


 Cite this: *RSC Adv.*, 2026, **16**, 27666

Deep learning-assisted SERS for detection of propoxate and isopropoxate in E-cigarettes

 Jiahao Teng,^a Shenggang Huang,^a Wenkai Zheng,^a Xuqing Wang,^b Yingsheng He,^c Jiye Wang^{*a} and Yazhou Qin^{*a}

The illicit use of new psychoactive substances in e-cigarettes has posed severe threats to human health and social security, urgently necessitating the development of targeted rapid and highly sensitive detection methods. In this study, we developed a highly sensitive detection approach for propoxate and isopropoxate, commonly illegally added drugs in e-cigarettes, by integrating deep learning-assisted SERS technology. First, the characteristic spectral peaks of the two isomeric compounds were identified through conventional Raman and SERS analysis of reference standards. Furthermore, DFT calculations were employed to interpret the vibrational modes in the Raman spectra corresponding to their molecular structures. Subsequently, a sample pre-treatment method was developed for spiked e-cigarette samples, enabling trace-level SERS detection of both substances. Finally, an innovative dual-branch deep learning network integrating time-domain and frequency-domain features was developed for high-precision classification and identification of two structurally similar substances, achieving an identification accuracy of 99.73%. This study provides a reference for the detection of structurally similar compounds.

Received 23rd January 2026

Accepted 18th May 2026

DOI: 10.1039/d6ra00614k

rsc.li/rsc-advances

1 Introduction

E-cigarettes have gained widespread popularity due to their diverse flavors and user-friendly features.^{1–3} However, criminals have been adding various prohibited substances to e-cigarettes for profit, evolving from early nicotine to new psychoactive substances such as propoxate and isopropoxate.^{4,5} Nicotine is highly addictive, and whether actively inhaled or passively exposed, it stimulates dopamine release, leading to central nervous system excitation.^{6,7} Excessive nicotine intake may cause severe health issues such as cardiovascular and pulmonary diseases.⁸ Propoxate and isopropoxate can interfere with the endocrine system by binding to estrogen receptors, impair fertility, and exposure to these substances affects neurodevelopment, posing potential threats to human health.^{9,10} Currently, China has classified propoxate and isopropoxate as Schedule II controlled psychoactive substances. However, due to their highly similar molecular structures with only differences in side chain substituents (isopropyl vs. propyl groups), and their low concentrations in e-cigarettes, existing detection

technologies struggle to achieve accurate identification. There is an urgent need to develop targeted rapid detection methods for these compounds.

Currently, the detection methods for illegal additives in e-cigarettes primarily include traditional techniques such as gas chromatography (GC), high-performance liquid chromatography (HPLC), and nuclear magnetic resonance spectroscopy (NMR).¹¹ Additionally, gas chromatography-mass spectrometry (GC-MS),^{12,13} liquid chromatography-mass spectrometry (LC-MS),^{13,14} and immunoassay methods are also widely employed. For example, Zhang *et al.* employed probe electrospray ionization quadrupole time-of-flight mass spectrometry (PESI-QTOF-MS) to achieve rapid identification and differentiation of etomidate and its structural analogs in e-liquids within 0.3 minutes, with LODs of 20 ng mL⁻¹ and a linear range of 50–5000 ng mL⁻¹. The method was successfully applied to the analysis of 38 positive samples and impurities.¹⁵ The team led by Klaudia Adels proposed an approach based on 80 MHz low-field nuclear magnetic resonance (LF-NMR), enabling simultaneous determination of nicotine and other e-cigarette components in various products without sample pre-treatment. In 37 real samples, they detected eight organic acids with concentrations as high as 56 mg mL⁻¹.⁴ The Robertson team employed a combined approach of gas chromatography, liquid chromatography, and mass spectrometry to study 25 different flavoured e-cigarettes from four popular brands, identifying six additives with concentrations exceeding known toxicity thresholds in multiple samples.¹³ While such methods enable highly sensitive

^aKey Laboratory of Drug Prevention and Control Technology of Zhejiang Province, Zhejiang Police College, 555 Binwen Road, Binjiang District, Hangzhou 310053, Zhejiang Province, P. R. China. E-mail: yazhouqin@zju.edu.cn

^bSchool of Pharmacy, Hangzhou Normal University, Hangzhou 311121, Zhejiang, China

^cKey Laboratory of Drug Prevention and Control Technology of Zhejiang Province, National Narcotics Laboratory Zhejiang Regional Center, 555 Binwen Road, Binjiang District, Hangzhou, 310053, Zhejiang Province, PR China



detection, they require complex pre-treatment steps such as solid-phase extraction and derivatization, involve lengthy single-test durations, and rely on large-scale equipment and specialized operators, making them unsuitable for on-site rapid screening needs.

Among various detection methods, surface-enhanced Raman scattering (SERS) technology has garnered widespread attention from researchers due to its advantages of being rapid, non-destructive, and highly sensitive. The core principle of SERS technology lies in the localized surface plasmon resonance effect generated by metal nanoparticles under light field excitation, which enhances the electric field intensity on the nanoparticle surface by millions of times, thereby achieving exponential amplification of molecular signals adsorbed on the particle surface.¹⁶ For instance, in biomedical detection, the Cardellini team developed a novel SERS probe called LipoGold, which achieved signal enhancement by several orders of magnitude and maintained stability for several weeks, enabling precise differentiation between cells from healthy donors and those from patients with GM1 gangliosidosis.¹⁷ In environmental science, Kim *et al.* prepared Au/Ag nano-coral SERS substrates through electrochemical deposition and silver etching techniques, achieving ultra-high sensitivity detection of the pesticide thiram. The detection limits reached 347.8 fM, 473.5 fM, 529.1 fM, and 549.3 fM in soil, streams, tap water, and drinking water, respectively.¹⁸ Additionally, in the field of food testing, Li *et al.* developed a one-step rapid (<1 minute per substrate, 100 samples per hour) SERS substrate fabrication method based on flame nanoparticle synthesis and deposition, which was successfully applied for label-free, rapid on-site detection of pesticide residues in fresh orange juice samples.¹⁹ Although SERS technology has demonstrated great potential in the aforementioned fields, traditional analysis methods based on original spectral peaks face challenges when dealing with isomers with highly similar chemical structures and overlapping spectral fingerprints. The subtle differences in local vibrations are concealed within the highly overlapping spectral background, resulting in severely insufficient feature distinguishability in SERS spectra. This makes it difficult for the classification accuracy of isomeric substances to meet the practical regulatory requirements for precise differentiation.

The rapid development of deep learning technology has provided possibilities to overcome the limitations of traditional spectral analysis, which relies heavily on manual expertise and struggles with complex overlapping spectra. Deep learning algorithms can automatically extract high-dimensional features from massive datasets and establish complex nonlinear mappings, significantly improving the automation level and accuracy of analysis.^{16,20–23} Among various deep learning architectures, the Transformer architecture has demonstrated outstanding performance in spectral classification tasks due to its pure attention-based mechanism that effectively captures long-range dependencies in sequences. It has achieved success in multiple SERS application scenarios such as biological diagnostics and food safety.^{24–28} For instance, Mao *et al.* achieved precise control over chiral gold nanostar structures to fabricate high-performance SERS substrates (with an EF of 5.49

$\times 10^6$). By integrating a Transformer model, they attained a diagnostic accuracy of 99.94% for urine samples from healthy individuals and patients with acute/non-acute interstitial nephritis.²⁶ Akshata *et al.* developed SERSFormer-2.0, a Transformer-based multi-task learning model, which combined with Au@Ag core-shell nanoparticles, enabled simultaneous identification (F1 score 0.992) and precise quantification (accuracy 0.999) of mixed pesticides in agricultural products.²⁷

