



## Near-infrared methane sensor with neural network filtering



Linbo Tian <sup>a,b,1</sup>, Jiachen Sun <sup>b,c,1</sup>, Sasa Zhang <sup>a,b,c,\*</sup>, Alexandre A. Kolomenskii <sup>d</sup>, Hans A. Schuessler <sup>d</sup>, Zhao Wang <sup>a,b</sup>, Jinbao Xia <sup>e,f</sup>, Jun Chang <sup>a,b,c,\*</sup>, Zhaojun Liu <sup>a,b,c,\*</sup>

<sup>a</sup> Key Laboratory of Education Ministry for Laser and Infrared System Integration Technology, Shandong University, 72 Binhai Road, Qingdao 266237, China

<sup>b</sup> Shandong Provincial Key Laboratory of Laser Technology and Application, Shandong University, 72 Binhai Road, Qingdao 266237, China

<sup>c</sup> School of Information Science and Engineering, Shandong University, 72 Binhai Road, Qingdao 266237, China

<sup>d</sup> Department of Physics and Astronomy, Texas A&M University, College Station, Texas 77843–4242, USA

<sup>e</sup> State Key Laboratory of Crystal Materials, Shandong University, Jinan 250100, China

<sup>f</sup> The State Key Laboratory of Applied Optics, Changchun 130000, China

### ARTICLE INFO

#### Keywords:

Near-infrared methane sensor  
Neural network filter  
Direct absorption spectroscopy  
High sensitivity and stability

### ABSTRACT

A novel methane sensor based on neural network filter (NNF) assisted direct absorption spectroscopy (DAS) technique was proposed and experimentally demonstrated. The developed detection device adds the benefits of a digital filter based on the neural network, thereby compensating the shortcomings of traditional DAS. We overcame the scarce data problem by using the simulated absorption spectra that are highly consistent with practical experimental conditions to construct and train the NNF. The proposed NNF showed the best performance compared with several widely used filtering algorithms. We performed a detailed assessment of the NNF-improved detection system. The proposed sensor shows more accurate concentration retrieval and better stability in a real-time measurement. The minimum detection limit of 2.93 ppm•m ( $1\sigma$ ) was obtained, which is a significant improvement compared to previous reports of near-infrared methane detection with the DAS technique. Finally, we systematically discuss the frequency principle underlying the NNF to explicitly interpret the mechanism of the generalized filtering. The improved methane sensor proves the feasibility of enhancing the performance of DAS technique with the neural network algorithm and broad applicability of this approach to the high-sensitivity measurements of methane and other trace gases.

### 1. Introduction

Methane is a greenhouse gas second only to carbon dioxide, which seriously affects the global climate and environment. Being the main component of natural gas, methane is a colorless, odorless, non-toxic, but flammable and explosive gas [1–3]. Therefore, in consideration of environmental protection and public safety, it is very important to monitor the in real time methane concentration (typical atmospheric concentration  $\sim 2$  ppm) and whether it is elevated and approaches the lower explosive limit (5–15%) [4–6]. Optical sensors have the advantages of high sensitivity, specificity and wide dynamic range; also, being non-invasive they provide fast response, and long life-time. Therefore, methane sensors based on absorption spectroscopy have become the first choice for many engineering applications. Methane sensors employing other principles, such as electrochemical [7], semiconductor [8], gas

chromatography [9], catalytic combustion [10] and surface acoustic wave [11] sensors often have insufficient sensitivity, specificity and dynamic range, limiting their applicability.

The sensitivity of sensors has always been of major priority for researchers. Several high-sensitivity detection techniques have been applied to the gas detection, among them cavity-enhanced absorption spectroscopy (CEAS) [12], cavity ring-down spectroscopy (CRDS) [13], Fourier transform infrared spectroscopy (FTIR) [14], photoacoustic spectroscopy (PAS) [15] and quartz-enhanced photoacoustic spectroscopy (QEPAS) [16]. However, due to the influence of temperature, humidity, vibration and noise, these technologies are hard to adapt to the full-range detection and hazardous site applications, such as coal mines and natural gas pipelines. In contrast, tunable diode laser absorption spectroscopy (TDLAS) is widely deployed in such applications, in particular, such techniques as direct absorption spectroscopy (DAS) [17]

\* Corresponding authors at: Key Laboratory of Education Ministry for Laser and Infrared System Integration Technology, Shandong University, 72 Binhai Road, Qingdao 266237, China

E-mail addresses: [sasazhang@sdu.edu.cn](mailto:sasazhang@sdu.edu.cn) (S. Zhang), [changjun@sdu.edu.cn](mailto:changjun@sdu.edu.cn) (J. Chang), [zhaojunliu@sdu.edu.cn](mailto:zhaojunliu@sdu.edu.cn) (Z. Liu).

<sup>1</sup> Linbo Tian and Jiachen Sun contribute equally to this work.

and wavelength modulation spectroscopy (WMS) [18]. Although WMS has the advantage of shifting the absorption signal to the high-frequency region, thereby greatly suppressing 1/f noise, it also has disadvantages, such as high cost, complexity of the system and nonlinearity at high concentrations. Therefore, in such application, as the full-range concentration measurement, triggering alarm at certain level of methane content and in consideration of the required costs, the DAS is more suitable. The concentration of the absorbing gas molecules can be directly derived by calculating the integral area of the absorption signal in appropriately selected spectral interval. However, the sensitivity and resolution of the DAS technique is more seriously affected by different types of noise (intrinsic laser, white noise, 1/f and interference noise, etc. [19]). Therefore, a highly precise and efficient processing of the measured spectra is crucial.

Software based filtering techniques have become a preferable choice because of their simplicity of implementation and low cost. Among a variety of filtering techniques that have been applied to process gas absorption spectra, only a few are considered effective for online filtering. Multi-signal average filtering (MAF) can decrease white noise by a factor of  $1/\sqrt{N}$  by averaging N spectra. Although widely used, this method is very time-consuming, and it can only suppress white noise [20,21]. Savitzky-Golay (S-G) filtering is a smoothing technique, requiring selection of the smoothing window size and the order of the polynomial. However, it is designed for filtering high frequency noise, but not low frequency noise [22,23]. Kaltman filtering (KF) algorithm uses state space equations to optimally estimate the system state from a series of data with known normal distributed noise covariance [24,25]. Its performance is satisfactory for white noise with known statistical characteristics, whereas KF may cause a serious distortion of the signal in a nonlinear system. Wavelet transform (WT) denoising technique is a signal time-frequency analysis method for processing local or transient signals. Such analysis method focuses on the low-frequency trend and high-frequency details of the signal and can effectively filter out Gaussian white noise and interference noise [26,27]. However, the WT method still has obvious shortcomings, including a large number of preset parameters, such as wavelet basis functions, decomposition level, threshold, etc., strong subjectivity in parameter selection, reduced by required calculations processing speed and weak protection against external noise interferences [28]. Currently, the state-of-the-art filtering method is the dual-optimized BP adaptive Kalman filter (BP-KF), which is based on the back propagation (BP) neural network and variance compensation [29,30]. Although this novel filtering method optimizes the parameters of the Kalman filter by using the BP neural network and eliminates the variations in the parameters of dynamic systems through variance compensation, it is still based on the accurate set of the state space equations and evaluation of the system noise. Therefore, the accuracy under different detection conditions and the applicability of different application scenarios are greatly limited.

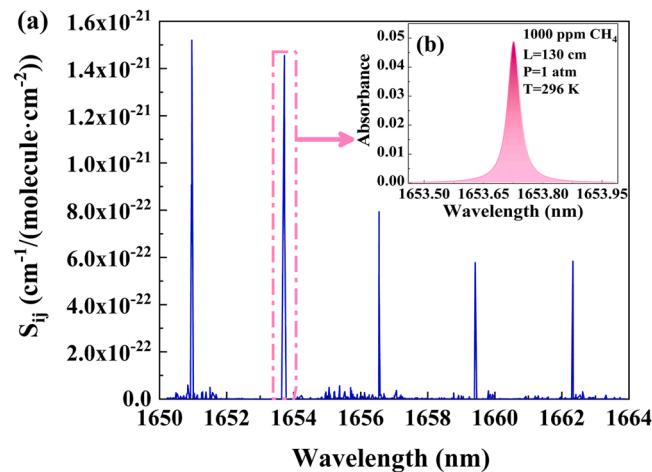
In addition, a kind of near-infrared spectrum preprocessing algorithm based on orthogonal signal correction (OSC) has been widely developed in chemometrics [31]. The OSC is mainly used for the preprocessing of near-infrared spectral matrix. Its basic idea is to filter out the information not related to the property Y to be measured in the original spectra matrix X by means of mathematical orthogonality. Therefore, the OSC method can ensure that the filtered information is independent of the property to be measured such as absorption peaks of non-target components or noise etc. It is reasonable that different chemical components should be orthogonal and OSC can effectively eliminate the interference of other non-target components and realize the decoupling of mixed spectra through the physical properties Y and spectra matrix X when spectral analysis is carried out for some chemical multi-components. However, the filtering performance of OSC and its variants, such as direct orthogonal signal correction (DOSC) [32], is barely satisfactory when the spectral signal-to-noise ratio (SNR) is very low and the noise is not strictly orthogonal to the concentration matrix.

