Sensors & Actuators: B. Chemical 467 (2026) 140467

Contents lists available at ScienceDirect

Sensors and Actuators: B. Chemical

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

Photoacoustic carbon monoxide gas sensor using deep neural networks for nonlinear noise suppression

Xukun Yin [a], Lei Liu [a], Xiu Yang [a], Yize Liang [a], Dacheng Zhang [a], Pietro Patimisco [b], Angelo Sampaolo [b], Vincenzo Spagnolo [b,*], Huailiang Xu [a,c,**], Hongpeng Wu [d,*]

[a] School of Optoelectronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China
[b] PolySense Lab, Dipartimento Interateneo di Fisica, University and Politecnico of Bari, Bari 70126, Italy
[c] College of Electronic Science & Engineering, Jilin University, Changchun 130012, China
[d] State Key Laboratory of Quantum Optics Technologies and Devices, Institute of Laser Spectroscopy, Shanxi University, Taiyuan, Shanxi 030006, China

ARTICLE INFO

Keywords:
Photoacoustic spectroscopy
Gas sensing
Signal denoising
Deep learning

ABSTRACT

In photoacoustic spectroscopy (PAS) based gas detection, measurement accuracy and sensitivity are frequently constrained by noise interference, particularly under low signal-to-noise ratio (SNR) conditions. Conventional filtering methods are generally unable to effectively suppress noise while preserving signal integrity and retaining characteristic features, thereby restricting further improvements in detection performance. To address this limitation, a novel signal-denoising framework, termed U-LSTM, is proposed by integrating the symmetric encoder-decoder architecture of a U-shaped neural network (U-Net) with Long Short-Term Memory (LSTM) modules. This hybrid architecture enables the simultaneous extraction of local signal features and the modeling of global temporal dependencies, thereby attaining superior suppression of complex, nonlinear noise. The results demonstrate that the proposed U-LSTM model substantially outperforms conventional filtering methods. For example, when processing the second-harmonic (2 *f*) signal of carbon monoxide (CO), the SNR increased from 15.4 for the original signal to 390.4 after denoising, demonstrating the excellent noise-suppression and signal-reconstruction capabilities. The proposed method provides a robust and effective denoising solution for PAS-based trace-gas detection and significantly improves detection accuracy and system reliability in noisy environments.

1. Introduction

Gas detection technology plays a vital role in diverse fields such as environmental monitoring [1], industrial safety [2], and medical diagnosis [3–8], particularly for the accurate detection of trace gases. Recently, laser-based spectroscopy techniques have attracted significant attention due to their high selectivity, exceptional sensitivity, long operational lifetime, and low maintenance requirements [9,10]. Common optical sensing techniques for greenhouse gas detection include tunable diode laser absorption spectroscopy (TDLAS), cavity ring-down spectroscopy (CRDS) [11], Fourier transform infrared spectroscopy (FTIR) [12], non-dispersive infrared spectroscopy (NDIR) [13], and photoacoustic spectroscopy (PAS). PAS is particularly noteworthy for its background-free detection capability [14,15], making it highly suitable for trace gas analysis. In PAS, target gas molecules absorb laser energy at specific wavelengths and are excited from the ground state to an excited state. According to the principle of energy conservation, the excited molecules relax to the ground state through non-radiative transitions, releasing energy in the form of acoustic waves. These acoustic signals are typically detected by a microphone or a quartz tuning fork, which converts them into the electrical current signal [16–20]. Since the amplitude of the acoustic signal is directly proportional to the gas concentration, the concentration can be accurately determined by measuring the corresponding electrical output. Moreover, PAS-based sensors possess the advantage of being insensitive to detector wavelength selectivity, as the signal is generated acoustically rather than optically. As a result, the detection sensitivity can be further enhanced when using mid- or far-infrared excitation sources, where molecular absorption cross-sections are significantly stronger. However, lasers operating in the mid- and far-infrared regions are typically expensive,

* Corresponding authors.
** Corresponding author at: School of Optoelectronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China.
E-mail addresses: vincenzoluigi.spagnolo@poliba.it (V. Spagnolo), xuhuailiang@xidian.edu.cn (H. Xu), wuhp@sxu.edu.cn (H. Wu).

https://doi.org/10.1016/j.snb.2026.140467
Received 5 June 2026; Received in revised form 28 June 2026; Accepted 5 July 2026
Available online 7 July 2026
0925-4005/© 2026 Published by Elsevier B.V.

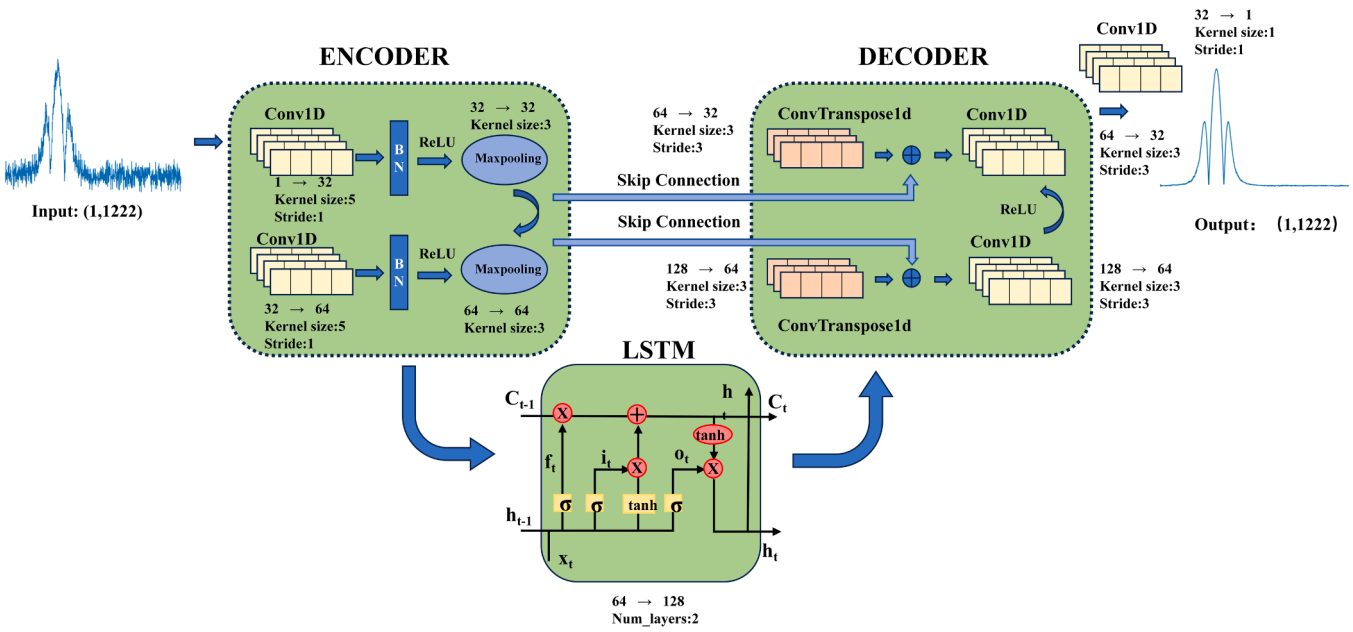


Fig. 1. Schematic of the model.

bulky, and often require water-cooling systems, which limits their suitability for practical applications. With the rapid advancements of fiber-optic communication technologies, near-infrared distributed feedback (DFB) lasers have become more compact, cost-effective, and well-suited for integration into photoacoustic sensors. Additionally, since the detection sensitivity of PAS sensors is directly proportional to the excitation optical power, it can be enhanced by optimizing the optical power while avoiding signal saturation. However, in real-world measurements, the background-free detection advantage of PAS can be compromised by environmental and window-induced noise, especially when high-power excitation sources are employed [21–24].

In recent years, digital filtering techniques have become increasingly prevalent in spectral signal processing due to their low cost, high compatibility, and simple system architecture. These filters enhance measurements qualities without altering the optical setup, thereby improving the adaptability and robustness of detection systems. Commonly used algorithms include adjacent average filtering (AF), Savitzky-Golay (SG) filtering, and wavelet denoising (WD). The AF algorithm reduces noise by averaging data points within a local neighborhood, making it effective for suppressing high-frequency noise. However, for nonlinear signals, it often leads to the loss of edge information [25]. The SG filter applies a local polynomial least-squares fit, preserving higher-order signal features and achieving a balance between smoothing and feature retention. However, its performance depends strongly on the stationarity and continuity of input data, requiring careful parameter optimization to achieve optimal denoising performance [26]. The WD algorithm, leveraging strong time-frequency localization capabilities, is particularly effective for noise suppression and feature extraction in non-stationary signals, as it can identify and retain transient and high-frequency components. Nonetheless, its effectiveness is highly sensitive to parameter choices, such as the wavelet basis, thresholding method, and decomposition level, which must be fine-tuned for specific applications [27]. Although these filtering techniques can effectively suppress noise through different mathematical mechanisms, addressing signal properties in the time, statistical, and frequency domains, their performance varies with the application conditions. As signal and noise complexity increases, traditional filtering approaches often fail to adequately handle nonlinearities, complex noise structures, or non-smooth spectral features.

