chemosensors

MDPI

Article

Ensemble Learning Combined with Laser-Induced Breakdown Spectroscopy for Detecting Pesticide Residues in Xinhui Dried Tangerine Peel

Wenhao Bi [1,†], Dongxin Shi [2,†], Feifei Wang [3], Yuxiao Song [3], Jing Sun [2,*] and Chenyu Jiang [1,2,*]

[1] School of Medicine and Information Engineering, Shandong University of Traditional Chinese Medicine, Jinan 250001, China; 2023110535@sdutcm.edu.cn
[2] Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, China; shidx@sibet.ac.cn
[3] Jinan Guoke Medical Technology Development Co., Jinan 250001, China; wangff@sibet.ac.cn (F.W.); songyuxiao@sibet.ac.cn (Y.S.)
* Correspondence: sunj@sibet.ac.cn (J.S.); jiangcy@sibet.ac.cn (C.J.); Tel.: +86-18253121729 (J.S.); +86-13402617742 (C.J.)
† These authors contributed equally to this work.

Abstract

In recent years, pesticides have been widely applied in the commercial cultivation of traditional Chinese medicinal plants to increase the yield of medicinal materials. *Xinhui* dried tangerine peel (*Citri Reticulatae Pericarpium*), a common ingredient in traditional Chinese medicine, utilizes the citrus peel as its medicinal part. During cultivation, the peel is directly exposed to pesticides, making it susceptible to pesticide residue accumulation. To enable the rapid identification of pesticide types and their targeted removal, this study integrated laser-induced breakdown spectroscopy with ensemble learning algorithms. Three lightweight neural network models—1D-CNN, Res-CNN, and LIBS-UNet—were developed and trained using either a single loss function or a composite loss function. The 1D-CNN, Res-CNN, and LIBS-UNet models achieved accuracies of 97.50% and 98.69%, 95.00% and 95.73%, and 74.06% and 76.88% for the single loss and composite loss functions, respectively. During the model ensemble stage, individual models were weighted according to their classification accuracy and test similarity matrices. Through this approach, the pesticide identification accuracy reached 99.99%. This study demonstrates that ensemble learning can effectively integrate the strengths of multiple weak classifiers, thereby significantly enhancing classification performance and providing a novel approach for the rapid detection of pesticide residues in traditional Chinese medicine ingredients.

Keywords: Xinhui dried tangerine peel; laser-induced breakdown spectroscopy; ensemble learning; lightweight neural network models; qualitative analysis

Check for updates

Received: 21 February 2026
Revised: 19 April 2026
Accepted: 28 April 2026
Published: 14 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

With the accelerated commercialization of traditional Chinese medicine, medicinal material sources have gradually shifted from the harvesting of wild plants to large-scale cultivation [1]. Herbicides, insecticides, and preservatives are widely applied during cultivation, storage, and transportation to enhance yields and resistance to pests and diseases. The heavy metals and toxic substances introduced through these practices, if not effectively identified and selectively removed, may severely compromise the quality of the finished medicinal products and pose significant risks to clinical safety. Currently,

analytical techniques such as chromatography and mass spectrometry are the primary methods used for drug safety assessment [2–4]. Although these methods provide high analytical accuracy, they are relatively costly and require complex sample pretreatment; therefore, they are difficult to implement for in situ and rapid detection.

Laser-induced breakdown spectroscopy (LIBS) enables the identification and quantification of elemental constituents by capturing plasma emissions generated from different materials [5]. Owing to its minimal sample preparation requirements and capability for real-time in situ atomic spectroscopic analysis, LIBS has been increasingly applied in recent years in fields such as geological exploration, pharmaceutical safety, and the detection of heavy metals in food products.

With the rapid development of deep learning techniques in recent years, numerous neural network models have been integrated with LIBS for material analysis. Castorena et al. [6] employed deep convolutional networks to achieve high-accuracy quantitative analysis of Martian soil LIBS data. Similarly, Yang et al. [7] utilized a five-layer convolutional neural network (CNN) to classify 12 types of Martian rock samples. Ma et al. [8] enhanced convolutional neural networks by incorporating long short-term memory (LSTM) units and attention mechanisms to quantitatively analyze LIBS data. However, most of these studies focused on increasing the complexity of individual models. Although such designs improve representational capacity to some extent, they often neglect computational efficiency and lightweight deployment capabilities, making them poorly suited for handheld LIBS devices.

In contrast, ensemble learning can integrate multiple lightweight weak classifiers and assign them different voting weights, thereby overcoming the limitations of single learners, such as structural complexity and susceptibility to local minima [9]. In existing studies combining LIBS with advanced analytical methods, Chu et al. [10] employed a random subspace (RS) strategy to integrate linear discriminant analysis (LDA) for the qualitative discrimination of blood cancer samples; Shih et al. [11] integrated LDA, support vector machines (SVM), partial least squares discriminant analysis (PLS-DA), and random forest (RF) models, achieving classification accuracies substantially higher than those of the individual sub-models; and Wang et al. [12] compared ensemble tree models (extra trees, ETs) with conventional machine learning approaches, including RF, SVM, and gradient boosting decision trees (GBDTs), and found that ensemble learning achieved higher accuracy than traditional machine learning methods in distinguishing different stages of Omicron infection. By leveraging machine learning to address the inherent instability of LIBS signals, ensemble learning reduces the reliance on extensive prior knowledge required in traditional analytical approaches. To a certain extent, this approach enhances the classification performance of neural network models while distributing the computational burden required for training.

In this study, Xinhui dried tangerine peel treated with different pesticides was selected as the research object. Elemental information on pesticide residues was analyzed using LIBS in combination with ensemble learning strategies. Three weak classifiers—1D-CNN, Res-CNN, and LIBS-UNet—were constructed, and an ensemble weighting method integrating test-result similarity matrices with individual model accuracies was proposed. A performance evaluation based on classification accuracy demonstrated that the proposed approach enables effective multi-model integration for the qualitative analysis of pesticide residues in traditional Chinese medicinal materials. The results indicate that ensemble learning can achieve higher pesticide residue discrimination accuracy without increasing the complexity of individual models, thereby providing a reliable analytical basis for the subsequent selection of appropriate processing methods.

