Scientific Reports

https://doi.org/10.1038/s41598-026-45433-0

Article in Press

A study for potential rapid discrimination of smokeless powders by near-infrared spectroscopy and chemometric modeling methods for forensic application

Received: 9 December 2025

Accepted: 18 March 2026

Published online: 03 April 2026

Cite this article as: Guo H., Shi H., Feng Y. *et al.* A study for potential rapid discrimination of smokeless powders by near-infrared spectroscopy and chemometric modeling methods for forensic application. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-45433-0

Hongling Guo, Haoyuan Shi, Yinghua Feng, Yiting Guo, Xiuli Zhang, Ping Wang, Can Hu, Hongcheng Mei, Yajun Li & Jun Zhu

A study for potential rapid discrimination of smokeless powders by near-infrared spectroscopy and chemometric modeling methods for forensic application

Hongling Guo[a*], Haoyuan Shi[b,c†], *Yinghua Feng* [d], Yiting Guo[a,e], Xiuli,Zhang[f], Ping wang[a,e], Can Hu[a], Hongcheng Mei[a], Yajun Li[a], Jun Zhu[a*]

[a] *Institute of Forensic Science, Ministry of Public Security of China, Beijing, China;*

[b] *State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China.*

[c] *Department of Urology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China*

[d] *Department of Epidemiology, School of Public Health, Shanxi Medical University*

Taiyuan, China

[e] *Chinese People's Public Security University, Beijing, China;*

[f] *Department of Rheumatology and Clinical Immunology, Peking University First Hospital, Beijing 100034, China*

† Hongling Guo and Haoyuan Shi are co-first authors.

* Correspondence to Hongling Guo and Jun Zhu, Institute of Forensic Science, Ministry of Public Security, China.

Address: Institute of Forensic Science, Ministry of Public Security of China, No. 17 Muxidi Nanli, Western District, Beijing, PRC.

Correspondence: guohongling1234@163.com; zhujun001cn@126.com

Phone No. 0086-10-66269495

ABSTRACT

Smokeless powder is the primary propellant in civilian and military ammunition, and in China, the use of propellants to make homemade ammunition and bombs is an emerging criminal practice. The identification and differentiation of the propellants used can provide forensic information about their sources. Depending upon the ammunition manufacturer and type, the recipe of propellants varies, and the characterization of smokeless powders in terms of their spectral components is useful for differentiating propellants. In this work, near-infrared spectroscopy (NIR) and chemometric modeling were used to explore the feasibility of differentiating and predicting smokeless powders from different sources. By comparison, the proposed neural network model in the study exhibited an average accuracy of over 80%. Furthermore, the potential for differentiating smokeless powders was well demonstrated via simple and rapid near-infrared spectroscopic analyses, and the employment of chemical agents and time-consuming chromatography and mass spectrometry could thereby be avoided.

KEYWORDS: Smokeless powder; Near-infrared spectroscopy; Chemometrics; neural network model；Propellant; Forensics

1. Introduction

Smokeless powder (SLP) is the primary propellant in civilian and military ammunition, and is the most common low explosive used to fabricate improvised explosive devices (IEDs) such as pipe bombs and home-made ammunition [1]. It can also act as an initiator in explosive mixtures when combined with other components such as coal, ammonium nitrate, sulfur, etc [2]. Additionally, in China, there are cases involving homemade guns with an unknown source of smokeless powder. Investigations into firearms and explosive-related crimes, alongside the inspection and characteristic profiling of smokeless propellants (SLPs), are of paramount significance; these processes aid investigators in determining the brand of the SLP incorporated into IEDs[3] and linking SLPs in different cases or locations[4]. Key compound classes present in smokeless propellants include energetics, stabilizers, plasticizers, flash suppressants, deterrents, opacifiers, and dyes [5]. According to the main chemical substances in the energetic class of compounds, SLPs can be classified into three different types: single-, double- and triple-base. Single-base powder mainly contains nitrocellulose(NC), NC and nitroglycerin(NG) are main component for double-base propellants, and a triple-base powder mainly contains NC, NG, and nitroguanidine [6].

Many studies have been performed where the intact SLPs and residues have been analyzed. Infrared spectroscopy, chromatography, and mass spectroscopy have been commonly used for the determination of different ingredients in SLPs [7,8,9]. Infrared spectroscopy has normally been used for the analysis of basic components of nitrocellulose, but it showed a limited application for other substances with low amounts. Gas chromatography-mass spectrometry (GC-MS) and liquid chromatography (LC) have been widely applied for the identification of stabilizers, plasticizers, and other organic additives, providing high chemical specificity and strong discriminatory capability[10,11,12]. Dennis[13] examined the organic constituents of six classes of smokeless powders via GC-MS and chemometric techniques. Low discrimination between samples was observed based on Fisher transform and Pearson correlations, and it was determined that probabilistic statements regarding smokeless powder comparisons had limited

strength. Lennert [14]analyzed 34 SLP samples from 6 classes by GC-MS and DART-TOFMS and classified them using K-nearest neighbors and random forest modeling. It demonstrated good discrimination and classification results

Despite their considerable analytical power, these chromatographic and mass spectrometric techniques are inherently destructive, depend on target-specific chemical information, and require sophisticated instrumentation, which limits their suitability for rapid, non-destructive, and routine forensic screening, particularly in time-sensitive casework. Nondestructive techniques with improved discrimination possibilities (such as Raman spectroscopy) have, therefore, been introduced for the analyses of intact SLPs and residue particles. Raman spectroscopy has reported to have been used for the identification of the unburnt smokeless powder particles [15].