Based on the above background, we build a neural network filter (NNF) to achieve the efficient and authentic signal extraction from the noise spectrum. To date, deep learning algorithms represented by neural networks have achieved exciting results in hyperspectral imaging classification [33], ultrashort pulse reconstruction [34], material discovery [35] and gas concentration retrieval [36]. Through a large number of input and output training samples, the neural network searches the mapping relationship between the input and output without prior knowledge. In this paper, the NNF follows a *Frequency Principle* [37], which is different from the artificially preset filtering algorithms. It extracts different frequency components in different extend in the order from the fundamental frequency to high frequency, and the extraction degree decreases in turn. As a result, NNF could retain the low-frequency components as much as possible, but not completely. At the same time, it keeps partial high-frequency components rather than discarding all the high-frequency parts. Subsequently NNF restores the extracted frequency components to the filtered signal. The reason why this is more effective than the direct high-frequency cut-off is that the real signal is not completely in the low-frequency region, and complex signals often have high-frequency components. NNF can restore the signal to the greatest extent to avoid signal distortion and loss of information. The sensor configuration is introduced in Section 2. We described in detail in Section 3 how we constructed and trained our NNF, and in Section 4 we showed the comparison of the filtering effect of the NNF and other filtering algorithms mentioned above. With the excellent filtering performance of the NNF, the methane sensor performance has been enhanced achieving lower detection limit of the column density of 2.93 ppm·m and better detection stability. We carried out a detailed evaluation of the performance of the methane sensor enhanced by the NNF in Section 4. In Section 5, from the perspective of the frequency domain, we discuss the advantages and underlying properties of our NNF, comparing it with other filtering algorithms.

## 2. Sensor configuration

### 2.1. Selection of the absorption line

The optimal target absorption line of methane should assure high sensitivity and avoid cross interference with other atmospheric species. The assessment of the interference of water vapor and carbon dioxide in the air to methane detection is contained in the Supplementary. The strong R(3) absorption line near 1653.7 nm, depicted in Fig. 1, satisfies these requirements and is selected for methane detection. We note that the cost of the required 1653 nm near-infrared laser is much lower than



**Fig. 1.** Absorption lines of  $\text{CH}_4$  in the spectral range of 1650–1664 nm based on HITRAN database. Inset: absorption spectra of 1000 ppm  $\text{CH}_4$  near 1653.7 nm with  $L = 130 \text{ cm}$ ,  $P = 1 \text{ atm}$  and  $T = 296 \text{ K}$ .

that of the mid-infrared laser. The parameters for the simulation in the inset of Fig. 1 as well as the conditions for the experiments described in this paper are as follows: temperature  $T = 296$  K, pressure  $P = 1$  atm and optical path length  $L = 130$  cm.

## 2.2. Sensor system configuration

The setup of the methane detection system is depicted as Fig. 2. The distributed feedback laser diode (DFB-LD) (SWLD-165310S22-01, All-wave Devices Inc., China) with a wavelength near 1653 nm was employed. The DFB-LD was controlled by a laser diode controller (LDC501, Stanford Research Systems Inc., USA), and its driving signal was sawtooth wave signal of 0.9 Hz generated by signal generator (FY2300A, Feel Tech, China). The output beam of the DFB-LD was coupled into an optical fiber and propagated to a fiber-coupled Herriott cell. 100 sets of methane gas of different concentrations diluted with high-purity nitrogen were produced with a gas mixing system and introduced into the Herriott cell, as shown in the dashed box in Fig. 2. Among them, the Mass Flow Controller 1 and the Mass Flow Controller 2 (CS200A, Beijing Sevenstar Flow Co.,Ltd., China) were employed to control the flows in the two gas paths to enable the configuration of samples with different  $\text{CH}_4$  concentrations. The errors of sample gas configuration are less than 2%. The method of concentration configuration and error evaluation of mass flow controller are shown in the Supplementary. The output light intensity from the Herriott cell was converted to a voltage signal by a photodetector (BF14-PD300-F-N, Wuhan 69 Sensor Technology, China) with a pre-amplifier. Finally, the experimental data processing was completed by the PC.

## 3. Methodology

### 3.1. Dataset

Despite excellent performance in a variety of fields, the deep learning is often under criticism, since the quality of deep learning model predictions often depends on the quantity of datasets. However, in the field

of gas absorption spectroscopy, it is extremely time-consuming and laborious to experimentally obtain a large number of datasets that can meet the requirements of deep learning training. In the present study, the method to solve scarce data problem is to generate a large simulated dataset from the same distribution as the target data, which are applied to train the NNF. Such model-agnostic, physics-informed data augmentation strategy has gradually become the mainstream method to solve the problem of data scarcity in the fields where the information is scarce, and has been proved effective [34,38]. Therefore, with simulated experimental conditions, we created a set of 10,000 spectra for different concentrations of methane and used it for the training and evaluation of the neural network. The simulated methane concentrations range from 0 to 1000 ppm and other parameters correspond to experimental conditions. In order to mimic the noisy distribution of real measurements, we added the Gaussian white noise (with the mean of 0 and the variance of  $9.71886 \times 10^{-5}$ ) and the interference noise (simulated by an Airy function) [39] to the generated pure absorption spectra to simulate the detector circuit noise and optical device noise in the sensor system. Therefore, in our study, the simulation dataset is composed of data pairs which consist of noisy absorption spectra (input data) and de-noised absorption spectra (ground truth). Each input data and ground truth are composed of 1111 sampling points, corresponding to the 1111 neurons in the input and output layers in the neural network. The more detailed description of the simulated spectral dataset is provided in the Supplementary.

In this study, we employed the 10-fold cross-validation (CV) and hold-out (HO) as approaches to tune and evaluate the architecture and hyperparameters with respect to the neural network. Specifically, we first execute the HO by randomly divide the whole dataset into 90% training set and 10% test set. Subsequently using the simulated training set, the neural network architecture with the optimal combination of hyper-parameters is obtained through CV, in which the training set is divided into 10 subsets evenly, and take 9 subsets as training data and 1 subset as validation data in turn to carry out the experiment. The final model performance metrics are computed to determine the optimal combination of hyperparameters of the neural network in this particular

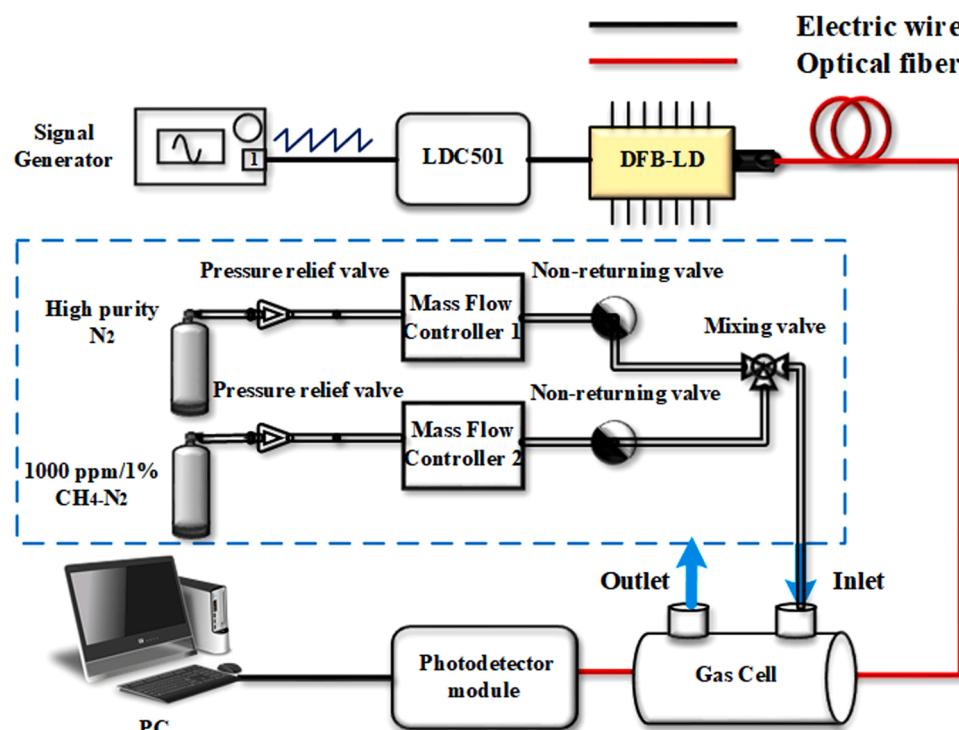


Fig. 2. Schematic diagram of the experimental setup.

task. In Section 3.3 we explicitly show the influence of different architectures on the de-noising effect in detail. Once the optimal architecture parameters as well as hyperparameters have been determined, we retrain the neural network through the complete training set to generate the NNF. In Section 4, we demonstrate the effect of several filtering algorithms (including our NNF) on the test set, as an objective evaluation of the filtering performance. In order to keep the correctness of the implementation of the BP-KF filtering model, we keep its' original architecture as described in [30]. For a fair comparison, the BP-KF is also trained on our training set and its filtering performance is evaluated on our test set. Furthermore, we experimentally collected 100 sets of methane absorption spectra for concentrations ranging from 0 to 1000 ppm with the sensor system shown in Fig. 2 to compare the filtering performance of various filtering algorithms under practical application conditions. The complete algorithm development process is shown in Fig. 3.

