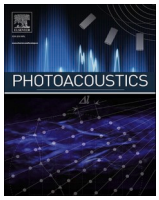
Photoacoustics 45 (2025) 100758

Contents lists available at ScienceDirect

Photoacoustics

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



Enhancing photoacoustic trace gas detection via a CNN–transformer denoising framework☆

Chen Zhang [a,b,1], Yan Gao [c,1], Ruyue Cui [a,b,1], Hanxi Zhang [a], Jinhua Tian [a], Yujie Tang [a], Lei Yang [c], Chaofan Feng [a,b], Pietro Patimisco [a,d,e], Angelo Sampaolo [a,d,e], Vincenzo Spagnolo [a,d,e,*], Xukun Yin [f,**], Lei Dong [a,b,d,*], Hongpeng Wu [a,b,d,f,*]

[a] State Key Laboratory of Quantum Optics and Quantum Optics Devices, Institute of Laser Spectroscopy, Shanxi University, Taiyuan 030006, China
[b] Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan 030006, China
[c] School of Physical and Electronic Sciences, Shanxi Datong University, Datong 037009, China
[d] PolySense Lab—Dipartimento Interateneo di Fisica, Politecnico and University of Bari, Bari, Italy
[e] PolySense Innovations Srl, Bari, Italy
[f] School of Optoelectronic Engineering, Xidian University, Xi'an 710071, China

ARTICLE INFO

Keywords:
Photoacoustic spectroscopy
Trace gas detection
Deep learning
CNN–transformer framework
Signal denoising

ABSTRACT

We present a novel approach for gas concentration measurement using a differential resonant photoacoustic cell combined with a deep learning-based signal denoising model. This method addresses the persistent challenge of noise interference in 2 *f* signals at low gas concentrations, where conventional processing methods struggle to maintain signal fidelity. To resolve this, we propose a deep learning model that integrates 1D Convolutional Neural Networks (1D CNNs) for local feature extraction and Transformer networks for capturing global dependencies. The model was trained using synthetic signals with added noise to simulate real-world conditions, ensuring robustness and adaptability. Applied to experimental 2 *f* signals, the model demonstrated excellent noise suppression capabilities, enhancing the signal-to-noise ratio (SNR) of 500 ppb acetylene signals by a factor of approximately 70. Furthermore, the determination coefficient (R^2) improved, reflecting better accuracy and linearity in signal reconstruction. These results underscore the model's potential for improving detection sensitivity and reliability in trace gas measurements, marking a significant advancement in spectroscopic signal processing for gas detection.

1. Introduction

Trace gas sensing is essential for environmental monitoring, industrial processes, and medical diagnostics, as it enables the detection of trace gas concentrations that can profoundly impact health and safety [1–7]. Among various gas sensing technologies, laser spectroscopy has gained significant attention for its superior sensitivity, exceptional selectivity, and extended operational lifespan, making it a preferred choice for high-precision applications [8–10]. Photoacoustic spectroscopy, a specific type of laser spectroscopy, detects acoustic signals produced by the interaction of modulated laser light with gas molecules to determine gas concentrations [11–13]. This technique not only retains the advantages of laser spectroscopy—such as high sensitivity and selectivity—but also offers unique benefits, including zero background noise and a direct correlation between detection sensitivity and laser power [14–16]. These features have driven its rapid advancement and adoption in recent years [17–19]. However, significant challenges emerge when applying this technique to

☆ Given their role as Guest Editor, Lei Dong, Vincenzo Luigi Spagnolo and Angelo Sampaolo had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor

* Corresponding authors at: State Key Laboratory of Quantum Optics and Quantum Optics Devices, Institute of Laser Spectroscopy, Shanxi University, Taiyuan 030006, China.
** Corresponding author.
E-mail addresses: vincenzoluigi.spagnolo@poliba.it (V. Spagnolo), xkyin@xidian.edu.cn (X. Yin), donglei@sxu.edu.cn (L. Dong), wuhp@sxu.edu.cn (H. Wu).
[1] These authors contributed equally to this work

https://doi.org/10.1016/j.pacs.2025.100758
Received 30 June 2025; Received in revised form 31 July 2025; Accepted 5 August 2025
Available online 6 August 2025
2213-5979/© 2025 The Authors. Published by Elsevier GmbH. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

low-concentration gases. The primary obstacle is the interference from various noise sources during measurement, including environmental noise (low-frequency), laser source noise (high-frequency), and thermal noise (Gaussian white noise). These noise components combined to produce a substantial background signal that can mask the weak photoacoustic signals at low gas concentrations, severely degrade the SNR and, in turn, the detection sensitivity. Consequently, mitigating or eliminating the noise impact to enhance detection performance remains a critical research focus and a major challenge in this field [20–23].

Deep learning, a cutting-edge technology in artificial intelligence, has shown remarkable potential in signal processing, particularly for signal denoising [24–32]. Unlike traditional denoising methods, deep learning techniques can automatically capture complex, nonlinear features within data, overcoming the limitations of manual feature extraction [33–36]. This ability is especially advantageous in high-noise environments and non-stationary signal scenarios [37]. End-to-end deep learning models can directly learn noise patterns and adapt to varying noise conditions, effectively filtering out unwanted signals while preserving essential signal features, thus significantly enhancing signal quality [38–40]. Furthermore, deep learning models exhibit strong generalization capabilities, allowing them to handle diverse noise types and signal variations, making them ideal for real-world applications [41,42]. These strengths provide innovative solutions to noise-related challenges in photoacoustic spectroscopy, enabling breakthrough improvements in the detection and analysis of low-concentration gases [43,44].

In this study, we present a deep learning model that integrates 1D CNNs and Transformers (CTNN) to effectively denoise photoacoustic signals. 1D CNNs excel at extracting local features, such as peaks and valleys, which are critical for precise gas concentration measurements [36]. In parallel, Transformers are adept at capturing long-range dependencies, enabling the integration of global context to preserve the overall signal structure during denoising. By combining these complementary strengths, the proposed model achieves significant enhancements in the SNR of low-concentration gas signals. This synergic approach not only improves denoising performance but also ensures high accuracy in gas detection. Notably, this work presents the first application of 1D CNNs and Transformers in tandem for photoacoustic spectroscopy, representing a novel contribution to the field and laying the foundation for further advancements in signal processing for trace gas detection [45–47].

This paper underscores the growing significance of deep learning in advancing signal processing techniques, particularly in addressing one of the most persistent challenges in photoacoustic spectroscopy: the accurate detection of low-concentration gases in high-noise environments. Our research achieved a remarkable 70-fold improvement in the SNR of acetylene 2*f* signals, alongside an enhancement in the R-squared value at low concentrations from 99.0 % to 99.3 %. Leveraging the unique capabilities of deep learning, combined with effective normalized preprocessing, our method demonstrates exceptional robustness and generalizability, making it suitable for denoising signals from other gas species. As deep learning technologies continue to evolve, their applications in scientific measurement and analysis are expected to expand, providing innovative and transformative solutions to long-standing challenges across various disciplines.

2. Algorithm design

2.1. Data preprocessing

2.1.1. Noise analysis

To train the deep learning model for denoising photoacoustic spectra, we employed theoretically simulated 2*f* signals comprising 501 sampling points. 501 sampling points were selected in this study to facilitate comparison between experimentally measured data and theoretical calculations. In future applications, the number of sampling points can be adjusted according to specific requirements. These signals are inherently smooth and free from noise artifacts. However, to emulate real-world experimental conditions and ensure the model's ability to effectively process noisy signals, we introduced various types of noise commonly encountered in photoacoustic spectroscopy.