With the development of instrumental technology and chemometrics, near-infrared (NIR) spectroscopy has been found capable of recording spectra of complicated compounds, with minor (or no) sample pretreatments. And it is not necessary to specify the target compounds. NIR has been widely used as a quality control method in the agriculture, food, petrochemical engineering, and pharmacy industries [16,17,18]. And it has also been reported as a fast quality control tool for the determination of ingredient concentrations in rock propellant fuel liquid pre-mixes [19]. Furthermore, NIR has been used for an in-process determination of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) in a solid propellant intermediate[20]. Reports about the use of NIR in propellants have mainly focused on a fast quantification of some specific compounds in the propellants and the accuracy of the used algorithm [21]. The possibility to use NIR as a non-target approach for the discrimination and specification of SLPs collected from different sources has been investigated in the present study. Chemometric approaches were applied to evaluate the spectra data obtained.

2. Materials and methods

2.1. Collection and preparation of samples

SLPs from 79 cartridges collected from different manufacturers were used as known source

samples to validate the method used. Detailed information was presented in Table 1. The propellants were taken out from the cartridges by using a cartridge puller. Prior to NIR analysis, they were equilibrated at room temperature (of about 25℃) for 24 h.

Table 1 Ammunition propellant sample information for measurement

Series Number	Manufacturer	Cartridge category	Production Year	Caliber (mm)	SLP type(single-, double- and triple-base)	Manufacturer ID	Series Number	Manufacturer	Cartridge category	Production Year	Caliber (mm)	SLP type9(single-, double- and triple-base)	Manufacturer ID
s1	121	51 pistol	1958	7.62	Double	A	s41	81	59 rifle	1966	9	Single	I
s2	121	51 pistol	1963	7.62	Double	A	s42	911	56 rifle	1968	5.6	Single	J
s3	121	51 pistol	1964	7.62	Single	A	s43	911	56 rifle	1979	5.6	Single	J
s4	121	51 pistol	1965	7.62	single	A	s44	C	Sporting rifle	unknown	5.6	Double	K
s5	121	51 pistol	1970	7.62	Single	A	s45	CJ	SS109 rifle	2000	5.6	Double	L
s6	121	51 pistol	1972	7.62	single	A	s46	Czech Republic	pistol	unknown	7.65	Single	M
s7	121	51 pistol	1979	7.62	Single	A	s47	KKJ	Nail	unknown	6.8x18	Double	N
s8	121	51 pistol	1981	7.62	single	A	s48	KKJ	Nail	unknown	6.3x16	Double	N
s9	121	51 pistol	1987	7.62	Single	A	s49	KKJ	Nail	unknown	6.3x16	Double	N
s10	121	51 pistol	1991	7.62	single	A	s50	LY	Parabellum pistol	1994	9	Double	O
s11	121	51 pistol	2018	7.62	single	A	s51	NS	Nail	unknown	6.8x11	Double	P
s12	121	Revolverpistol	2005	9	Double	A	s52	NS	Nail	unknown	6.8x11	Double	P
s13	121	DAP92 pistol	2019	9	Double	A	s53	NS	Nail	unknown	6.8x18	Double	P
s14	121	DAP92 pistol	2018	9	Double	A	s54	NS	Nail	unknown	6.8x18	Double	P
s15	121	64 pistol	2007	7.62	Double	A	s55	NS	Nail	unknown	5.6x16	Double	P
s16	121	64 pistol	1990	7.62	Double	A	s56	NS	Nail	unknown	5.6x16	Double	P
s17	121	64 pistol	1990	7.62	Double	A	s57	NS	Nail	unknown	6.3x16	Double	P
s18	121	64 pistol	1992	7.62	Double	A	s58	NS	Nail	unknown	5.6x16	Double	P
s19	121	64 pistol	1995	7.62	Double	A	s59	NS	Nail	unknown	6.8x11	Double	P
s20	121	64 pistol	1996	7.62	Double	A	s60	NS	Nail	unknown	5.6x16	Double	P
s21	121	64 pistol	1990	7.62	Double	A	s61	NS	Nail	unknown	6.8x11	Double	P
s22	121	92 rifle	2002	7.62	Double	A	s62	NS	Nail	unknown	6.8x11	Double	P
s23	301	64 pistol	1980	7.62	Double	B	s63	NS	Nail	unknown	6.8x11	Double	P
s24	301	64 pistol	1987	7.62	Double	B	s64	Ω	Sporting rifle	unknown	5.6	Double	Q
s25	311	64 pistol	1989	7.62	Double	C	s65	Double Ring	Sporting rifle	unknown	5.6	Double	R
s26	311	64 pistol	1992	7.62	Double	C	s66	△	Sporting rifle	Unknown	5.6	Single	S
s27	311	64 pistol	1994	7.62	Double	C	s67	△	Sporting rifle	unknown	5.6	Single	S
s28	311	64 pistol	2004	7.62	Double	C	s68	YRD	Nail	unknown	6.8x11	Double	T
s29	611	56 rifle	1967	7.62	single	D	s69	YRD	Nail	unknown	6.8x11	Double	T
s30	671	53 pistol	2013	7.62	single	E	s70	YRD	Nail	unknown	6.8x11	Double	T
s31	71	56 rifle	1956	7.62	single	F	s71	YRD	Nail	unknown	6.8x18	Double	T
s32	724	DAP92 pistol	2007	5.8	Double	G	s72	YRD	Nail	unknown	6.8x18	Double	T
s33	791	9mm pistol	2016	9	Double	H	s73	YRD	Nail	unknown	5.6x16	Double	T
s34	791	56 pistol	1988	7.62	Double	H	s74	YRD	Nail	unknown	5.6x16	Double	T
s35	791	SS109 rifle	2001	5.56	Double	H	s75	YRD	Nail	unknown	5.6x16	Double	T
s36	791	DCV05 pistol	2008	5.8	Double	H	s76	YRD	Nail	unknown	6.3x16	Double	T

s37	791	95 pistol	2016	5.8	Double	H	s77	YRD	Nail	unknown	6.3×16	Double	T
s38	791	pistol	2014	5.8	Double	H	s78	YRD	Nail	unknown	5.6×16	Double	T
s39	791	DVP88 A pistol	2017	5.8	Double	H	s79	YRD	Nail	unknown	5.6×16	Double	T
s40	81	56 rifle	1964	7.62	Double	I							