Traditional single time-domain analysis methods often struggle to comprehensively capture the complete characteristics of molecular vibration modes. In contrast, the time-frequency dual-domain fusion mechanism achieves complementary feature extraction by jointly analysing signal evolution across both temporal and frequency dimensions, thereby enabling more refined extraction of discriminative information obscured by noise.^{29–34} For instance, the Ordered Twin-Branch Network with Unimodal Binomial Distribution (OTBN-UBD) model proposed by Jin *et al.* demonstrates outstanding performance in multi-classification tasks. This model features adjustable weights to mitigate extreme misclassification errors, with its effectiveness validated using real-world fault case data collected from operational industrial rotating machinery.³² In response to the current situation where existing fault diagnosis models have large network scales and parameter redundancy, making it difficult to meet platform real-time requirements, Duan *et al.* proposed a lightweight time-frequency feature fusion model. This model demonstrates excellent diagnostic accuracy and robustness across multiple datasets, achieving a diagnostic accuracy of over 99.8% with only 10K model parameters and an inference time of 68 ms.³³ The dual-branch time-frequency feature fusion network (DTFFNet) proposed by Zhu's team can accurately predict the remaining useful life (RUL) of mechanical equipment. Experimental analysis on widely used turbofan engine datasets and the PHM2012 bearing dataset demonstrates that DTFFNet not only significantly improves prediction accuracy compared to mainstream RUL prediction methods, but also exhibits stronger robustness in noisy environments.³⁴ However, the application of time-frequency feature fusion techniques to SERS spectroscopy, particularly for the simultaneous precise detection of isomeric compounds (propoxate and isopropoxate) in the complex matrix of e-cigarettes, has not been reported yet.

This study innovatively proposes a dual-branch interpretable classification network-based detection strategy with time-frequency feature fusion, aiming to achieve highly sensitive and accurate classification and identification of propoxate and isopropoxate in e-cigarettes. Firstly, Raman and SERS detection were performed on the standard samples of propoxate and isopropoxate isomers, and their characteristic vibration modes were assigned through DFT theoretical calculations. Secondly, a pre-treatment method was developed for the spiked samples of propoxate and isopropoxate in e-cigarettes, enabling their highly sensitive SERS detection. Finally, for the labelled SERS sample data, a dual-branch classification network with time-frequency feature fusion was developed, achieving a verification accuracy of 99.73% for two prohibited additives. This



provides reliable technical support for rapid screening and precise regulation of prohibited additives in e-cigarettes.

2 Materials and methods

2.1 Materials and instruments

The reference standards propoxate and isopropoxate (both with purity of 99.9%) used in this study were provided by the National Narcotics Laboratory. The main chemical reagents, including sodium chloride (AR, $\geq 99.5\%$), ethyl acetate (AR, $\geq 99.5\%$), sodium hydroxide (ACS, $\geq 98\%$), and hydrochloric acid (37%), were purchased from Aladdin Reagent Co., Ltd. Commercially available e-cigarette samples were obtained through commercial channels. Ultrapure water (resistivity 18.2 M Ω cm) prepared by the MilliQLab water purification system (Merck Millipore) was used as the experimental water. To avoid potential contamination, all glassware was sequentially soaked in aqua regia (a 3 : 1 mixture of hydrochloric acid and nitric acid) for 30 minutes, thoroughly rinsed with ultrapure water, and then air-dried in a clean environment for subsequent use. Gold nanoparticles (AuNPs) were prepared as an enhanced substrate using the sodium citrate reduction method. A 100 mL 0.01% HAuCl₄ solution was heated to vigorous boiling, followed by rapid injection of 1 mL 1% sodium citrate solution. The mixture was continuously heated under reflux for 1 hour before stopping the heating and allowing it to cool naturally to room temperature, resulting in the formation of gold nanoparticles. The obtained AuNPs were stored in the dark at 4 °C.

Characterization and spectral acquisition of AuNPs: the morphology and size of AuNPs were characterized by transmission electron microscopy (TEM, Hitachi HT7700, 100 kV) and dynamic light scattering (DLS, Malvern Zetasizer Nano ZS90), respectively. Surface-enhanced Raman scattering (SERS) spectra were collected using a Raman spectrometer (model PERS-RZ1710) from Push Nano Technology Co., Ltd, equipped with a 785 nm tunable power laser. The Raman tests were performed using a Thermo Fisher DXR2 xi confocal Raman microscope, equipped with multi-wavelength laser sources (455 nm, 532 nm, 633 nm, 785 nm) and various magnification objectives (10 \times , 20 \times , 50 \times , 100 \times).

2.2 SERS detection of standard substances

First, we used methanol as the base to prepare standard solutions of propoxate and isopropoxate with a concentration of 10 mg mL⁻¹. Subsequently, gradient dilutions were performed to obtain standard solutions with concentrations of 10 mg mL⁻¹, 1 mg mL⁻¹, 0.1 mg mL⁻¹, 10 μ g mL⁻¹, 1 μ g mL⁻¹, 0.1 μ g mL⁻¹, 10 ng mL⁻¹, and 1 ng mL⁻¹, totalling eight concentration levels. We then conducted SERS testing on the propoxate and isopropoxate standard solutions. First, 70 μ L of the standard solution was added to a test tube, followed by the addition of 70 μ L of gold nanoparticle solution and 70 μ L of 2 M sodium chloride solution, respectively. After vortex oscillation mixing for 1 minute, set the collection times to 3 and acquisition time to 3 seconds for SERS testing. A total of 455 data sets for propoxate and 482 data sets for isopropoxate were obtained,

resulting in 937 data points collected. The theoretical Raman spectra of propoxate and isopropoxate were calculated using density functional theory (DFT) with the basis set B3LYP/6-311++G(d, p).

2.3 SERS detection of spiked samples

First, take 100 μ L of e-liquid, dilute it 100-fold, and transfer it into a centrifuge tube. Then, add 100 μ L of propoxate standard solutions with different concentrations (gradient concentrations: 10 mg mL⁻¹, 5 mg mL⁻¹, 1 mg mL⁻¹, 0.5 mg mL⁻¹, 0.1 mg mL⁻¹, 0.05 mg mL⁻¹, 0.01 mg mL⁻¹, 5 μ g mL⁻¹, 1 μ g mL⁻¹) respectively to prepare spiked samples of propoxate at various concentrations in the presence of nicotine. Next, add 800 μ L of pure water, 1 mL of ethyl acetate, and 0.5 g of sodium chloride to the centrifuge tube to form a biphasic extraction system. Vortex for 3 minutes, let stand for 1 minute, transfer the upper organic phase to a new centrifuge tube, and centrifuge at 5000 rpm for 1 minute. After standing for another 1 minute, transfer 600 μ L of the supernatant to a new centrifuge tube, add 600 μ L of 1 mol per L HCl solution, vortex mix for 3 minutes, and let stand for phase separation. Subsequently, take 400 μ L of the upper organic phase, add 400 μ L of 1 mol per L NaOH solution, vortex mix for 3 minutes, and let stand for 1 minute. Finally, take 70 μ L of the upper organic phase for SERS detection, obtaining a total of 473 sets of real propoxate sample data. Following the same procedure, 493 sets of real isopropoxate sample data were obtained.

2.4 Data processing and analysis

To reduce the interference of testing environments on spectral data, we implemented a standardized pre-processing pipeline. First, all raw spectra were truncated to the 500–1800 cm⁻¹ range. Next, baseline drift was eliminated using asymmetric least squares smoothing with polynomial order 5, maximum iterations of 300, and tolerance of 0.001. A Savitzky–Golay filter was then applied with polynomial order 2 and window length of 5 for data smoothing and denoising. All spectral intensities were normalized to the [0, 1] range. Finally, each spectrum was resampled into a standardized sequence of 2048 points through linear interpolation to ensure consistent dimensionality for subsequent model input.

2.5 Classification network construction

By combining short-time Fourier transform, convolutional neural networks, and attention mechanisms, a dual-branch SERS spectral classification model was constructed. As shown in Fig. 1, this network architecture primarily consists of five core modules: the time-frequency feature transformation module, CNN frequency-domain feature extraction branch module, Transformer time-domain feature extraction branch module, gated fusion module, and multi-granularity classifier module.