The rapid advancement of artificial intelligence, particularly deep learning (DL) [28], has yielded remarkable progress in various optical applications, including ultrashort pulse reconstruction, hyperspectral imaging classification, and gas concentration inversion [29–31]. Deep learning models, especially those based on neural network architectures, possess powerful nonlinear mapping and modeling capabilities, enabling effective self-regulation and adaptive feedback. These strengths have opened new possibilities for enhancing optical measurement technologies. Currently, most research efforts focus on integrating deep learning with traditional filtering methods to optimize key parameters, thereby reducing human bias and improving the adaptability and stability of conventional algorithms. However, these hybrid approaches still depend heavily on traditional filters, with neural networks serving primarily as auxiliary tools rather than independently performing noise reduction on spectral signals. This dependency constrains the potential of neural networks, limiting their ability to fully exploit their denoising and feature extraction capabilities [32–35]. A few studies have explored end-to-end deep learning models for noise reduction [36,37]. However, these efforts generally employ simplified network architectures and remain in early development stages. Therefore, further research into integration of deep neural networks with optical measurement systems is essential for fully leveraging spectral information and advancing gas detection technologies.

Building on these insights, this study proposes a novel deep learning framework that combines a U-shaped convolutional neural network (U-Net) [38] with a long short-term memory (LSTM) architecture to mitigate noise-induced correlation degradation in photoacoustic spectroscopy (PAS). The proposed U-LSTM model achieves outstanding denoising performance, enabling precise detection of low signal-to-noise ratio and providing a robust solution for accurate gas detection in complex environments.

2. U-LSTM model

The convolutional neural network (CNN) is classified as a feed-forward neural network in which convolutional operations are employed within a deep architecture. It is particularly suited for processing data with grid-like structures, such as images, videos, and speech. A typical CNN includes an input layer, one or more hidden layers, and an output layer. The hidden layers typically contain three types of components, *i.e.* convolutional layers, pooling layers, and fully

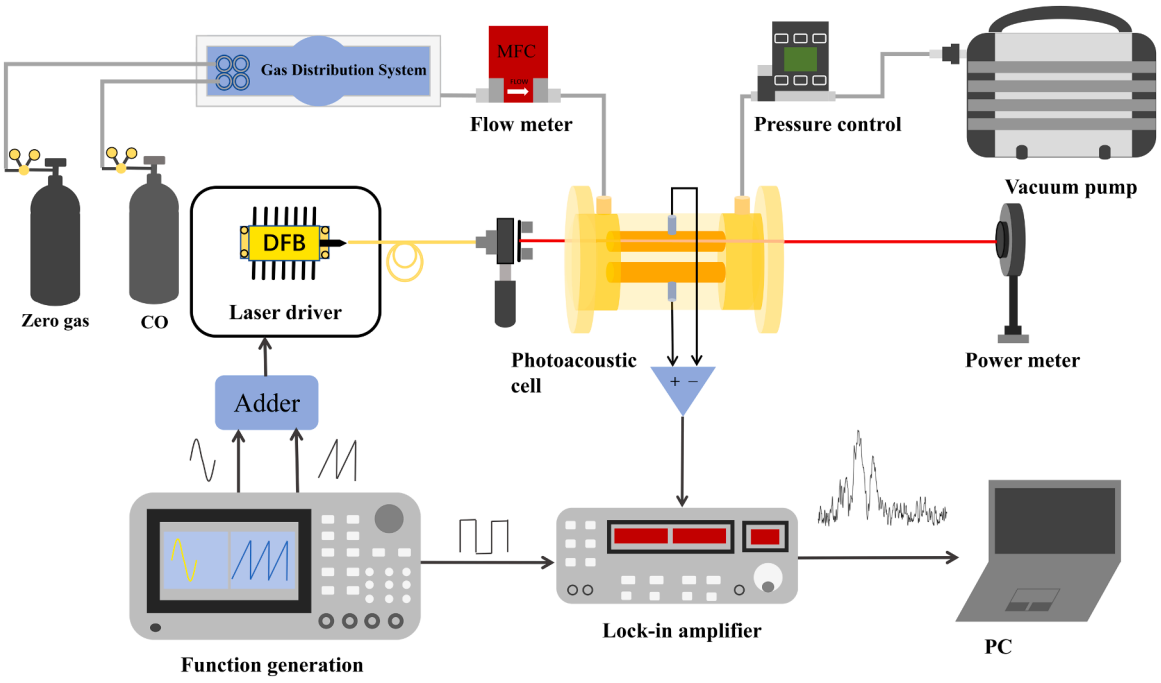


Fig. 2. Schematic diagram of CO gas sensor.

connected layers. In CNNs, initial features are extracted through convolutional operations, after which activation functions are applied to introduce nonlinear transformations. This nonlinear mapping enables the model to capture complex features, thereby enhancing its expressive power. As a result, the model performs deeper abstractions and achieves a more sophisticated understanding of the data. This ability to capture intricate patterns and relationships within the data substantially improves the model performance, allowing it to extract more meaningful and valuable information. The signal discussed in this work is the second-harmonic (*2 f*) signal, represented as a one-dimensional time series. One-dimensional CNNs (1D CNNs) are particularly effective for processing sequential data, as they capture local dependencies by applying convolutional filters across the sequence, making them an ideal choice for time-series data analysis.

The long short-term memory (LSTM) model is a type of recurrent neural network (RNN) that excels in sequence learning. It addresses the limitations of traditional RNNs, particularly the difficulty in capturing long-term dependencies caused by gradient vanishing or explosion. The structure of the LSTM cell is illustrated in Fig. 1. At the core of the LSTM is the introduction of gating mechanisms and cell states. The forget gate f_t determines which historical information is discarded, the input gate i_t controls the storage of new information, and the output gate o_t regulates the final output state. The cell state serves as the memory backbone across time steps, enabling the LSTM to selectively retain important information while filtering out redundant data. This design allows the LSTM to effectively model temporal dependencies and capture long-term relationships in sequential data. The formulation for LSTM is as Eqs. (1–6):

$$i_t = \sigma(W_i \cdot [x_t, h_{t-1}] + b_i) \tag{1}$$

$$f_t = \sigma(W_f \cdot [x_t, h_{t-1}] + b_f) \tag{2}$$

$$o_t = \sigma(W_o \cdot [x_t, h_{t-1}] + b_o) \tag{3}$$

$$\widetilde{c}_t = \tanh(W_c \cdot [x_t, h_{t-1}] + b_c) \tag{4}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \widetilde{c}_t \tag{5}$$

$$h_t = o_t \odot \tanh(c_t) \tag{6}$$

.

σ denotes the sigmoid activation function, $\odot$ denotes element-wise accumulation, W_i, W_f, W_o, W_c and b_i, b_f, b_o, b_c represent the weight matrix and bias term of each module, respectively, x_t denotes the input value at time t, h_t denotes the hidden state vector at time t, c_t denotes the updated cell state at time t, $\widetilde{c}_t$ denotes the candidate cell state at time t, and tanh denotes the hyperbolic tangent function.

Several state-of-the-art models are based on network architectures such as autoencoders, residual networks, dense networks, or their combinations. In this study, the U-Net deep learning framework is implemented. The U-Net model is a specific type of autoencoder architecture. The encoding part (contraction path) follows the typical architecture of a CNN. Its detailed structure is illustrated in Fig. 1 and includes two convolutional layers. Each convolutional layer performs convolutions on the feature maps, with the operations determined by the number of kernels and the stride. After these operations, the corresponding new feature maps are generated. Subsequently, a batch normalization (BN) operation is applied to enhance the model's generalization ability. The ReLU activation function is then used to introduce a nonlinear transformation to the output feature maps. Max pooling is applied for downsampling, with the number of feature channels doubled at each step. The output of the final convolutional layer is then passed to the bottleneck LSTM layer to capture time-dependent relationships and establish global temporal dependencies within the signal. In the proposed U-LSTM architecture, the LSTM module is inserted only at the bottleneck layer between the encoder and decoder to capture global temporal dependencies, rather than being applied to other feature extraction stages. The decoding section (expansion path) restores the size of the feature maps and maintains a structure that is symmetric to the encoding section. This symmetry allows the reuse of feature maps by concatenating those at the same level, thereby reducing information loss during the encoding-decoding process.