### 3.2. Neural network concept

Many applications of deep learning use feed forward neural network architectures, which aim to map a fixed-size input to a fixed-size output. Such property is suitable for us to achieve the de-noising function while maintaining the consistency of the input and output spectra. To go from one layer to the next, a set of units in each layer computes a weighted sum of their inputs from the previous layer and passes the results through a non-linear function as expressed in the following equations:

$$z^{(l)} = W^{(l)} \bullet a^{(l-1)} + b^{(l)}, \quad (1)$$

$$a^{(l)} = g(z^{(l)}), \quad (2)$$

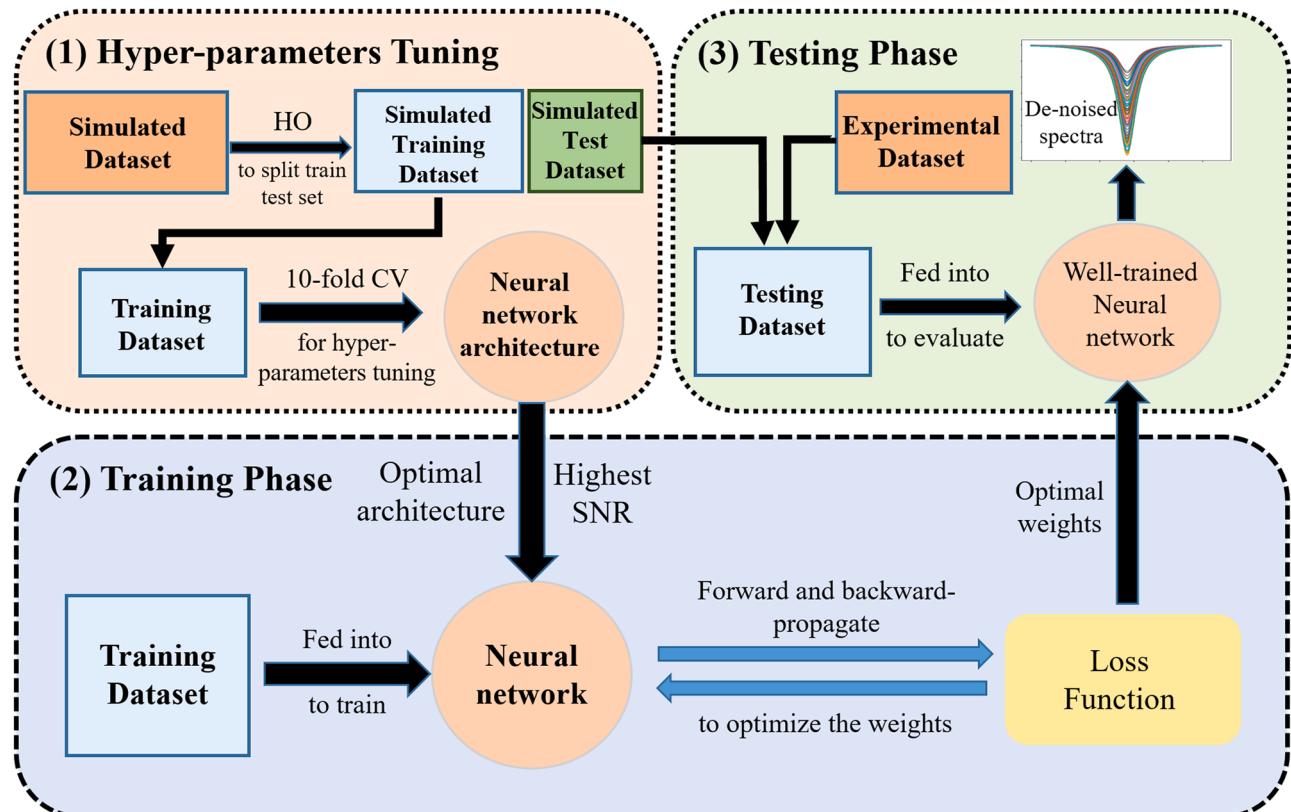
where  $z^{(l)}$  represents the hidden vectors that are linearly related to the weights matrix  $W^{(l)}$ , the bias vectors  $b^{(l)}$  with respect to the  $l$ -th layer and the activation vector  $a^{(l-1)}$  from the previous layer;  $g(\bullet)$  stands for the non-linear activation function. At present, the most popular non-linear function is the rectified linear unit (ReLU), which is simplified as the half-wave rectifier  $g(z^{(l)}) = \max(z^{(l)}, 0)$ . The adjustable parameters ( $W^{(l)}, b^{(l)}$ ), also called weights, are the keys to achieve the optimal input-output function of the neural network. To properly adjust the weight vector, the learning algorithm is introduced to backward propagate the calculation error of the objective function from the output layer all the way to the input layer. To optimize the trainable parameters the gradient of weights with respect to each layer is computed. The process is iteratively performed during training until the objective function converges [40]. The objective function (loss function) in our de-noising task is the mean squared error function (MSE) expressed as

$$\begin{aligned} \mathcal{L} &= \underset{W}{\operatorname{argmin}} \{ \text{loss}(y^{(i)}, \hat{y}^{(i)}) \} \\ &= \underset{W}{\operatorname{argmin}} \left\{ \frac{1}{m} \sum_{i=1}^m [y^{(i)} - \text{NN}(x^{(i)}; W)]^2 \right\}, \end{aligned} \quad (3)$$

where  $x^{(i)}$ ,  $y^{(i)}$ ,  $\hat{y}^{(i)}$  represent the  $i$ -th input data, label and neural network prediction respectively.  $\text{NN}(\bullet)$  represents the neural network mapping function and  $m$  is the index of samples.

### 3.3. Neural network optimization

The noisy signal is formed as the interference of noises from the



**Fig. 3.** The schematic diagram of the developed algorithm framework. (1) Hyper-parameters tuning: the simulated data set is divided into training set and test set by HO. Using the simulated training set, the neural network architecture of the optimal combination of hyper-parameters is obtained through 10-fold CV. (2) Training phase: once the optimal architecture is determined, the simulated training set is used to retrain the model and realize the learning of trainable parameters by continuously reducing the loss function through forward and backward-propagation. (3) Testing phase: the generalization performance of the well-trained neural network filter is finally verified on the test dataset.

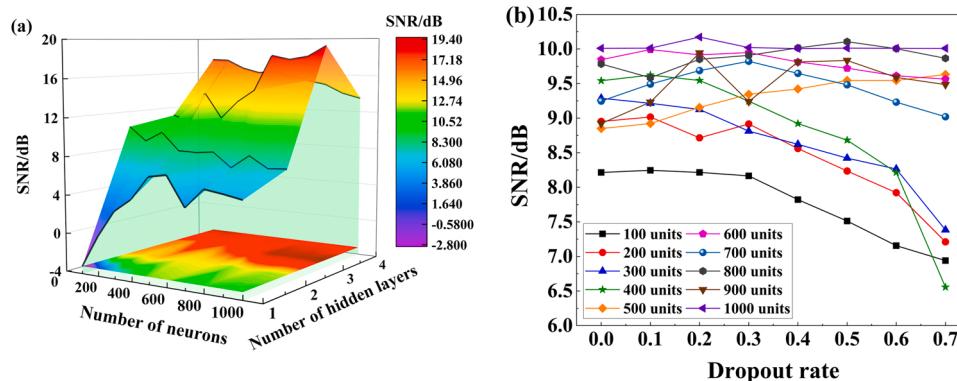
optical devices as well as electronic noises from the detector superimposed on the initial pure signal. In this study, we aim to achieve an accurate approximation of the mapping function between noisy signal and de-noised signal through the construction of an adequate neural network architecture.

Generally, our NNF is based on a neural network composed of multiple consecutive layers of neurons. To find the best model for the filtering performance in terms of the SNR of the de-noised spectrum obtained with NNF, we carry out an extensive search for the best neural network architecture including optimal hyperparameters.

In this search, we start with architecture of one hidden layer and incrementally increase the depth (the number of layers) to improve the fitting capacity of our model until it converges. We also gradually increase the width (the number of neurons) in each layer along with the increase of the model's depth. The SNRs of the predicted outputs of the NNF corresponding to the different combinations of the number of neurons per layer and the number of network layers are shown in Fig. 4(a). In the experiment, the observation is that a three hidden layer architecture is able to achieve good enough generalization model in terms of SNR. Of most concern, however, is the trade-off between the fitting power of the NNF and the overfitting. In this case, the NNF tends to overfit when more than three hidden layers are used, which leads to the risk of a decline in generalization capability. Fig. 4(a) explicitly shows this trend: with the increase of the network depth, the performance of the NNF that extracts signal components from noise has improved, whereas the SNR of predicted signals decrease gradually with the increase of the number of neurons when the NNF reaches four layers. On the other hand, concerning each hidden layer, the number of neurons, which are variables rather than constants has significant effect on the SNR of the output signals predicted by the NNF. Based on the grid search optimization, the numbers of neurons in the hidden layer were finally determined as 500, 100, and 1000 respectively.

A dropout layer was added whenever the number of neurons between consecutive layers changed to avoid overfitting. We explored different dropout rates on the last hidden layer as shown in Fig. 4(b). Increasing dropout rate significantly damages the SNR performance of the NNF and results in under fitting when the model is narrow. By contrast, when the NNF begins to widen, the generalization of the NNF is improved by randomly deactivating some neurons, which forces the network to reduce the dependence on specific neurons. The results of different combinations of the number of neurons and the dropout rate in the last hidden layer in terms of SNR are compared in Fig. 4(b). The optimal effect is achieved for the combination of 1000 units and 0.2 dropout rate. Therefore, the final model architecture was determined with the total of five layers, including three hidden layers as shown in Fig. 5.

The other hyperparameters of NNF model were also tuned by performing an extensive search explained below rather than a grid search that would be unfeasible due to time and computational resource cost.



**Fig. 4.** (a) The effect of the model width and depth on the SNR of output signals. (b) The effect of the dropout in the last hidden layer on the SNR of output signals.