Time frequency feature transformation module: in the data pre-processing stage, the original time series is converted into a 2D time-frequency spectrogram through Short-Time Fourier Transform (STFT) with a window length of 128 and an overlap length of 32. The amplitude values undergo logarithmic



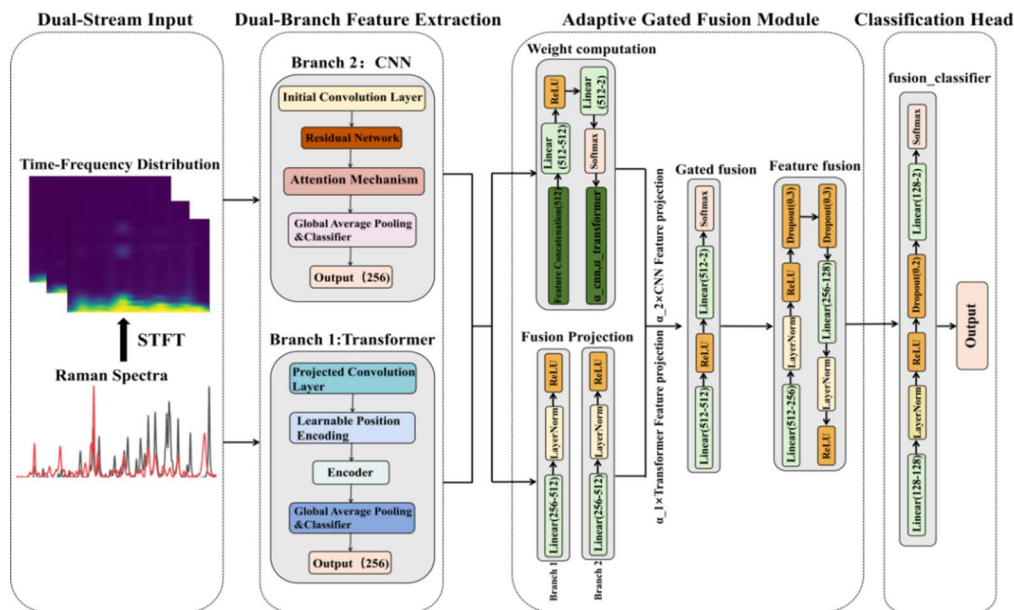


Fig. 1 Architecture of the dual-branch classification network model.

transformation to enhance feature contrast, followed by min–max normalization. For time-frequency spectrograms with insufficient dimensions, bilinear interpolation is applied to resize them to a fixed dimension (64×64). The final output is a standardized single-channel time-frequency image, providing time-frequency feature information for subsequent CNN branch processing.

CNN frequency domain feature extraction branch: for local feature extraction of time-frequency images, a four-level progressive feature extraction architecture is adopted. The initial convolutional layer employs a 7×7 kernel with stride-2 down sampling to effectively expand the receptive field. This is followed by three residual module groups with progressively increasing channel numbers, where skip connections ensure effective gradient propagation. Additionally, at the high-level feature stage, the model incorporates a dual attention mechanism. Channel attention learns the importance weights of feature channels through global average pooling and fully connected layers, while spatial attention learns spatial region saliency by concatenating features from max pooling and average pooling. Finally, global average pooling compresses the 2D feature maps into fixed-dimensional feature vectors, forming the feature representation of the frequency domain branch.

Transformer temporal feature extraction branch: to fully exploit the global dependencies in the original time series and complement the CNN branch, a temporal feature extraction branch based on the attention mechanism was constructed. First, at the input stage, the original sequence is mapped to a high-dimensional feature space through a convolutional layer. Then, a learnable parameter matrix is employed for positional encoding, generating unique encoding vectors for each sequence position before entering the encoder. The encoder consists of multiple stacked layers, each containing a spectral-specific attention mechanism and a feedforward neural network. The attention mechanism incorporates relative

positional bias to effectively model the relative positional relationships among sequence elements, while the feedforward network utilizes a GELU activation function combined with a two-layer linear transformation with an expansion rate, further enhancing the capability to mine deep-level relationships. Finally, temporal feature representations are generated through layer normalization and global pooling.

Gated adaptive fusion module: to achieve complementary advantages of time-frequency domain features, the model incorporates a gated mechanism-based adaptive fusion module that receives CNN frequency-domain features and Transformer time-domain features. First, both features are mapped to a unified fusion space through independent linear projection layers. The gating weight network adopts a fully connected architecture, generating normalized weight coefficients for both branches *via* nonlinear transformation of concatenated features, enabling adaptive learning of feature importance for input samples. The fused features are further refined through a compression network, employing layer normalization, activation functions, and Dropout regularization techniques to ensure discriminative power and generalization capability of the fused features.

Multi-granularity classification decision module: the classification module adopts a hierarchical design to provide classification outputs for feature representations at different levels. The fusion classifier receives fused features and gradually compresses them to the target number of categories through fully connected layers, with regularization embedded in between. Each branch classifier processes its respective feature vector separately, employing a similar compression architecture.

3 Results and discussion

3.1 Reference standard spectral analysis

The size and morphology of the prepared gold nanoparticles were characterized by transmission electron microscopy (TEM),



and the results are shown in Fig. S1. TEM images revealed that the synthesized gold nanoparticles were spherical in shape with a relatively uniform size distribution, primarily ranging from 60 to 100 nm. Studies have demonstrated that the size and density of gold nanoparticles significantly influence the detection performance of SERS.^{35–38} Therefore, we employed a sodium chloride solution to enrich the gold nanoparticles. According to our previous research,³⁹ NaCl can induce the aggregation of gold nanoparticles through electrostatic interactions, thereby achieving signal enhancement effects.

To quantitatively evaluate the SERS activity of the as-prepared substrate, as shown in Table S1, we calculated the enhancement factor (EF) using isopropoxate as model molecule (molecular formula: $C_{15}H_{18}N_2O_2$, molecular weight: 258.32 g mol⁻¹). The average Raman intensity of 50 solid-state powder spectra at 1028 cm⁻¹ was 267.15. Based on a solid density of approximately 1.1 g cm⁻³, the effective molar concentration of the powder was calculated to be $C_{RS} = 4.26$ mol L⁻¹. Under identical instrumental conditions, the average SERS intensity of 50 spectra from a 1 μg mL⁻¹ isopropoxate solution at the same characteristic peak was 35 431.64, and the corresponding molar concentration was $C_{SERS} = 3.87 \times 10^{-6}$ mol L⁻¹. Substituting these values into the formula $EF = I_{SERS} \times C_{RS} / I_{RS} \times C_{SERS}$, the calculated EF is 1.46×10^8 . This result demonstrates that the substrate exhibits excellent SERS enhancement capability.

Fig. 2 displays the spectral test results of propoxate and isopropoxate standards. Fig. 2A and A1 show the optimized molecular structures of propoxate and isopropoxate, respectively. Fig. 2B and B1 present the Raman spectra (black lines), SERS spectra (red lines), and DFT calculation results (yellow lines) of the two standard substances. It can be observed that the similar molecular structures of the two compounds result in numerous identical characteristic peaks. Fig. 2C and C1 illustrate the SERS detection results of the two substances at different concentrations, with a detection limit reaching 1 ng mL⁻¹. To verify the method's stability, SERS detection was performed at 20 randomly selected spots, and the results are shown in Fig. 2D and D1. The relative standard deviations of propoxate at 1398 cm⁻¹ and 1720 cm⁻¹ were 3.06% and 2.43%, respectively, while those of isopropoxate at 616 cm⁻¹ and 1720 cm⁻¹ were 4.59% and 3.55% (Fig. 2E and E1), indicating the good stability of this method. The corresponding characteristic peak intensity projection diagrams (Fig. 2F and F1) further demonstrated the excellent stability and reproducibility of this method.