First, thermal noise, modeled as Gaussian white noise, was introduced to simulate random fluctuations caused by electronic components and other thermal sources within the measurement system. Mathematically, the thermal noise is expressed as:

$$Thermal\ Noise = N(0\ ,\ \sigma^2)$$

where $N(0\ ,\ \sigma^2)$ denotes a normal distribution with a mean of zero and a variance of σ^2, set to 0.000003 in this context. We incorporated pink noise, also known as 1/f noise, which commonly occurs in low-frequency environmental noise scenarios. Its spectral density is inversely proportional to frequency, and can be modeled as follows:

$$Pink\ Noise(f) = \frac{1}{f} \bullet Random\ Signal$$

To generate the pink noise, a random signal was initially transformed into the frequency domain using the Fast Fourier Transform (FFT). The spectrum was then modified to achieve a $1/f$ frequency distribution, and the signal was converted back to the time domain using the Inverse Fast Fourier Transform (IFFT). Additionally, environmental noise was introduced, represented by a low-frequency sinusoidal waveform to simulate the background noise typically encountered in experimental settings. This noise component can be expressed as:

$$Environmental\ Noise = A \bullet \sin(2\pi f t)$$

where A is the amplitude, set to 0.000003, and f is the frequency, with a slight random variation between 0.4 and 0.5 Hz, to mimic real-world variability. To simulate fluctuations caused by the laser during the photoacoustic measurement process, high-frequency noise was also introduced. This noise can be mathematically expressed as:

$$Laser\ Noise = N(0\ ,\ \sigma^2) + A \bullet \sin\left(2\pi f_{high} t\right)$$

Here, $N(0\ ,\ \sigma^2)$ represents the Gaussian white noise, while f_{high} denotes the high-frequency component, randomly varying between 40 and 60 Hz.

By superimposing these noise components onto the theoretically smooth signals, we created a dataset that closely replicates real experimental conditions, as illustrated in Fig. 1. The similarity between the noisy simulated signal and the real experimental data, particularly in the increased fluctuations, indicates that the synthetic signal with added noise can effectively mimic real-world scenarios. This highlights the utility of the synthetic dataset for training deep learning models, as it exposes them to diverse and realistic noise conditions. Such exposure enhances the model's robustness and effectiveness in denoising tasks, particularly for photoacoustic spectroscopy applied to low-concentration gas detection.

2.1.2. Normalization

The spectral 2*f* data used for training are derived from theoretical simulations of specific gases. However, in real measurements, signal values for different gases at varying concentrations can span multiple orders of magnitude. To ensure the algorithm's robustness across a wide range of gases, it is essential to maintain consistency in the scale of both the training data and the denoised outputs. Without proper scaling, the model's performance could be compromised, leading to suboptimal results.

To ensure consistency, we applied a normalization process to the training data. This process involves subtracting the mean of the signal and dividing it by its standard deviation, effectively standardizing the data to have a mean of zero and a standard deviation of one. Mathe-

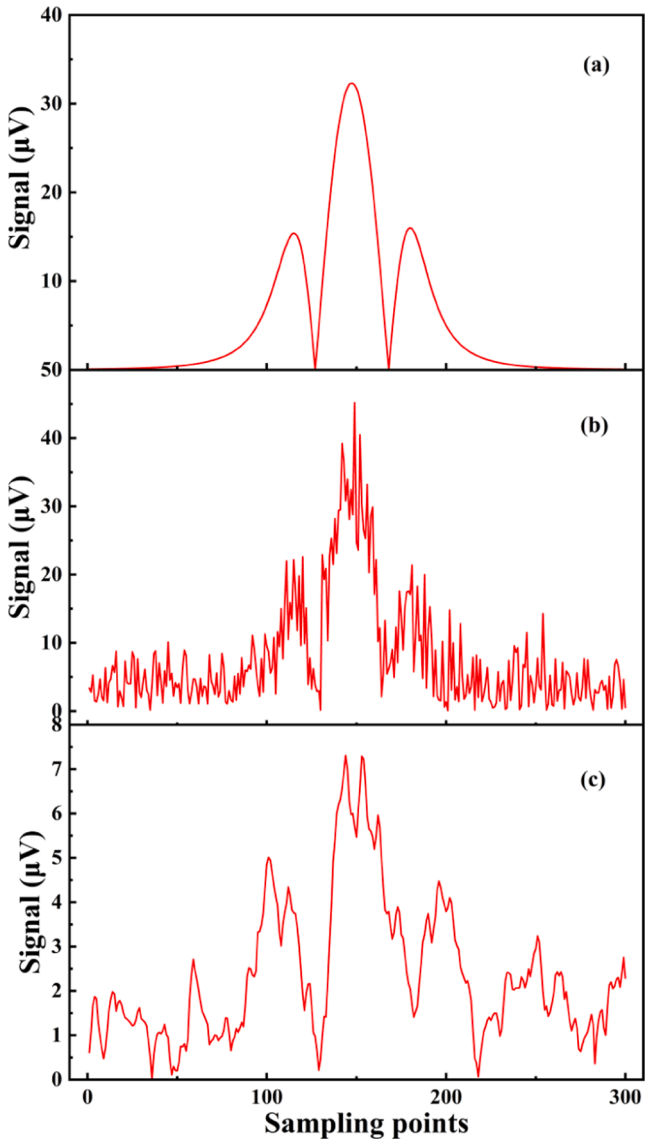


Fig. 1. Comparison of simulated and real acetylene signals: (a) Ideal, noise-free simulated signal; (b) Simulated signal with added noise components, including environmental, laser source, and thermal noise; and (c) Experimentally measured acetylene signal at a concentration of 100 ppb.

matically this can be expressed as:

$$Normalized\ signal = \frac{signal - \mu}{\sigma}$$

where μ represents the mean of the signal, and σ represents the standard deviation. This normalization ensures that the training data are appropriately scaled, enabling the model to generalize effectively across various gas species and concentrations when applied to real-world scenarios.

We selected z-score normalization over min-max scaling primarily because it preserves the relative morphological features of photoacoustic signals across varying concentrations. Since gas absorption peaks exhibit critical shape characteristics essential for quantitative analysis, z-score's mean-centering and standard-deviation scaling maintain these intrinsic patterns without distorting waveform structures—unlike min-max normalization, which forcibly compresses all amplitudes into a fixed range (e.g., [0,1]), disproportionately flattening low-concentration signals and amplifying noise in near-zero regions. Furthermore, z-score's robustness to outliers (common in experimental noise) prevents localized anomalies from skewing the entire dataset's transformation, whereas min-max's dependence on extreme values would catastrophically distort outputs if sporadic high-amplitude artifacts exist. Crucially, by retaining original data distributions, z-score enables consistent feature extraction for gases spanning multiple orders of magnitude—ensuring the model generalizes across trace to high concentrations without amplitude-induced bias.

