Niu, B. Gao, et al., Ultra-high-temperature sensing using fiber grating sensor and demodulation method based on support vector regression optimized by a genetic algorithm, Opt Express 31 (2023) 3401–3414.
[99] S. Dhanalakshmi, P. Nandini, S. Rakshit, P. Rawat, R. Narayanamoorthi, R. Kumar, et al., Fiber bragg grating sensor-based temperature monitoring of solar photovoltaic panels using machine learning algorithms, Opt. Fiber Technol. 69 (2022) 102831.
[100] M.S.E. Djurhuus, S. Werzinger, B. Schmauss, A.T. Clausen, D. Zibar, Machine learning assisted fiber bragg grating-based temperature sensing, IEEE Photon. Technol. Lett. 31 (2019) 939–942.
[101] N. Ushakov, A. Markvart, L. Liokumovich, Singlemode-multimode-singlemode fiber-optic interferometer signal demodulation using MUSIC algorithm and machine learning, Photonics 9 (2022) 879.
[102] A.G. Leal-Junior, V. Campos, C. Díaz, R.M. Andrade, A. Frizera, C. Marques, A machine learning approach for simultaneous measurement of magnetic field position and intensity with fiber bragg grating and magnetorheological fluid, Opt. Fiber Technol. 56 (2020) 102184.
[103] S. Sarkar, D. Inupakutika, M. Banerjee, M. Tarhani, M. Shadaram, Machine learning methods for discriminating strain and temperature effects on FBG-based sensors, IEEE Photon. Technol. Lett. 33 (2021) 876–879.
[104] J. Liu, Y. Hou, J. Wang, G. Zhong, L. Zhang, F. Zhuang, et al., Multi-parameter demodulation for temperature, salinity and pressure sensor in seawater based on the semi-encapsulated microfiber mach-zehnder interferometer, Measurement 196 (2022) 111213.
[105] M. Kikuchi, T. Ogasawara, S. Fujii, S.-I. Takeda, Application of machine learning for improved accuracy of simultaneous temperature and strain measurements of carbon fiber-reinforced plastic laminates using an embedded tilted fiber bragg grating sensor, Compos. A Appl. Sci. Manuf. 161 (2022) 107108.
[106] O. Costilla-Reyes, P. Scully, K.B. Ozanyan, Temporal pattern recognition for gait analysis applications using an "intelligent carpet" system, IEEE Sens. J. 16 (2015) 8815–8822.
[107] J.L. Haiming Huang, L. Wu, B. Fang, Z. Wen, F. Sun, Machine learning-based multi-modal information perception for soft robotic hands, Tsinghua Sci. Technol. 25 (2020) 255–269.
[108] Y.-N. Pang, B. Liu, J. Liu, S.-P. Wan, T. Wu, J. Yuan, et al., Singlemode-multimode-singlemode optical fiber sensor for accurate blood pressure monitoring, J. Lightwave Technol. 40 (2022) 4443–4450.
[109] T. Li, F. Qiao, P.A. Huang, Y. Su, L. Wang, X. Li, et al., Flexible optical fiber-based smart textile sensor for human-machine interaction, IEEE Sensors J. 22 (2022) 19336–19345.
[110] L.G. Gruionu, A.L. Udristoiu, A.V. Iacob, C. Constantinescu, R. Stan, G. Gruionu, Feasibility of a lung airway navigation system using fiber-bragg shape sensing and artificial intelligence for early diagnosis of lung cancer, PLoS One 17 (2022) e0277938.
[111] A.M. Caravaca-Aguirre, E. Niv, D.B. Conkey, R. Piestun, Real-time resilient focusing through a bending multimode fiber, Opt Express 21 (2013) 12881–12887.
[112] P. Caramazza, O. Moran, R. Murray-Smith, D. Faccio, Transmission of natural scene images through a multimode fibre, Nat Commun 10 (2019) 2029.
[113] Y. Li, Y. Xue, L. Tian, Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media, Optica 5 (2018) 1181–1190.