3.2 SERS detection and analysis of real samples

Since the main components of e-liquid are organic compounds such as glycerol, which has high viscosity and is unfavourable for detection, commercially available e-liquid was first diluted with pure water. Subsequently, standard solutions of the analytes at different concentrations were added to prepare spiked samples. During the sample pre-treatment process, an equal volume of ethyl acetate was first added and thoroughly shaken for extraction to achieve the separation of the aqueous and organic phases. To further eliminate interference from impurities such as nicotine in the organic phase, equal volumes of

hydrochloric acid solution and sodium hydroxide solution were sequentially added to the organic phase for purification. Finally, SERS detection of the target substances was performed on the organic phase.

To analyse the mechanism of the pre-treatment process, we conducted SERS monitoring at each step. First, we examined whether the extraction step could effectively eliminate background interference signals under different PG/VG ratios. As shown in Fig. S2, the black curve represents the SERS spectrum of the blank e-liquid, the red curve represents the SERS spectrum of the organic phase after extraction, and the gray curve represents the SERS spectrum of the spiked sample after pre-treatment. The results indicate that the blank e-liquid still exhibits certain interference signals near the target characteristic peaks (black line). If these interferences are not effectively eliminated, they will significantly impact the specificity and accuracy of detection. It is noteworthy that after the liquid-liquid extraction process established in this study, the relevant interference signals in the organic phase blank were essentially eliminated (red line), thereby ensuring the reliability of the test results.

In Fig. 3A, the planar structural features of propoxate and isopropoxate are labelled as the main structure (cyan) and substituents (orange). To further achieve the assignment of characteristic peak vibration modes, DFT calculations were employed to compute the Raman spectra (black), which were then compared and analysed with the SERS spectra of the labelled samples (red curves). For propoxate (Fig. 3B), the SERS spectra showed high consistency with the DFT-calculated spectra at wavenumbers of 616 cm⁻¹, 1002 cm⁻¹, and 1196 cm⁻¹.

The peak at 616 cm⁻¹ corresponds to the stretching vibration of the C=C bond in the imidazole ring skeleton, coupled with the out-of-plane bending vibration of the C-H bond on the imidazole ring. This characteristic peak is one of the signature signals indicating the presence of the imidazole ring, reflecting the conjugated vibration characteristics of the heterocyclic skeleton. The peak at 783 cm⁻¹ is attributed to the out-of-plane bending vibration of the three C-H bonds on the benzene ring. The peak at 910 cm⁻¹ arises from the combined contributions of two types of vibrations: one is the out-of-plane bending vibration of the three C-H bonds on the benzene ring, and the other is the out-of-plane bending vibration of one C-H bond in the propyl side chain. Additionally, the strong peak at 1002 cm⁻¹ is characteristically assigned to the stretching vibration of the C-O bond connecting the imidazole ring and the propyl group. The peak at 1028 cm⁻¹ corresponds to the breathing vibration of the benzene ring. The peak at 1196 cm⁻¹ is a superposition of multiple vibrational modes, primarily including the stretching vibration of the C-O bond, the stretching vibration of the C-N bond linking the imidazole ring and the benzene ring, and the in-plane bending vibration of the two C-H bonds on the benzene ring. The peak at 1398 cm⁻¹ arises from the stretching vibration of the C=C bonds in the imidazole ring skeleton, coupled with the out-of-plane bending vibration of the four C-H bonds connecting the imidazole ring and benzene ring regions. Finally, the strong peak at 1720 cm⁻¹



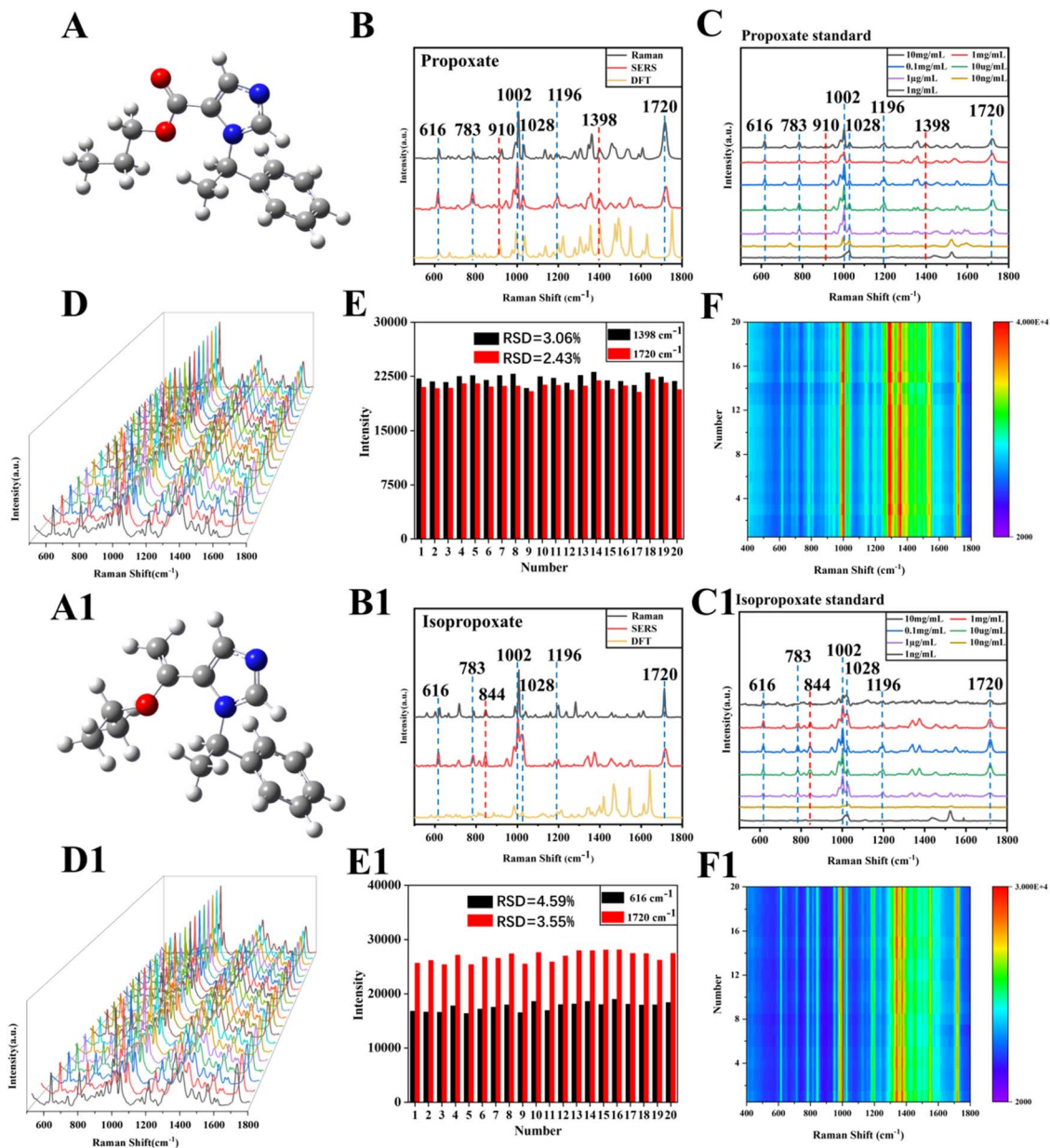


Fig. 2 (A and A1) The optimized molecular structures of propoxate and isopropoxate. (B and B1) Raman (black), SERS (red), and DFT calculated spectra (yellow) of propoxate and isopropoxate. (C and C1) SERS spectra of propoxate and isopropoxate standards at different concentrations. (D and D1) SERS spectra from 20 different positions. Characteristic peak intensity distribution map (E and E1) and projection (F and F1).

is attributed to the stretching vibration of the C=O bond in the molecule, with its high intensity resulting from the significant change in vibrational dipole moment due to the high polarity of the C=O bond.