2. Experiments

2.1. Experimental Design and Theoretical Basis

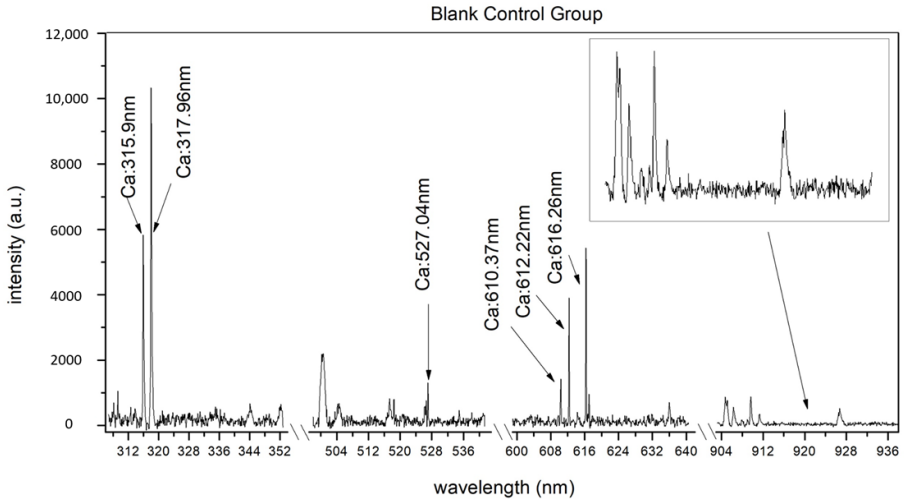
2.1.1. Selection of Excitation Wavelength Band

Based on the NIST database and previous studies [13], analytical emission lines were selected by comprehensively considering factors such as large spontaneous transition probabilities (A_{ki}), relatively well-defined spectral line shapes, stable peak intensities, and the avoidance of detector overexposure or damage. Accordingly, the analytical emission lines selected for Cu, Ca, S, and Zn in this study, based on Refs. [14,15], are summarized in Table 1. Based on the existing equipment, the emission spectra in this study were obtained in four separate spectral bands. Each spectral band contains 500 dimensions (spectral data points), resulting in a total of 2000 dimensions.

Table 1. Selection results for the wavelength bands containing the excitation peaks of the elements to be analyzed.

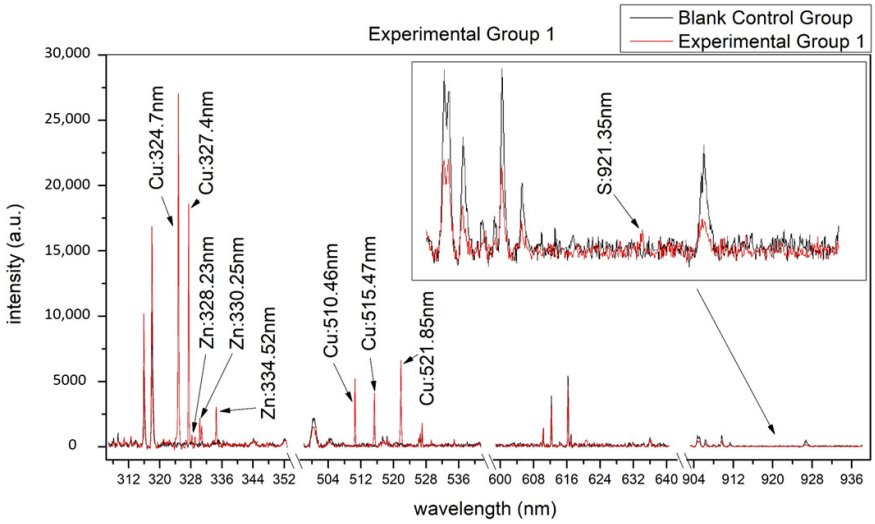
Element	Wavelength (nm)	Spontaneous Transition Probability ($A_{ki}/10^8$ s^{-1})
Cu	324.7	1.395
	327.4	1.376
	510.46	0.020
	515.47	0.600
	521.85	0.750
Zn	328.23	0.900
	330.25	1.200
	334.52	1.700
	636.21	0.470
Ca	315.9	3.100
	527.04	0.500
	535.07	0.560
	612.22	0.287
	616.26	0.477
S	921.35	0.279
	922.8	0.277
	923.8	0.277

Using the LIBS system to excite the spectral regions containing the target peaks of each sample, the observed LIBS spectra are shown in the Figure 1 below.

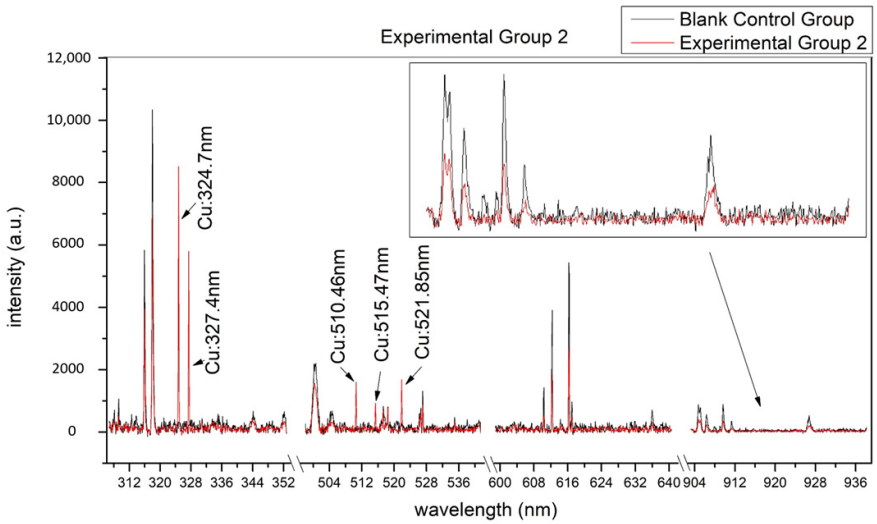


(**a**)

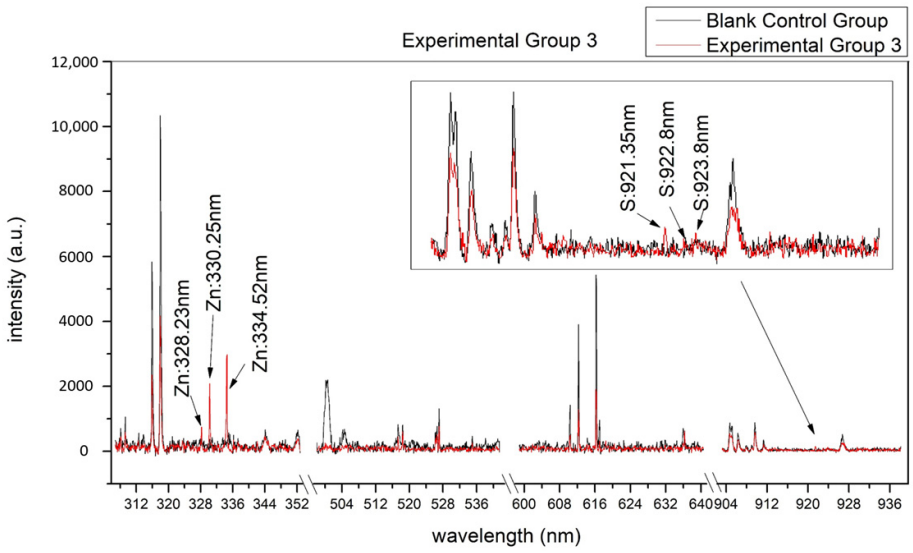
Figure 1. *Cont.*



(**b**)