We started with a small range of values for each hyperparameters based on preliminary trials. The search space comprised of different values of momentum, learning rate, optimization algorithms and other hyperparameters. Learning rate values from 0.1 to  $10^{-6}$  were tried, decreasing by a factor of 10. In terms of momentum, we tested values [0.05, 0.9, 0.99] and 0.9 performed the best. The Adam (Adaptive moment estimation) outperformed other optimization algorithms in our study. Similarly, we experimented with a range of values for other hyperparameters.

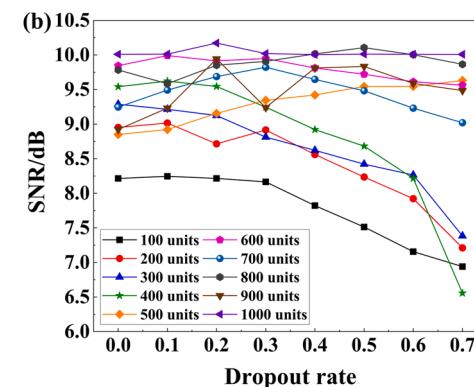
## 4. Experimental results and analysis

### 4.1. Results and analysis of simulated data

As described in Section 2, we first demonstrate and compare the performance of our NNF with various widely used filtering algorithms on the simulated test set. We first show the filtering effect qualitatively in time domain and compare the SNR of de-noised spectra quantitatively.

Concerning the hyperparameter selection of the filter algorithms with respect to the S-G filter, we optimized the window length as well as the given polynomial order [23]. Similarly, the selection of wavelet basis, decomposition level and threshold in the process of wavelet de-noising were also optimized, as reported in [26]. We established the state space equations of the system for the KF and BP-KF based on [30], and evaluated the covariance matrix of noise according to the simulated and experimental spectra. The DOSC is selected as the representative of the OSC variants and has been trained on the same dataset. The number of components regarding to the DOSC has been selected carefully. Although it was suggested in Ref. [32] that one or two components were sufficient, we found in practice that fewer components could hardly filter out noise on our dataset. When it reaches 300, the performance is the best, and the filtering performance starts to decline when it is less than or more than that number. It is also worth noting that we have adjusted the original dataset, that is, we have reduced the number of simulated spectra set from 10,000–1000 to meet the  $N$  (number of samples)  $< K$  (the number of X-variables) requirement. We trained the OSC as well as our NNF using the reduced version of the dataset for fair comparison. (For the full code implementation of the algorithms mentioned above, see the Supplementary).

As shown in Fig. 6, the S-G filter shows the worst filtering performance, even though it improves the low SNR original signal by around 3.3303 dB. As a direct method that is solely based on least squares steps, DOSC can only remove the projection in the direction orthogonal to the concentration matrix, and there are still some noise residues. The SNR of DOSC is 4.3155 dB. By contrast, the filtering effects of KF and BP-KF are much more obvious, and the SNR is increased to 4.4424 dB and 6.0022 dB respectively. Despite its better performance with added neural network to dual-optimize the spectra rather than using plain KF,



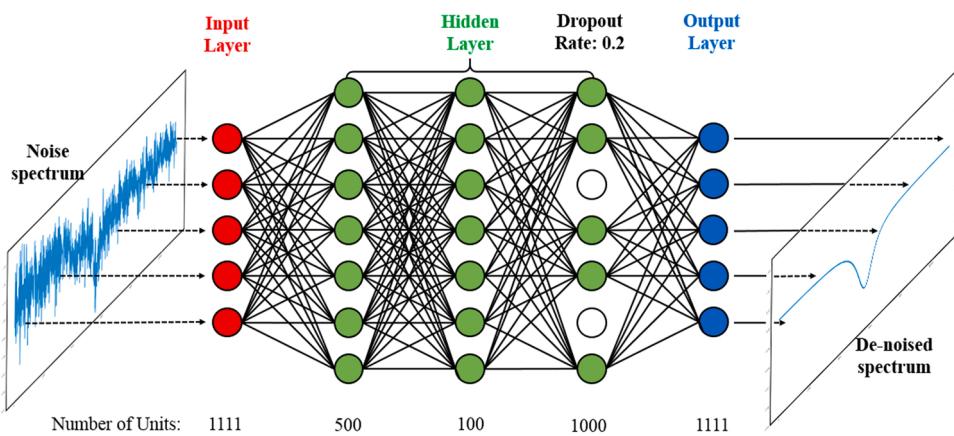


Fig. 5. The schematic architecture of neural network filter.

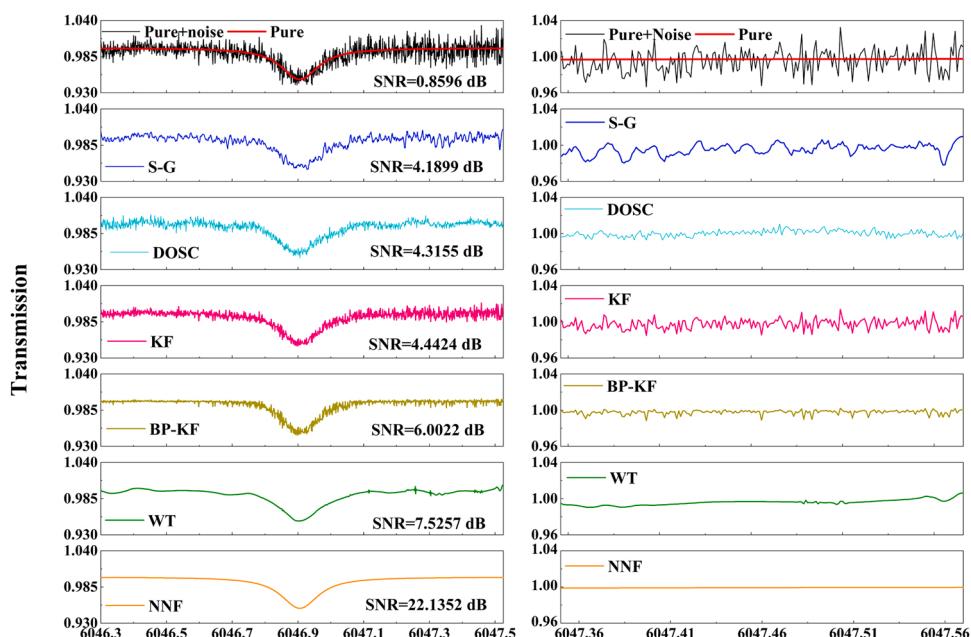


Fig. 6. Filtering effect of several filtering algorithms (for the background comprised of a Gaussian white and an interference noises).

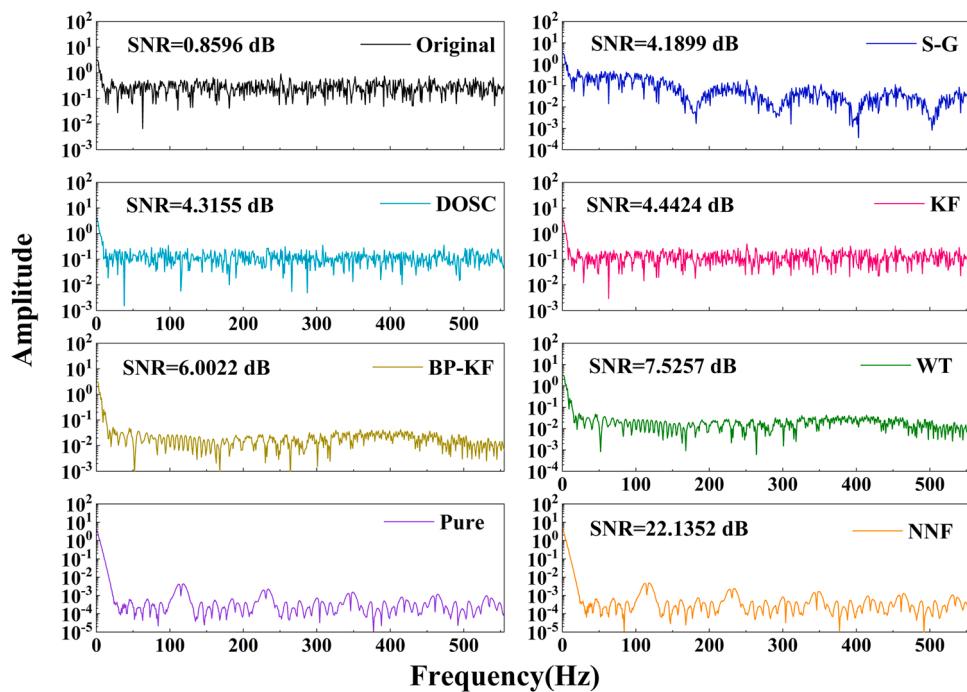
BP-KF is still not as good as reported in [30]. It is worth noting that the performance of the optimized WT is better than that of BP-KF. Although the original waveform is distorted to a certain extent, the SNR of de-noised spectra reach 7.5257 dB after WT. Standing out from the above algorithms, our proposed NNF shows the best performance and the highest SNR among all filter algorithms. The SNR of the original spectrum is improved by about 21.2756 dB. The differences are statistically important ( $P < 0.001$ ), thus the filtered results of NNF significantly outperformed other filters. Thus, Fig. 6 depicts the effects of the various filtering algorithms, showing that NNF largely retains the original signal waveform and extracts the informative signals from the noise.