For isopropoxate (Fig. 3C), the experimental SERS data showed excellent agreement with the DFT-calculated spectra at wavenumbers of 616 cm^{-1} , 1028 cm^{-1} , and 1196 cm^{-1} . The peak at 616 cm^{-1} corresponds to the in-plane bending vibration of four C-H bonds on the benzene ring, while the peak at 783 cm^{-1} represents the out-of-plane bending vibration of three C-H bonds on the benzene ring. The peak at 844 cm^{-1} indicates the stretching vibration of the C-C bond in the propyl group. The peak at 1002 cm^{-1} corresponds to the stretching vibration

of the C-C bond at the junction between the benzene ring and the imidazole ring, as well as the breathing vibration of the benzene ring. The peak at 1028 cm^{-1} represents the in-plane bending vibration of four C-H bonds on the benzene ring. The peak at 1196 cm^{-1} represents the in-plane rocking vibration of the two C-H bonds on the imidazole ring. The peak at 1720 cm^{-1} indicates the stretching vibration of the C=C bond at the junction between the propyl group and the imidazole ring. Tables S2 and S3 summarize the assignments of characteristic peak vibration modes for propoxate and isopropoxate. The raw data of propoxate and isopropoxate standards as well as spiked samples are shown in Fig. S3–S6.

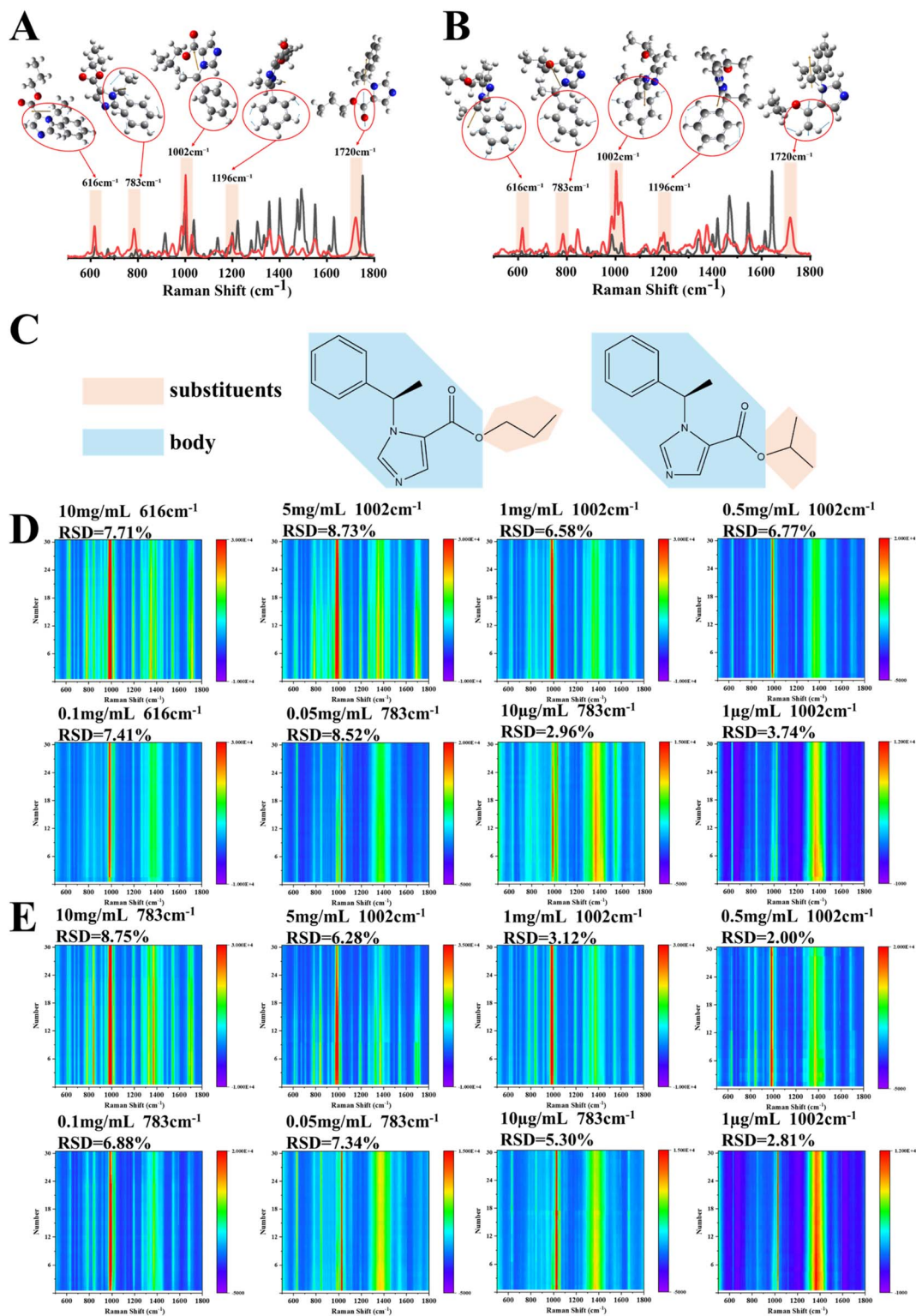


Fig. 3 (A) Molecular structures of propoxate and isopropoxate. (B and C) DFT-calculated Raman spectra (black lines) and SERS characteristic peaks (red lines) of propoxate and isopropoxate. (D and E) SERS detection results of different concentrations of propoxate and isopropoxate in e-cigarettes, along with the corresponding RSD values of characteristic peaks.

Fig. 3D and E present the SERS detection results of propoxate and isopropoxate at different concentrations, with 30 randomly selected sites for each concentration. LODs for both substances

reached $1 \mu\text{g mL}^{-1}$. For propoxate, the relative standard deviations (RSD) were calculated using data from three strong peaks at 616 cm^{-1} , 783 cm^{-1} and 1002 cm^{-1} . The RSD range across



eight concentrations was from 8.73% to 2.96%. At high concentrations of 10 mg mL⁻¹ and 5 mg mL⁻¹, the RSD values were relatively large, reaching 7.71% and 8.73%, respectively. This is because at high concentrations, the saturation of molecules adsorbed on the surface of gold nanoparticles may lead to a decrease in SERS peak intensity.^{40,41} The relative standard deviations were all less than 9%, with the smallest relative standard deviation of 2.96% observed at 10 μg mL⁻¹. As shown in Fig. 3E, the relative standard deviations of isopropoxate were calculated using data from two strong peaks at 783 cm⁻¹ and 1002 cm⁻¹. The relative standard deviations across eight concentrations ranged from 8.75% to 2.00%. Similarly, at higher concentrations of 10 mg mL⁻¹, the relative standard deviations were larger, reaching 8.75%, respectively. In contrast, at concentrations of 5 mg mL⁻¹ and below, the relative standard deviations were all less than 8%, with the smallest relative standard deviation of 2.00% observed at 0.5 mg mL⁻¹. It demonstrates the stability of this method within a certain concentration range. Existing studies on the detection of propoxate and isopropoxate in e-cigarettes report concentration ranges of 1–50 ppm.⁴² The limit of detection of our method is 1 μg mL⁻¹ (equivalent to 1 ppm), which corresponds to the lower limit of the above range. Therefore, it fully meets the practical requirements for on-site rapid detection.

3.3 Classification performance evaluation

To thoroughly explore the complementary effects of time-domain and frequency-domain features and enhance the classification accuracy of SERS spectra, we constructed a dual-branch gated fusion model based on STFT time-frequency transformation, CNN residual networks, and an interpretable Transformer architecture. By integrating time-frequency feature extraction, dual-branch parallel processing, and an adaptive gated fusion mechanism, this model significantly improves the synergistic capture capability of global spectral dependencies and local frequency-domain features. All experiments were conducted on a unified hardware platform equipped with GPUs and a software environment running Python 3.9 and Pytorch 2.1 to ensure comparability and reproducibility of results.