2.2. Acquisition of NIR spectra

The NIR analyses of the SLPs were carried out with a Fourier transform NIR spectrometer (Nicolet Antaris-II, ThermoFisher Scientific, U.S.A.), which was equipped with an integrating sphere module and an InGaAs detector. The sample cup was loaded with about 20 mg of propellant sample and rotated continuously at a constant speed. The NIR spectra were then collected in a diffuse reflectance mode at room temperature (25℃±2℃), and each sample was analyzed three times. All spectra were recorded as an absorbance with a resolution of 4 cm^{-1} and at a wavenumber range from 4000 cm^{-1} to 10,000 cm^{-1}. Also, each spectrum had an average of 32 scans.

2.3 Algorithmic statistics

2.3.1 Basic algorithmic statistics

79 smokeless powder samples were received from 20 different manufacturers. Basic descriptive statistics were performed to evaluate the distribution of NIR data for all samples. The calculation of SSSVs (specific significant statistical values) followed standard protocols. UMAP (Uniform Manifold Approximation and Projection) was performed using the UMAP-learn Python package, version 0.5.1. Prior to modeling, the NIR spectral data were inspected for quality and consistency. No explicit spectral pre-processing procedures, such as baseline correction, normalization, or scatter correction, were applied. Raw spectra were used to assess the intrinsic discriminatory capability of the NIR data under minimal processing conditions.

A deep learning neural network model (NN) was also developed for a more comprehensive and accurate classification prediction. Furthermore, comparisons between the NN model and four classical traditional machine learning models were made, including the Stochastic Gradient Descent Classifier (SGDC), Support Vector Machine (SVM), K-Nearest Neighbors Classifier (KN), and Random Forest Classifier (RF).

We constructed a neural network model using the keras module in TensorFlow, consisting

of six main components: a feature extraction network that takes data with dimensions identical to the input and outputs a one-dimensional vector of length 784; a feature map transformation module that converts the one-dimensional vector into a three-dimensional tensor of size 28 x 28 x 1; a feature extraction framework based on VGG19 for deep feature extraction; an embedding network that maps the extracted feature maps to the classification layer for classification discrimination. Throughout the training process, we utilized cross-entropy minimization as the loss function and employed the SGD optimization method with a momentum parameter of 0.9, lasting for 1000 epochs. All network layers used ReLU as the activation function. All machine learning methods are executed according to functions provided in the python package Scikit-Learn, with all parameters at default values. We have trained our models using a split of 90% training data and 10% validation data. For each model, we conducted a grid search to tune the hyperparameters, selecting the combination maximizing the validation accuracy. Accuracy was used to evaluate the performance of different models. Accuracy measured the proportion of correctly predicted samples out of the total samples. Given the limited number of samples available for certain manufacturers, the dataset was divided into training and validation sets using a 9:1 ratio to ensure sufficient data for model training. To reduce evaluation bias and mitigate potential overfitting, cross-validation was implemented during model development. Although alternative data-splitting strategies and the use of independent test sets may further enhance model robustness, these approaches will be explored in future studies with expanded datasets.

2.3.2 Network architecture and hyperparameter settings

The optimized neural network was designed as a deep multilayer perceptron (MLP). The architecture consists of an input layer with 3112 units, followed by two hidden layers with 128 and 64 neurons, and an output layer for target classification. The Sigmoid activation function was employed in the hidden layers to manage non-linear spectral features. For model optimization, the Stochastic Gradient Descent (SGD) optimizer was used with a learning rate of 0.01 and a momentum of 0.9. The model was trained for 300 epochs to ensure full

convergence and stability.

In addition to deep learning, traditional machine learning models including Stochastic Gradient Descent Classifier (SGDC), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF) were implemented using the Scikit-learn library. To ensure optimal performance and methodological transparency, hyperparameters for these models were optimized via grid search with five-fold cross-validation. For the SVM model, an RBF kernel was used with the penalty parameter C optimized within the range of [0.1, 1, 10, 100]. For KNN, the number of neighbors (k) was tuned between 1 and 20. For Random Forest, the number of trees (n_estimators) was evaluated at [50, 100, 200]. The parameter combinations yielding the highest validation accuracy were selected for final model construction.

2.3.3 Model benchmarking and interpretability analysis

To evaluate the scientific rigor and effectiveness of the proposed model, a benchmarking study was conducted against two established methods: Partial Least Squares Discriminant Analysis (PLS-DA), representing traditional linear chemometrics, and 1D-ResNet, representing state-of-the-art deep learning architectures. The performance of traditional machine learning baselines (SGDC, SVM, KNN, RF) was also integrated into this benchmarking to ensure a comprehensive evaluation. Furthermore, SHAP (SHapley Additive exPlanations) analysis was implemented to provide a robust quantification of feature importance. This analysis identifies the specific near-infrared bands (6700-7100 cm-1 and 5700-6000 cm-1) that drive the differentiation, ensuring the model's decisions are based on chemically relevant molecular vibrations rather than background noise.