2.2. Algorithm architecture

This section describes the architecture of the deep learning model developed for denoising photoacoustic spectra. The model combines 1D CNNs [36] with Transformer networks [48,49], called CTNN, to effectively leverage both local and global information. Local features extracted by 1D CNNs directly correspond to the second-derivative characteristics of 2 f signals (governed by Voigt profiles). Global dependencies captured by Transformers address long-range noise correlations inherent in resonant photoacoustic systems. Specifically, self-attention mechanisms suppress low-frequency environmental noise exhibiting high temporal coherence across wavelength-scanned 2 f cycles. This integration optimizes the denoising process while preserving critical peak signal features. The input to the model consists of simulated signal with 501 data points corrupted by noise, and the output is the corresponding smooth, noise-free signal.

The proposed denoising model consists of three main components: an encoder, a Transformer encoder layer, and a decoder. The encoder extracts the local features from the input signal, the Transformer captures long-range dependencies to retain global context, and the decoder reconstructs the clean signal from the encoded representation.

The encoder consists of a series of 1D convolutional layers that progressively extract features from the noisy input signal. Each layer is followed by a ReLU [50] activation function, introducing non-linearity to enhance the model's ability to learn complex patterns. As illustrated in Fig. 2. 1D CNNs are particularly well-suited for processing sequential data, such as time-series signals, as they effectively capture local dependencies by applying convolutional filters across the sequence. These filters identify critical patterns, such as peaks and troughs, which are essential for accurately representing the signal structure. By stacking multiple convolutional layers, the encoder extracts increasingly abstract features, enabling the model to effectively differentiate between noise and the true signal.

After the encoder, the extracted features are fed into a Transformer encoder layer, designed to capture global dependencies within the sequence. The Transformer layer employs self-attention mechanisms to evaluate the relevance of each data point relative to the others. This enables the model to preserve critical signal features while effectively suppressing noise. The structure of the Transformer encoder is as follows:

$$TransformerLayer(d_{model} = 256,\quad n_{head} = 4) \quad \times \quad 2$$

Originally developed for natural language processing tasks, the Transformer architecture has demonstrated exceptional proficiency in capturing long-range dependencies within sequential data. At its core, the self-attention mechanism enables each element of the input sequence to assess the relevance of other elements, allowing the model to focus on critical parts of the signal, even those separated by large temporal distances. This capability is particularly advantageous in signal processing applications, where preserving global context is essential for accurate and reliable denoising.

The decoder mirrors the encoders' structure but employs 1D transposed convolutional layers (also known as deconvolution layers) to reconstruct the denoised signal from the features extracted by the encoder and enhanced by the Transformer. These layers upsample the encoded feature map, enhanced by the Transformer, to match the original signal's dimensions, producing a smooth and noise-free output, as shown in Fig. 3. The use of transposed convolutional layers enables

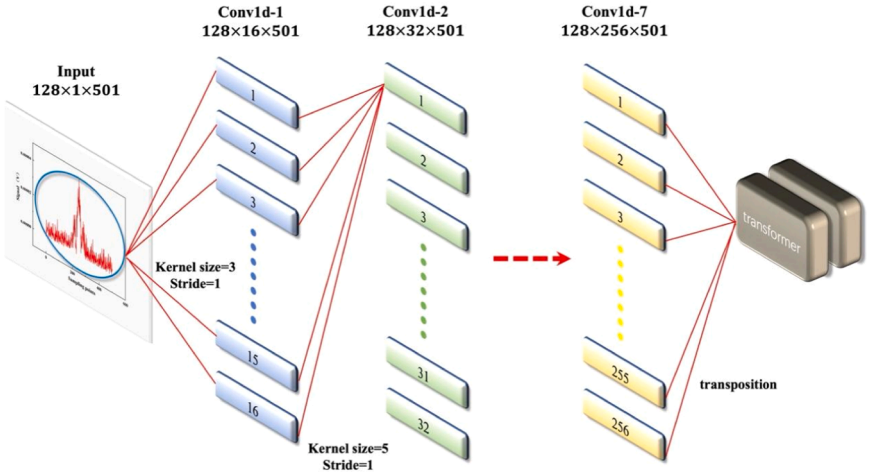


Fig. 2. Structure of the proposed denoising model, consisting of multiple 1D convolutional layers followed by a transformer module. The model processes 128 input signals, extracting local features through convolution, while the Transformer module captures the global context for enhanced signal representation.

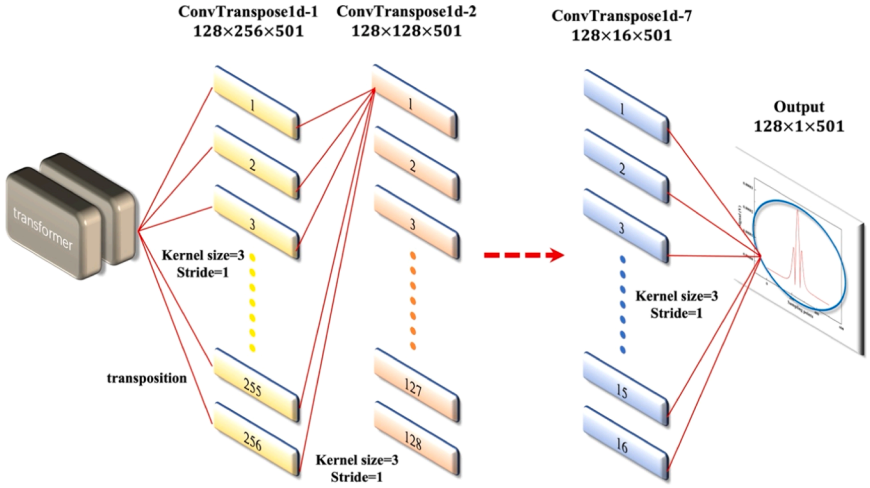


Fig. 3. The second stage of the proposed denoising model, illustrating the transposed convolutional layers responsible for signal reconstruction. Following feature extraction by the Transformer module, these layers upsample the encoded feature map to match the original signal's dimensions.

precise restoration of the input's spatial resolution, preserving the essential characteristics of the original signal. This architecture is particularly well-suited for denoising photoacoustic signals with complex noise profiles, as it effectively retains both local signal details and global dependencies, crucial for accurate reconstructing the 501-point signal shown in Fig. 1.

By combining the strengths of 1D CNNs for local feature extraction and Transformers for capturing global context, the model efficiently analyzes fine-grained and broad patterns within the noisy input. This synergy ensures that critical peak information is preserved while extraneous noise is effectively canceled, resulting in a robust and highly accurate denoising process.

2.3. Implementation detail

The CTNN was implemented using the Pytorch deep learning framework. The model employs a modular architecture, as outlined in Table 1. The encoder, composed of a series of 1D convolutional layers, progressively captures local dependencies and extracts increasingly abstract features from the noisy input signal. These features are further refined by a Transformer encoder, which utilizes four Transformer layers with self-attention mechanisms to preserve global context and suppress noise effectively. The decoder uses transposed convolutional

Table 1
Detailed information of the CTNN.