[114] L. Gu, H. Gao, H. Hu, Demonstration of a learning-empowered fiber specklegram sensor based on focused ion beam milling for refractive index sensing, Nanomaterials (basel) 13 (2023) 768.
[115] S. Lu, Z. Tan, G. Li, Y. Jingya, A sensitized plastic fiber sensor for multi-point bending measurement based on deep learning, IEEE Photonics J. 13 (2021) 1–7.
[116] A.R. Cuevas, M. Fontana, L. Rodriguez-Cobo, M. Lomer, J.M. Lopez-Higuera, Machine learning for turning optical fiber specklegram sensor into a spatially-resolved sensing system. proof of concept, J. Lightwave Technol. 36 (2018) 3733–3738.
[117] Z. Yang, L. Gu, H. Gao, H. Hu, Demodulation of fiber specklegram curvature sensor using deep learning, Photonics 10 (2023) 169.
[118] Z. Sun, K. Liu, J. Jiang, T. Xu, S. Wang, H. Guo, et al., Optical fiber distributed vibration sensing using grayscale image and multi-class deep learning framework for multi-event recognition, IEEE Sens. J. 21 (2021) 19112–19120.
[119] W. Xiong, C.W. Hsu, H. Cao, Long-range spatio-temporal correlations in multimode fibers for pulse delivery, Nat Commun 10 (2019) 2973.
[120] J. Pauwels, G. Van der Sande, G. Verschaffelt, Space division multiplexing in standard multi-mode optical fibers based on speckle pattern classification, Sci Rep 9 (2019) 17597.
[121] B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, C. Moser, Multimode optical fiber transmission with a deep learning network, Light Sci Appl 7 (2018) 69.
[122] Z. Liu, L. Wang, Y. Meng, T. He, S. He, Y. Yang, et al., All-fiber high-speed image detection enabled by deep learning, Nat Commun 13 (2022) 1433.
[123] W. Naku, C. Zhu, A.K. Nambisan, R.E. Gerald, J. Huang, Machine learning identifies liquids employing a simple fiber-optic tip sensor, Opt Express 29 (2021) 40000–40014.
[124] W. Naku, A.K. Nambisan, M. Roman, C. Zhu, R.E. Gerald 2nd, J. Huang, Identification of volatile organic liquids by combining an Array of fiber-optic sensors and machine learning, ACS Omega 8 (2023) 4597–4607.
[125] X. Wang, Y. Wang, K. Zhang, K. Althoefer, L. Su, Learning to sense three-dimensional shape deformation of a single multimode fiber, Sci Rep 12 (2022) 12684.
[126] S. Razmyar, M.T. Mostafavi, Deep learning for estimating deflection direction of a multimode fiber from specklegram, J. Lightwave Technol. 39 (2021) 1850–1857.
[127] L. Kong, K. Cui, J. Shi, M. Zhu, S. Li, 1D phase unwrapping based on the quasi-gramian matrix and deep learning for interferometric optical fiber sensing applications, J. Lightwave Technol. 40 (2022) 252–261.
[128] L. Kong, X. Pan, Z. Ren, K. Cui, Robust one-dimensional phase unwrapping algorithm based on LSTM network with reduced parameter number, J. Lightwave Technol. (2022) 1–8.
[129] N.L. Saleh, B. Faisal, M.S. Yusri, A.H. Sulaiman, M.F. Ismail, N.A.H.A. Nik Zulkefli, et al., Human activities classification based on ϕ-OTDR system by utilizing gammatone filter cepstrum coefficient envelope using support vector machine, Optics Laser Technol. 164 (2023) 109417.
[130] Y. Pan, T. Wen, W. Ye, Time attention analysis method for vibration pattern recognition of distributed optic fiber sensor, Optik 251 (2022) 168127.
[131] L. Huang, Y. Li, S. Chen, Q. Zhang, Y. Song, J. Zhang, et al., Building safety monitoring based on extreme gradient boosting in distributed optical fiber sensing, Opt. Fiber Technol. 55 (2020) 102149.
[132] H. Liu, J. Ma, T. Xu, W. Yan, L. Ma, X. Zhang, Vehicle detection and classification using distributed fiber optic acoustic sensing, IEEE Trans. Veh. Technol. 69 (2020) 1363–1374.