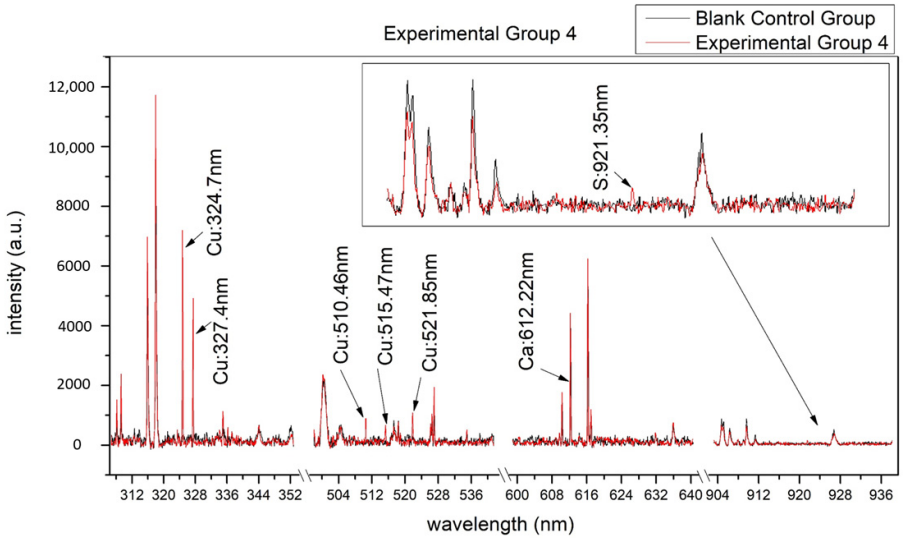


(**c**)



(**d**)

Figure 1. *Cont.*



(**e**)

Figure 1. (**a**) Blank control group LIBS spectrum; (**b**) Sample containing copper-based pesticide residues (basic copper chloride) (Experimental Group 1); (**c**) Sample containing copper-based pesticide residues (copper hydroxide) (Experimental Group 2); (**d**) Sample containing zinc-based pesticide residues (thiazole zinc) (Experimental Group 3); (**e**) Sample containing copper- and calcium-based pesticide residues (copper calcium sulfate) (Experimental Group 4).

2.1.2. Weak Classifier Design

When designing the weak classifier models, we aimed to improve generalization performance with fewer network layers, thereby attaining relatively high training accuracy and low loss values while minimizing computational resource consumption on a single device. In contrast, many models developed for natural language processing and image processing struggle to satisfy the requirements of ensemble learning for LIBS data in terms of both complexity and practicality. As a commonly used neural network in the LIBS field, LIBS-UNet has inherent advantages in analyzing LIBS data. On this basis, we developed a structurally simpler 1D-CNN model, consisting of two convolutional layers. After the first convolution, the output is activated by the softmax function and fed into the second convolutional layer for deeper feature extraction. The extracted features are then passed through a regularization layer to prevent the outputs of the second convolution from exceeding the sensitive activation range. Finally, a classifier comprising the softmax activation function, a flattening layer, and a fully connected layer performs nonlinear mapping to classify different Xinhui dried tangerine peel samples. Beyond the linear network architecture, a residual structure was incorporated into the layer combination, yielding a model called Res-CNN. In this architecture, features extracted by the convolutional layers are fused at the feature level with the regularized original input, and the integrated features are subsequently classified by a unified classifier. The overall structure of the Res-CNN model is illustrated in Figure 2.

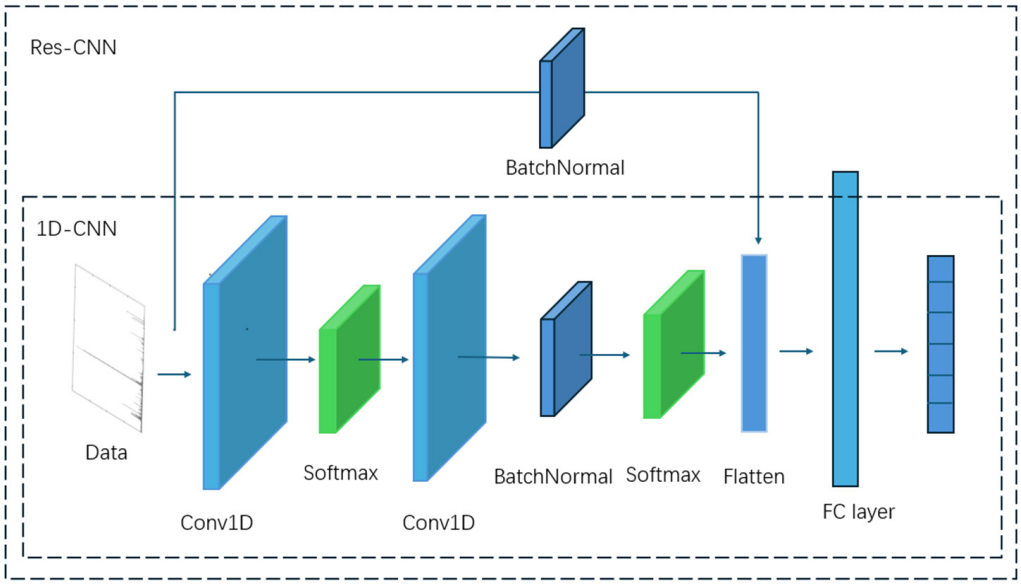


Figure 2. Schematic of the 1D-CNN and Res-CNN model structures.

2.1.3. Training Strategy Design and Selection of Evaluation Metrics

Regarding the configuration of the loss functions and backpropagation strategies, two loss-handling schemes were separately applied to the three models. Considering the similar physicochemical properties of individual tangerine peel, in order to better distinguish the characteristics related to trace and ultrafine elements, the first training strategy employed the mean squared error (MSE) loss function, which directly optimizes inter-class distances, in combination with backpropagation to train the models. The MSE loss is calculated as follows:

$$MSE = \frac{\sum_{i=1}^{n} (\mathrm{f(x)} - \mathrm{y})^2}{n} \tag{1}$$

where n denotes the number of samples, $f(x)$ represents the predicted value, and y denotes the ground-truth value. The predicted value is the output result calculated by the model based on the input data (denoted as $f(x)$ in this paper), and the true value is the standard answer or actual observation value of the data (denoted as y in this paper). In the process of model training and evaluation, the model is continuously optimized by comparing the difference between the predicted value and the true value, so that the predicted result is as close to the true value as possible.

The second strategy builds upon the first by incorporating an additional cross-entropy loss term. The cross-entropy loss imposes a nonlinear penalty on high-confidence misclassifications and emphasizes the correctness of classifications rather than the magnitude of prediction errors. The five-class cross-entropy loss is computed as follows:

$$L_{CE} = -\sum_{i=1}^{n} \mathrm{y}_i \ln \left(\frac{\exp(\mathrm{z_i})}{\sum_{\mathrm{j}=1}^{n} \exp(\mathrm{z_j})} \right) \tag{2}$$

where z_i denotes the logit value of the i-th class, representing the network's output score for that class, and y_i denotes the corresponding ground-truth label.

A weighted combination of the two losses was adopted as the evaluation criterion [16]. The weighting hyperparameters were selected via grid search, with a ratio of 2:8 assigned to the two loss functions in the composite loss used for ensemble training, thereby further

optimizing model accuracy and specificity. During backpropagation, only the weighted composite loss propagates through the network, rather than propagating each individual loss function separately. The overall loss is calculated as follows:

$$\mathrm{Loss_{all}} = \mathrm{Weight_1} \times \mathrm{Loss_1} + \mathrm{Weight_2} \times \mathrm{Loss_2} \tag{3}$$

where $Loss_{all}$ denotes the overall loss propagated at each epoch; $Weight_1$ and $Weight_2$ represent the weights assigned to the first and second loss functions, respectively; and $Loss_1$ and $Loss_2$ are the loss values computed from the first and second loss functions, respectively.