Encoder	Transformer	Decoder
Conv1d(1 → 16, kernel_size=3, stride=1, padding=1)	TransformerLayer (d_model=256, n_head=4) × 2	ConvTranspose1d(256→256, kernel_size=3, stride=1, padding=1)
Conv1d(16 → 32, kernel_size=5, stride=1, padding=2)		ConvTranspose1d(256→128, kernel_size=3, stride=1, padding=1)
Conv1d(32 → 64, kernel_size=7, stride=1, padding=3)		ConvTranspose1d(128→64, kernel_size=5, stride=1, padding=2)
Conv1d(64 → 64, kernel_size=7, stride=1, padding=3)		ConvTranspose1d(64→64, kernel_size=7, stride=1, padding=3)
Conv1d(64 → 128, kernel_size=5, stride=1, padding=2)		ConvTranspose1d(64→32, kernel_size=7, stride=1, padding=3)
Conv1d(128 → 256, kernel_size=3, stride=1, padding=1)		ConvTranspose1d(32→16, kernel_size=5, stride=1, padding=2)
Conv1d(256 → 256, kernel_size=3, stride=1, padding=1)		ConvTranspose1d(16→1, kernel_size=3, stride=1, padding=1)

layers to reconstruct the denoised signal, restoring the original dimensions while retaining critical spatial and spectral characteristics.

During the training process, the model was trained using the Adam optimizer with a learning rate of 0.0001, a batch size of 128, and over 50 epochs. The training and testing are conducted on an NVIDIA T4 GPU.

To improve the fidelity of the denoised model in preserving both the peak regions and the global morphology of spectral signals, a custom loss function, named Peak MSE With Correlation Loss, was designed. This loss function consists of two components: the Peak-Weighted Mean Squared Error (Peak-MSE) and the Correlation Loss. The Peak-MSE assigns higher weights to the peak regions of the target signal, emphasizing the accuracy of denoising in these critical areas, while calculating the mean squared error for all data points. Correlation Loss ensures the preservation of the overall trend and global structure of the signal by maximizing the correlation between the model's output and the target signal. The total loss function balances these two components through the following equation:

$$L_{total} = L_{Peak-MSE} + \lambda \bullet L_{Correlation}$$

Here, λ is the weight of the correlation loss, ensuring both local and global fidelity in the denoised output.

3. Experimental setup

The experimental setup can be divided into three main sections: the laser control circuit, the gas dilution system and the photoacoustic cell, as illustrated in Fig. 4.

Using the HITRAN database, a strong absorption peak for acetylene was identified at a wavelength of 6530.39 cm^{-1}. The distributed feedback (DFB) laser employed for the experiment was precisely tuned to this wavelength. Optimal performance was achieved at a temperature of 27.1 °C and a current of 310 mA, yielding a laser power output of 46.3 mW. The experimental platform and gas line were then assembled as illustrated in Fig. 4.

Mixtures of 10 ppm acetylene in pure nitrogen and high-purity nitrogen were fed to the gas mixer to generate samples with varying acetylene concentrations. Acetylene concentrations ranging from 100 ppb to 2 ppm in nitrogen were produced. The obtained gas mixture was directed through a needle valve into the sealed differential resonant photoacoustic cell, the core element of the gas sensing system. After passing through the photoacoustic cell, the gas was subsequently evacuated via a second needle valve connected to a vacuum pump. The flow rate and pressure within the gas line were regulated by a flow meter and a pressure controller, ensuring stable conditions throughout the experiment.

A low-noise differential resonance photoacoustic cell, shown in Fig. 5, was utilized to enhance detection sensitivity and suppress external noise [6,15]. The cell features two parallel resonators, each 90 mm in length and 8 mm in inner diameter, designed with fully symmetrical geometry to reduce noise from gas flow, window absorption, and electromagnetic interference. Buffer volumes and λ/4 filters further mitigate external acoustic noise. Two microphones with identical frequency sensitivity (-32 dB) were centrally placed in each resonator, enabling differential amplification to enhance the photoacoustic signal while canceling coherent noise.

The resonant frequency of the differential photoacoustic resonator resulted to be 1766 Hz. When the laser was switched on, a sine wave from a function generator was applied at half the resonant frequency, i.e., 883 Hz. A triangular wave was used to scan the laser wavelength across the selected absorption feature, fully encompass one complete cycle of the 2 *f* signal. This continuous scanning process allowed real-time observation of the 2 *f* signal on a computer via LabVIEW, interfaced with a lock-in amplifier. The corresponding 2 *f* signals for the gas sample with different acetylene concentrations were then recorded and analyzed.

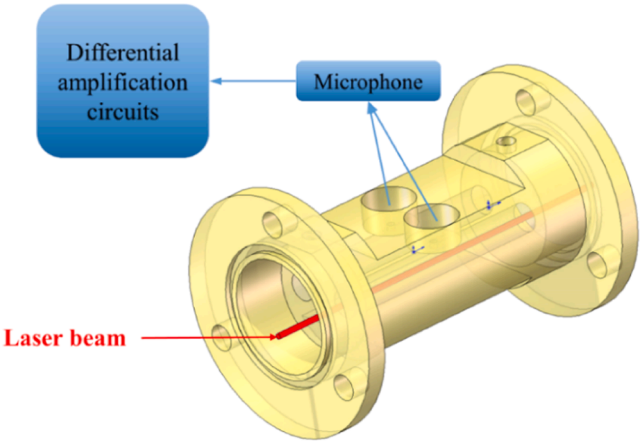


Fig. 5. The structure of low-noise differential resonant photoacoustic cell.

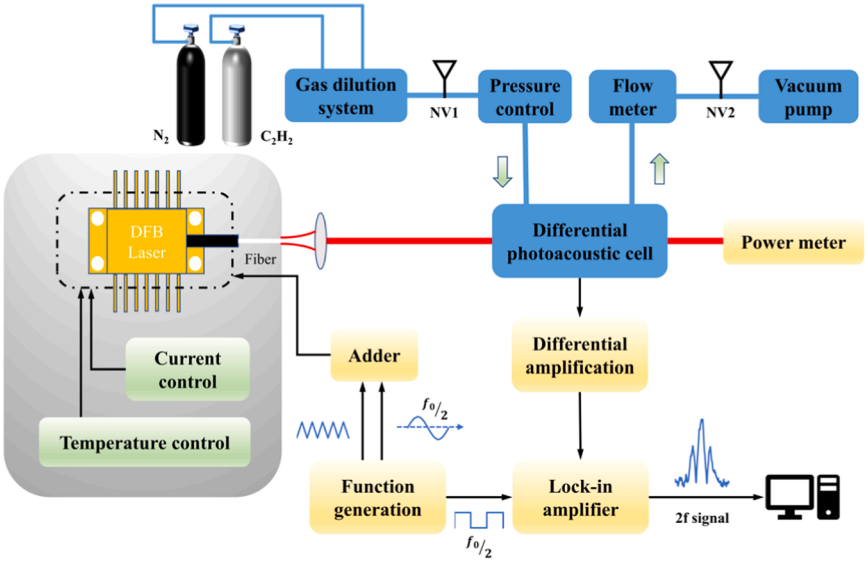


Fig. 4. Schematic diagram of the experimental setup for acetylene gas detection using differential photoacoustic spectroscopy.

4. Results and denoising analysis

In order to verify the cross-gas applicability of the algorithm used in this article, the simulated absorption line 2 *f* signal of CH_4 was specifically used as the training data, while the real test and evaluation data used the 2 *f* signal of C_2H_2 collected in the laboratory. Therefore, the subsequent result analysis proves the performance of the algorithm by comparing its performance on training data and real data, which can demonstrate the robustness of the denoising effect on absorption line changes.

The model training and testing process begins by taking the absolute value of the signals, as the actual measured 2 *f* signals are positive, whereas the simulated signals, include two small negative lobes. This preprocessing step ensures consistency between the simulated data and real measurements.