3. Results and Discussion

3.1 NIR analyses

The smokeless powder samples exhibited broadly similar macroscopic appearances, with differences primarily observed in particle size, shape, and color. Most samples consisted of granular or flake-like particles, and visual discrimination among powders from different manufacturers was generally challenging. This limited macroscopic variability underscores the

need for spectroscopic techniques combined with chemometric analysis to achieve reliable source differentiation. Samples with series number of s1–s10 and s30–s39 were shown in Figure 1. The samples selected covered samples from different manufacturers and types. Giving that the spectral region 4000-4100 cm^{-1} mainly shows noise signal, this region was eliminated from the spectra data sets. The major peaks of the NIR spectra for smokeless powder samples were broad and similar. No significant difference was identified between them. The major absorption features observed in the NIR spectra are primarily associated with overtone and combination vibrations of functional groups present in nitrocellulose-based materials. Absorption bands in the region of approximately 5200-5300 cm^{-1} are mainly attributed to O-H stretching overtones and combination bands, whereas features near 5800 cm^{-1} predominantly correspond to first overtones of C-H stretching vibrations. The broad absorption observed around 7000 cm^{-1} is commonly associated with combination bands involving both O–H and C-H vibrations [20,22]. Owing to the inherently broad and overlapping nature of NIR absorption bands, direct assignment of individual chemical components is limited, necessitating the application of multivariate and machine learning approaches to extract subtle discriminatory information from the spectral data.

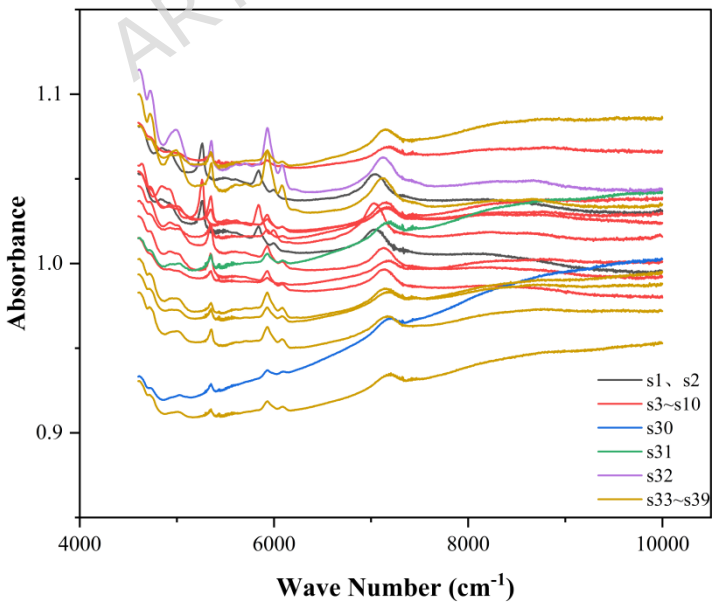

Fig.1 NIR spectra of Samples with series number of s1–s10 and s30–s39.

The majority of propellants from different sources in this study showed absorbance values that were predominantly distributed in the range of 0.9-1.1 (Fig. 2a). However, some samples from the manufacturers B, F, K, Q, R, and T showed some spectral absorbance values exceeding 1.2. It was interesting to find that the absorbance values of rifle-type cartridges were higher than the absorbance values for other types of cartridges. Correlations between SLPs produced by different manufacturers were also calculated. Notably, the similarity of most of the SLPs from different manufacturers was remarkably high, with correlation coefficients larger than 0.8 (Fig. 2b). However, SLPs from manufacturers D, E, and F showed a lower NIR spectral similarity with other samples. In a small number of cases, particularly among manufacturers D, E, and F, markedly lower and even negative correlation values were observed.