Thus, the U-LSTM model was developed to perform noise reduction on photoacoustic spectroscopy signals. This architecture bridges dimensional expansion and reduction by first projecting the input data into a higher-dimensional feature space and subsequently mapping it

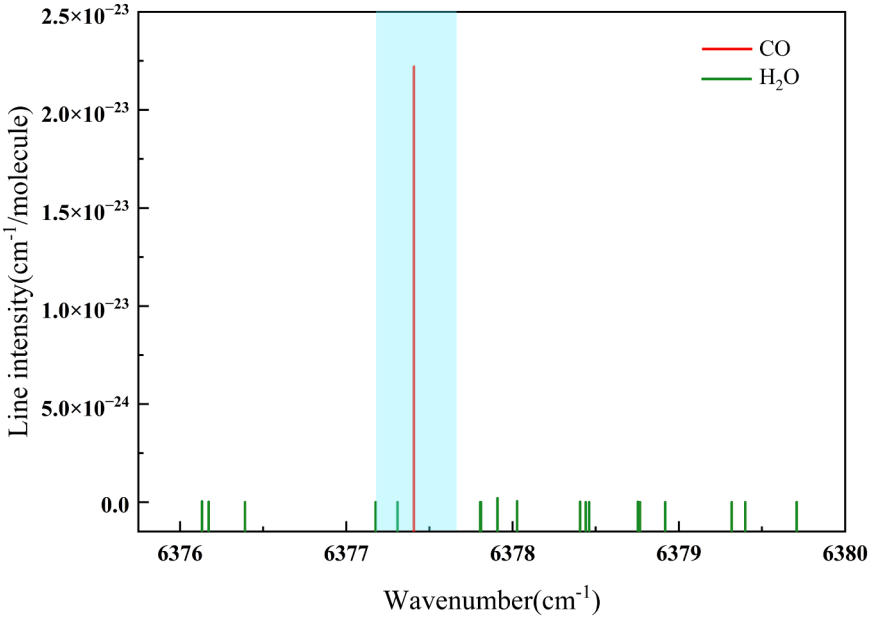


Fig. 3. Absorption lines of CO and H_2O between 6376 cm^{-1} and 6380 cm^{-1}.

back to a lower dimensional representation. Through this process, noise filtering is achieved without attenuating the signal peak amplitude. By analyzing spectral data on a global scale, the model captures dependencies and correlations among all data points while preserving the essential peak characteristics of the signal. The model input consists of a noisy signal with 1222 data points, and the output is the corresponding smoothed, noise-free signal reconstruction.

3. Experimental system and data collection

3.1. Experimental setup

The experimental system is shown in Fig. 2. A gas distribution system was employed to mix CO with zero gas in different proportions. In this work, zero gas refers to synthetic air, which is mainly composed of nitrogen and oxygen and contains no target gas components. The CO mixtures used in this work were generated from 1% CO balanced in zero gas and were further diluted to the desired concentrations using mass flow controllers. To optimize signal quality, the gas flow rate was maintained at 50 standard cubic centimeters per minute (sccm) to minimize airflow-induced noise. The sample gases were kept at room temperature (approximately 296–298 K, ~23–25°C) without additional heating or cooling. The prepared gas mixtures at different concentrations were introduced into the sealed differential resonant photoacoustic cell through a mass flow controller (MFC). After passing through the cell, the gas was then evacuated through a pressure controller by using a vacuum pump. Both the gas flow rate and pressure within the sensing system were precisely regulated using flow meters and pressure controllers to ensure stable operating conditions throughout the experiments. A low-noise differential resonant photoacoustic cell was employed to enhance detection sensitivity and suppress external noise. The cell contains two identical cylindrical cavities serving as resonant chambers, each 90 mm in length and 4 mm in radius, and was designed with fully symmetrical geometry to minimize noise generated by airflow, window absorption, and electromagnetic interference. The resonant frequency of the differential photoacoustic resonator was 1786 Hz. A signal generator was used to produce two types of waveform signals, sine waves and sawtooth waves. These signals modulated the frequency and intensity of the laser. When the emitted laser passes through the photoacoustic cell, an interaction occurs between the laser and gas molecules. The resulting photoacoustic signal was detected by two Primo EM258 electret condenser microphones integrated into the detection cavity. The EM258 is an omnidirectional microphone with a sensitivity of -32 ± 3 dB at 1 kHz (0 dB = 1 V/Pa), an impedance of 1.6 kΩ ± 30% at 1 kHz, an S/N ratio of 74 dB at 1 kHz, and a maximum input sound pressure level of 115 dB SPL. Its frequency response covers approximately 100 Hz−20 kHz. The microphones converted the photoacoustic signals into electrical signals, which were then transmitted to the differential amplification circuit and lock-in amplifier to extract the 2 *f* signal. A personal computer was used for data acquisition and control via an RS232 serial port, and the acquired data were displayed in real time using a LabVIEW program. The corresponding *2 f* signals from CO samples at different concentrations were recorded and analyzed.

Based on the spectral line intensity, the HITRAN database was used to identify a strong absorption peak of CO at a wavenumber of 6377.4 cm^{-1} (Fig. 3). The system uses a DFB diode laser as the excitation source, with a center wavelength of 1567.38 nm. This wavelength was coarsely adjusted by changing the temperature and finely tuning the injection current.

3.2. Dataset description

The deep learning model for photoacoustic spectroscopy denoising was trained using theoretically simulated second-harmonic (*2 f*) signals derived from the HITRAN database. The spectral line parameters were obtained from the HITRAN database using the SpectraPlot Line Survey tool. For the CO absorption line near 6377.4 cm^{-1}, the air-broadened half-width γ_{air}, self-broadened half-width γ_{self}, and temperature-dependence exponent n_{air} were 0.0606 cm^{-1} atm^{-1}, 0.066 cm^{-1} atm^{-1}, and 0.75, respectively. Spectral data corresponding to 200 sets of CO/synthetic-air mixtures at varying concentrations were generated. Specifically, the concentration points were distributed in three ranges, including 55 sets from 40 to 95 ppm, 135 sets from 110 to 400 ppm, and 10 sets from 410 to 500 ppm. Instead of using a single fixed step size, a non-uniform concentration distribution was adopted to cover different concentration regions. Based on the simulated spectral absorbance, the *2 f* signal was obtained by first defining key modulation parameters, such as the center wavelength and modulation depth. The wavelength of the laser was modulated as Eq. (7)

$$\lambda(t) = \lambda_0 + \Delta\lambda \cdot \sin(2\pi f t) \tag{7}$$

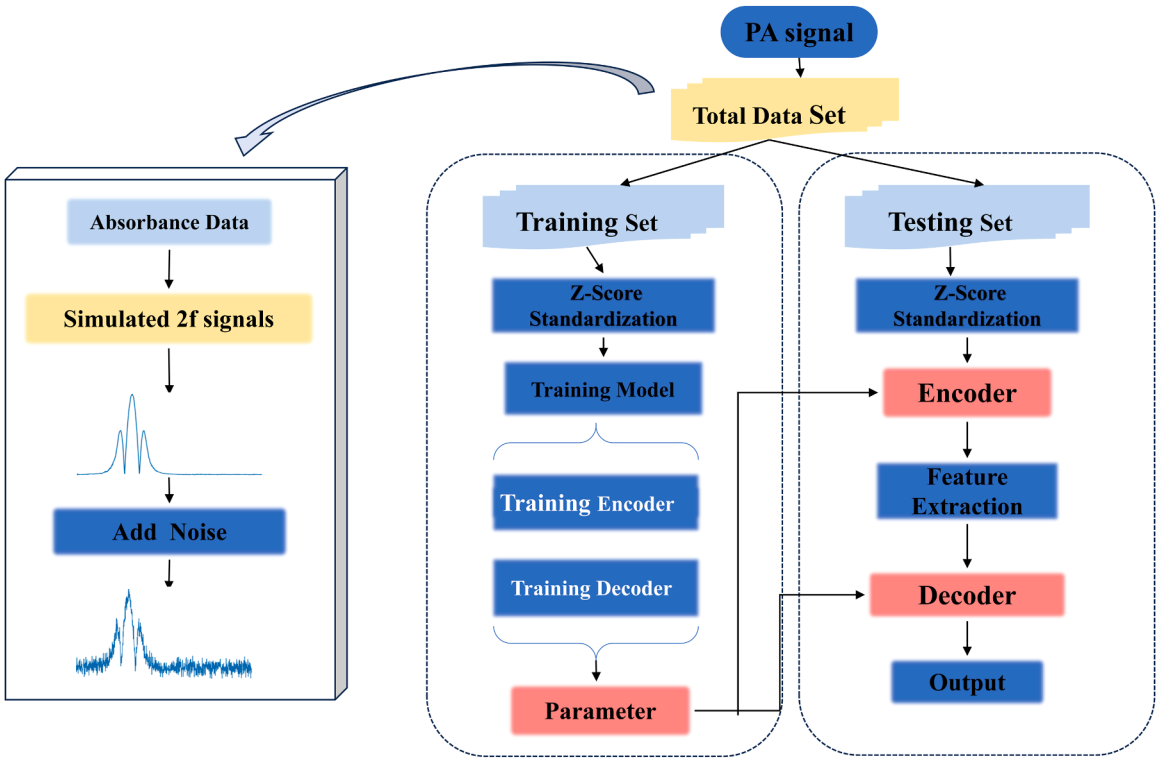


Fig. 4. Detailed learning procedure of U-LSTM algorithm.

where f is the modulation frequency, λ_0 represents the center wavelength, and $\triangle\lambda$ denotes the modulation depth. Numerical calculations indicated that the modulation depth corresponding to the maximum *2 f* signal was approximately 2.2 times the half width at half maximum (HWHM). The photoacoustic signal $P(t)$ is proportional to the absorbed optical power, as shown in Eq. (8)

$$P(t) \propto I(t) \cdot (A(\lambda(t)) \tag{8}$$

where $I(t)$ is the laser intensity and $A(\lambda(t))$ is the instantaneous absorbance. The *2 f* signal originates from the second-harmonic term in the Fourier expansion of $A(\lambda(t))$.