In addition, we performed discrete-time Fourier transform (DFT) on the noisy and de-noised transmission spectra to observe the changes of different frequency components in the spectra before and after filtering through the frequency domain distribution as shown in Fig. 7. The original noisy spectra are composed of the dominant low-frequency absorption signal and a large amount of irregular low-power broad frequency range noise. The S-G filter obviously suppresses the high frequency regions on the basis of the original frequency distribution, while the effect is not satisfactory, which corresponds to the poor filtering effect in the time domain. The spectral frequency response after

DOSC filtering is similar to that of S-G and the filtering performance is very limited, which is mainly because the noise condition is complicated and is only partially orthogonal to the target property. By contrast, the KF, BP-KF and WT gradually improve the effect of noise suppression. Compared with KF and BP-KF, WT filter not only simply eliminates the high-frequency components, but also makes an effort to retain the frequency distribution characteristics of the informative signals. Same as the explicit results in time domain, the NNF shows the best filtering performance in the process of extracting the signals from noise, and maintains the frequency distribution characteristics of the signal to the maximum extent. The features of NNF are discussed in more details in Section 5.

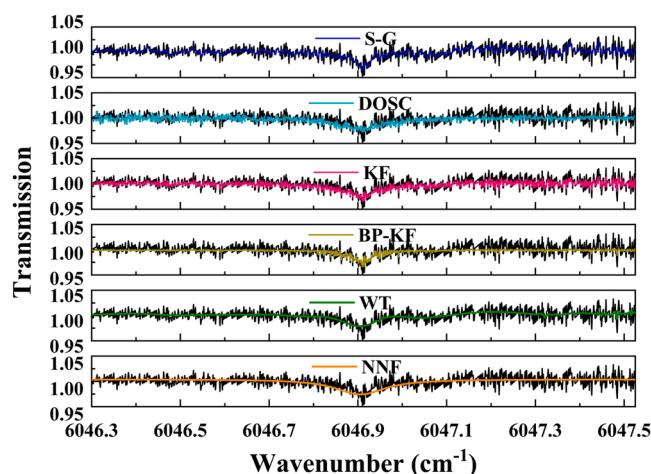
#### 4.2. Results and analysis of experimental data

In order to verify the performance of our NNF under practical experimental conditions, we experimentally collected 100 sets of methane transmitted signals with concentration ranging from 0 to 1000 ppm with the plain methane sensor shown in Fig. 2 and then converted the transmitted signals to transmission spectra to verify the effect of the filtering algorithms.

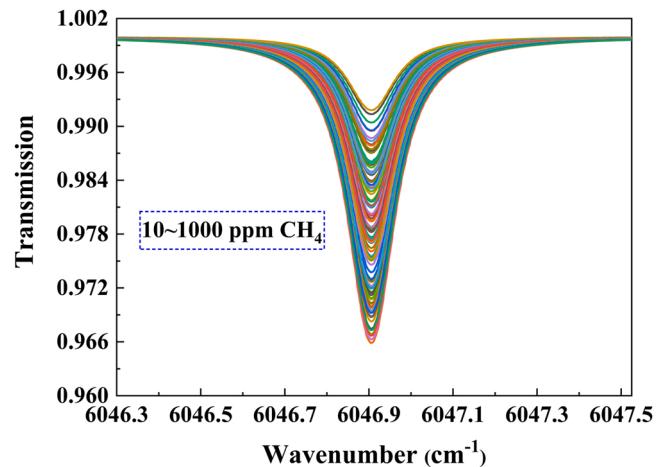


**Fig. 7.** Frequency domain distributions of the original transmission spectrum and filtered transmission spectra.

Here, we selected the typical methane transmission spectrum with a concentration of 500 ppm to intuitively compare the noise suppression effects of different filtering algorithms, as shown in Fig. 8. Although KF, BP-KF, DOSC and S-G filters can restrain the noise in the absorption region, the distortions of the peaks caused by noise still exist. This will undoubtedly greatly reduce the accuracy of identification and fitting of absorption spectrum, thus affecting the precision and stability of concentration detection. In contrast, WT and our NNF significantly smoothed the fluctuating original noisy spectrum, especially in the region of the absorption peak. Furthermore, although our NNF is trained by simulation data, it still outperforms all other listed filters in processing of experimental data and Wilcoxon signed rank tests are statistically significant at  $P < 0.005$  for all cases, which proves the feasibility and generalization power of the model in the practical application scenario. Fig. 9 shows the filtering effect of NNF on the transmission spectra with respect to the complete 100 sets under experimental conditions.



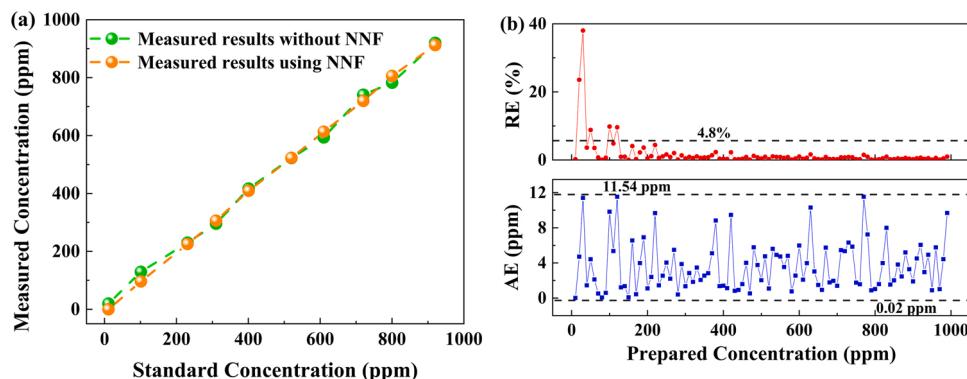
**Fig. 8.** Effects of various filtering algorithms on processing experimental data.



**Fig. 9.** Filtered by NNF experimental transmission spectra.

#### 4.3. Assessment of NNF-assisted methane sensor

After comparison, we retrained the NNF using 90% of complete simulation data set (10,000 spectra) and integrated the well-trained NNF with plain methane sensor and assessed this improvement in a series of experiments. In order to evaluate the detection accuracy of the proposed NNF-assisted methane sensor in concentration retrieval task, we retrieved the corresponding methane concentration from 100 sets of transmission spectra filtered by NNF and verified with the standard gas concentration preset by the mass flow controller (MFC). The measured concentration results of the plain methane sensor (without the NNF enhancement) are considered as the baseline to compare. As illustrated in Fig. 10, the linearity of the methane sensor optimized by the NNF is relatively higher than that of the plain methane sensor without the NNF in terms of the coefficient of determination  $R^2 = 1 - \sum (y - \hat{y})^2 / \sum (y - \bar{y})^2$  (0.9995 and 0.9971 respectively), indicating that the performance as well as the detection accuracy of the methane sensor are significantly improved with the NNF support, where



**Fig. 10.** (a) The coefficients of determination comparison for results with and without NNF (for the sake of clarity, only 10 are plotted out of 100 results). (b) The error analysis of the measured results using NNF: relative error (RE) and absolute error (AE) vs. concentration.

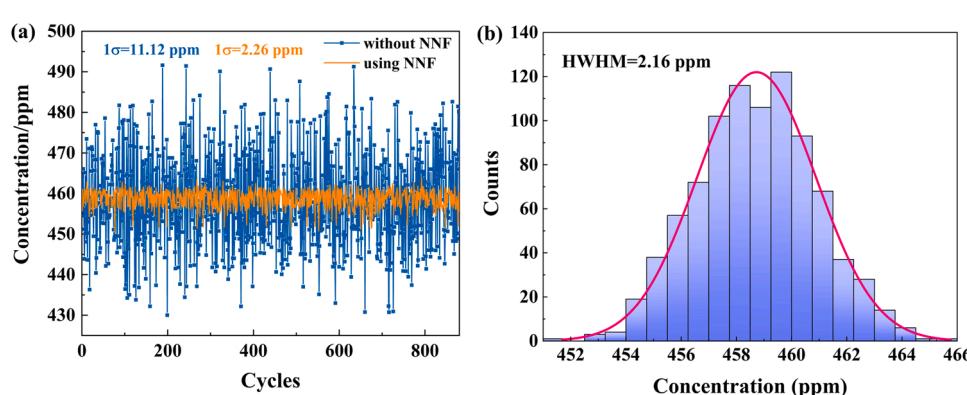
$\bar{y}$  is the mean of the standard concentration  $y$  and  $\hat{y}$  represents the measured concentration. Moreover, the results of error analysis also prove the high detection precision of our methane sensor: despite the relatively poor performance in the extreme low concentration region (meaning that even with the NNF it is difficult to extract the weak absorption signal from the noise), the overall errors remain in a satisfactory range (under 4.8%). On the other hand, the absolute errors of the system are also maintained in a satisfactory range from 0.02 ppm to 11.54 ppm, which shows that the enhancement of methane sensor by NNF is consistent within different concentrations.