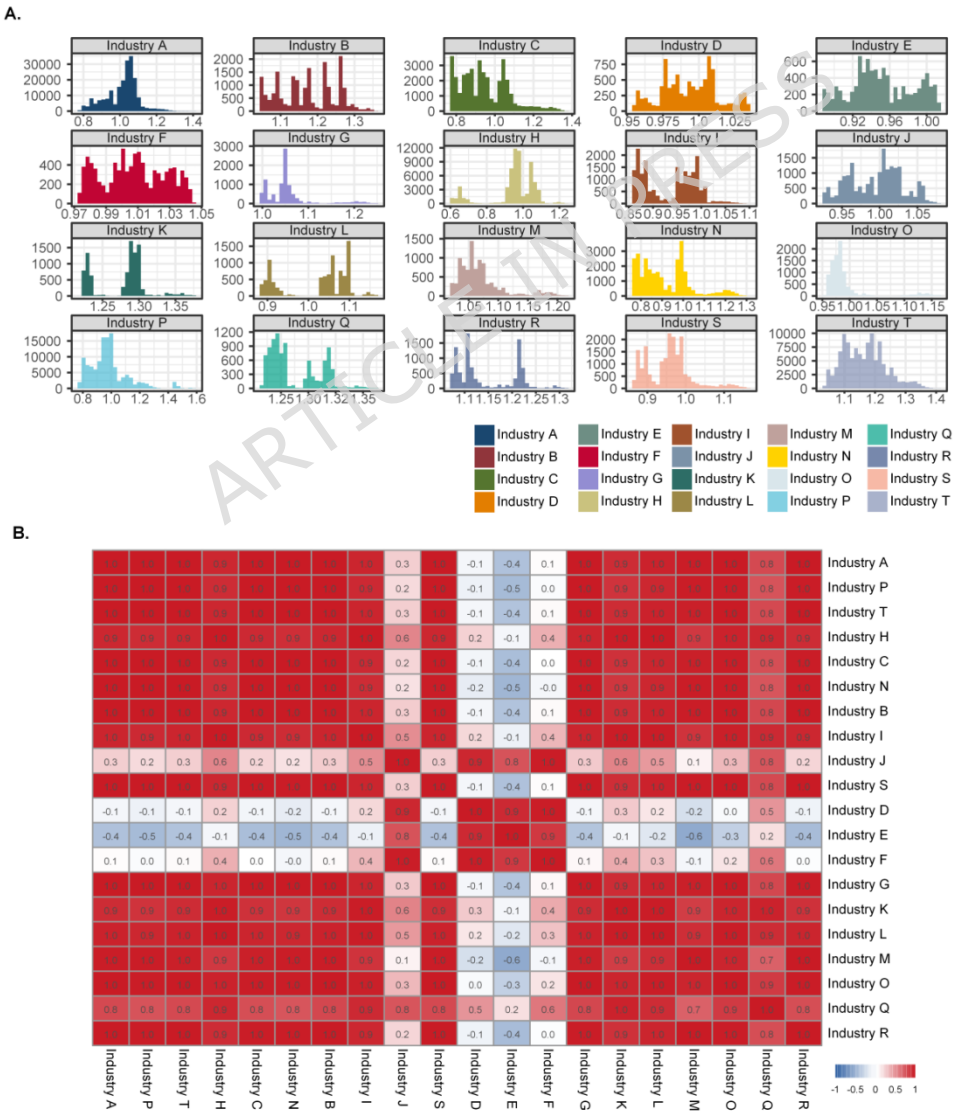

Fig.2. (A) Distribution of NIR absorbance values for SLPs from different manufacturers, and

(B) correlation coefficients between averaged NIR spectra of cartridges from different manufacturers.

3.2 Descriptive statistics

Descriptive statistical analyses were performed to characterize the distribution, class imbalance, and spectral similarity of the samples, thereby providing a rationale for adopting complex models such as neural networks. As shown in Figures 2-4, the NIR spectra of powders from different manufacturers exhibit high overlap and severe sample imbalance. These observations directly explain why traditional linear models struggle to achieve satisfactory discrimination and underscore the necessity of deep learning to capture subtle non-linear features.

The results revealed that the samples involved 20 different manufacturers, with a highly uneven distribution of sample quantities. Specifically, the samples produced by manufacturers A, P, T, and H accounted for over 60% (162/237) (Fig. 3A). Among these samples, there were 14 different cartridge types. 60% (156/237) of the total data were comprised by nails and pistols with 7.62mm caliber from 64 pistol and 51 pistol model of cartridges (Fig. 3B).

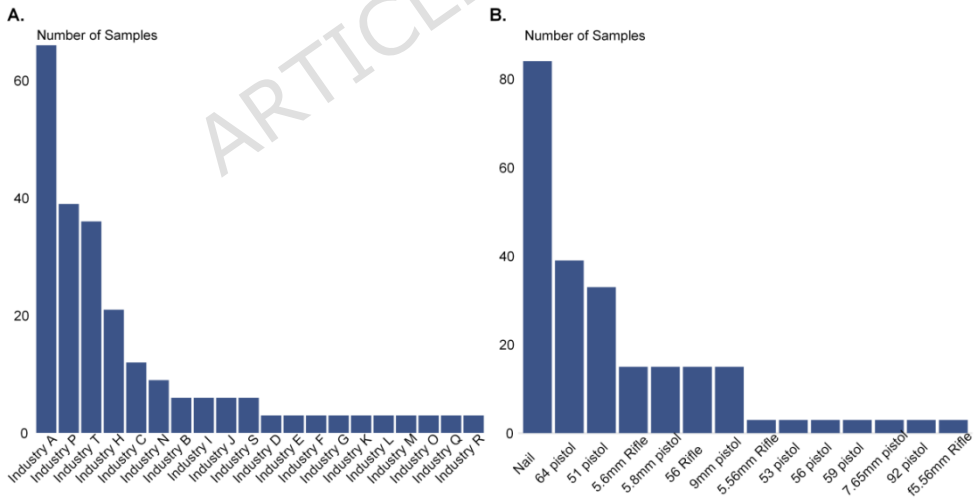

Fig. 3 Distribution of (A) manufacturers and (B) cartridge types of the collected samples

The manufacturers produced one or more types of cartridges. Manufacturers A and H produced the most diverse range of cartridge types, each producing four different types. Manufacturer A produced cartridges of the following types: cartridges with 7.62 mm caliber for Chinese 51 model pistol, 64 model pistol, 92 model rifle and pistols with 9mm caliber.

Manufacturer H, on the other hand, produced cartridges of the following types: cartridges with 5.8 mm caliber for Chinese pistols, 56 model pistol, and SS109 model rifle and pistols with 9mm calibers (Fig. 4A). The spectral absorbance of cartridges from different manufacturers have also been analyzed in the present study. As a result, the cartridges from 16 of the 20 manufacturers showed near-infrared characteristic distributions in the wavenumber range of 2750 cm^{-1} to 3500 cm^{-1}. The manufacturer K and Q produced primarily cartridges of the 5.6 mm rifle type, with the strongest characteristic peaks at 3972 cm^{-1} and 4007 cm^{-1}, respectively. However, the manufacturer R and S produced 5.6 mm rifle-type cartridges, which exhibited the strongest near-infrared characteristic peaks at 3646 cm^{-1} and 3072 cm^{-1}, respectively (Fig. 4B). The extreme imbalance of all NIR data between samples poses a significant challenge when trying to accurately identify the cartridge sources.