[133] H. Jia, S. Lou, S. Liang, X. Sheng, Event identification by F-ELM model for φ-OTDR fiber-optic distributed disturbance sensor, IEEE Sens. J. 20 (2020) 1297–1305.
[134] S. Xu, Z. Qin, W. Zhang, X. Xiong, H. Li, Z. Wei, et al., A novel method of recognizing disturbance events in Φ-OTDR based on affinity propagation clustering and perturbation signal selection, IEEE Sens. J. 21 (2021) 13272–13282.
[135] X. Zhao, H. Sun, B. Lin, H. Zhao, Y. Niu, X. Zhong, et al., Markov transition fields and deep learning-based event-classification and vibration-frequency measurement for φ-OTDR, IEEE Sens. J. 22 (2022) 3348–3357.
[136] Z. Zhou, W. Jiao, X. Hu, D. Zhang, J. Qu, X. Zheng, et al., Open-set event recognition model using 1-D RL-CNN with OpenMax algorithm for distributed optical fiber vibration sensing system, IEEE Sens. J. 23 (2023) 12817–12827.
[137] A.M. Rizzo, L. Magri, D. Rutigliano, P. Invernizzi, E. Sozio, C. Alippi, et al., Known and unknown event detection in OTDR traces by deep learning networks, Neural Comput. Applic. 34 (2022) 19655–19673.
[138] H. Wu, B. Zhou, K. Zhu, C. Shang, H.Y. Tam, C. Lu, Pattern recognition in distributed fiber-optic acoustic sensor using an intensity and phase stacked convolutional neural network with data augmentation, Opt Express 29 (2021) 3269–3283.

[139] Y. Shang, J. Wang, S. Huang, S. Qu, Q. He, M. Wang, et al., Fault identification method based on generative adversarial network in distributed acoustic sensing, Meas. Sci. Technol. 34 (2023) 115117.
[140] W. Wei, S. Xu, L. Zhang, J. Zhang, Y. Zhang, Boosting hyperspectral image classification with unsupervised feature learning, IEEE Trans. Geosci. Remote Sens. 60 (2022) 1–15.
[141] Z. Peng, H. Wen, J. Jian, A. Gribok, M. Wang, S. Huang, et al., Identifications and classifications of human locomotion using rayleigh-enhanced distributed fiber acoustic sensors with deep neural networks, Sci Rep 10 (2020) 21014.
[142] J. Zhang, X. Zhao, Y. Zhao, X. Zhong, Y. Wang, F. Meng, et al., Unsupervised learning method for events identification in φ-OTDR, Opt. Quant. Electron. 54 (2022) 457.
[143] J. Wang, Y. Li, J. Liao, Temperature extraction for brillouin optical fiber sensing system based on extreme learning machine, Opt. Commun. 453 (2019) 124418.
[144] Y. Zhang, Y. Lu, J. Peng, C. He, Z. Zhang, Fast and accurate brillouin optical time-domain sensing by sparse frequency sampling and ANN-based recover method, IEEE Sens. J. 23 (2023) 6892–6900.
[145] Y. Zhang, L. Yu, Z. Hu, L. Cheng, H. Sui, H. Zhu, et al., Ultrafast and accurate temperature extraction via kernel extreme learning machine for BOTDA sensors, J. Lightwave Technol. 39 (2021) 1537–1543.
[146] H. Wu, Y. Wan, M. Tang, Y. Chen, C. Zhao, R. Liao, et al., Real-time denoising of brillouin optical time domain analyzer with high data Fidelity using convolutional neural networks, J. Lightwave Technol. 37 (2019) 2648–2653.
[147] N.D. Nordin, M.S.D. Zan, F. Abdullah, Comparative analysis on the deployment of machine learning algorithms in the distributed brillouin optical time domain analysis (BOTDA) fiber sensor, Photonics 7 (2020) 79.
[148] H. Zhu, L. Yu, Y. Zhang, L. Cheng, Z. Zhu, J. Song, et al., Optimized support vector machine assisted BOTDA for temperature extraction with accuracy enhancement, IEEE Photonics J. 12 (2020) 1–14.