These signals were inherently ideal and free from noise interference. To simulate real-world experimental conditions and enhance the model's ability to process noisy data, Gaussian white noise was added together with non-Gaussian components, including pink noise and low-frequency sinusoidal noise. The standard deviation of the Gaussian white noise was set in the range of approximately 1×10^{-6} to 1×10^{-5}, while the amplitude of the pink noise was approximately 1×10^{-6} to 3×10^{-6} and that of the low-frequency sinusoidal noise was approximately 3×10^{-6}. By adjusting these noise parameters, the overall noise level was varied to expand the dataset, ensuring that the spectral characteristics closely resembled those of the experimental spectra. This approach improves the model robustness and enables it to perform reliably under unforeseen disturbances.

As a result, a comprehensive dataset comprising 836 sets of CO PAS *2 f* signals was constructed. Each signal contained 1222 data points, and all ideal CO PAS *2 f* signals were treated as ground-truth references (model labels).

To ensure the stability and generalization of the trained model when applied to experimentally measured data, the comprehensive dataset was normalized. This preprocessing step involved subtracting the mean of each signal and dividing it by its standard deviation, thereby standardizing the data such that the mean equals zero and the standard deviation equals one. Mathematically, this normalization process can be expressed as Eq. (9)

$$y = \frac{x - \mu}{\sigma} \tag{9}$$

where x represents the original *2 f* signal, μ represents the normalized *2 f* signal, μ is the mean of all x, and σ is the standard deviation of x. The network parameters were then trained using the normalized dataset, and the entire learning procedure is illustrated in Fig. 4.

3.3. Training the U-LSTM model

The U-LSTM model was implemented using the PyTorch 2.4.1 deep learning framework under Python 3.8 and CUDA 12.4. As shown in Fig. 1, the U-LSTM model adopted an encoder-LSTM-decoder architecture. The input was a one-dimensional 2 *f* signal with a size of 1×1222. The encoder contained two Conv1D blocks, where the feature channels were mapped from 1 to 32 and then from 32 to 64. Both Conv1D layers used a kernel size of 5, a stride of 1, and same padding, followed by batch normalization and ReLU activation. Max pooling with a kernel size of 3 and a stride of 3 was applied after each encoder block for downsampling. The extracted high-level features were then fed into a two-layer LSTM module at the bottleneck, where the feature dimension was increased from 64 to 128 to capture global temporal dependencies. In the decoder, two transposed convolutional layers were used for upsampling, with channel transitions of 128–64 and 64–32, respectively. Skip connections were introduced between the encoder and decoder to preserve local feature information. Finally, a 1×1 Conv1D layer mapped the 32-channel feature representation back to a single-channel denoised output. The dataset was randomly divided into training and testing sets at an 8:2 ratio. During training, hyperparameters were tuned within a predefined range to optimize model performance and ensure robust convergence. The mean square error (MSE) was employed as the loss function in the model, as shown in Eq. (10)

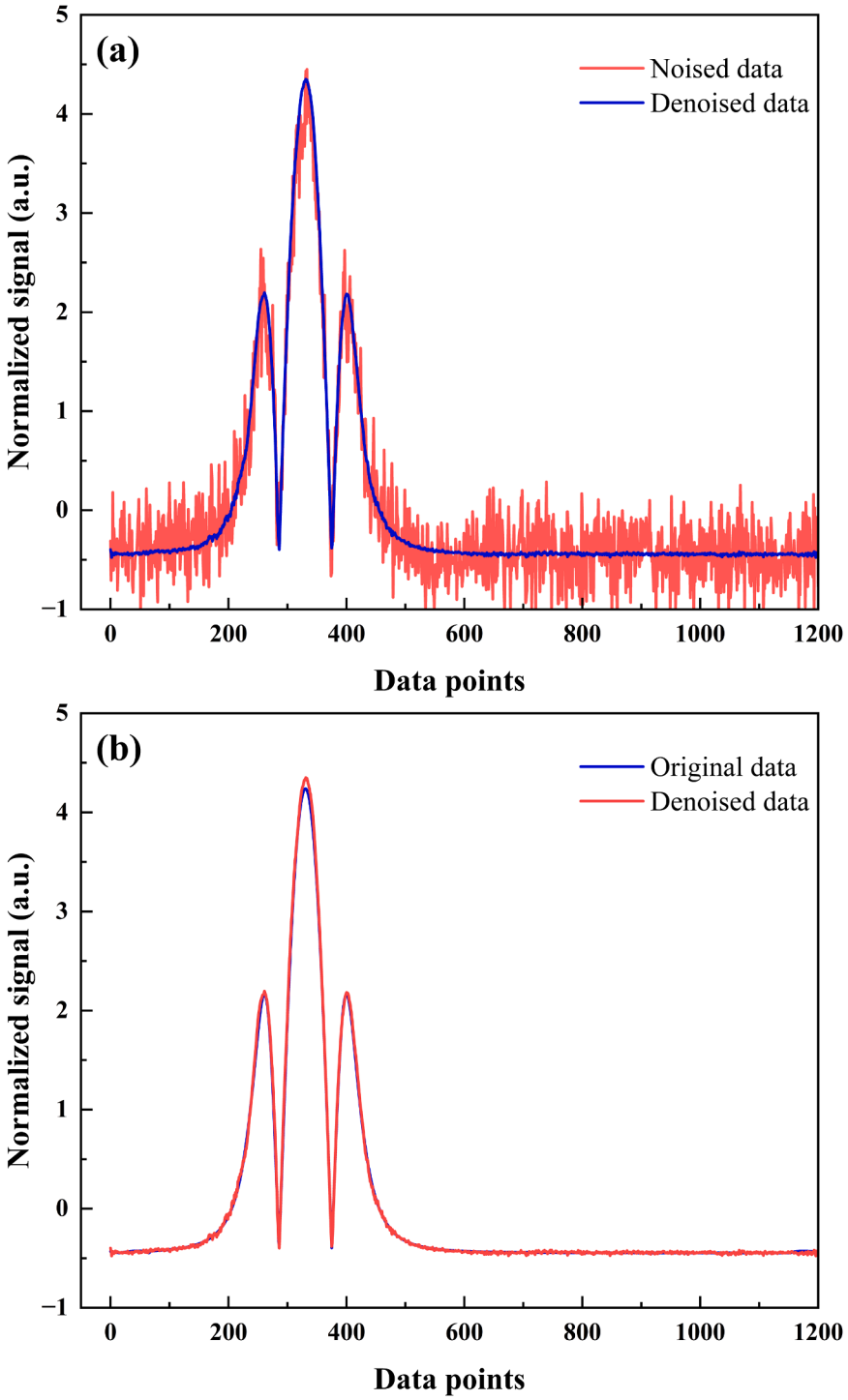


Fig. 5. (a) Comparison of noisy signal and denoised signal. (b) Comparison of denoised signal and original signal.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y(i) - \hat{y}(i))^2 \tag{10}$$

where i is the index corresponding to the whole output data or label of the input data, and n is the number of the data points of the truth label. $y(i)$ represents the $2f$ signal output signal value of the model at the position i. $\hat{y}(i)$ denotes the $2f$ signal label value at the position i. To accelerate training and prevent gradient vanishing or explosion, the ReLU activation function was used in the model, as shown in Eq. (11)

$$\mathrm{Re}LU(x) = \begin{cases} x, & x > 0 \\ 0, & otherwise \end{cases} \tag{11}$$

.