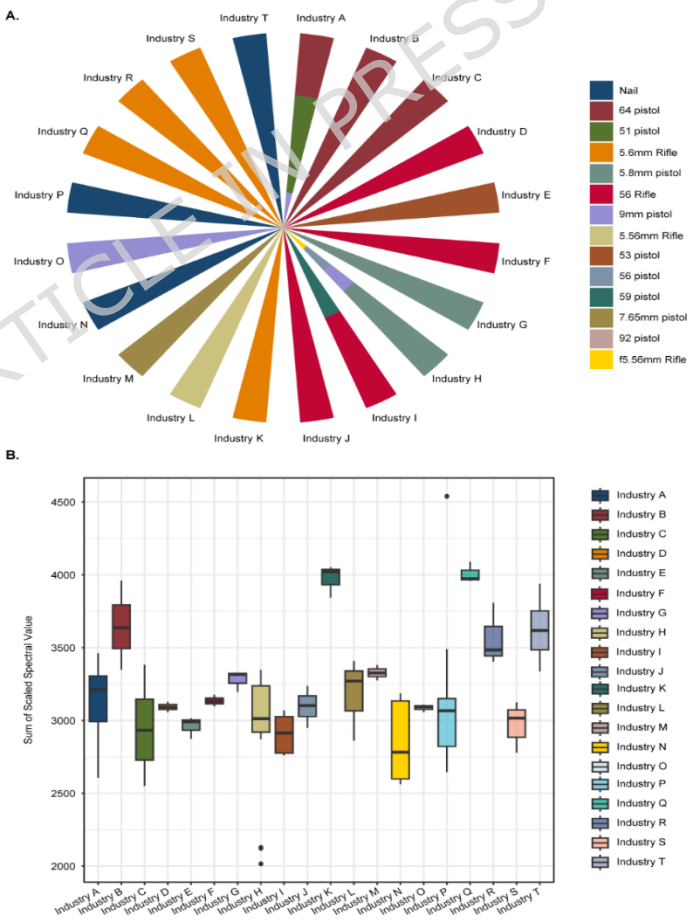

Fig.4 Distributions of (A) cartridge types and (B) characteristic peaks for different cartridge manufacturers

3.3 Model analysis of data

The Uniform Manifold Approximation and Projection (UMAP) algorithm has here been used for the dimension reduction and clustering analyses. The results showed that cartridges from manufacturers C, P, and N, which showed the largest variations in averaged absorbance value, were closely clustered in the low-dimensional projection space. However, the cartridges from other manufacturers were clustered together (Fig. 5), which made them difficult to distinguish. Furthermore, the NIR spectra of cartridges from different manufacturers showed some differences, but the distinguishing features were not clear. It was, therefore, a challenge to establish clear classification boundaries for an accurate and systematic categorization of samples for statistical and dimensionality reduction clustering analyses. With the large similarity and imbalance of NIR data for all cartridges, it was impossible to make an accurate differentiation based only on a simple comparison of NIR spectra and descriptive statistics. It was, instead, necessary to use more complex algorithmic methods for the NIR data analyses.

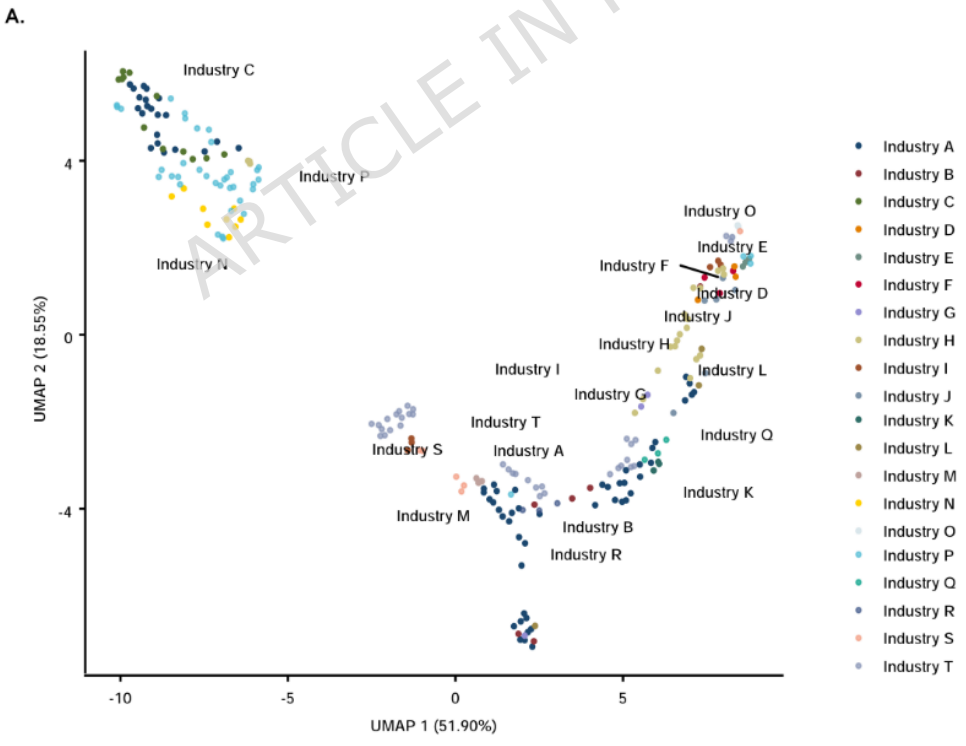

Fig 5. (A)UMAP-based clustering of NIR spectral data for smokeless powder samples from different manufacturers. Each point represents one sample, and colors indicate manufacturer

identity.

3.4 Algorithmic modeling

With the aim of developing a comprehensive and accurate classification model, a deep-learning neural network method (NN) has in the present study been used to analyze the spectra of different cartridges from various manufacturers. For evaluation purposes, the performance of this model was then compared with four classical traditional machine learning models: Stochastic Gradient Descent Classifier (SGDC), Support Vector Machine (SVM), K-Nearest Neighbors Classifier (KN), and Random Forest Classifier (RF). The training and test data were divided with a ratio of 9:1, and 10-fold cross-validation was used to mitigate the evaluation error.