During the data processing stage, we innovatively transformed SERS spectral data from the time domain to the frequency domain using Short-Time Fourier Transform (STFT), clearly revealing frequency variations over time as shown in Fig. 4. We treated the Raman shift of spectral data as the time axis to analyse frequency information at different Raman shifts. Fig. 4A displays the pre-processed (cropped, interpolated, and normalized) spectral curve of the original Raman spectrum, with the horizontal axis representing Raman shift and the vertical axis indicating signal intensity. Fig. 4B presents the STFT results in a heatmap format, where the horizontal axis is treated as time corresponding to Raman shift, the vertical axis represents frequency, and colour denotes intensity. In Fig. 4C, we present the STFT results using a 3D surface plot, where the X-axis represents time, the Y-axis represents frequency, and the Z-axis represents amplitude. This visualization allows for a more intuitive observation of the fluctuations in the time-frequency

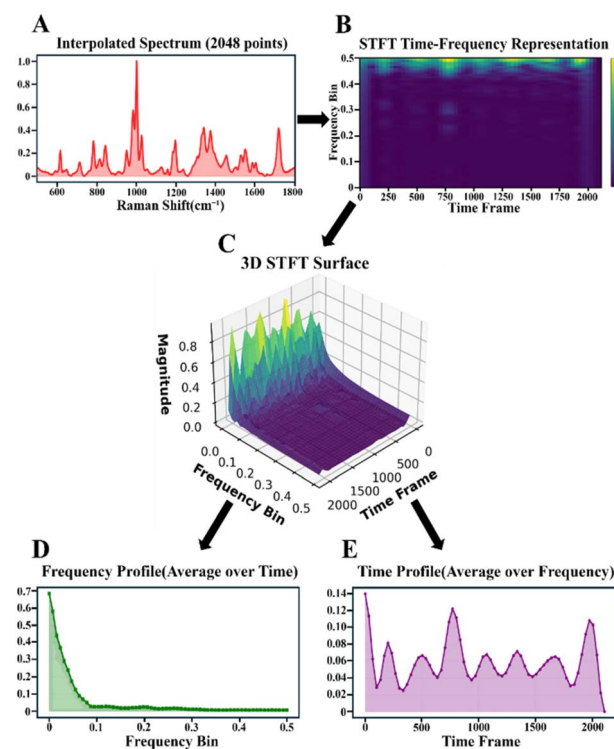


Fig. 4 (A) Pre-processed spectrum. (B) Time frequency diagram. (C) Contour plot. (D) Frequency profile. (E) Time profile.

representation. Fig. 4D displays the frequency profile, obtained by averaging along the time dimension to derive the mean amplitude at each frequency point, which reflects which frequency components are more prominent overall. Fig. 4E shows the time profile, generated by averaging the STFT results along the frequency dimension to obtain the mean amplitude at each time point, illustrating the overall frequency intensity variations over time (Raman shift).

In terms of optimization strategies, a hierarchical learning rate configuration is implemented for the dual-branch characteristics of the model. The learning rate for CNN branch parameters is set to a baseline of 0.001, while Transformer branch parameters adopt twice the baseline rate (0.002), with fusion parameters maintaining the baseline rate. The AdamW optimizer is selected with a weight decay coefficient of 1×10^{-4} . A cosine annealing learning rate scheduling algorithm is employed to dynamically adjust the learning rate, with the training epoch fixed at 100 and a batch size of 32. To enhance model generalization, Dropout regularization is applied at multiple levels. The dropout rate for the attention mechanism layer is set to 0.1, while the fully connected layers use dropout rates ranging from 0.2 to 0.3. Gradient clipping technique ($\text{max_norm} = 1.0$) is employed to stabilize the training process.

To address the characteristics of SERS spectral data, we established a standardized data pre-processing pipeline. The raw spectral sequences were first truncated according to the Raman shift range (500–1800 cm⁻¹), then uniformly interpolated to 2048 data points *via* linear interpolation. After min-max normalization, short-time Fourier transform was applied to



generate time-frequency images (STFT parameters: $n_{\text{perseg}} = 128$, $n_{\text{overlap}} = 32$), with logarithmic scaling of amplitude values to enhance contrast. The dataset was randomly partitioned into training, validation, and test sets at a ratio of 7 : 2 : 1 to ensure objective model evaluation.

As shown in Fig. 5, the dual-branch fusion model achieved optimal classification performance with an average accuracy of 99.86% (Fig. 5A, blue line). The Transformer branch exhibited significant training fluctuations before the 20th epoch, while its accuracy and loss fluctuations diminished after 50 epochs, reaching full convergence (Fig. 5B, green line). The Transformer branch demonstrated pronounced instability during initial training phases, likely attributable to its sensitivity to initialization parameters and insufficient learning of attention weights in early stages, leading to unstable gradient updates. In contrast, the CNN branch leveraged its inherent local connectivity structure to capture fundamental data patterns more rapidly, enabling both the CNN and fusion models to converge around the 20th epoch (Fig. 5B, orange line). It is noteworthy that the loss curves of CNN and the fusion model exhibit a high degree of similarity. This phenomenon indicates that during the adaptive weighting process of the gated fusion mechanism, frequency-domain features derived from time-frequency images are assigned higher weights. This further confirms the

effectiveness of the STFT transformation, which decouples and reconstructs the implicit, highly overlapping molecular vibration modes in the one-dimensional spectrum into two-dimensional time-frequency images, thereby preserving the dynamic information of signal frequency evolution over time (Raman shift). This pattern of amplifying local vibrational features provides the CNN branch with a more distinguishable feature representation than the original sequence.

The stability of the model was confirmed through multiple independent training validations. In ten training sessions with random seed initializations as shown in Table 1, the fused model maintained a final validation accuracy standard deviation within 0.0225, with precision standard deviation of 0.0194, recall standard deviation of 0.0219, and F1-score standard deviation of 0.0226, demonstrating the convergence stability of the training process. Meanwhile, the three-branch comparison curves revealed that the fused model exhibited superior loss convergence speed and accuracy improvement trends compared to individual branches, validating the effectiveness of feature fusion.

We constructed an independent test set comprising 297 reference samples, including 140 propoxate samples and 157 isopropoxate samples, none of which were involved in model training or cross-validation phases. As shown in Fig. 5C, both