During algorithm training, a batch size of 128 was used, together with the Adam optimizer with a learning rate of 0.0001, for a total of 500 epochs. All training and testing were performed on an NVIDIA 3050 Ti GPU. After training was completed, the parameters and structure of the optimal model were saved. The trained model was further evaluated in inference mode, and the average inference time for a single

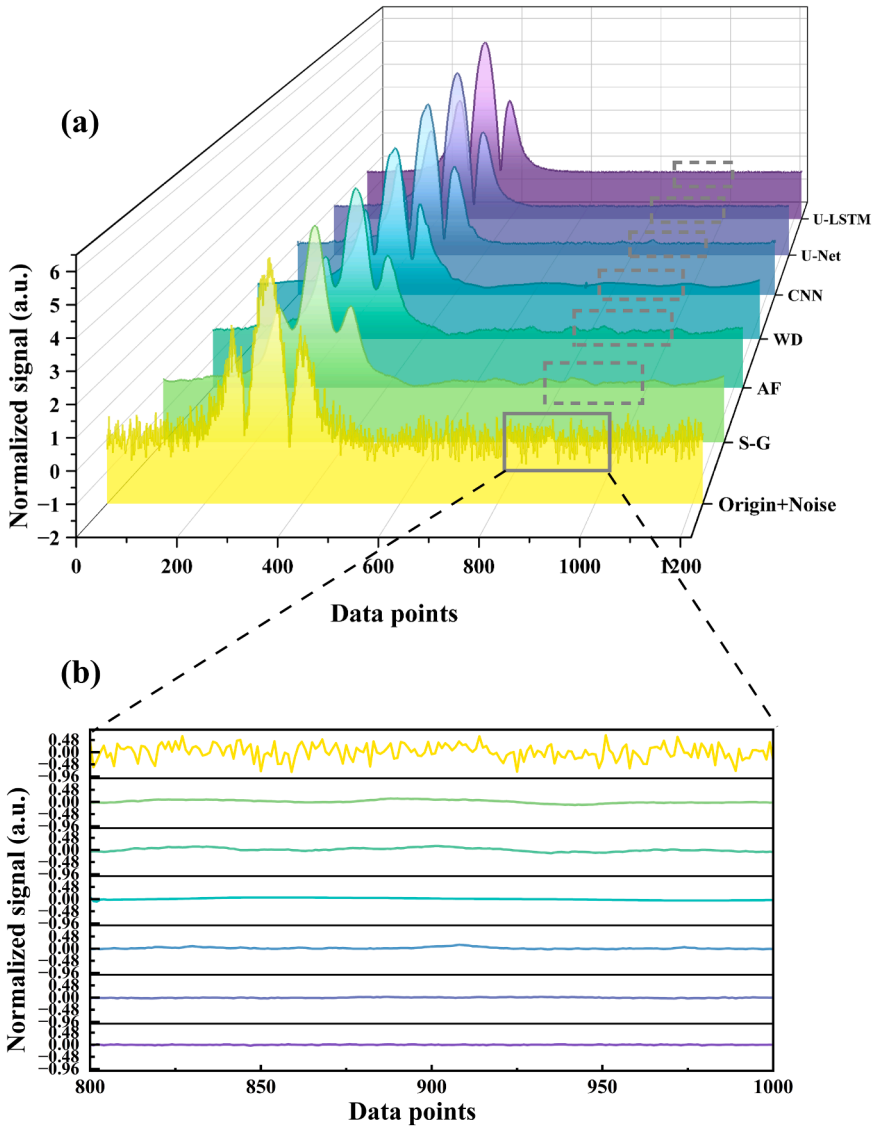


Fig. 6. (a) Filtering effectiveness of different filtering algorithms. (b) Detailed comparison of denoising performance in the non-absorptive region for different filtering algorithms. The curves are vertically offset for clarity.

normalized 1222-point 2 *f* signal was approximately 175 ms on the same NVIDIA 3050 Ti GPU.

4. Results and discussion

Based on the test dataset, Fig. 5(a) compares the noisy and denoised signals, demonstrating that the model effectively suppresses noise. Fig. 5 (b) presents a comparison between the denoised signal and the ideal signal, confirming that the model preserves the fundamental characteristics of the original waveform. Collectively, these results validate the effectiveness of the proposed model in photoacoustic signal denoising.

Subsequently, the denoising performance of the proposed model was systematically compared with those of conventional filtering methods and deep-learning-based approaches. The 2 *f* signals were processed using the S-G, AF, WD, CNN, U-Net, and U-LSTM algorithms, and the corresponding filtering results, together with the original noisy signal, are presented in Fig. 6(a). To more intuitively evaluate the denoising capability of each method, an enlarged view of the non-absorptive region is provided in Fig. 6(b). The results show that the S-G and AF algorithms exhibit relatively limited noise suppression capability, as evidenced by the residual fluctuations observed in both the absorptive and non-absorptive regions. Although partial suppression of high-frequency noise is achieved, obvious baseline oscillations still remain, suggesting that these methods provide limited denoising performance under complex noise conditions. In comparison, improved smoothing performance is achieved by the WD algorithm owing to its superior time–frequency localization capability. However, slight distortions of local signal details are still observed after processing.

Compared with conventional filtering methods, the deep-learning-based CNN, U-Net, and U-LSTM approaches exhibit significantly enhanced denoising performance. Random noise is effectively suppressed by the CNN model; however, local fluctuations near the absorption peaks are still retained, indicating limited capability in modeling long-range temporal dependencies. By employing the encoder–decoder architecture and skip-connection mechanism, the U-Net model further improves signal smoothness while enhancing feature extraction and information recovery during the reconstruction process. Among all investigated methods, the proposed U-LSTM model exhibits the best overall denoising performance. This improvement can be attributed to the synergistic effect of the U-Net and LSTM modules. The U-Net encoder-decoder structure with skip connections enhances local spectral feature extraction and helps preserve the main waveform

Table 1
Comparison of the signal-to-noise ratio (SNR) and mean squared error (MSE) obtained using different filtering algorithms for simulated 2 *f* photoacoustic signals.

Filtering method	SNR	MSE
Noisy signal	19.119	$9.853e^{-12}$
S-G	82.205	$4.686e^{-12}$
AF	71.043	$3.032e^{-12}$
WD	123.856	$1.71e^{-12}$
CNN	154.319	$1.082e^{-12}$
U-Net	368.625	$8.66e^{-13}$
U-LSTM	585.839	$1.192e^{-12}$

morphology, while the LSTM layer captures the global sequential correlations of the 2 *f* signal and further suppresses residual noise fluctuations. Therefore, the superior denoising performance of the U-LSTM model can be attributed to the synergistic integration of local feature extraction and temporal dependency modeling, rather than to either component individually. In particular, the smallest baseline fluctuation is observed in the non-absorptive region after U-LSTM processing. SNR and MSE serve as primary metrics for evaluating denoising performance, defined as Eqs. (12) and (13)

$$SNR = \frac{Signal_{peak}}{std(Signal_{noise})} \tag{12}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Signal_{denoised} - Signal_{noise-free})^2 \tag{13}$$

where *std* denotes the standard deviation, $Signal_{peak}$ is the peak amplitude of the *2 f* signal, $Signal_{noise}$ is the noise component of the *2 f* signal, $Signal_{denoised}$ is the denoised signal, and $Signal_{noise-free}$ is the ideal analog *2 f* signal.

Detailed comparison parameters are summarized in Table 1. When the U-LSTM model was applied for signal denoising, the SNR of the noisy 2 *f* signal increased significantly from 19.119 to 585.839, while the MSE decreased from 9.853×10^{-12} to 1.192×10^{-12}. In the non-absorptive region, the U-LSTM algorithm effectively suppresses noise fluctuations, whereas in the absorption region, the principal spectral characteristics and peak morphology are well preserved while residual noise is efficiently removed, resulting in smoother and clearer signal profiles.

Compared with conventional filtering methods, including S-G, AF, and WD, the deep-learning-based CNN, U-Net, and U-LSTM approaches exhibit superior denoising performance. The CNN model increases the SNR to 154.319 and reduces the MSE to 1.082×10^{-12}, indicating improved noise suppression capability compared with traditional methods. By employing the encoder-decoder architecture and skip connections, the U-Net model further enhances signal reconstruction performance, achieving an SNR of 368.625 and the lowest MSE of 8.66×10^{-13}. Although the MSE obtained by the U-LSTM model is slightly higher than that of U-Net, the U-LSTM model achieves the highest SNR of 585.839, demonstrating superior capability in suppressing noise fluctuations and enhancing signal quality. These results indicate that the U-Net model provides better point-wise reconstruction accuracy, whereas the proposed U-LSTM model exhibits stronger overall denoising capability and signal enhancement performance for simulated 2 *f* signals. Although MSE provides a useful measure of the overall point-by-point reconstruction error, SNR is considered a more critical evaluation metric for the present photoacoustic denoising task. In photoacoustic gas sensing, the practical detection capability is mainly determined by the amplitude of the absorption-related signal relative to the residual noise background. Therefore, SNR is directly associated with signal detectability, measurement precision, and the potential limit of detection. In contrast, MSE reflects the average deviation over the entire waveform and may not fully represent the improvement in sensing performance, especially when slight smoothing or amplitude distortion occurs. For this reason, MSE was used as a complementary reconstruction metric, whereas SNR was emphasized as the primary indicator for evaluating the denoising performance.