The results showed that the average accuracy in the discrimination of sources of different cartridges from 20 different manufacturers was all below 60%. This result was based on their simulated near infrared spectra and by using the four traditional machine learning models. In contrast, the neural network modeling approach resulted in a corresponding average accuracy of over 80% (Fig. 6). Although the neural network model achieved an average classification accuracy exceeding 80%, a subset of samples remained misclassified. These misclassifications were primarily associated with manufacturers whose NIR spectral profiles exhibited a high degree of similarity, suggesting that overlapping chemical characteristics and formulation similarities pose challenges, even for deep learning-based approaches. No systematic bias toward a specific cartridge caliber or production year was observed among the misclassified samples, indicating that spectral similarity, rather than metadata-related factors, predominantly contributed to the classification difficulty.

Compared with traditional machine learning models, including SGDC, SVM, KNN, and RF, the neural network model demonstrated superior discriminatory performance. Conventional models typically rely on linear decision boundaries or distance-based metrics, which are less effective when applied to highly correlated and overlapping NIR spectral data. In contrast, neural networks are capable of capturing complex non-linear relationships and subtle spectral

variations, thereby enabling improved classification of SLPs from different manufacturers. These findings are consistent with the clustering trends observed in the UMAP analysis, in which only limited separation was achieved using linear dimensionality reduction alone.

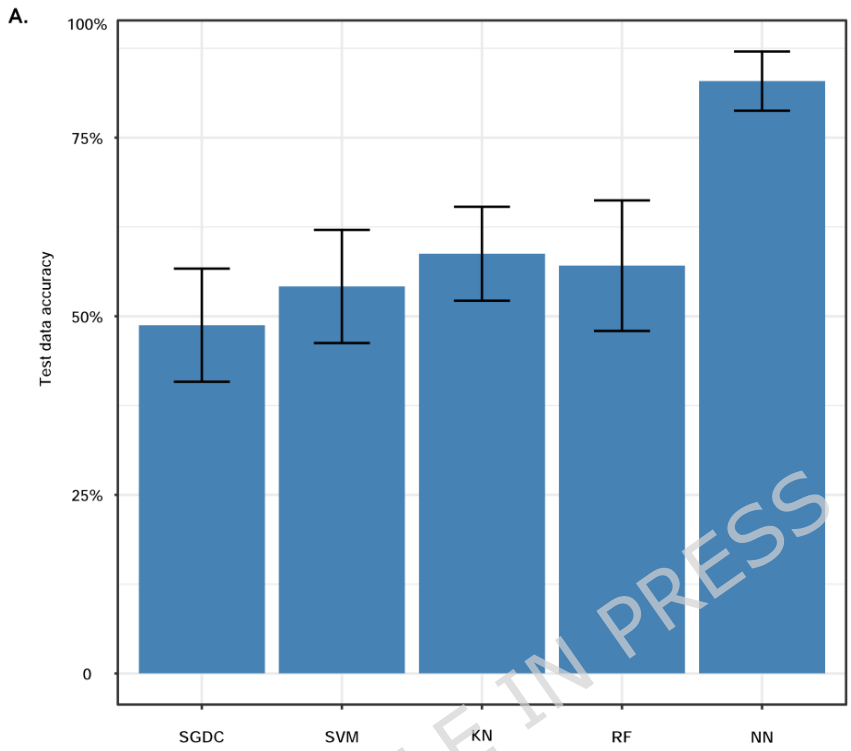

Fig.6 (A)Comparison of average classification accuracies obtained using the neural network (NN) model and four traditional machine learning models (SGDC, SVM, KNN, and RF). Error bars represent the standard deviation of accuracy across cross-validation runs.

To further enhance the scientific transparency and interpretability of the proposed model, two key analytical results are provided in the Supplementary Information. Supplementary Figure 1 presents a comparative accuracy assessment between the proposed neural network model and benchmark methods including PLS-DA and 1D-ResNet, demonstrating the superior performance of our approach in discriminating smokeless powders from different manufacturers. Supplementary Figure 2 illustrated the SHAP-based feature importance analysis, which highlights the near-infrared spectral regions most influential to model decisions-particularly the overtone and combination bands of O-H and C-H stretching vibrations-thereby confirming that the model's learning process aligns with established chemical knowledge.

4. Conclusion

In this study, near-infrared spectroscopy combined with chemometric and deep learning approaches was investigated for the differentiation of SLPs from cartridges produced by different manufacturers. The results demonstrate that, despite the high similarity and imbalance of the NIR spectral data, the neural network model achieved an average classification accuracy exceeding 80%, outperforming several traditional machine learning methods. These findings indicate that deep learning is capable of capturing subtle non-linear spectral features that are not readily resolved by conventional statistical or linear classification approaches.

It should be noted that the present study focused on intact SLPs extracted from cartridges, and the applicability of the proposed method to fired propellant residues was not assessed. In addition, limitations related to sample size, manufacturer imbalance, and the absence of spectral pre-processing may influence model robustness. Future work will therefore focus on expanding the dataset, incorporating appropriate spectral pre-processing strategies, conducting blind validation using independent samples, and extending the approach to the analysis of fired residues. Overall, the integration of near-infrared spectroscopy and deep learning provides a promising, rapid, and non-destructive screening tool that complements existing chromatographic and mass spectrometric methods for the forensic source differentiation of SLPs.

Data availability statement

The data that supports the findings of this study are available in the supplementary material of this paper.

CRediT authorship contribution statement:

Hongling Guo: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft.

Haoyuan Shi: Methodology, Formal analysis, Investigation, Writing - original draft.

Yinghua Feng: Formal analysis, Investigation

Yiting Guo: investigation

Xiuli zhang: investigation

Ping Wang: investigation

Can Hu: investigation

Hongcheng Mei: investigation

Yajun Li: investigation

Jun Zhu: Writing -review & editing.