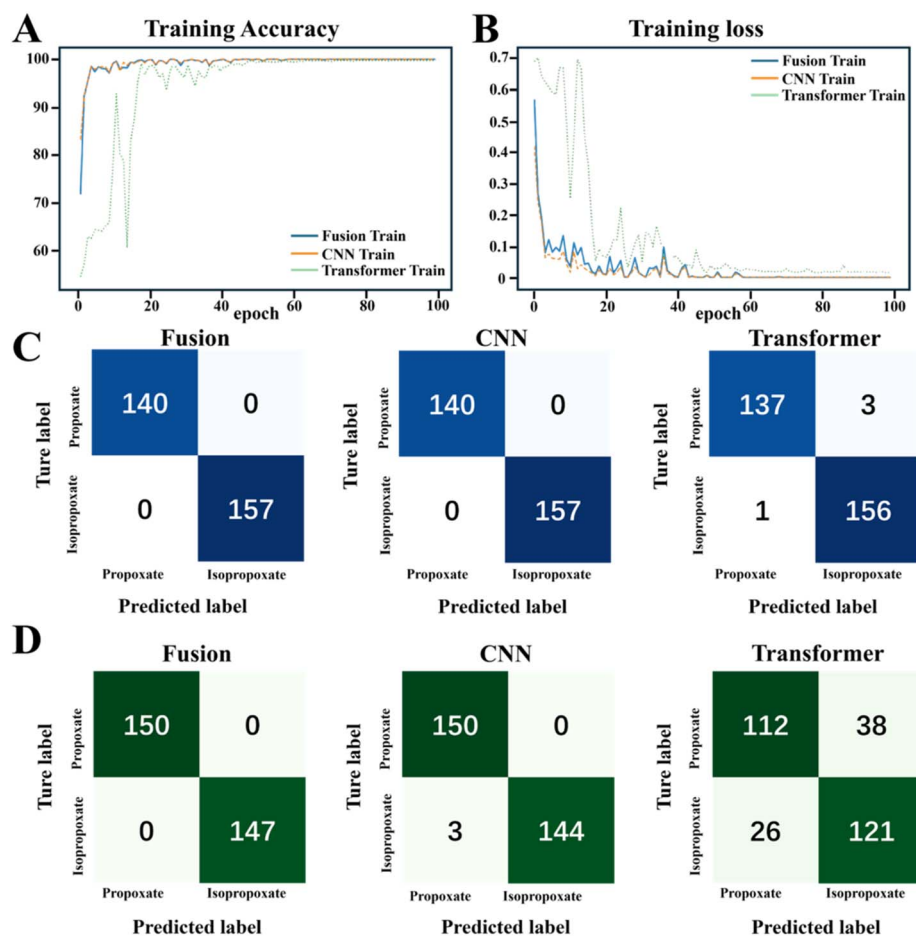


Fig. 5 The training accuracy (A) and loss (B) of the model. The confusion matrix of the model test on standard sample (C) and real samples (D).



Table 1 Ten-fold cross validation results

Fold	Accuracy	Precision	Recall	F1-score
1	0.9149	0.9259	0.9167	0.9145
2	0.9894	0.9898	0.9891	0.9894
3	0.9788	0.9800	0.9783	0.9787
4	0.9894	0.9898	0.9891	0.9895
5	0.9894	0.9894	0.9896	0.9894
6	0.9788	0.9804	0.9778	0.9787
7	0.9787	0.9787	0.9787	0.9787
8	0.9892	0.9898	0.9889	0.9893
9	1.0000	1.0000	1.0000	1.0000
10	0.9892	0.9898	0.9889	0.9892
SD	0.0225	0.0194	0.0219	0.0226
Mean	0.9798	0.9814	0.9797	0.9797

the fusion model and CNN branch accurately identified all 297 samples. The Transformer branch model correctly identified 293 samples, with 1 misclassified as propoxate and 3 misclassified as isopropoxate. To further evaluate the model's generalization capability and robustness in practical applications, we established another independent test set containing 297 real-world samples, consisting of 150 propoxate samples and 147 isopropoxate samples, none of which participated in model training or cross-validation stages. The prediction results on the independent test set demonstrated an overall classification accuracy of 99.73%. As shown in Fig. 5D, the fusion model correctly identified 297 samples with zero misclassifications. The Transformer branch model accurately recognized 233 samples, with 26 misclassified as propoxate and 38 misclassified as isopropoxate. The CNN branch model correctly identified 294 samples, with 3 misclassified as propoxate and none misclassified as isopropoxate, achieving an accuracy of 98.98%, precision of 98.03%, recall of 100.00%, and an F1-score of 99.00%. These results indicate the excellent generalization capability and robustness of the dual-branch network model.

For SERS data of stable reference standards, both branch models and the fusion model achieve high classification accuracy. However, in rigorous testing with real-world samples, the performance differences among the three further highlight the necessity of time-frequency feature fusion. The Transformer branch, which directly processes raw spectral sequences, sees its ability to capture global dependencies susceptible to background noise interference in complex e-cigarette matrices, leading to degraded classification performance. In contrast, the CNN branch processes time-frequency images where overlapping frequency-domain information is pre-separated *via* STFT, granting it stronger robustness against matrix interference and thus maintaining exceptionally high accuracy. Most critically, the fusion model does not merely combine features through simple addition but achieves adaptive feature complementarity *via* a gating mechanism. For samples with minimal matrix interference and distinct sequence features, the model assigns higher weights to the temporal domain branch. Conversely, for samples exhibiting highly overlapping features and complex noise patterns, it adaptively prioritizes the spectral domain branch with greater feature weighting. This dynamic trade-off capability enables the model to simultaneously

address the dual challenges of precise differentiation of structural analogs and anti-interference in complex matrices. Consequently, it achieves superior generalization performance beyond single-branch approaches in real-world scenarios, reaching an accuracy rate of 99.73%.

Notably, in our previous work, the target analytes nicotine and etomidate exhibited well-separated SERS characteristic peaks, allowing a standard CNN model to achieve high classification accuracy with relative ease. In contrast, propoxate and isopropoxate are structural isomers with highly overlapping SERS spectral features. Conventional deep learning methods, such as a Transformer directly processing raw one-dimensional spectral sequences, struggle to capture the subtle differences between these isomers in complex e-cigarette matrices, achieving only 78.45% accuracy on the same real-sample test set. To address this challenge, we employed STFT to convert the one-dimensional spectral sequences into two-dimensional time-frequency images, effectively amplifying the minor spectral differences between the isomers and boosting the CNN model's accuracy to 98.98%. Furthermore, through the adaptive fusion of the temporal and spectral branches, the model robustness was significantly enhanced, ultimately achieving classification accuracy of 99.73%. Compared with our previous work, where the spectral features were distinctly different and did not require such sophisticated modelling, the proposed time-frequency fusion strategy represents a substantial advancement, providing a generalizable solution for SERS-based identification of highly similar isomers in complex real-world matrices.

4 Conclusions

In this study, we first prepared gold nanoparticles as enhancement materials and achieved a LODs of 1 ng mL⁻¹ for both propoxate and isopropoxate through SERS detection of their standards. Simultaneously, the vibrational modes of characteristic fingerprint peaks for propoxate and isopropoxate were determined *via* DFT calculations, enabling precise assignment of each substance's characteristic SERS peaks to their corresponding vibrational modes. Subsequently, a sample pre-treatment method was developed for spiked e-liquid samples, effectively isolating propoxate and isopropoxate to achieve trace detection in real samples. Finally, the precise classification of SERS detection results for propoxate and isopropoxate in e-cigarette liquids was successfully achieved using a dual-branch gated fusion model with time-frequency feature integration, reaching an identification accuracy of 99.73%.

Author contributions

Yazhou Qin and Jiye Wang: conceptualization, methodology, software. Jiahao Teng, Shenggang Huang, Wenkai Zheng, Xuqing Wang, Yingsheng He: data curation, validation, formal analysis, investigation. Jiahao Teng, Yingsheng He and Yazhou Qin: writing— original draft preparation. Jiye Wang and Yazhou Qin: supervision, funding acquisition. Jiahao Teng, Xuqing Wang and Yazhou Qin: writing— reviewing and editing.



Conflicts of interest

There are no conflicts to declare.

Data availability

Data will be availability on request.

Supplementary information (SI) is available. See DOI: <https://doi.org/10.1039/d6ra00614k>.

Acknowledgements

This research was supported by Zhejiang Provincial Natural Science Foundation of China (LMS25H230001 and LTGC24B050008), National Natural Science Foundation of China (62541210), Central Guidance on the Development of Local Science and Technology (2025ZY01109).