The commonly used accuracy metric (Acc) was adopted as the evaluation criterion for model training, defined as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where TP denotes the number of true positives, FN the number of false negatives, FP the number of false positives, and TN the number of true negatives.

2.1.4. Ensemble Method Design

In the multi-model voting decision process of ensemble learning, we comprehensively considered model accuracy as well as the influence of inter-model similarity on the voting process and final decision outcomes. To assess the weight of each weak classifier, the proportion of its classification accuracy relative to the sum of the accuracies across all weak classifiers was calculated, and this proportion was used as the initial weight. Subsequently, the prediction results of each weak classifier on the test dataset were analyzed. For each pair of weak classifiers, instances in which the same misclassification occurred were recorded. The pairwise error rates corresponding to identical mistakes between weak classifiers were computed based on the proportion of such misclassified samples relative to the total number of test samples. An error-rate matrix representing identical misclassification occurrences among the classifiers was constructed, and the average classification error rate of each weak classifier was used as a penalty term to reduce its assigned weight. Finally, the penalty term was subtracted from the initial weight to obtain the final assigned weight for each weak classifier. The weight assignment process is expressed as follows:

$$\mathrm{w_j} = \frac{Acc_{m_j}}{\sum_{i=1}^{n} Acc_{m_i}} - \frac{\sum_{i=1}^{n} Sim_{m_{ji}}}{n} \tag{5}$$

where w_j denotes the weight assigned to the j-th model, Acc_{m_j} represents the test-set accuracy of the j-th model, n denotes the total number of models participating in the ensemble, and $Sim_{m_{ji}}$ denotes the similarity value between the j-th and i-th models.

2.2. Experimental Apparatus and Procedure

2.2.1. Experimental Equipment

A schematic of the LIBS prototype used in this study is shown in Figure 3. The system was equipped with a water-cooled Nd:YAG Q-switched pulsed laser operating at a wavelength of 1064 nm. Maintaining a constant operating temperature of 25 °C. The power supply voltage was set to 740 V, corresponding to a laser pulse energy of 150 mJ (Power density: 2.12×10^{10} W/cm^2). The laser emitted 1 Hz pulsed radiation with a spot diameter of 4.8 mm. The pulse width is 10 ns and delay time is 2.75 μs. The laser beam is focused on the sample surface through a plano-convex lens with a coating of 1064 nm

and a focal length of 65 mm. The excited plasma enters the light collector at an Angle of 70° from the direction of the incident light. The experimental samples were placed on an x–y–z three-axis motorized translation stage. After irradiation by the high-energy pulsed laser, the plasma emission was collected by the light collector (diameter of 25.4 mm and a focal length of 100 mm) and transmitted via optical fiber to a spectrometer (Zolix Omni-λ500i, Resolution:0.11nm@435.83nm with a slit width of 10 μm, Manufacturer: Zolix from Beijing, China). A DField-VS-A2 cooled charge-coupled device (CCD, Manufacturer: Zolix from Beijing, China) was employed as the data acquisition device for spectral signal collection. Synchronization between the CCD and the Nd:YAG Q-switched laser (Manufacturer: Harbin Institute of Technology from Heilongjiang China)was achieved via a signal generator and the integration time is controlled by the gating unit. To ensure high signal quality, the CCD was cooled to −60 °C, with an integration time of 1 ms and a slit width of 100 μm. Finally, the acquired spectral data were converted into output files using proprietary optoelectronic software developed by Zolix (ZolixSoftware_v1.0.20.1).

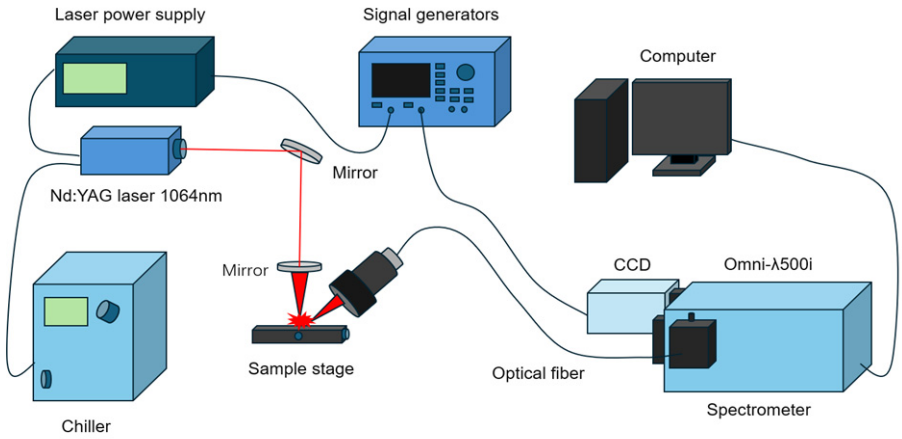


Figure 3. LIBS experimental prototype diagram.

2.2.2. Sample Preparation

Considering the potential variability among citrus samples, 15 citrus samples were used in each experimental group for subsequent sample preparation. All samples were tested for their compositional content prior to the experiments (constant volume of 5 mL in the test) to ensure compliance with the DB44/T 604-2009 standard [17]. According to the standard test, the content of heavy metal elements in the samples selected by us is within the scope of the standard "DB44/T 604-2009", and the content of As, Pb and Cd is 0.000012%, 0.000022% and 0, respectively. The results are shown in Table 2. Untreated Xinhui dried tangerine peel was used as the blank control group. For the experimental groups, the control samples were separately sprayed with the following pesticide formulations: 47% cupric oxychloride–mancozeb wettable powder (pesticide registration No. Shengxu (Su) 0079), 77% copper hydroxide wettable powder (pesticide registration No. Shengxu (Yu) 0031) [18], 30% thiazole zinc solution (pesticide registration No. Shengxu (Zhe) 0033), and 77% copper calcium sulfate wettable powder (pesticide registration No. Shengxu (Su) 0067). The dilution ratios were 1:500, 1:400, 1:500, and 1:400, respectively. Each formulation was sprayed once every ten days for a total of three treatments. The spray was applied every 10 days for a total of three times. The conventional pressure spraying device was used, the nozzle was about 1 m away from the sample, and the spraying time was about 0.5 s. The oranges in different positions and different occlusions were sprayed from multiple angles. The blank control group used the same spraying device with clean water to ensure proper control of the variables. They are then labeled as experimental groups 1, 2, 3, and 4 in turn.

Table 2. Sample qualification test table (constant volume 5 mL).