Following this adjustment, the output dataset for the model was generated by adding random noise to the smooth, simulated signals. To account for the significant role of data scaling in effective denoising, normalization was applied to the dataset to enhance the model's robustness. The model was trained using a processed dataset, achieving an SNR enhancement of nearly two orders of magnitude on test data with initially low SNR. The denoising performance on a representative dataset is illustrated in Fig. 6, demonstrating the model's effectiveness in noise reduction. A comparison between the original noisy signal, the clean signal, and the denoised output highlights the model's ability to reduce noise while preserving essential signal characteristics.

The application of the previously trained model was extended to real experimental data. Following promising results on the simulated dataset, the trained CTNN model was applied to 2 *f* signals obtained from acetylene gas mixtures measurements. The model was tested on acetylene signals with concentrations ranging from 100 ppb and 2 ppm, and the denoising results are presented in Fig. 7. The trained model exhibited strong performance, effectively reducing noise while preserving the essential features of the acetylene signals across this concentration range.

A detailed analysis of the denoising performance using various methods is presented. For each concentration level, approximately five 2 *f* signals were recorded. For the 100 ppb measurements, both raw and denoised signals were isolated and displayed, as shown in Fig. 8. To further evaluate the improvement, composite signals were generated by averaging each data point across the five signals, highlighting the significant enhancement in SNR achieved through denoising.

Given the relatively low amplitude of the 100 ppb signals, which complicates accurate SNR calculation, the 500 ppb signals were used for this analysis. The denoising results are shown in Fig. 9. The original signals exhibited an SNR of ~ 29, corresponding to a detection limit of ~17 ppb. After denoising, the SNR increased dramatically to ~2044, with a corresponding detection limit of 245 ppt. We use the method of dividing the gas concentration by the system signal-to-noise ratio to calculate the detection limit. This represents a 70-fold enhancement in both SNR and detection limit, underscoring the effectiveness of the denoising process.

For actual trace gas sensing, calculation speed is a very important parameter. The algorithm proposed in this paper has been tested and can perform denoising within 2 s, which basically meets the requirements of

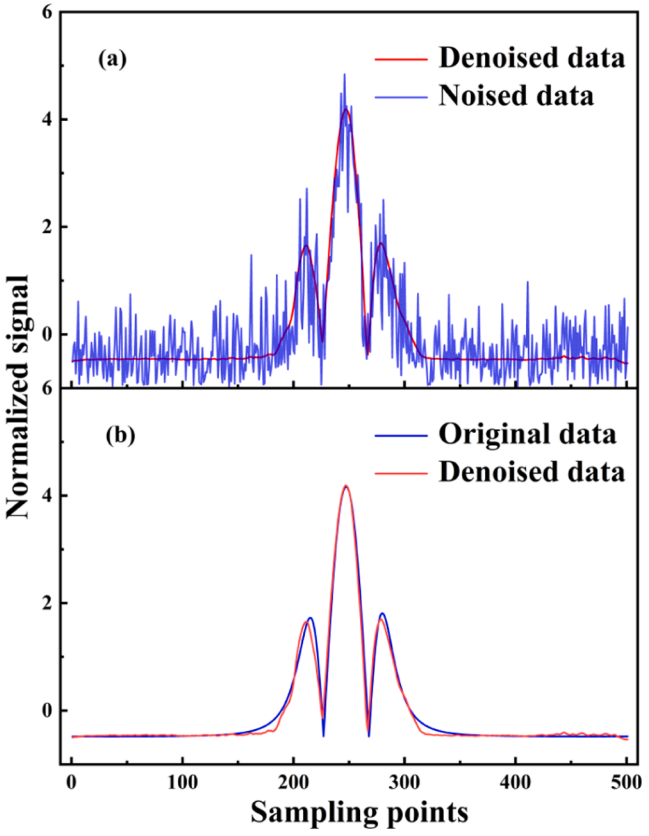


Fig. 6. (a) Comparison between noisy signals (blue) and denoised signals (red). (b) Comparison between the original smooth signal (blue) and the denoised output (red), demonstrating the model's ability to recover key features.

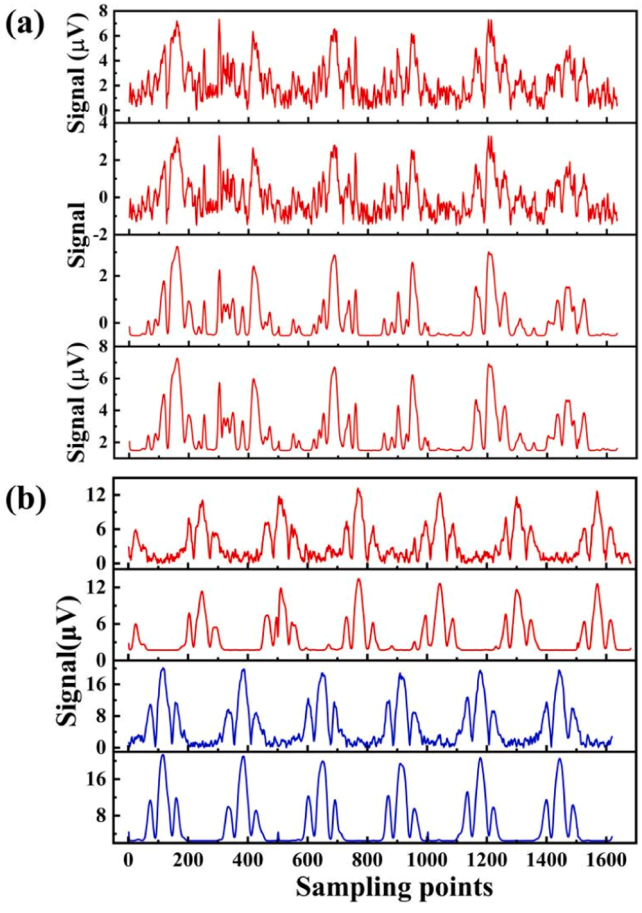


Fig. 7. Denoising results for acetylene 2 *f* signals at various concentrations using the trained CTNN model.(a) Acetylene concentration of 100 ppb. Each plot includes four signals: the original measured signal, the normalized signal, the denoised normalized signal, and the final denoised signal after inverse normalization.(b) The red line represents the signal before and after denoising of acetylene gas with a concentration of 500ppb. The blue line represents the signal before and after denoising of acetylene gas with a concentration of 1 ppm.