In addition, by setting appropriate MFC flow rates we produced a gas mixture with 460 ppm methane concentration for real-time measurements to further evaluate the proposed sensor. The methane concentrations are computed from the fitted peak region of the acquired spectra. Fig. 11 (a) illustrates the concentration retrieval of the methane sensor with and without NNF enhancement. The mean concentration of the plain sensor is 458.58 ppm with the standard deviation of 11.12 ppm, while the mean concentration of methane sensor optimized by NNF is 458.69 ppm with the standard deviation of 2.26 ppm. The unstable power of laser output, interference noise in optical devices and the electric noise in the detector and subsequent circuit greatly affect the stability of measurement, which shows irregular large-scale fluctuations in Fig. 11 (a). In contrast, the NNF-assisted methane sensor is more robust to the noisy conditions and system instability resulting in more stable detection. It should be noted that some deviation between the measured with optimized detection and the preset concentrations was also observed, which is partly due small variations of the gas flow rate controlled by the MFC. Nevertheless, Fig. 11 (b) shows the statistical histogram of the real-time measurements, which can be well fitted by a Gaussian distribution with the HWHM of 2.16 ppm. This demonstrates the good stability of our NNF-assisted methane sensor in the real-time detection. Since the minimum detection limit of the NNF-assisted

methane sensor is 2.26 ppm ( $1\sigma$ ) at room temperature and the optical path length is about 1.3 m, a minimum detectable column density of  $2.26 \text{ ppm} \times 1.3 \text{ m} = 2.93 \text{ ppm m}$  is achieved. In Table 1 we compared our results with three other methane sensors that have been reported and were using the R(3) absorption line near 1653.7 nm. We indicated the detectable column density, which is the most important evaluation parameter, and it can be clearly seen that the NNF-based methane sensor proposed in this paper has the lowest detectable column density. Moreover, our sensor employing the NNF shows ultra-high filtering performance, thus improving the DAS technique and realizing the full-range concentration detection with low cost and simpler system compared with the three other sensors employing WMS technology. In addition, to make our conclusions even more justified, we have added in-situ measurement results. Since our original design intention of methane sensor is to deal with scenarios such as mine and natural gas pipeline transmission leakage. Therefore, we made a special trip to a coal mine in Yan'an city, Shaanxi Province, China for one-day underground real-time measurement. Fig. S5 shows the filtering results of NNF on the transmission spectrum of methane in underground coal mine and the concentration results of real-time measurement. The results show

**Table 1**  
Comparison of our sensor with three other reported sensors.

Technique	Detectable range	Optical path length	Detectable column density	Refs.
WMS	$0-5 \times 10^4$ ppm	40 cm	$\sim 11.8 \text{ ppm m}$	[41]
DAS+WMS	Full range	10 cm	$\sim 6.48 \text{ ppm m}$	[42]
WMS	Unavailable	52.2 m	$\sim 5.22 \text{ ppm m}$	[43]
DAS	Full range	130 cm	$\sim 2.93 \text{ ppm m}$	This paper



**Fig. 11.** (a) Real-time measured and filtered methane concentrations. (b) Statistical histogram for real-time measurements.

that NNF can indeed filter noise from the real-world methane transmission spectra, as shown in Fig. S5 (a). The methane concentration detection results optimized by NNF show more stability and less variance than those unprocessed as shown in Fig. S5 (b), the difference are also statistically significant at  $P < 0.005$ . We believe that the above results are sufficient to prove that the performance of NNF and NNF-assisted methane sensor proposed in our manuscript is feasible for practical application requirements rather than just exist in the laboratory.

## 5. Discussion of the frequency behavior underlying the NNF

Here we analyze in detail the underlying mechanisms of various filtering algorithms to systematically discuss why our NNF shows much better results in signal filtering compared with other filtering algorithms. S-G filter is a filtering algorithm based on local polynomial least square fitting in time domain. Therefore, the selection of the moving window width and polynomial order of the algorithm determine the smoothing effect of the filtering results. If the polynomial order is set to be higher, the S-G filter would be more inclined to fit the abrupt noise points, while if the order is set to be lower, it will be difficult to accurately reproduce the long-term trend of the signal. The OSC method is almost always used together with a latent variable method such as partial least squares (PLS) or principal component regression (PCR) to build the calibration model. When the background noise of our spectral matrix is too large (the SNR of the unprocessed noisy transmission spectrum is as low as 0.86 dB, which is worse at lower concentration level), the spectral loading matrices corresponding to the first several main components usually are not the concentration matrix information, but the spectral signals that are independent of the concentration matrix. When the spectral matrix filters out this part of information, it distorts the useful absorption information. We suspect that this is the reason why the noise reduction effect of OSC is not ideal and even makes worse the absorption signal when filtering our spectral data. Moreover, we believe that the noise with respect to the gas absorption spectra obtained by such kind of methane sensor presented in this manuscript, in fact, is not strictly orthogonal to the concentration matrix. Therefore, the noise conditions of such kind of spectra are more complex than those obtained by high-resolution precision spectrometers, such as Fourier transform spectrometer commonly used in chemometrics. KF is an optimal recursive data processing algorithm based on the idea of data fusion, and its accuracy depends on the accurate representation of the system state space equation and the accurate evaluation of the measurement noise. Although BP-KF applies neural network and variance compensation to dually optimize the output of KF filter, its optimization premise still depends on the output of KF, and thereby it is limited by the assessment of the hyperparameters, such as state matrix, control matrix and noise covariance matrix. Moreover, BP-KF optimizes the filtering result of KF step by step, that is, the neural network only takes the output of KF in each time step rather than predicting the trend of the signal and noise of the complete spectra, which greatly limits its practical applicability and robustness at different experimental conditions. For this reason, in our study the BP-KF did not perform as well as reported in [30]. In contrast to the above algorithm based on time domain data optimization, WT is a time-frequency signal analysis method, which has the features of multi-resolution analysis and the ability to address the local characteristics of the signal in both time and frequency domains. However, the selection of wavelet basis, decomposition level and threshold in the process of WT de-noising are the key factors affecting the final effect and limiting the range of practical applications. While the WT filtering results in superior de-noising, this comes at the price of the increased measurement time caused by the selection of the optimal wavelet basis functions, the decomposition level and the wavelet threshold.

By contrast with the filtering algorithms above, NNF is based on the global learning of the signal behavior. Specifically, we trained the NNF

model with the dataset comprised of the pure spectra with noise, testing the ability of this model to learn and realize the best transformation function from the noisy signal to the de-noised one. Although the architecture of the neural network, including the specific number of layers and the number of nodes in each layer as well as the weight values of each node and the activation results are completely transparent, the neural network is unable to exactly express the mapping function it has learned. Therefore, below we describe an approach for the better interpretability of our NNF and explain the underlying mechanism of the filtering function.

For a particular training sample  $\{x_j; y_j\}$ , where  $x_j \in [0, 1]^{N_{in}}$  is the  $j$ -th transmission spectrum with noise as input data in the training dataset,  $N_{in}$  is the dimension (the number of sampling points) of the transmission spectrum, and the corresponding ground truth  $y_j \in [0, 1]^{N_{in}}$  is the  $j$ -th de-noised spectrum in the dataset, using the DFT we obtain

$$\mathcal{F}(y_j)[k] = y[j](j) = \frac{1}{N_{in}} \sum_{n=0}^{N_{in}-1} y_j \exp\left(-\frac{2\pi i}{N_{in}} nk\right), \quad (4)$$

where  $k$  is the *frequency index*. For the NNF output  $D(x_j)$ , similarly, we obtain

$$\mathcal{F}(D(x_j))[k] = D[k](j) = \sum_{n=0}^{N_{in}-1} D(x_j) \exp\left(-\frac{2\pi i}{N_{in}} nk\right). \quad (5)$$

To examine the convergence behavior of different frequency components during the training of an NNF, we compute the residual of  $\mathcal{F}(y_j)$  and  $\mathcal{F}(D(x_j))$  at each recording epochs, i.e.,

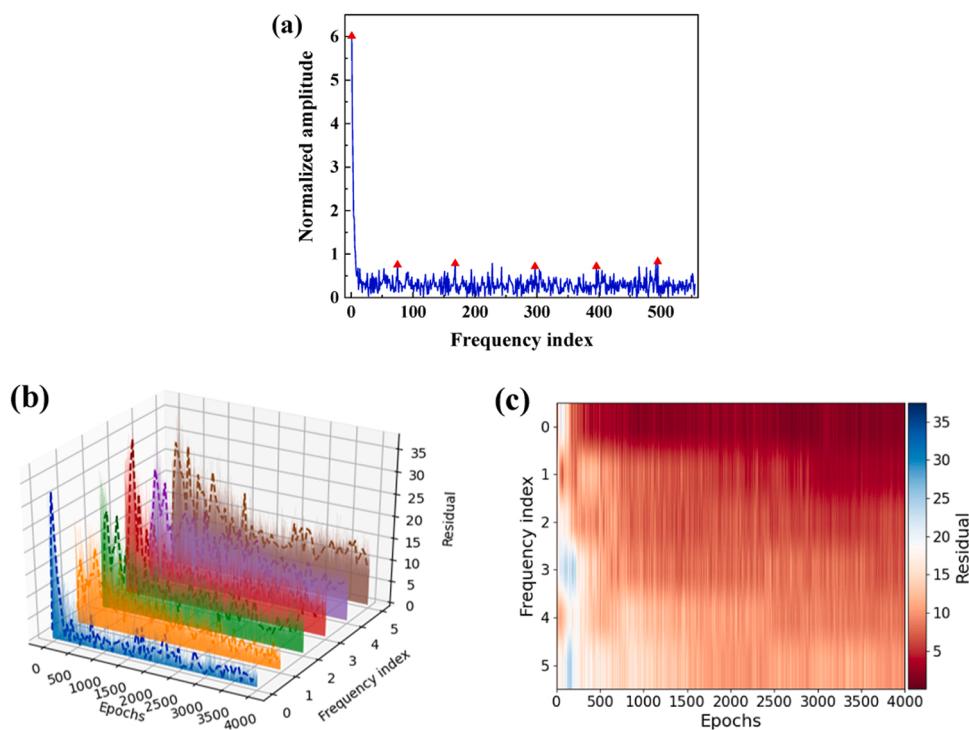
$$\Delta \mathcal{F}(k) = |\mathcal{F}(y_j)[k] - \mathcal{F}(D(x_j))[k]|, \quad (6)$$

where  $|\bullet|$  denotes the absolute value.