Funding declaration

The work was funded by the grant of Ministry of Public Security under Grant Basic Work Plan for Strengthening Police Force of Ministry of Public Security(2022JC12),Grant Central Public-Interest Scientific Institution Basal Research Fund(2024JBGS002),Grant Double Ten Project of Ministry of Public Security, China (2021SSGG028) and Grant Technology Research Project of the Ministry of Public Security, China (2021JSYJC08).

References

[1] Heramb, R. M. & McCord, B. R. The manufacture of smokeless powders and their forensic analysis: A brief review. Forensic Sci. Commun. **4,** 1-7 (2002).

[2] Álvarez, Á., Yáñez, J., Contreras, D., Saavedra, R., Sáez, P. & Amarasiriwardena, D. Propellant's differentiation using FTIR-photoacoustic detection for forensic studies of improvised explosive devices. *Forensic Sci. Int.* **280**, 169-175 (2017).

[3] Lennert, E. & Bridge, C. Analysis and classification of smokeless powders by GC-MS and DART-TOFMS. Forensic Sci. Int. **292,** 11-22 (2018).

[4] Van den Hurk, R. S., Abdulhussain, N., Van Beurden, A. S. A., Dekker, M. E., Hulsbergen, A., Peters, R. A. H., Pirok, B. W. J. & Van Asten, A. C. Characterization and comparison of smokeless

powders by on-line two-dimensional liquid chromatography. J. Chromatogr. A **1672,** 463072 (2022).

[5] Heramb, R. M. & McCord, B. R. The manufacture of smokeless powders and their forensic analysis: A brief review. *Forensic Sci. Commun*. **4**, 1-7 (2002).

[6] Goudsmits, E., Sharples, G. P. & Birkett, J. W. Preliminary classification of characteristic organic gunshot residue compounds. Sci. Justice **56**, 421-425 (2016).

[7] ASTM E2999-17. Standard test method for analysis of organic compounds in smokeless powder by gas chromatography-mass spectrometry and Fourier transform infrared spectroscopy. ASTM International (2017).

[8] López-López, M., Ferrando, J. L. & García-Ruiz, C. Comparative analysis of smokeless gun powders by Fourier transform infrared and Raman spectroscopy. Anal. Chim. Acta **717**, 92-99 (2012).

[9]Thomas, J. L., Lincoln, D. & McCord, B. R. Separation and detection of smokeless powder additives by ultra performance liquid chromatography with tandem mass spectrometry (UPLC/MS/MS). J. Forensic Sci. **58**, 609-615 (2013).

[10] Assner, A. L. & Weyermann, C. LC-MS method development and comparison of sampling materials for the analysis of organic gunshot residues. Forensic Sci. Int. **264**, 47-55 (2016).

[11] Lemmink, I. B., van Schaik, J. R., van den Hurk, R. S. & Van Asten, A. C. Characterizing nitrocellulose by nitration degree and molecular weight: A user-friendly yet powerful forensic comparison method for smokeless powders. Forensic Chem. **42**, 100637 (2025).

[12] Kesic, B., McCann, N., Bowerbank, S. L., Standley, T., Liechti, J., Dean, J. R. & Gallidabino, M. D. Forensic profiling of smokeless powders (SLPs) by gas chromatography-mass spectrometry (GC-MS): A systematic investigation into injector conditions and their effect on the characterization of samples. Anal. Bioanal. Chem. **416**, 1907-1922 (2024).

[13] Dennis D. M. K., Williams M. R., Sigman M. E. Assessing the evidentiary value of smokeless powder comparisons. Forensic Sci. Int. **259**, 179-187 (2016).

[14] Emily Lennert , Candice Bridge. Analysis and classification of smokeless powders by GC–MS and DART-TOFMS. Forensic Sci. Int. **292,** 11-22(2018).

[15] López-López, M., Fernández de la Ossa, M. Á. & García-Ruiz, C. Fast analysis of complete macroscopic gunshot residues on substrates using Raman imaging. Appl. Spectrosc. **69**, 889-893

(2015).

[16] Dou, Y., Qu, N., Wang, B., Chi, Y. Z. & Ren, Y. L. Simultaneous determination of two active components in compound aspirin tablets using principal component artificial neural networks (PC-ANNs) on NIR spectroscopy. Eur. *J. Pharm. Sci.* **32,** 193-199 (2007).

[17] Ouyang, Q., Zhao, J. & Chen, Q. Measurement of non-sugar solids content in Chinese rice wine using near infrared spectroscopy combined with an efficient characteristic variables selection algorithm. *Spectrochim. Acta Part A*. **151,** 280-285 (2015).

[18] Reboucas, M. V., Santos, J. B. D., Domingos, D. & Massa, A. R. C. G. Near-infrared spectroscopic prediction of chemical composition of a series of petrochemical process streams for aromatics production. *Vib. Spectrosc.* **52**, 97-102 (2010).

[19] Judge, M. D. The application of near-infrared spectroscopy for the quality control analysis of rocket propellant fuel pre-mixes. *Talanta.* **62**, 675-679, (2004).

[20] Zou, Q., Deng, G., Guo, X., Jiang, W. & Li, F. A green analytical tool for in-process determination of RDX content of propellant using the NIR system. *ACS Sustain. Chem. Eng.* **1**, 1506-1510 (2013).

[21]Zhou, S., Wang, Z. Q., Lu, L. M., Yin, Q. S., Yu, L. H. & Deng, G. D. Rapid quantification of stabilizing agents in single-base propellants using near infrared spectroscopy. *Infrared Phys. Technol.* **77**, 1-7 (2016).

[22] Guo, Z. Q., Ren, Q., Huang, Y. Z. & Dong, S. L. Application of near infrared spectroscopy in determination of components of detonator. Chin. J. Spectrosc. Lab. 23, 187-190 (2006).