Building upon the promising results obtained from the simulated dataset, the trained U-LSTM model was further applied to the experimentally measured 2 *f* signal of 1% CO balanced in zero gas. The corresponding denoising results are presented in Fig. 7, demonstrating that the trained model exhibits strong robustness, effectively suppressing

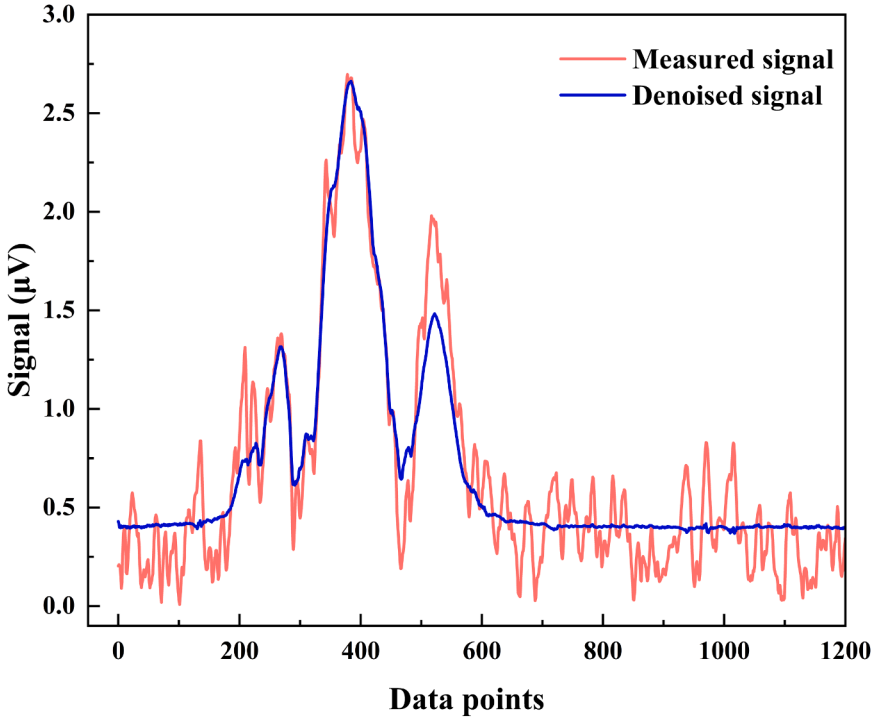


Fig. 7. Denoising results of the 2 *f* signal measured for 1% CO in zero gas using the U-LSTM model.

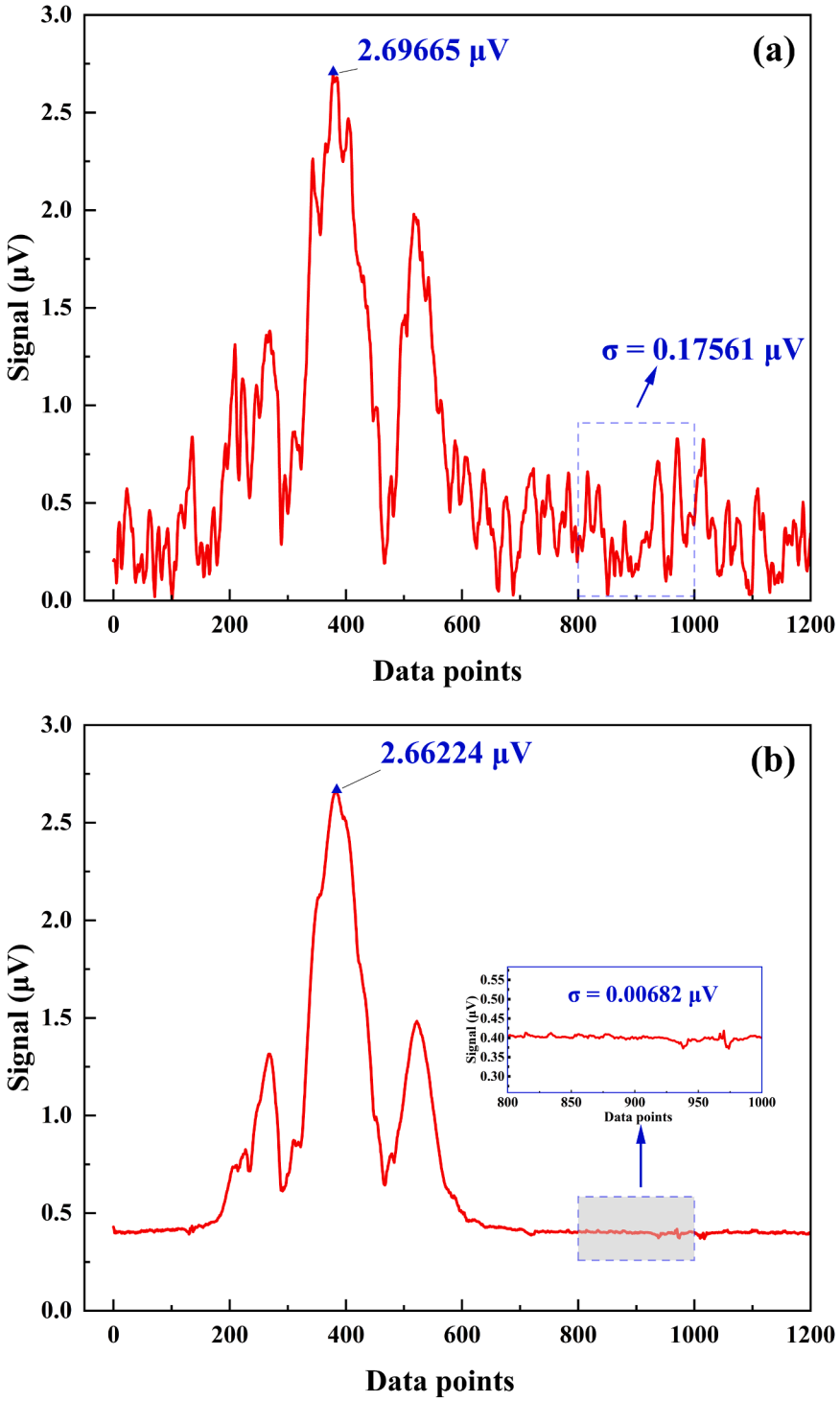


Fig. 8. (a) Experimentally measured peak amplitude and noise standard deviation of the *2 f* signal. (b) Corresponding values after U-LSTM denoising, showing a significant enhancement in SNR and overall signal clarity.

noise while preserving the essential features of the CO signal at this concentration.

Fig. 8 displays the peak amplitude and noise standard deviation of both the experimentally measured and denoised *2 f* signal. The SNR of the original signal was ~15.4, and after denoising, it increased to ~390, representing an approximately 25-fold improvement.

5. Conclusions

This study proposes a U-LSTM model designed to reduce noise in photoacoustic spectroscopy signals under complex noise conditions. By integrating a U-shaped convolutional network architecture with LSTM modules, the model overcomes the feature extraction limitations of shallow networks, substantially enhancing both noise reduction performance and system stability. Experiments with simulated data demonstrate that the U-LSTM model outperforms traditional filtering

methods, achieving an SNR of ~ 585.8 and an MSE of 1.2×10^{-12}. Analysis of experimental data further reveals that, for a 1% concentration of CO in zero gas, the model increases the SNR by approximately 25-fold. These findings confirm that the proposed U-LSTM-based denoising algorithm exhibits high robustness and excellent denoising performance. The proposed U-LSTM model learns the mapping relationship between noisy and clean photoacoustic $2f$ signals, and therefore has potential applicability to other gas species or photoacoustic sensing platforms. Nevertheless, because different gases and sensing systems may exhibit different absorption line shapes, modulation conditions, resonance characteristics, and noise distributions, additional training data, transfer learning, or model fine-tuning may be required for practical implementation in other systems. From an engineering perspective, the proposed method can serve as a software-based denoising strategy without significantly modifying the optical or acoustic hardware, which is beneficial for improving signal quality and detection stability in practical photoacoustic gas sensing applications. The proposed approach provides a promising framework for high-precision gas detection in challenging environments and contributes to the advancement of deep learning applications in optical and photoacoustic gas sensing.

CRediT authorship contribution statement

Dacheng Zhang: Project administration. **Yize Liang:** Formal analysis. **Xiu Yang:** Investigation. **Huailiang Xu:** Writing – review & editing, Investigation. **Vincenzo Spagnolo:** Writing – review & editing, Supervision, Funding acquisition. **Angelo Sampaolo:** Software. **Pietro Patimisco:** Resources. **Lei Liu:** Methodology. **Xukun Yin:** Writing – original draft, Methodology, Conceptualization. **Hongpeng Wu:** Writing – review & editing, Supervision.

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.

Acknowledgements

This work was partially funded by the National Natural Science Foundation of China (62475204, 62475137); Hangzhou Science and Technology Development Project (202204T04); The authors from the Physics Department of the Polytechnic University and the University of Bari acknowledge support from the project MUR—Dipartimenti di Eccellenza 2023–2027 (QuaSiModO, "Quantum Sensing and Modelling for One-Health"); the project PNC 0000001 (D3–4Health, "Digital Driven Diagnostics, Prognostics and Therapeutics for Sustainable Health Care," CUP: B83C22006120001); and the National Plan for Complementary Investments to the NRRP, funded by the European Union—NextGenerationEU, with the NRRP project "BRIEF—Biorobotics Research and Innovation Engineering Facilities" (CUP: J13C22000400007).

Data availability

Data will be made available on request.

References

[1] I. Bamberger, J. Stieger, N. Buchmann, W. Eugster, Spatial variability of methane: attributing atmospheric concentrations to emissions, Environ. Pollut. 190 (2014) 65–74, https://doi.org/10.1016/j.envpol.2014.03.028.