[149] J. Chen, H. Jiang, T. Liu, X. Fu, Wavelength detection in FBG sensor networks using least squares support vector regression, J. Opt. 16 (2014) 045402.
[150] B. Li, Z.W. Tan, P.P. Shum, C. Wang, Y. Zheng, L.J. Wong, Dilated convolutional neural networks for fiber bragg grating signal demodulation, Opt Express 29 (2021) 7110–7123.
[151] S. Ren, S. Chen, J. Yang, J. Wang, Q. Yang, C. Xue, et al., High-efficiency FBG array sensor interrogation system via a neural network working with sparse data, Opt Express 31 (2023) 8937–8952.
[152] A. Usman, N. Zulkifli, M.R. Salim, K. Khairi, Fault monitoring in passive optical network through the integration of machine learning and fiber sensors, Int. J. Commun Syst 35 (2022) e5134.
[153] Y.C. Manie, P.-C. Peng, R.-K. Shiu, Y.-T. Hsu, Y.-Y. Chen, G.-M. Shao, et al., Enhancement of the multiplexing capacity and measurement accuracy of FBG sensor system using IWDM technique and deep learning algorithm, J. Lightwave Technol. 38 (2020) 1589–1603.
[154] A. Kokhanovskiy, N. Shabalov, A. Dostovalov, A. Wolf, Highly dense FBG temperature sensor assisted with deep learning algorithms, Sensors (basel) 21 (2021) 6188.
[155] Y. Shao, C. Chen, Z. Lu, Y. Zheng, Y. Zhang, An intelligent leakage detection method for diaphragm wall joints based on fiber bragg grating sensors and intelligent algorithms, Measurement 197 (2022) 111339.
[156] K.P. Nascimento, A. Frizera-Neto, C. Marques, A.G. Leal-Junior, Machine learning techniques for liquid level estimation using FBG temperature sensor array, Opt. Fiber Technol. 65 (2021) 102612.
[157] F. Liu, Y. Yang, B. Fu, Q. Qi, Uncertain fiber bragg grating sensor data analysis based on sparse bayesian learning, 2012 fifth international symposium on computational intelligence and design, IEEE 2 (2012) (2012) 85–88.
[158] J. Jin, Y. Zhu, Y. Zhang, D. Zhang, Z. Zhang, Micrometeoroid and orbital debris impact detection and location based on FBG sensor network using combined artificial neural network and mahalanobis distance method, IEEE Trans. Instrum. Meas. 70 (2021) 1–10.
[159] B. Lee, K. Yu, J. Jeon, E.J. Choi, Machine learning analysis of broadband optical reflectivity of semiconductor thin film, J. Korean Phys. Soc. 80 (2022) 347–351.
[160] S. Sridevi, T. Kanimozhi, N. Ayyanar, S. Chugh, M. Valliammai, J. Mohanraj, Deep learning based data augmentation and behavior prediction of photonic crystal fiber temperature sensor, IEEE Sens. J. 22 (2022) 6832–6839.
[161] A.N. Nithish, S.K. Patel, N. Ayyanar, J. Surve, S. Rajaram, S.N. Deepa, et al., Terahertz women reproductive hormones sensor using photonic crystal fiber with behavior prediction using machine learning, IEEE Access 11 (2023) 75424–75433.
[162] K. Dey, V. Nikhil, P.R. Chaudhuri, S. Roy, Demonstration of a fast-training feed-forward machine learning algorithm for studying key optical properties of FBG and predicting precisely the output spectrum, Opt. Quant. Electron. 55 (2022) 16.
[163] Y.S. Dwivedi, R. Singh, A.K. Sharma, A.K. Sharma, Enhancing the performance of photonic sensor using machine-learning approach, IEEE Sens. J. 23 (2023) 2320–2327.
[164] H.-T. Kim, M. Yu, On-fiber multiparameter sensor based on guided-wave surface plasmon resonances, J. Lightwave Technol. 40 (2022) 2157–2165.