Element	Converted Concentration (mg/kg)	Mass Concentration (%)
Sn	0.2184	0.000022
Sb	0.0170	0.000002
Cu	3.7849	0.000378
Zn	6.0764	0.000608
Cl	10,224.92	1.022492
Ca	570.5274	0.057053

2.2.3. Dataset

During dataset acquisition, 15 samples were used in each group and 40–60 spectra were collected for each sample. To ensure data quality, only the first five spectra from each ablation crater were selected and included in the dataset. Ultimately, the raw dataset consisted of 3638 data records, comprising 732 data records from the blank control group, 626 from the copper calcium sulfate group, 654 from the cupric oxychloride–mancozeb group, 806 from the copper hydroxide group, and 820 from the thiazole zinc group. Each data record contains four spectral bands, with each band having a dimensionality of 500, resulting in a total of 2000 dimensions used for training. Subsequently, the alternating least squares (ALS) algorithm [19] was applied to the raw data for background subtraction and the processed data were used as the final training dataset. The background removal effect is illustrated in Figure 4.

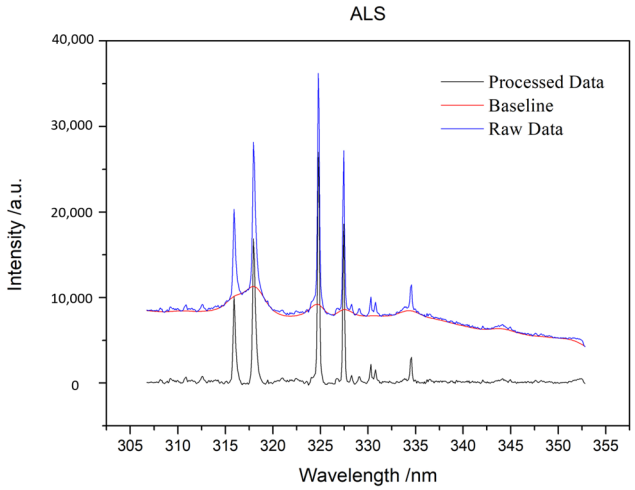


Figure 4. ALS algorithm baseline fitting removal effect.

2.2.4. Training Procedure

The three model architectures were trained for 20 epochs using two training strategies: a single loss function and a composite loss function. The initial learning rate was set to 0.01, and a cosine annealing schedule with a period of 10 was employed to dynamically adjust the learning rate during training. Considering the high-dimensional characteristics of LIBS data, the stochastic gradient descent (SGD) optimizer, which is well suited for handling high-dimensional sparse features, was selected to optimize the training process. The training dataset was partitioned using a 7:2:1 ratio through six rounds of independent random sampling, and the resulting datasets were correspondingly fed into six weak classifiers for training. Considering a spectral wavelength sampling interval of 0.0921 nm, we employed a convolution kernel with a length of 15 and a stride of 1 for the convolution operation. As it moves across the signal data, its receptive field (approximately 1.5 nm) can effectively cover most of the emission line profiles of the analyzed elements in LIBS spectra. The model was trained on a system equipped with four RTX 3090 GPUs.

3. Experimental Results

Training Performance

Each sub-model was trained independently, and for each model, the checkpoint that achieved the highest training accuracy within 20 epochs was selected. The corresponding training accuracy was recorded, and the selected model was subsequently evaluated on the test dataset. The results are summarized in the Table 3:

Table 3. Training accuracy and resource usage of weak classifiers.

Model Name	Accuracy Rate			Time/ Epoch	GPU Occupancy	Model Parameters
	Training Set	Validation Set	Test Set			
LIBS-Unet	87.46%	84.32%	74.06%	9.46 s	33.44 MB	4,246,599
1D-CNN	96.39%	95.47%	97.50%	0.3 s	0.28 MB	10,039
Res-CNN	99.73%	100%	95.00%	0.35 s	0.28 MB	10,039
Eloss-LIBS-Unet	76.65%	78.82%	76.88%	10.65 s	33.4 MB	4,246,599
Eloss-1D-CNN	99.96%	100%	98.69%	0.39 s	0.3 MB	10,039
Eloss-Res-CNN	98.24%	96.52%	95.73%	0.48 s	0.3 MB	10,039

From the training results, the minimum accuracy among the weak classifiers was 74.06%, whereas the maximum accuracy was 98.69%. Under these conditions, the distribution of model performance was relatively concentrated, and the inter-model noise levels remained within a reasonable range. A comparison of different training strategies applied to the same model architecture indicates that the models trained with the composite loss function generally achieved higher accuracy than those trained with the single loss function. This improvement can be attributed to the multi-perspective optimization objectives introduced by the composite loss, as well as the richer gradient information it provides for backpropagation during learning.

To further validate the rationality of the weak classifier selection, we selected several neural network models from general domains and appropriately adapted them to meet the input requirements of LIBS data. These models were trained under the same training framework, and their accuracies are presented in Table 4.

Table 4. Training accuracy and resource usage of the traditional network.

Model Name	Accuracy Rate			Time/ Epoch	GPU Occupancy	Model Parameters
	Training Set	Validation Set	Test Set			
AlexNet	50.67%	52.94%	59.83%	15.73 s	370.96 MB	24,254,853
ELoss-AlexNet	59.23%	58.68%	60.26%	15.7 s	370.96 MB	24,254,853
GhostNet	25.33%	27.77%	27.42%	12.23 s	64.16 MB	8,197,284
Eloss-GhostNet	30.86%	32.24%	30.68%	12.46 s	64.20 MB	8,197,284

From the observed training accuracies, it can be seen that neural network models adapted from other domains exhibit a certain degree of incompatibility when applied to LIBS data. Whether trained using an ensemble loss function or a single loss function, their classification accuracies are significantly lower than those of the three aforementioned models. In the weighted ensemble of weak classifiers, such models contribute less to the overall accuracy and may introduce additional inter-model noise. Therefore, only the 1D-CNN, Res-CNN, and LIBS-UNet models were selected for the ensemble.

At the same time, by comparing the number of parameters of the model, training time and GPU occupancy, we can see that neural network models adapted from other domains need to consume more computing power resources during training, which also puts forward relatively higher requirements for the deployment environment of the model. However, 1D-CNN, Res-CNN, LIBS-UNet are relatively more suitable for various lightweight deployment environments.

A weighted ensemble learning approach was employed to integrate 1D-CNN, Res-CNN, and LIBS-UNet, followed by comparative analysis. As shown in Table 5, the soft independent modeling approach employing the proposed weighting strategy achieved an ensemble accuracy of 99.99%, which was higher than the classification accuracy of any individual weak classifier. At the same time, it is also higher than the 95.51% accuracy of XGBoost and 93.68% accuracy of LightBGM. In contrast, under the hard independent modeling scheme, the ensemble classification accuracy was only 88.98%, which was even lower than the accuracies achieved by the 1D-CNN and Res-CNN models when used individually (97.50% and 95.00%, respectively).

Table 5. Accuracy performance of ensemble learning and machine learning.

Modeling Methods	Test Accuracy
Hard independent modeling	88.98%
Soft independent modeling	99.99%
GBDT	95.15%
XGBoost	95.51%
LightBGM	93.68%

Further analysis indicates that in multi-model ensemble learning, soft independent modeling offers advantages in terms of uncertainty quantification and tolerance of fuzzy classification boundaries, yielding higher accuracy in multiclass classification tasks compared with hard independent modeling. The lower accuracy observed for hard independent modeling may be attributed to the comparatively poor performance of the LIBS-UNet and Eloss-LIBS-UNet models within this framework. Although these two models account for one-third of all network models, they are assigned voting weights equal to those of the other classifiers, which likely dilutes the overall decision quality and degrades the ensemble performance. In contrast, soft independent modeling performs classification based on probability-weighted contributions from different models, which partially explains why the ensemble accuracy obtained with hard independent modeling is lower than that achieved using soft independent modeling.