[2] W. Zhang, Y. Tang, A. Shi, L. Bao, Y. Shen, R. Shen, Y. Ye, Recent developments in spectroscopic techniques for the detection of explosives, Materials 11 (8) (2018) 1364, https://doi.org/10.3390/ma11081364.

[3] M. Jahjah, W. Ren, P. Stefański, R. Lewicki, J. Zhang, W. Jiang, J. Tarkaab, F. K. Tittel, A compact QCL based methane and nitrous oxide sensor for environmental and medical applications, Analyst 139 (9) (2014) 2065–2069, https://doi.org/10.1039/C3AN01452E.

[4] P. Patimisco, G. Scamarcio, F. Tittel, V. Spagnolo, Quartz-enhanced photoacoustic spectroscopy: a review, Sensors 14 (4) (2014) 6165–6206, https://doi.org/10.3390/s140406165.

[5] J. Wang, H. Wu, A. Sampaolo, P. Patimisco, V. Spagnolo, S. Jia, L. Dong, Quartz-enhanced multiheterodyne resonant photoacoustic spectroscopy, Light. Sci. Appl. 13 (1) (2024) 77, https://doi.org/10.1038/s41377-024-01425-1.

[6] S.M. Cristescu, S.T. Persijn, S. Te Lintel Hekkert, F.J.M. Harren, Laser-based systems for trace gas detection in life sciences, Appl. Phys. B 92 (3) (2008) 343, https://doi.org/10.1007/s00340-008-3127-y.

[7] K. Kinjalk, F. Paciolla, B. Sun, A. Zifarelli, G. Menduni, M. Giglio, H. Wu, L. Dong, D. Ayache, D. Pinto, A. Vicet, A. Baranov, P. Patimisco, A. Sampaolo, V. Spagnolo, Highly Selective and Sensitive Detection of Volatile Organic Compounds Using Long Wavelength InAs-Based Quantum Cascade Lasers through Quartz-Enhanced Photoacoustic Spectroscopy, Appl. Phys. Rev. 11 (2) (2024) 021427, https://doi.org/10.1063/5.0189501.

[8] H. Zheng, L. Dong, X. Yin, X. Liu, H. Wu, L. Zhang, W. Ma, W. Yin, S. Jia, Ppb-Level QEPAS NO_2 Sensor by Use of Electrical Modulation Cancellation Method with a High Power Blue LED, Sens. Actuators B Chem. 208 (2015) 173–179, https://doi.org/10.1016/j.snb.2014.11.015.

[9] B. Sun, A. Zifarelli, H. Wu, S. Dello Russo, S. Li, P. Patimisco, L. Dong, V. Spagnolo, Mid-Infrared Quartz-Enhanced Photoacoustic Sensor for Ppb-Level CO Detection in a SF_6 Gas Matrix Exploiting a T-Grooved Quartz Tuning Fork, Anal. Chem. 92 (20) (2020) 13922–13929, https://doi.org/10.1021/acs.analchem.0c02772.

[10] H. Wu, L. Dong, H. Zheng, Y. Yu, W. Ma, L. Zhang, W. Yin, L. Xiao, S. Jia, F. K. Tittel, Beat Frequency Quartz-Enhanced Photoacoustic Spectroscopy for Fast and Calibration-Free Continuous Trace-Gas Monitoring, Nat. Commun. 8 (1) (2017) 15331, https://doi.org/10.1038/ncomms15331.

[11] J. Tong, R. Hu, C. Hu, X. Liu, H. Cai, C. Lin, L. Zhong, J. Wang, P. Xie, Development of a net ozone production rate detection system based on dual-channel cavity ring-down spectroscopy, J. Environ. Sci. 149 (2025) 419–430, https://doi.org/10.1016/j.jes.2024.01.035.

[12] B.K. Babu, M.M. Yadav, S. Singh, V.K. Yadav, Fuel forensics: Recent advancements in profiling of adulterated fuels by ATR-FTIR spectroscopy and chemometric approaches, Spectrochim. Acta Part. A Mol. Biomol. Spectrosc. 312 (2024) 124049, https://doi.org/10.1016/j.saa.2024.124049.

[13] L. Fu, S. You, G. Li, Z. Fan, Enhancing Methane Sensing with NDIR Technology: Current Trends and Future Prospects, Rev. Anal. Chem. 42 (1) (2023) 20230062, https://doi.org/10.1515/revac-2023-0062.

[14] A.A. Kosterev, F.K. Tittel, D.V. Serebryakov, A.L. Malinovsky, I.V. Morozov, Applications of quartz tuning forks in spectroscopic gas sensing, Rev. Sci. Instrum. 76 (4) (2005) 043105, https://doi.org/10.1063/1.1884196.

[15] H. Zheng, L. Dong, H. Wu, X. Yin, L. Xiao, S. Jia, R.F. Curl, F.K. Tittel, Application of Acoustic Micro-Resonators in Quartz-Enhanced Photoacoustic Spectroscopy for Trace Gas Analysis, Chem. Phys. Lett. 691 (2018) 462–472, https://doi.org/10.1016/j.cplett.2017.12.006.

[16] X. Yin, C. Zhu, X. Yang, L. Liu, D. Zhang, K. Xu, H. Wang, S. Zhang, N. Ardito, P. Patimisco, A. Sampaolo, V. Spagnolo, H. Wu, H. Xu, Compact Differential Photoacoustic Exhaled Gas Sensor for Online $ETCO_2$ and ETO_2 Monitoring, ACS Sens. 11 (2) (2026) 1476–1483, https://doi.org/10.1021/acssensors.5c03835.

[17] Y. Liu, S. Qiao, C. Fang, Y. He, H. Sun, J. Liu, Y. Ma, A Highly Sensitive LITES Sensor Based on a Multi-Pass Cell with Dense Spot Pattern and a Novel Quartz Tuning Fork with Low Frequency, Opto-Electron. Adv. 7 (3) (2024) 230230, https://doi.org/10.29026/oea.2024.230230.

[18] W. Chen, S. Qiao, Y. He, J. Zhu, K. Wang, L. Xiao, Y. Ma, Quasi-Distributed Quartz Enhanced Photoacoustic Spectroscopy Sensing Based on Hollow Waveguide Micropores, Opt. Lett. 49 (10) (2024) 2765, https://doi.org/10.1364/OL.525188.

[19] H. Sun, S. Qiao, Y. He, Y. Liu, Y. Ma, Highly Sensitive CH_4, C_2H_2 and CO Simultaneous Measurement LITES Sensor Based on Multi-Pass Cell with Overlapped Spots Pattern and QTFs with Low Resonant Frequency, Opt. Express 32 (16) (2024) 28183, https://doi.org/10.1364/OE.531925.

[20] X. Yang, B. Li, X. Geng, T. Zhao, Q. Yu, J. Hou, D. Zhang, Y. Liang, K. Xu, H. Wu, X. Yin, Ppb-Level Formaldehyde Sensor Utilizing a Compact 3D-Printed Differential Photoacoustic Cell and a 320 nm UV Laser, Photoacoustics 46 (2025) 100780, https://doi.org/10.1016/j.pacs.2025.100780.

[21] C. Zhang, Y. He, S. Qiao, Y. Liu, Y. Ma, High-Sensitivity Trace Gas Detection Based on Differential Helmholtz Photoacoustic Cell with Dense Spot Pattern, Photoacoustics 38 (2024) 100634, https://doi.org/10.1016/j.pacs.2024.100634.

[22] L. Fu, J. Zhang, Y. Pan, P. Lu, Differential Photoacoustic Spectroscopy for Flow Gas Detection Based on Single Microphone, Photoacoustics 38 (2024) 100624, https://doi.org/10.1016/j.pacs.2024.100624.

[23] X. Zhao, K. Chen, D. Cui, M. Guo, C. Li, H. Qi, G. Zhang, Z. Gong, Z. Zhou, W. Peng, Ultra-High Sensitive Photoacoustic Gas Detector Based on Differential Multi-Pass Cell, Sens. Actuators B Chem. 368 (2022) 132124, https://doi.org/10.1016/j.snb.2022.132124.

[24] X. Yin, Q. Hu, X. Yang, K. Xu, Y. Liang, J. Hou, A. Sampaolo, P. Patimisco, V. Spagnolo, D. Zhang, Z. Liu, H. Xu, H. Wu, Ppb-Level HF Sensor in GeF_4 Gas Matrices with a 76 m TDLAS Cell, Sens. Actuators B Chem. 449 (2026) 139141, https://doi.org/10.1016/j.snb.2025.139141.

[25] B. Turan, A New Approach for SPN Removal: Nearest Value Based Mean Filter, PeerJ Comput. Sci. 8 (2022) e1160, https://doi.org/10.7717/peerj-cs.1160.