As shown in the Fig. 12 (a), the transmission spectrum is dominated by low-frequency components along the frequency index. In the following, we only focus on the convergence behavior of the frequency peaks during the training. By examining the residual of the certain selected key frequency components shown in Fig. 12 (a)), one can clearly observe that NNF tends to capture the training data in an order from low to high frequencies, as stated by *Frequency-Principle* [37] and is shown in Fig. 12 (b) and Fig. 12 (c). In this case, NNF is equivalent to a multi-channel band-pass filter, which first captures the dominant low frequency components in the noisy spectrum, and then gradually captures the high frequency part in the efficient frequency range. From this perspective, NNF can retain the effective information in the spectra from the frequency domain and keep the complete data trend spontaneously without the influence of manual intervention. Meanwhile, NNF does not simply conduct a high frequency truncation to eliminate the frequency components outside the cutoff frequency, but in contrast, the residual of the high frequency region is also gradually reduced during training, which explains the reason why the NNF filter can realize the full frequency range filtering. Based on such an excellent filtering performance of the NNF, the methane sensor has been enhanced in the detectable column density as well as in the detection stability.

## 6. Conclusion

In this paper, a novel methane sensor based on DAS assisted by NNF is proposed. We overcome the scarce data problem by training the NNF on the simulated data set which is highly consistent with the experimental conditions. The developed digital filter based on the neural network outperforms other widely applied filtering algorithms. Even under the situation of non-orthogonal random noise or interference effects in the practical application scenarios, the NNF can still effectively extract informational signals under the condition of extreme low SNR, helping sensors to have significant improvements in analytical performance. With the NNF assistance, the enhanced sensor achieves higher



**Fig. 12.** Frequency analysis of NNF output function showing characteristics of frequency components during the training. (a) The selected frequency peaks in the frequency domain. (b) Evolution of residuals of the selected frequency components during training for increasing numbers of epochs. (c) Similar to (b), but shown as a density plot of the residuals of the frequency components for different epoch numbers.

precision of the concentration detection and more stable real-time detection performance compared to the plain methane sensor. Considering other reported DAS applications of near-infrared methane sensors, our proposed technique reveals the further improvement of the detectable column density, showing the prospect of practical implementation in various application scenarios. Therefore, the NNF is expected to perform well in signal processing of various gas species and sensors. Although neural network is often criticized as a black box, we analyze the NNF systematically from the perspective of frequency domain, which adds to the interpretability of the inner workings of the NNF. Based on the successes of deep learning in concentration retrieval and filtering in the field of absorption spectroscopy, we believe that deep learning can effectively contribute to the research of identification and quantitative determination of concentrations of gas components in a gas mixture. We plan to carry out such work in the future.

#### CRediT authorship contribution statement

**Linbo Tian:** Conceptualization; Investigation; Writing – original draft; Methodology; Software. **Jiachen Sun:** Conceptualization; Investigation; Writing – original draft; Methodology; Software. **Sasa Zhang:** Supervision; Funding acquisition; Project administration. **Jinbao Xia:** Writing – review & editing; Funding acquisition. **Alexandre A. Kolomenskii:** Writing – review & editing. **Hans A. Schuessler:** Funding acquisition; Writing – review & editing. **Zhao Wang:** Writing – review & editing. **Jun Chang:** Supervision; Funding acquisition; Project administration. **Zhaojun Liu:** Supervision; Funding acquisition; Project administration.

#### 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.

#### Acknowledgments

This work was supported by National Natural Science Foundation of China (61475085), The Key Research and Development Program of Shandong Province (No. 2020CXGC010104), Qingdao science and technology demonstration and guidance project (No. 21–1–4-sf-1-nsh), Open Fund of State Key Laboratory of Applied Optics (SKLA02020001A12), The Robert A. Welch Foundation, grant No. A1546, Key Technology Research and Development Program of Shandong (2020CXGC010204), and Natural Science Foundation of China (No. 62075116).

#### Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.snb.2021.131207.

#### References

- [1] Yifan Song, Bouras Nassim, Qi Zhang, Explosion energy of methane/deposited coal dust and inert effects of rock dust, *Fuel* 228 (2018) 112–122.
- [2] Dong He, Zhimin Peng, Yanjun Ding, Ignition-delay-time/time-resolved CO<sub>2</sub>-concentration measurements during the combustion of iC<sub>4</sub>H<sub>10</sub>/H<sub>2</sub> mixtures, *Fuel* 284 (2021), 118980.
- [3] V. Wittstock, L. Scholz, B. Bierer, A.O. Perez, J. Wöllensteiner, S. Palzer, Design of a LED-based sensor for monitoring the lower explosion limit of methane, *Sens. Actuators B: Chem.* 247 (2017) 930–939.
- [4] L. Shi, J. Wang, G. Zhang, X. Cheng, X. Zhao, A risk assessment method to quantitatively investigate the methane explosion in underground coal mine, *Process Saf. Environ. Prot.* 107 (2017) 317–333.
- [5] J. Dang, L. Kong, C. Zheng, Y. Wang, Y. Sun, H. Yu, An open-path sensor for simultaneous atmospheric pressure detection of CO and CH<sub>4</sub> around 2.33 μm, *Opt. Lasers Eng.* 123 (2019) 1–7.
- [6] X. Cui, F. Dong, Z. Zhang, P. Sun, H. Xia, E. Fertein, W. Chen, Simultaneous detection of ambient methane, nitrous oxide, and water vapor using an external-cavity quantum cascade laser, *Atmos. Environ.* 189 (2018) 125–132.
- [7] P.K. Sekhar, J. Kysar, E.L. Brosha, C.R. Kreller, Development and testing of an electrochemical methane sensor, *Sens. Actuators B: Chem.* 228 (2016) 162–167.
- [8] George Fedorenko, et al., Semiconductor gas sensors based on Pd/SnO<sub>2</sub> nanomaterials for methane detection in air, 12.1, *Nanoscale Res. Lett.* (2017) 1–9.