Notes and references

- 1 Y. Yang, E. N. Lindblom, R. G. Salloum and K. D. Ward, *Tob. Control*, 2023, **32**, e23–e30.
- 2 J. Y. Chien, Y. C. Gu, C. H. Liu, H. M. Tsai, C. N. Lee, A. C. Yang, J. Huang, Y. L. Wang, J. K. Wang and C. H. Lin, *J. Pharm. Biomed. Anal.*, 2023, **233**, 115456.
- 3 D. Hammond, J. L. Reid, M. L. Goniewicz, A. McNeill, R. J. O'Connor, D. Corsetti, A. C. Block, L. S. Brose and D. Robson, *JAMA Netw. Open*, 2025, **8**, e2462544.
- 4 Y. Z. Qin, X. Q. Wang, M. Z. Fu, J. K. Mao, Y. S. He and J. Y. Wang, *Anal. Chim. Acta*, 2026, **1406**, 345514.
- 5 E. A. Cowan, H. Tran, N. Gray, J. J. Perez, C. Watson, B. C. Blount and L. Valentin-Blasini, *Talanta*, 2022, **238**, 122985.
- 6 C. Liu, A. J. Tose, J. P. H. Verharen, Y. Zhu, L. W. Tang, J. W. de Jong, J. X. Du, K. T. Beier and S. Lammel, *Neuron*, 2022, **110**, 3018–3035.
- 7 C. V. Weiger, S. K. Gratale, O. Ganz, M. LaVake, E. M. Talbot and O. A. Wackowski, *Tob. Control*, 2025, DOI: [10.1136/tc-2024-058876](https://doi.org/10.1136/tc-2024-058876).
- 8 J. J. Rose, S. Krishnan-Sarin, V. J. Exil, N. M. Hamburg, J. L. Fetterman, F. Ichinose, M. A. Perez-Pinzon, M. Rezk-Hanna and E. Williamson, *Circulation*, 2023, **148**, 703–728.
- 9 Y. Guo, Y. Yang, Z. Zhou, C. Zhao, Y. Li, H. Zhou, S. Ren, Y. Gu and Z. Gao, *Environ. Health*, 2024, **2**, 301–310.
- 10 C. Merola, G. Caioni, C. Bertolucci, T. Lucon-Xiccato, B. B. Savaşçı, S. Tait, M. Casella, S. Camerini, E. Benedetti and M. Perugini, *Sci. Total Environ.*, 2024, **912**, 168925.
- 11 G. E. Cozier, M. Gardner, S. Craft, M. Skumlien, J. Spicer, R. Andrews, A. Power, T. Haines, R. Bowman, A. E. Manley, P. Sunderland, O. B. Sutcliffe, S. M. Husbands, L. Hines, G. Taylor, T. P. Freeman, J. Scott and C. R. Pudney, *Addiction*, 2025, **120**, 1995–2004.
- 12 E. F. Valenzuela, I. F. Simões, F. W. R. Camelier, A. de Freitas Santos Júnior, Z. das Graças Guimarães Viola and G. C. da Fonseca Andrade, *Talanta*, 2025, **294**, 128255.
- 13 N. E. Robertson, H. C. Hunsaker, M. Yamamoto, K. Cheung, B. A. Poulin and T. B. Nguyen, *ACS Omega*, 2025, **10**, 29615–29627.
- 14 K. Kanamori, S. M. Ahmad, A. Hamid and K. Lutfy, *Drug Metab. Dispos.*, 2024, **52**, 171–179.
- 15 M. T. Lin, Z. Zhang, Q. He, H. Y. Hao, P. Xiang and J. Zhao, *J. Pharm. Biomed. Anal.*, 2025, **256**, 116677.
- 16 Z. K. Huang, J. P. Peng, L. G. Xu and P. J. Liu, *Nanomaterials*, 2024, **14**, 1417.
- 17 J. Cardellini, C. Dallari, I. De Santis, L. Riccio, C. Ceni, A. Morrone, M. Calamai, F. S. Pavone, C. Credi, C. Montis and D. Berti, *Nat. Commun.*, 2024, **15**, 7975.
- 18 W. Kim, K. Chai, E. Park, H. Park, G. Kim, J. Park and J. Park, *Chem. Eng. J.*, 2025, **518**, 164507.
- 19 H. P. Li, E. Dumont, R. Slipets, T. Thersleff, A. Boisen and G. A. Sotiriou, *Chem. Eng. J.*, 2023, **470**, 144023.
- 20 D. C. Zhang, X. L. Chen, J. Lin, S. Y. Jiang, M. Fan, N. R. Liu, Z. F. Huang and J. Wang, *Anal. Chem.*, 2025, **97**, 4101–4110.
- 21 X. Y. Bi, L. Lin, Z. Chen and J. Ye, *Small Methods*, 2024, **8**(1), 2301243.
- 22 X. Y. Liu, H. L. An, W. S. Cai and X. G. Shao, *Trends Anal. Chem.*, 2024, **172**, 117612.
- 23 Y. Z. Qin, H. Zhang, W. Wang and Y. S. He, *Spectrochim. Acta, Part A*, 2026, **348**, 127086.
- 24 Y. M. Tseng, K. L. Chen, P. H. Chao, Y. Y. Han and N. T. Huang, *ACS Appl. Mater. Interfaces*, 2023, **15**, 26398–26406.
- 25 W. Q. Guo, S. C. Gao, Y. H. Ding and D. M. Dong, *Comput. Electron. Agric.*, 2025, **237**, 110507.
- 26 T. Y. Xing, P. Y. Mao, S. X. Wang, H. Ye, S. C. Yang, W. K. Jiang, X. Z. Qiu, Y. L. Shi and L. J. Wu, *Anal. Chem.*, 2025, **97**, 20076–20087.
- 27 A. Hegde, M. Hajikhani, J. Snyder, J. Cheng and M. Lin, *ACS Appl. Mater. Interfaces*, 2025, **17**, 2018–2031.
- 28 H. Z. Li, S. H. Xu, J. H. Teng, H. Zhang, Y. Z. Qin, Y. S. He and L. Fan, *Microchem. J.*, 2025, **212**, 113224.
- 29 T. Y. Lei, J. C. Li and K. W. Yang, *Expert Syst. Appl.*, 2024, **252**, 124155.
- 30 J. D. Yang, Q. W. Wu, W. H. Weng, S. M. Wang and H. X. Yan, *Biomed. Signal Process. Control*, 2026, **112**, 108862.
- 31 R. C. Ma, J. L. Chen, Y. Feng, Z. T. Zhou and J. S. Xie, *Knowledge-Based Syst.*, 2025, **316**, 113410.
- 32 Z. L. Jin, Q. F. Xu, C. X. Jiang, Z. Liu, T. M. Xie and X. X. Wang, *J. Intell. Manuf.*, 2025, **37**, 1909–1929.
- 33 Z. L. Duan, W. B. Zhang, H. F. Zhang and F. Y. Yang, *Mech. Syst. Signal Process.*, 2025, **237**, 113101.
- 34 J. Y. Zhu, J. Ma, J. D. Wu and L. K. Fan, *Mech. Syst. Signal Process.*, 2025, **236**, 113006.
- 35 C. Ji and Y. X. Yue, *Opto-Electron. Eng.*, 2025, **52**, 250225.
- 36 Z. Y. Pei, C. Ji, M. R. Shao, Y. Wu, X. F. Zhao, B. Y. Man, Z. Li, J. Yu and C. Zhang, *Opto-Electron. Sci.*, 2025, **4**, 250015.
- 37 Y. Wu, T. Sun, Z. Pei, C. Ji, X. Zhao, M. Shao, Z. Li, J. Yu and C. Zhang, *Light: Sci. Appl.*, 2026, **1**, 202601.
- 38 R. Q. Wang, S. D. Qiao, Y. He, C. G. Yang, Z. Wang, Y. Liu and X. K. Yin, *Opto-Electron. Adv.*, 2025, **8**, 240275.
- 39 J. H. Teng, Y. H. Xu, Y. L. Wu, X. Q. Wang, H. Song, Y. S. He and Y. Z. Qin, *Microchem. J.*, 2026, **221**, 116990.



Paper

- 40 H. N. Shi, S. J. Feng, J. Z. Song, D. T. Han and G. Q. Liu, *Opt. Mater.*, 2023, **146**, 114491.
- 41 C. A. Visbal, W. R. Cervantes, L. Marín, J. Betancourt, A. Pérez, J. E. Diosa, L. A. Rodríguez and E. Mosquera-Vargas, *Nanomaterials*, 2024, **14**, 1525.
- 42 Y. Mo, X. P. Zhang, K. Zou, W. Xing, X. Y. Hou, Y. Zeng, Y. G. Cai, R. X. Xu, H. W. Zhang and W. P. Cai, *Nanomaterials*, 2024, **14**, 1958.