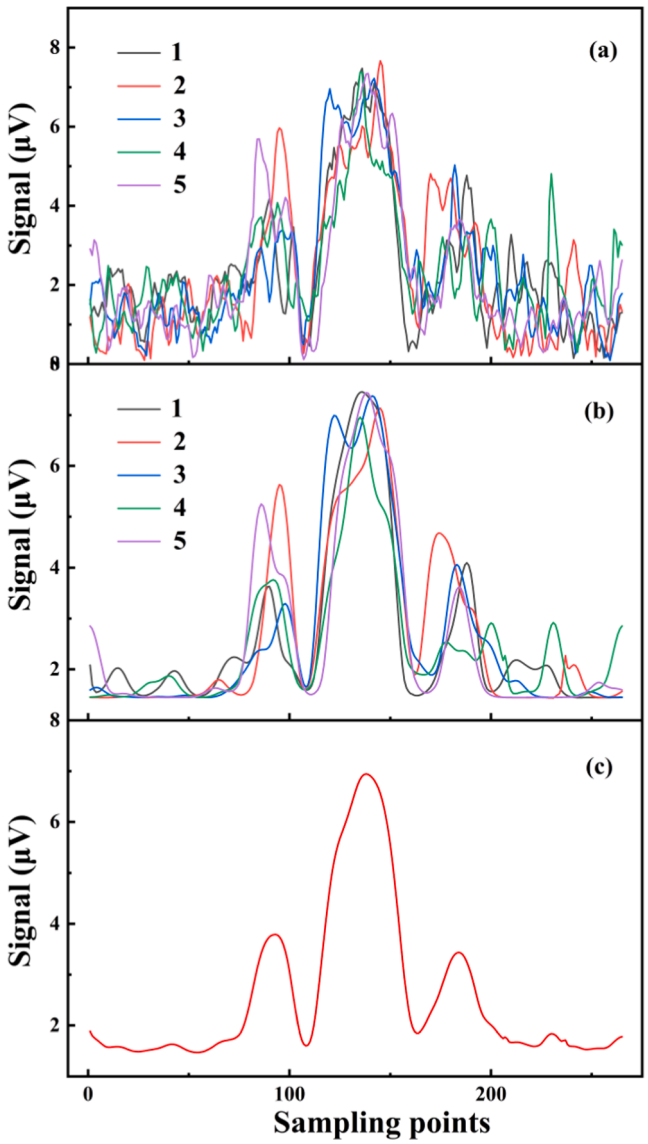


Fig. 8. Analysis of denoising performance for 100 ppb acetylene signals. (a) Five recorded 2*f* signals prior to denoising. (b) Corresponding signals after denoising. (c) Composite signal obtained by averaging the five recorded signals after denoising.

real-time detection in the field of gas detection.

The denoising results for signals ranging from 100 to 1000 ppb are presented in Fig. 10, with R-squared values calculated to evaluate the effectiveness of the CTNN model. The R-squared value increased from 99.0 % to 99.3 %, indicating an improvement in linearity allowing more accurate gas concentration measurements, particularly in challenging environments.

5. Conclusion

This study presents the development of a deep learning model that integrates 1D CNN and Transformers to effectively denoise photoacoustic spectral signals. Systematic analysis of both simulated and experimental data confirmed the model's ability to significantly reduce noise while preserving essential signal features. By incorporating various noise types during training to mimic real-world conditions, the model achieved substantial improvements in SNR.

When applied to actual acetylene gas measurements in the concentration range of 100 ppb to 2 ppm, the model demonstrated excellent at

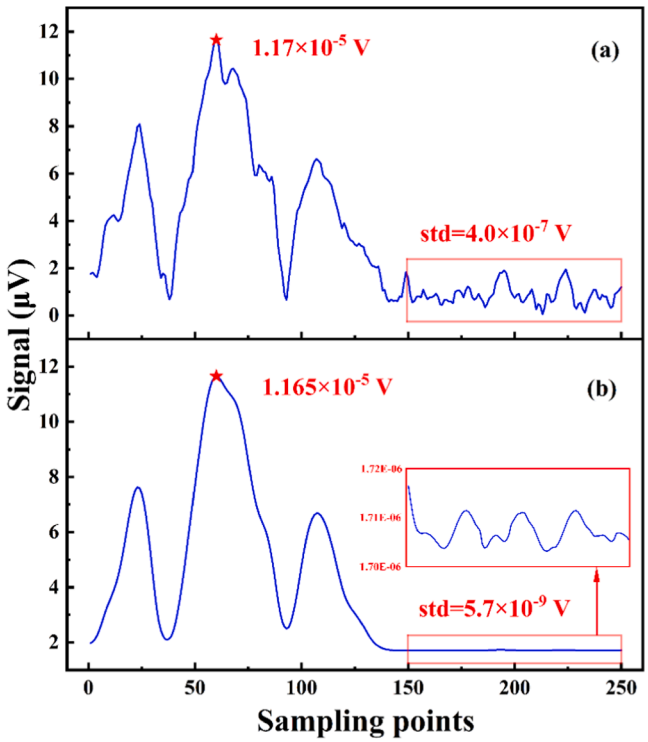


Fig. 9. Comparison of 2*f* signals for 500 ppb acetylene before and after denoising. (a) Original signal with a peak amplitude of 1.17 × 10^{-5} V and a noise standard deviation of 4.0 × 10^{-7} V. (b) Denoised signal with a preserved peak amplitude of 1.165 × 10^{-5} V and a significantly reduced noise standard deviation of 5.7 × 10^{-9} V.

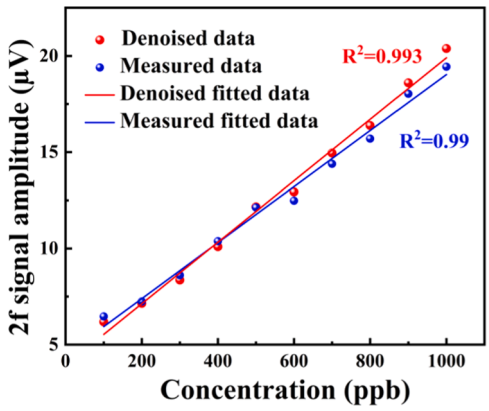


Fig. 10. Linear fitting of 2*f* signal amplitudes for acetylene concentrations ranging from 100 to 1 ppm, before and after denoising with the CTNN model.

low concentrations, enhancing SNR and improving detection sensitivity. For 500 ppb signals, the denoising process achieved a remarkable 70-fold increase in SNR, underscoring its practical potential. An improvement in R^2 values denotes an enhancement also in linearity of the signals, contributing to more accurate concentration measurements. Overall, this deep learning model represents a significant advancement in the denoising of photoacoustic spectroscopy signals, offering a reliable and effective tool for future spectral analysis in challenging environments.

CRediT authorship contribution statement

Lei Dong: Project administration, Funding acquisition, Conceptualization. **Yan Gao:** Investigation, Funding acquisition. **Ruyue Cui:**

Software, Resources, Funding acquisition. **Pietro Patimisco:** Writing – review & editing, Supervision. **Angelo Sampaolo:** Writing – review & editing. **Vincenzo Spagnolo:** Writing – review & editing. **Xukun Yin:** Software, Funding acquisition. **Chen Zhang:** Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jinhua Tian:** Formal analysis. **Yujie Tang:** Validation. **Lei Yang:** Investigation. **Chaofan Feng:** Visualization, Conceptualization. **Hongpeng Wu:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **Hanxi Zhang:** Data curation.

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.

Acknowledgment

The project is supported by National Natural Science Foundation of China (NSFC) [grant numbers 62475137, 62235010, 62105252]; Shanxi Provincial Special Fund for Scientific and Technological Cooperation and Exchange [202404041101022, 202304041101019]; The Shanxi Science Fund for Distinguished Young Scholars (20210302121003) and the National Key Research and Development Program of China (2023YFF0715700). The authors from Dipartimento Interateneo di Fisica acknowledge financial support through Project PNC 0000001 D3–4-Health—Digital Driven Diagnostics, prognostics and therapeutics for sustainable Health care (CUP: B83C22006120001), the National Plan for Complementary Investments to the NRRP, funded by the European Union– NextGenerationEU, the National Recovery and Resilience Plan (NRRP) project "BRIEF—Biorobotics Research and Innovation Engineering Facilities" (CUP: J13C22000400007), funded by the European Union—NextGenerationEU, the Italian National Recovery and Resilience Plan of NextGenerationEU MUR project PE0000023-NQSTI, project MUR – Dipartimenti di Eccellenza 2023–2027 – Quantum Sensing and Modeling for One-Health (QuaSi-ModO) and THORLABS GmbH within the PolySenSe joint research laboratory.