[26] X. Liu, S. Qiao, Y. Ma, Highly Sensitive Methane Detection Based on Light-Induced Thermoelastic Spectroscopy with a 2.33 μm Diode Laser and Adaptive Savitzky-

Golay Filtering, Opt. Express 30 (2) (2022) 1304, https://doi.org/10.1364/OE.446294.
[27] L. Gu, Z. Fei, X. Xu, Enhancement Method of Weak Lidar Signal Based on Adaptive Variational Modal Decomposition and Wavelet Threshold Denoising, Infrared Phys. Technol. 120 (2022) 103991, https://doi.org/10.1016/j.infrared.2021.103991.
[28] L. Alzubaidi, J. Zhang, A.J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M.A. Fadhel, M. Al-Amidie, L. Farhan, Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions, J. Big Data 8 (1) (2021) 53, https://doi.org/10.1186/s40537-021-00444-8.
[29] J. Sun, J. Chang, Y. Wei, Z. Zhang, S. Lin, F. Wang, Q. Zhang, Dual Gas Sensor with Innovative Signal Analysis Based on Neural Network, Sens. Actuators B Chem. 373 (2022) 132697, https://doi.org/10.1016/j.snb.2022.132697.
[30] P. Fu, L. Hou, Tomographic Measurements of Temperature and Species Concentration via Deep Learning and Wavelength Modulated Absorption Spectroscopy, AIP Adv. 13 (7) (2023) 075228, https://doi.org/10.1063/5.0138504.
[31] Z. Gong, Y. Fan, Y. Guan, G. Wu, L. Mei, Empirical Modal Decomposition Combined with Deep Learning for Photoacoustic Spectroscopy Detection of Mixture Gas Concentrations, Anal. Chem. 96 (46) (2024) 18528–18536, https://doi.org/10.1021/acs.analchem.4c04479.
[32] C. Zuo, J. Qian, S. Feng, W. Yin, Y. Li, P. Fan, J. Han, K. Qian, Q. Chen, Deep Learning in Optical Metrology: A Review, Light. Sci. Appl. 11 (1) (2022) 39, https://doi.org/10.1038/s41377-022-00714-x.
[33] 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, https://doi.org/10.1016/j.snb.2021.129982.
[34] X. Liu, S. Qiao, G. Han, J. Liang, Y. Ma, Highly Sensitive HF Detection Based on Absorption Enhanced Light-Induced Thermoelastic Spectroscopy with a Quartz Tuning Fork of Receive and Shallow Neural Network Fitting, Photoacoustics 28 (2022) 100422, https://doi.org/10.1016/j.pacs.2022.100422.
[35] L. Tian, J. Xia, A.A. Kolomenskii, H.A. Schuessler, F. Zhu, Y. Li, J. He, Q. Dong, S. Zhang, Gas Phase Multicomponent Detection and Analysis Combining Broadband Dual-Frequency Comb Absorption Spectroscopy and Deep Learning, Commun. Eng. 2 (1) (2023) 54, https://doi.org/10.1038/s44172-023-00105-z.
[36] C. Liu, G. Wang, C. Zhang, P. Patimisco, R. Cui, C. Feng, A. Sampaolo, V. Spagnolo, L. Dong, H. Wu, End-to-End Methane Gas Detection Algorithm Based on Transformer and Multi-Layer Perceptron, Opt. Express 32 (1) (2024) 987, https://doi.org/10.1364/OE.511813.
[37] Y. Wang, Y. Jiang, H. Hao, B. Yu, J. Li, S. Zhou, Pressure Calibration- and Profile Fitting-Free Spectroscopy Technology Based on Deep Neural Network for Gas Sensing, Measurement 204 (2022) 112077, https://doi.org/10.1016/j.measurement.2022.112077.
[38] L. Wang, H. Lv, Y. Zhao, C. Wang, H. Luo, H. Lin, J. Xie, W. Zhu, Y. Zhong, B. Liu, J. Yu, H. Zheng, Sub-Ppb Level HCN Photoacoustic Sensor Employing Dual-Tube Resonator Enhanced Clamp-Type Tuning Fork and U-Net Neural Network Noise Filter, Photoacoustics 38 (2024) 100629, https://doi.org/10.1016/j.pacs.2024.100629.

Xukun Yin received his Ph.D. degree in atomic and molecular physics from Shanxi University, China, in 2020. From 2018–2019, he studied as a research associate in the electrical and computer engineering department, Rice University, Houston, USA. Currently he is an associate professor in the School of Optoelectronic Engineering of Xidian University. His research interests include optical sensors, laser spectroscopy techniques and insulation fault diagnosis of electrical equipment.

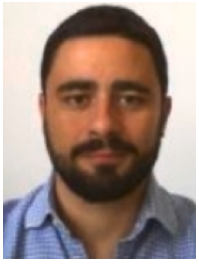
Lei Liu is currently pursuing a master's degree in the School of Physics and Optoelectronic Engineering, Xidian University. His research focuses on high-sensitivity trace gas sensing technology and neural network-based denoising.

Xiu Yang is now pursuing his doctoral degree in Electronic Information from Xidian University. His research interests include photoacoustic spectroscopy and gas sensing technology.

Yize Liang received his B.S. degree of optoelectronic information science and engineering and doctorate in optical engineering from Huazhong University of Science and Technology, in 2018 and 2023. He is now an assistant professor in the School of Optoelectronic Engineering, Xidian University. His research field includes structured light, multi-mode fiber and free-space optics.

Dacheng Zhang is a Professor and Ph.D. Supervisor at Xidian University. He serves as a member of several academic committees under the Chinese Society for Optical Engineering and the Chinese Optical Society, and as a Young Editorial Board Member of Modern Applied Physics, Metallurgical Analysis, and Uranium Mining and Metallurgy. His research focuses on laser plasma spectroscopy, laser resonance ionization, and tunable laser technology, with an emphasis on developing specialized light sources for spectroscopic measurements. He has led more than 20 national and provincial research projects, published over 80 SCI/EI papers, and holds more than 10 invention patents.

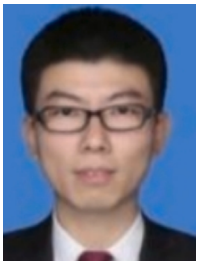
Pietro Patimisco obtained the Master degree in Physics (cum laude) in 2009 and the Ph.D. Degree in Physics in 2013 from the University of Bari. Since 2018, he is Assistant professor at the Technical University of Bari. He was a visiting scientist in the Laser Science Group at Rice University in 2013 and 2014. Dr. Patimisco's scientific activity addressed both micro-probe optical characterization of semiconductor optoelectronic devices and optoacoustic gas sensors. Recently, his research activities included the study and applications of trace-gas sensors, such as quartz-enhanced photoacoustic spectroscopy and cavity enhanced absorption spectroscopy in the mid infrared and terahertz spectral region, leading to several publications, including a cover paper in Applied Physics Letter of the July 2013 issue.



Angelo Sampaolo obtained his Master degree in Physics in 2013 and the Ph.D. Degree in Physics in 2017 from University of Bari. He was an associate researcher in the Laser Science Group at Rice University from 2014 to 2016 and associate researcher at Shanxi University since 2018. Since May 2017, he was a Post-Doctoral Research associate at University of Bari and starting from December 2019, he is Assistant Professor at Polytechnic of Bari. His research activity has included the study of the thermal properties of heterostructured devices via Raman spectroscopy. Most recently, his research interest has focused on the development of innovative techniques in trace gas sensing, based on Quartz-Enhanced Photoacoustic Spectroscopy and covering the full spectral range from near-IR to THz. His achieved results have been acknowledged by a cover paper in Applied Physics Letter of the July 2013 issue.

Vincenzo Spagnolo obtained the Ph.D. in physics in 1994 from University of Bari. From 1997–1999, he was researcher of the National Institute of the Physics of Matter. Since 2004, he works at the Technical University of Bari, formerly as assistant and associate professor and, starting from 2018, as full Professor of Physics. Since 2019, he is vice-rector of the Technical University of Bari, deputy to technology transfer. He is the director of the joint-research lab PolySense between Technical University of Bari and THORLABS GmbH, fellow member of SPIE and senior member of OSA. His research interests include optoacoustic gas sensing and spectroscopic techniques for real-time monitoring. His research activity is documented by more than 220 publications and two filed patents. He has given more than 50 invited presentations at international conferences and workshops.

Huailiang Xu received his Ph.D. degree in physics from Lund University, Lund, Sweden, in 2004. He was a Postdoctoral Researcher at the Laval University of Canada. In January 2008, he became an Assistant Professor at the University of Tokyo, Tokyo, Japan. Since September 2009, he has been a Full Professor at Jilin University, Changchun, China. His research interests include ultrafast strong laser sensing, strong-field laser physics, and atomic and molecular spectroscopy.



Hongpeng Wu received his Ph.D. degree in atomic and molecular physics from Shanxi University, China, in 2017. From September, 2015 to October, 2016, he studied as a joint p.D. student in the Electrical and Computer Engineering Department and Rice Quantum Institute, Rice University, Houston, USA. Currently he is a professor in the Institute of Laser Spectroscopy of Shanxi University. His research interests include gas sensors, photoacoustic spectroscopy, and laser spectroscopy techniques.