4. Analysis and Discussion

Next, a class-gradient-weighted activation-mapping algorithm was employed to analyze the regions of focus in feature learning. Taking samples from Experimental Group 1 as an example, the first convolutional layer, which provides stronger semantic interpretability, was selected for comparative analysis.

As shown in Figures 5 and 6, during the learning process, the LIBS-UNet network assigned substantially higher weights to the emission peaks in the 315–330 nm wavelength range than to peaks in other spectral regions. These weights were primarily concentrated on the Ca emission lines at 315.9 nm and 317.95 nm, the Cu emission lines at 324.7 nm and 327.4 nm, and the Zn emission line at 328.21 nm. The emission peaks within this band exhibit relatively strong signal intensities across the entire spectrum. Outside of this range, the Cu emission lines at 510.46 nm, 515.47 nm, and 521.85 nm, as well as the Ca emission line at 616.35 nm, were assigned marginal weights, while the remaining spectral peaks were assigned weights of approximately zero. Regarding the discrimination between cupric oxychloride–mancozeb and copper calcium sulfate treatments, although Ca is present at relatively low concentrations in Xinhui dried tangerine peel, it is an essential trace metallic element for the growth of citrus plants and is therefore widely distributed in the untreated samples. Therefore, classification based predominantly on the emission peaks of Cu and Ca suffers from limited elemental specificity. Moreover, LIBS signals are influenced by factors such as laser energy fluctuations and variations in

ablation craters, which introduce peak intensity instability and interference. Combined with observations from the confusion matrix, the Grad-CAM algorithm provides a clear explanation for why LIBS-UNet performs poorly on Experimental Group 4, achieving only 28% classification accuracy. It also explains why 60% of samples from Experimental Group 1 and 12% from Experimental Group 2 are misclassified as belonging to Experimental Group 4. Additionally, the weight analysis suggests that one reason for the relatively low accuracy of the LIBS-UNet model is the assignment of excessively high weights to specific spectral bands, which may cause the model to become trapped in local minima. Consequently, integrating additional models to capture more diverse feature representations is necessary to increase discrimination accuracy. As shown in Figure 6, although we trained the LIBS-UNet network using an ensemble loss function, the recognition accuracy for Experimental Group 4 remains relatively low. Specifically, 63.16% of samples from Experimental Group 4 are misclassified as Experimental Group 1, and 36.84% are misclassified as Experimental Group 2. This further indicates that training with an ensemble loss function can only enhance the diversity among models but cannot fully overcome the structural limitations of a single model. In contrast, ensemble learning leverages the differences among models to compensate for these limitations, thereby improving overall classification accuracy.

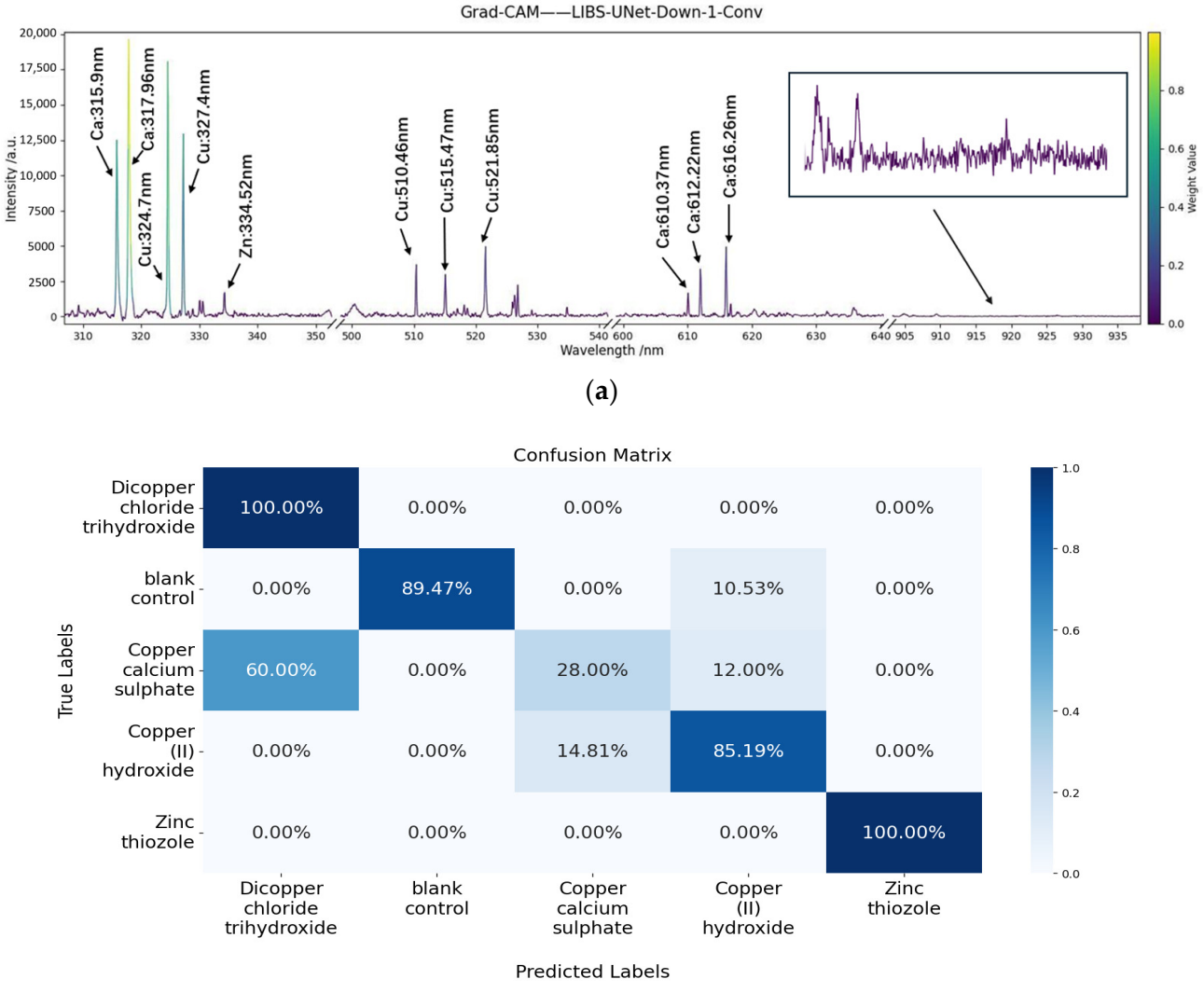


Figure 5. (**a**) LIBS-UNet model Grad-CAM region of interest map; (**b**) LIBS-UNet model confusion matrix diagram.