Data availability

The data that has been used is confidential.

References

[1] P. Patimisco, G. Scamarcio, F.K. Tittel, V. Spagnolo, Quartz-enhanced photoacoustic: a review, Sensors 14 (2014) 6165–6206.
[2] J.P. Wang, H.P. Wu, Angelo Sampaolo, et al., Quartz-enhanced multiheterodyne resonant photoacoustic spectroscopy, Light. Sci. Appl. 13 (2024) 77.
[3] S.M. Cristescu, S.T. Persijn, S.L. Hekkert, F.J.M. Harren, Laser-based systems for trace gas detection in life sciences, Appl. Phys. B 92 (2008) 343.
[4] K. Kinjalk, F. Paciolla, B. Sun, et al., 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 (2024) 021427.
[5] Raffaele De Palo, Annalisa Volpe, Pietro Patimisco, Andrea Zifarelli, Angelo Sampaolo, Antonio Ancona, Hongpeng Wu, Vincenzo Spagnolo, Femtosecond laser fabrication of black quartz for infrared photodetection applications, Light. Adv. Manuf. (2025), https://doi.org/10.37188/lam.2025.026.
[6] C.F. Feng, Marilena Giglio, B. Li, et al., Detection of hydrogen sulfide in sewer using an erbium-doped fiber amplified diode laser and a gold-plated photoacoustic cell, Molecules 27 (19) (2022) 6505.
[7] T.M. Sun, B. Feng, J.P. Huo, et al., Artificial intelligence meets flexible sensors: emerging smart flexible sensing systems driven by machine learning and artificial synapses, NanoMicro Lett. 16 (14) (2024).
[8] B. Sun, A. Zifarelli, H. Wu, S, et al., Mid-infrared quartz-enhanced photoacoustic sensor for ppb-level CO detection in SF6 gas matrix exploiting a T-grooved quartz tuning fork, Anal. Chem. 92 (2020) 13922–13929.
[9] B. Sun, P. Patimisco, A. Sampaolo, et al., Light-induced thermoelastic sensor for ppb-level H2S detection in a SF6 gas matrices exploiting a mini-multi-pass cell and quartz tuning fork photodetector, Photoacoustics 33 (2023) 100553.
[10] Hongpeng Wu, Lei Dong*, Huadan Zheng, et al., Beat frequency quartz-enhanced photoacoustic spectroscopy for fast and calibration-free continuous trace-gas monitoring, Nat. Commun. 8 (2017) 15331.
[11] Y.F. Ma, T.T. Liang, S.D. Qiao, et al., Highly sensitive and fast hydrogen detection based on light-induced thermoelastic spectroscopy, Ultra Sci. 3 (2023) 0024.
[12] R.Y. Cui, L. Dong, H.P. Wu, et al., Three-dimensional printed miniature fiber-coupled multipass cells with dense spot patterns for ppb-level methane detection using a near-IR diode laser, Anal. Chem. 92 (19) (2020) 13034–13041.
[13] P. Patimisco, A. Sampaolo, L. Dong, et al., Recent advances in quartz enhanced photoacoustic sensing, Appl. Phys. Rev. 5 (2018) 011106.
[14] C.F. Feng, B. Li, Y.J. Jing, et al., Enrichment-enhanced photoacoustic spectroscopy based on vertical graphene, Sens. Actuators B Chem. 417 (2024) 136204.
[15] L.H. Wang, H.H. Lv, Y.H. Zhao, et al., 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.
[16] H.P. Wu, X.K. Yin, L. Dong, et al., Ppb-level nitric oxide photoacoustic sensor based on a mid-IR quantum cascade laser operating at 52 ℃, Sens. Actuators B Chem. 290 (2019) 426–433.
[17] C.F. Feng, X.W. Shen, B. Li, et al., Carbon monoxide impurities in hydrogen detected with resonant photoacoustic cell using a mid-IR laser source, Photoacoustics 36 (2024) 100585.
[18] J. Ma, E.B. Fan, H.J. Liu, et al., Microscale fiber photoacoustic spectroscopy for in situ and real-time trace gas sensing, Adv. Photonics 6 (2024) 066008.
[19] S.Y. Zhang, Toby Bi, P. Del'Haye, On-the-Fly precision spectroscopy with a dual-modulated tunable diode laser and Hz-level referencing to a cavity, Adv. Photonics 6 (2024) 066008.
[20] T.G. Li, P.C. Zhao, P. Wang, et al., Miniature optical fiber photoacoustic spectroscopy gas sensor based on a 3D micro-printed planar-spiral spring optomechanical resonator, Photoacoustics 40 (2024) 100657.
[21] X.Y. Zhao, C.X. Li, H.C. Qi, et al., Integrated near-infrared fiber-optic photoacoustic sensing demodulator for ultra-high sensitivity gas detection, Photoacoustics 33 (2023) 100560.
[22] Andrea Zifarelli, Giuseppe Negro, Lavinia A. Mongelli, et al., Effect of gas turbulence in quartz-enhanced photoacoustic spectroscopy: a comprehensive flow field analysis, Photoacoustics 38 (2024) 100625.
[23] C. Zhang, Y. He, S.D. Qiao, et al., High-sensitivity trace gas detection based on differential helmholtz photoacoustic cell with dense spot pattern, Photoacoustics 38 (2024) 100634.
[24] C. Zuo, J.M. Qian, S.J. Feng, et al., Deep learning in optical metrology: a review, Light. Sci. Appl. 11 (1) (2022) 39.
[25] X.F. Pan, Z. Zhang, H. Zhang, et al., 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.
[26] J.X. Zhu, S.L. Ji, Z.H. Ren, et al., Triboelectric-induced ion mobility for artificial intelligence-enhanced mid-infrared gas spectroscopy, Nat. Commun. 14 (2023) 2524.
[27] G.Y. Guan, A.Q. Liu, X.Y. Wu, et al., Near-infrared off-axis cavity-enhanced optical frequency comb spectroscopy for CO2/CO dual-gas detection assisted by machine learning, ACS Sens. 9 (2024) 820–829.
[28] X.N. Liu, S.D. Qiao, G.W. Han, et al., Highly sensitive hf detection based on absorption enhanced light-induced thermoelastic spectroscopy with a quartz tuning for of receive and shallow neural network fitting, Photoacoustics 28 (2022) 100422.
[29] Y.T. Yang, J.C. Jiang, J.F. Zeng, et al., CH4, C2H6, and CO2 multi-gas sensing based on portable mid-infrared spectroscopy and pca-bp algorithm, Sensors 23 (3) (2023) 1413.
[30] L.B. Tian, J.B. Xia, A.A. Kolomenskii, et al., Gas phase multicomponent detection and analysis combining broadband dual-frequency comb absorption spectroscopy and deep learning, Commun. Eng. 2 (2023) 54.
[31] M. Arshad, Zahangir Chowdhury, Timothy E. Rice, Matthew A. Oehlschlaeger, TSMC-Net: deep-learning multigas classification using thz absorption spectra, ACS Sens. 8 (2023) 1230–1240.
[32] X.R. Meng, H.Q. Chang, X.Q. Wang, Methane concentration prediction method based on deep learning and classical time series analysis, Energies 25 (2022) 2262.
[33] P. Zhao, D. Ding, K.T. Li, et al., A tunable diode laser absorption spectroscopy (tdlas) signal denoising method based on LSTM-DAE, Opt. Commun. 567 (2024) 130327.
[34] Q.S. Wen, T. Zhou, C.L. Zhang, et al., Transformers in time series: A survey. in 32th International Joint Conference on Artificial Intelligence (IJCAI) 2023, pp. 6778-6786.
[35] L.B. Tian, J.C. Sun, J. Chang, et al., Retrieval of gas concentrations in optical spectroscopy with deep learning, Measurement 182 (2021) 109739.
[36] Y.M. Zhou, Y.L. Heng, J.T. Zhu, et al., Enhanced gas recognition of electronic nose using 1-D convolutional neural network with savitzky-golay filter, IEEE Sens. J. 24 (7) (2024) 10769.
[37] David Tomeček, Henrik Klein Moberg, Sara Nilsson, et al., Neural network enabled nanoplasmonic hydrogen sensors with 100 ppm limit of detection in humid air, Nat. Commun. 15 (2024) 1208.
[38] C. Liu, G. Wang, C. Zhang, et al., End-to-end methane gas detection algorithm based on transformer and multi-layer perceptron, Opt. Express 32 (1) (2024) 987–1002.