- [9] M. Kamiński, R. Kartanowicz, D. Jastrzebski, M.M. Kamiński, Determination of carbon monoxide, methane and carbon dioxide in refinery hydrogen gases and air by gas chromatography, 989.2, *J. Chromatogr. A* 989 (2003) 277–283.
- [10] J. Su, L. Cao, L. Li, J. Wei, G. Li, Y. Yuan, Highly sensitive methane catalytic combustion micro-sensor based on mesoporous structure and nano-catalyst., 5.20, *Nanoscale* 5 (2013) 9720–9725.
- [11] B. Shen, P. Yang, X.L. Liu, H. Zhang, S. Cao, Fabrication and characterizations of SAW methane sensor based on cryptophane-E membrane. *J. Ambient. Intell. Humaniz. Comput.* (2018) 1–10.
- [12] Sven E. Fiedler, Achim Hese, Albert A. Ruth, Incoherent broad-band cavity-enhanced absorption spectroscopy., 371.3-4, *Chem. Phys. Lett.* 371 (2003) 284–294.
- [13] Giel Berden, Rudy Peeters, Gerard Meijer, Cavity ring-down spectroscopy: Experimental schemes and applications, 19.4, *Int. Rev. Phys. Chem.* 19 (2000) 565–607.
- [14] Brian C. Smith, Fundamentals of Fourier transform infrared spectroscopy, CRC press., 2011.
- [15] Qinduan Zhang, et al., Acousto-optic Q-switched fiber laser-based intra-cavity photoacoustic spectroscopy for trace gas detection, 18.1, *Sensors* (2018) 42.
- [16] Qinduan Zhang, et al., QEPAS sensor for simultaneous measurements of  $H_2O$ ,  $CH_4$ , and  $C_2H_2$  using different QTFs, 10.6, *IEEE Photonics J.* (2018) 1–8.
- [17] M. Raza, L. Ma, C. Yao, M. Yang, Z. Wang, Q. Wang, R. Kan, W. Ren, MH-rate scanned-wavelength direct absorption spectroscopy using a distributed feedback diode laser at 2.3  $\mu m$ , *Opt. Laser Technol.* 130 (2020), 106344.
- [18] Gregory B. Rieker, Jay B. Jeffries, Ronald K. Hanson, Calibration-free wavelength-modulation spectroscopy for measurements of gas temperature and concentration in harsh environments, 48.29, *Appl. Opt.* 48 (2009) 5546–5560.
- [19] B. Lins, P. Zinn, R. Engelbrecht, B. Schmauss, Simulation-based comparison of noise effects in wavelength modulation spectroscopy and direct absorption TDLAS, *Appl. Phys. B* 100, 2 100 (2010) 367–376.
- [20] P.O. Werle, R. Mücke, F. Slemr, The limits of signal averaging in atmospheric trace-gas monitoring by tunable diode-laser absorption spectroscopy (TDLAS), 57.2, *Appl. Phys. B* 57 (1993) 131–139.
- [21] L. Shao, B. Fang, F. Zheng, X. Qiu, Q. He, J. Wei, C. Li, W. Zhao, Simultaneous detection of atmospheric CO and  $CH_4$  based on TDLAS using a single 2.3  $\mu m$  DFB laser, *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* 222 (2019), 117118.
- [22] S.V. Kireev, A.A. Kondrashov, S.L. Shnyrev, Improving the accuracy and sensitivity of  $^{13}C$  online detection in expiratory air using the TDLAS method in the spectral range of 4860–4880  $cm^{-1}$ , 15.10, *Laser Phys. Lett.* 15 (2018), 105701.
- [23] T. Zhang, J. Kang, D. Meng, H. Wang, Z. Mu, M. Zhou, X. Zhang, C. Chen, Mathematical methods and algorithms for improving near-infrared tunable diode-laser absorption spectroscopy, 18.12, *Sensors* 18 (2018) 4295.
- [24] F. Zheng, X. Qiu, L. Shao, S. Feng, T. Cheng, X. He, Q. He, C. Li, R. Kan, C. Fittschon, Measurement of nitric oxide from cigarette burning using TDLAS based on quantum cascade laser, *Opt. Laser Technol.* 124 (2020), 105963.
- [25] J. Xia, C. Feng, F. Zhu, S. Zhang, A.A. Kolomenskii, J. Dong, H.A. Schuessler, Real-time measurement of methane in a turbulent atmosphere over kilometer long paths with denoising methods, *Measurement* 168 (2021), 108467.
- [26] C.T. Zheng, W.L. Ye, J.Q. Huang, T.S. Cao, M. Lv, J.M. Dang, Y.D. Wang, Performance improvement of a near-infrared  $CH_4$  detection device using wavelet-denoising-assisted wavelength modulation technique, *Sens. Actuators B: Chem.* 190 (2014) 249–258.
- [27] J. Xia, C. Feng, F. Zhu, S. Ye, S. Zhang, A. Kolomenskii, Q. Wang, J. Dong, Z. Wang, W. Jin, H.A. Schuessler, A sensitive methane sensor of a ppt detection level using a mid-infrared interband cascade laser and a long-path multipass cell, *Sens. Actuators B: Chem.* 334 (2021), 129641.
- [28] M.S. Niu, P.G. Han, L.K. Song, D.Z. Hao, J.H. Zhang, L. Ma, Comparison and application of wavelet transform and Kalman filtering for denoising in  $\delta^{13}CO_2$  measurement by tunable diode laser absorption spectroscopy at 2.008  $\mu m$ , *Opt. Express* 25 (20) (2017) A896–A905.
- [29] S. Zhou, N. Liu, C. Shen, L. Zhang, T. He, B. Yu, J. Li, An adaptive Kalman filtering algorithm based on back-propagation (BP) neural network applied for simultaneously detection of exhaled CO and  $N_2O$ , *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* 223 (2019), 117332.
- [30] S. Zhou, C.Y. Shen, L. Zhang, N.W. Liu, T.B. He, B.L. Yu, J.S. Li, Dual-optimized adaptive Kalman filtering algorithm based on BP neural network and variance compensation for laser absorption spectroscopy, 27.22, *Opt. Express* 27 (2019) 31874–31888.
- [31] S. Wold, H. Antti, F. Lindgren, J. Öhman, Orthogonal signal correction of near-infrared spectra, 44.1-2, *Chemom. Intell. Lab. Syst.* 44 (1998) 175–185.
- [32] Johan A. Westerhuis, Sijmen De Jong, Age K. Smilde, Direct orthogonal signal correction, 56.1, *Chemom. Intell. Lab. Syst.* 56 (2001) 13–25.
- [33] B. Rasti, D. Hong, R. Hang, P. Ghamisi, X. Kang, J. Chanussot, J.A. Benediktsson, Feature extraction for hyperspectral imagery: the evolution from shallow to deep: overview and toolbox., 8.4, *IEEE Geosci. Remote Sens. Mag.* 8 (2020) 60–88.
- [34] T. Zahavy, A. Dikopoltsev, D. Moss, G.I. Haham, O. Cohen, S. Mannor, M. Segev, Deep learning reconstruction of ultrashort pulses., 5.5, *Optica* 5 (2018) 666–673.
- [35] Felipe Oviedo, et al., Fast and interpretable classification of small X-ray diffraction datasets using data augmentation and deep neural networks, 5.1, *npj Comput. Mater.* (2019) 1–9.
- [36] L. Tian, J. Sun, J. Chang, J. Xia, Z. Zhang, A.A. Kolomenskii, H.A. Schuessler, S. Zhang, Retrieval of gas concentrations in optical spectroscopy with deep learning, *Measurement* 182 (2021), 109739.
- [37] Xu, Zhi-Qin John, Yaoyu Zhang, and Yanyang Xiao. Training behavior of deep neural network in frequency domain. *International Conference on Neural Information Processing*. Springer, Cham, 2019.
- [38] X. Pan, Z. Zhang, H. Zhang, Z. Wen, W. Ye, Y. Yang, J. Ma, X. Zhao, A fast and robust mixture gases identification and concentration detection algorithm based on attention mechanism equipped recurrent neural network with double loss function, *Sens. Actuators B: Chem.* 342 (2021), 129982.
- [39] Y. Ma, H. Li, Y. Yu, Y. Yao, Y. Fang, B. Zhou, J. Huang, J. Min, K. Liang, Experimental analysis on calibration of instrument broadening in a lidar system with Fabry-Perot etalon, 60.21, *J. Mod. Opt.* 60 (2013) 1967–1975.
- [40] Yann LeCun, Yoshua Bengio, Geoffrey Hinton, Deep learning, 521.7553, *Nature* 521 (2015) 436–444.
- [41] B. Li, C. Zheng, H. Liu, Q. He, W. Ye, Y. Zhang, J. Pan, Y. Wang, Development and measurement of a near-infrared  $CH_4$  detection system using 1.654  $\mu m$  wavelength-modulated diode laser and open reflective gas sensing probe, *Sens. Actuators B: Chem.* 225 (2016) 188–198.
- [42] L. Zhang, T. Pang, Z. Zhang, P. Sun, H. Xia, B. Wu, Q. Guo, M. W. Sigrist, C. Shu, A novel compact intrinsic safety full range Methane microprobe sensor using ‘trans-world’ processing method based on near-infrared spectroscopy, *Sens. Actuators B: Chem.* 334 (2021), 129680.
- [43] L. Zhang, Z. Zhang, P. Sun, T. Pang, H. Xia, X. Cui, Q. Guo, M.W. Sigrist, C. Shu, Z. Shu, A dual-gas sensor for simultaneous detection of methane and acetylene based on time-sharing scanning assisted wavelength modulation spectroscopy, *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* 239 (2020), 118495.

**Linbo Tian** received his B.S. degree from the School of Information Science and Engineering, Shandong University in 2016 and then graduated from the Faculty of Engineering of the University of Sydney in 2019 and received his M.S. degree. Now he is a PhD student at Shandong University and his recent research interests include optical sensing and the application of deep learning in identification and concentration detection of multi-component blended gases.

**Jiachen Sun** received the B.S. degree from Shandong University in 2018. He is currently pursuing the Ph. D degree in the Department of Optical Engineering at the Institution of Information Science and Engineering, Shandong University, China. He is mainly engaged in research in the field of optical fiber sensing and gas sensing.

**Sasa Zhang** is currently a Professor and Doctoral Supervisor at School of Information Science and Engineering, Shandong University. He graduated from Optics Department of Shandong University in 1986 and received his PhD from School of Information Science and Engineering, Shandong University in 2008. His main research interests focus on laser technology, optical sensing, trace gas analysis with neural network and deep learning, remote sensing for gas detection, spectrum technology and application, nano-materials simulation and fabrication.

**Jinbao Xia** received his master's degree from Ocean University of China in 2010. He obtained his PhD degree from the Institute of Science and Technology, Shandong University and also performed his research visiting Texas A&M University. Now he works as a postdoc. His scientific interests include trace gas detection and laser remote sensing.

**Alexandre A. Kolomenskii** received his PhD degree from P. N. Lebedev Physical Institute of the Russian Academy of Sciences in 1980. Now he is a Research Scientist at Texas A&M University. His fields of scientific interests include applications of lasers in acoustics, spectroscopy, and ultrafast laser interactions with atomic and molecular systems.

**Hans A. Schuessler** is a Professor of Physics at the department of Physics & Astronomy of Texas A&M University. He received his PhD degree at Heidelberg University in 1964. His research fields include ion traps, laser spectroscopy and its applications, ultrafast laser phenomena and high sensitivity detection of trace gases.

**Zhao Wang** received the B.S. degree from Dongbei University of Finance and Economics in 2019. She is currently pursuing the master's degree in the major of Electronic Information at Key Laboratory of Education Ministry for Laser and Infrared System Integration Technology, Shandong University, China. She is mainly engaged in research in the field of deep learning and artificial neural network.

**Jun Chang** received the B.S., M.S. and PhD degrees in Shandong University in 1988, 1991, 2006 respectively. He was a visiting fellow in school of Electrical Engineering and Telecommunications in University of New South Wales during 2004 and 2005. He is currently a professor with School of Information Science and Engineering of Shandong University. His research interest includes optical fiber sensor and fiber laser.

**Zhaojun Liu** received the B.S. and Ph.D. degrees in optical engineering from Shandong University, Jinan, China, in 2003 and 2008, respectively. In 2008, he joined Shandong University. From 2011–2012, he was with Lehigh University, Bethlehem, PA, USA, as a Postdoctoral researcher. In 2017, he was promoted to a Full Professor with the School of Information Science and Engineering, Shandong University. His current research interests include single-crystal fiber manufacturing and their applications on compact single-frequency lasers, high-power fiber lasers, Raman lasers, and fiber sensing system.