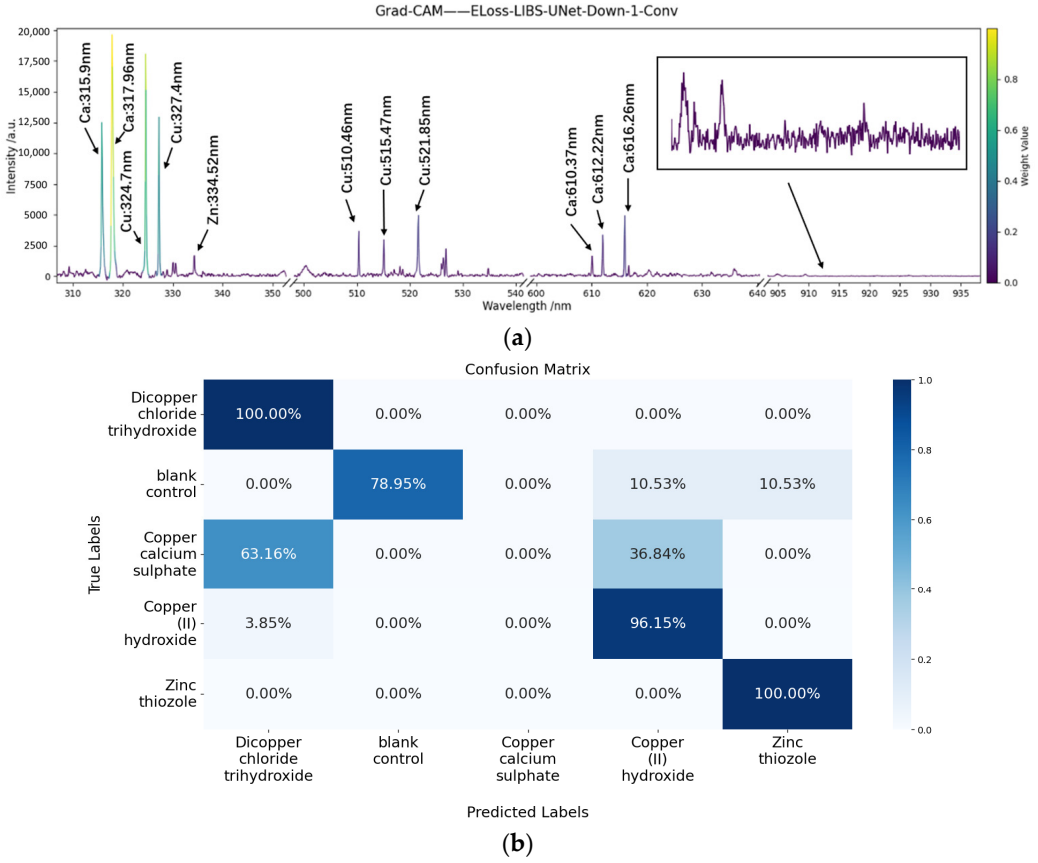


Figure 6. (**a**) Eloss-LIBS-UNet model Grad-CAM region of interest map; (**b**) ELoss-LIBS-UNet model confusion matrix diagram.

The analysis of the regions of interest in the 1D-CNN model (as shown in Figures 7 and 8) indicates that this relatively uniform weight allocation strategy achieves good training performance on local LIBS spectral data with relatively low dimensionality. Due to the similar weight distribution across spectral bands, the 1D-CNN exhibits comparable classification accuracy for each class under both training schemes, all exceeding 92%. However, there are still cases where certain classes cannot be correctly identified when using a single model. Moreover, this weighting strategy imposes stricter requirements on data quality and generalization performance. Therefore, for high-dimensional data, it is necessary to incorporate the discriminative features learned by LIBS-UNet into the ensemble to achieve more robust classification.

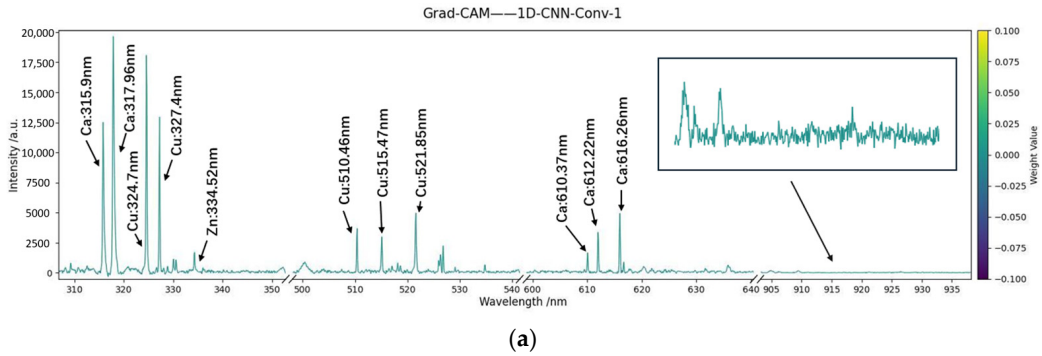
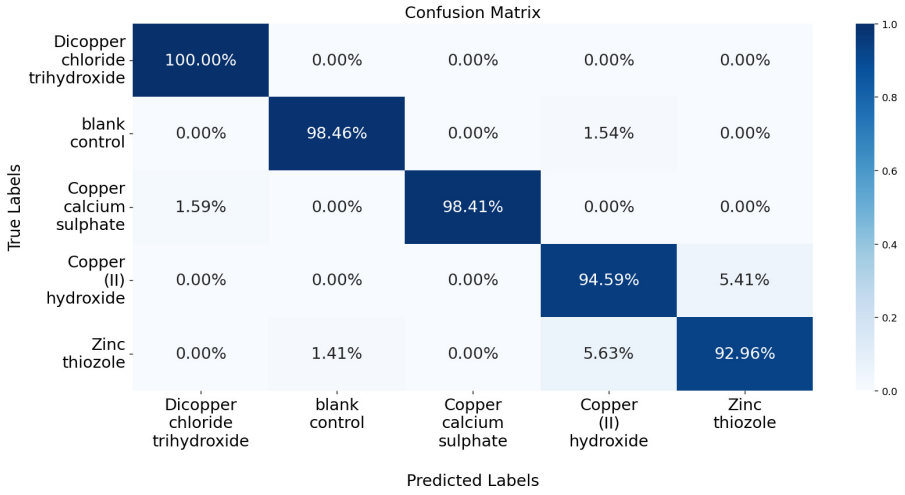
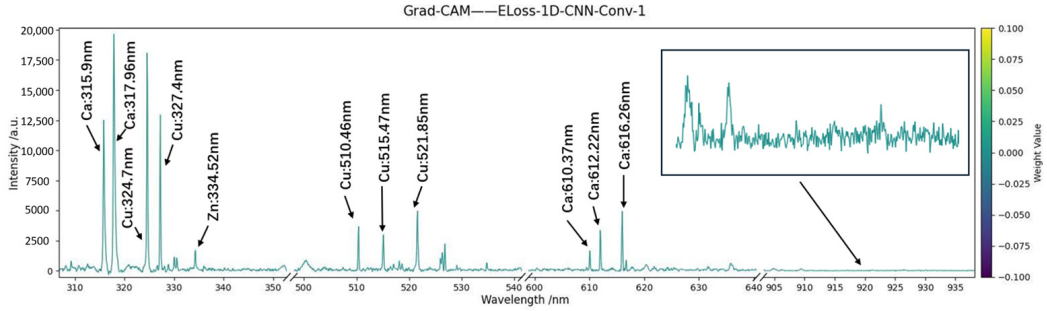