[39] Y.C. Wang, Y. Jiang, H. Hao, et al., Pressure calibration- and profile fitting-free spectroscopy technology based on deep neural network for gas sensing, Measurement 204 (2022) 112077.
[40] S. Zhou, C.Y. Shen, L. Zhang, et al., Dual-optimized adaptive kalman filtering algorithm based on bp neural network and variance compensation for laser absorption spectroscopy, Opt. Express 27 (2019) 31874–31888.
[41] R.Q. Xu, L.B. Tian, J.B. Xia, et al., Leveraging deep learning for optimal methane gas detection: residual network filter assisted direct absorption spectroscopy, Sens. Actuators A. Phys. 369 (2024) 115195.
[42] Z.W. Liu, C.T. Zheng, T.Y. Zhang, et al., High-precision methane isotopic abundance analysis using near-infrared absorption spectroscopy at 100 torr, Analyst 146 (2) (2021) 698–705.
[43] Y.C. Kim, H.G. Yu, J.H. Lee, et al., Hazardous gas detection for FTIR-based hyperspectral imaging system using DNN and CNN. Proc. SPIE 2017, 10433, 1043317.
[44] E. D'Andrea, S. Pagnotta, E. grifonin, et al., An artificial neural network approach to laser-induced breakdown spectroscopy quantitative analysis, Spectrochim. Acta Part B At. Spectrosc. 99 (2014) 52–58.
[45] H.Y. Sun, S.D. Qiao, Y. He, et al., Parts-per-quadrillion level gas molecule detection: CO-LITES sensing, Light Sci. Appl. 14 (2025) 180.
[46] Y.H. Liu, S.D. Qiao, C. Fang, et al., A highly sensitive LITES sensor based on a multi-pass cell with dense spot pattern and a novel quartz tuning fork with low frequency, OptoElectron. Adv. 7 (2024) 230230.
[47] R.Q. Wang, X.Y. Guan, S.D. Qiao, et al., Ultrahigh sensitive LITES sensor based on a trilayer ultrathin perfect absorber coated T-head quartz tuning fork, Laser Photonics Rev. 19 (2025) 2402107.
[48] J. Nalepa, M. Myller, M. Kawulok, Transfer learning for segmenting dimensionally reduced hyperspectral images, IEEE Geosci. Remote Sens. Lett. 17 (7) (2020) 1228–1232.
[49] A. Vaswani, N. Shazeer, N. Parmar, et al., Attention is all you need. In: Proceedings of the 31th Advances in Neural Information Processing Systems (NeurIPS) 2017, pp. 5998-6008.
[50] Xavier Glorot, Antoine Bordes, Yoshua Bengio, Deep sparse rectifier neural networks. In: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics 2011, 15:315-323.

Zhang Chen is currently pursuing a Master's degree in Optoelectronic Information Engineering at the Institute of Laser Spectroscopy, Shanxi University, China. His research focuses on photoacoustic spectroscopy for gas detection and the integration of deep learning techniques with optical sensing methodologies.

Yan Gao received his Ph.D. degree in optics from Shanxi University, China in 2015. He is an associate professor at Institute of intelligent Optoelectronics at Shanxi Datong University. His research interests encompass ultrafast optics, optical sensors and laser spectroscope techniques.

Ruyue Cui received her dual Ph.D. degrees in physics from Shanxi University, China, and Université du Littoral Côte d'Opale, France, in 2023. Currently, she is a lecturer at the Institute of Laser Spectroscopy at Shanxi University. Her research interests encompass optical sensors and laser spectroscopy techniques.

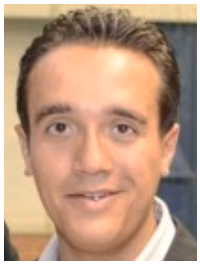
Hanxi Zhang is pursuing her bachelor's degree in Optoelectronic Information Science and Engineering at Shanxi University. Currently, she is a sophomore at Shanxi University.

Jinhua Tian is currently pursuing an engineering degree in Electronic Information Science and Technology at Shanxi University. His research interests cover optical sensors and laser spectroscopy techniques.

YuJie Tang Currently, he is pursuing a bachelor's degree in Electronic and Information Engineering at Shanxi University. His research interests encompass optical sensors and laser spectroscopy techniques.

Lei Yang received his Master degree in detection technology and automation device from Northwestern Polytechnical University, China, in 2015. Currently, he is a lecturer at the College of Physical and Electronic Sciences at Shanxi Datong University. His research interests include machine learning algorithms, fault diagnosis and other artificial intelligence fields.

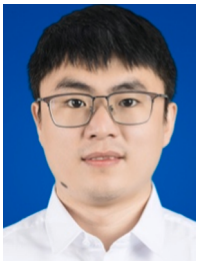
Chaofan Feng is now pursuing a Ph.D. degree in atomic and molecular physics in the Institute of Laser Spectroscopy of Shanxi University, China. His research interests include gas sensors, photoacoustic spectroscopy, and laser spectroscopy techniques.



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.



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 assistant 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 Dong received his Ph.D. degree in optics from Shanxi University, China, in 2007. From June, 2008 to December, 2011, he worked as a post-doctoral fellow 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 activities research activities are focused on research and development in laser spectroscopy, in particular photoacoustic spectroscopy applied to sensitive, selective and real-time trace gas detection, and laser applications in environmental monitoring, chemical analysis, industrial process control, and medical diagnostics. He has published more than 100 peer reviewed papers with> 2200 positive citations.

Hongpeng Wu received his Ph.D. degree in atomic and molecular physics from Shanxi university, China, in 2017. From 2015–2016, he studied as a joint Ph.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 optical sensors and laser spectroscopy techniques.