Figure 7. *Cont.*

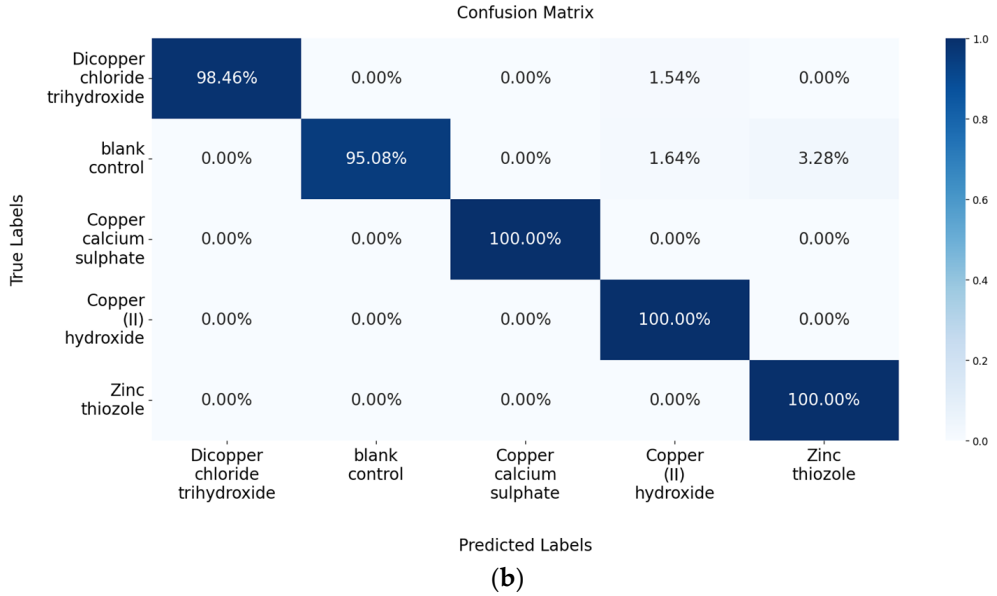


(**b**)

Figure 7. (**a**) 1D-CNN model Grad-CAM region of interest map; (**b**) 1D-CNN model confusion matrix diagram.



(**a**)



(**b**)

Figure 8. (**a**) ELoss-1D-CNN model Grad-CAM region of interest map; (**b**) ELoss-1D-CNN model confusion matrix diagram.

As illustrated in Figures 9 and 10, under both loss-function settings, the Res-CNN model exhibited a weighting pattern for LIBS signals that was highly similar to that of the 1D-CNN model, with weight magnitudes distributed within ±0.03. This phenomenon can be partially attributed to the identical layer composition of the two networks and the

relatively weak initial influence of the skip connections. The accuracy distribution reflected in the confusion matrix shows a clear difference from that of the 1D-CNN. Under the two loss training strategies, the Res-CNN achieves an accuracy of 84.51% for the blank control group and 97.67% for Experimental Group 3 when trained with the ensemble loss. In contrast, under a single loss function, the accuracy for Experimental Group 3 is 80%, while the accuracy for the blank control group reaches 100%. These results form a strong complementary relationship in terms of accuracy. While maintaining an overall accuracy comparable to that of the 1D-CNN, Res-CNN provides additional discriminative perspectives for the ensemble classification process.

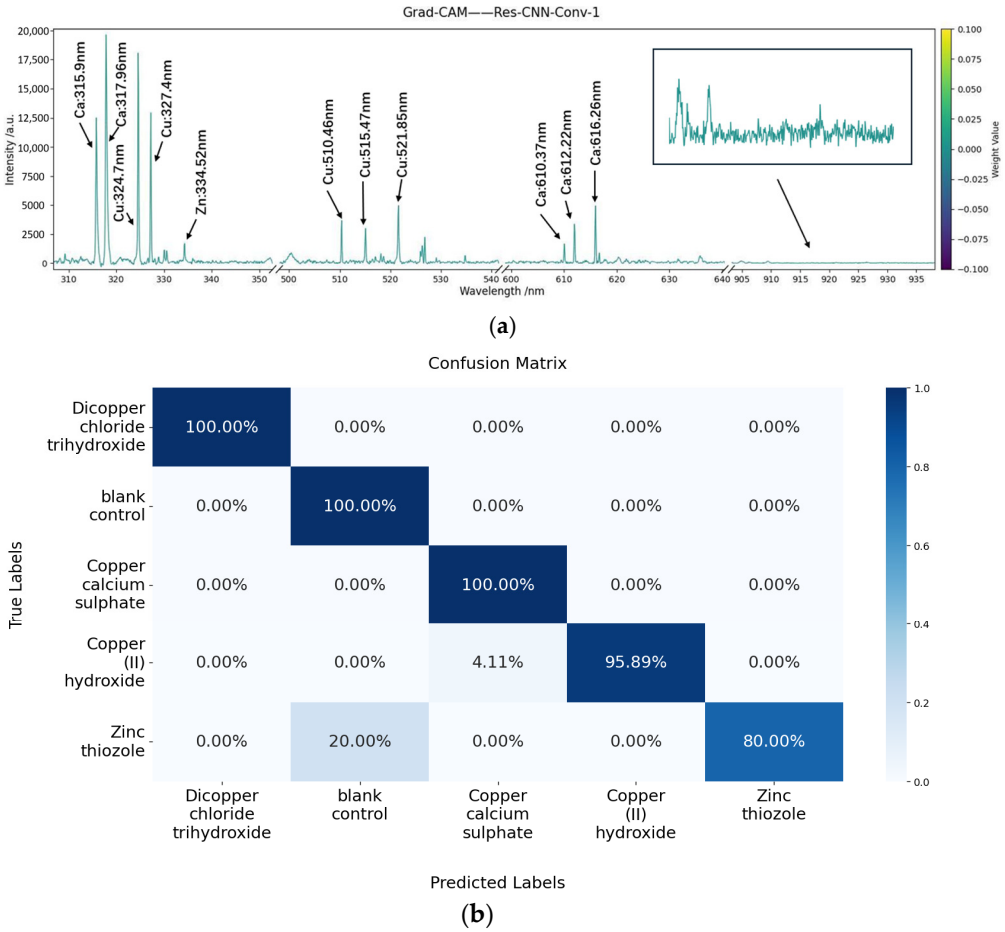


Figure 9. (**a**) Res-CNN model Grad-CAM region of interest map; (**b**) Res-CNN model confusion matrix diagram.

The limitations of the three types of weak classifiers can be mitigated to some extent through ensemble training. When the test dataset contained relatively high noise levels, the concentrated feature analysis capability of the LIBS-UNet network partially offset the influence of noise on the decisions of other weak classifiers, whereas the 1D-CNN and Res-CNN models assigned weights to excitation peaks in other spectral regions and learned features that were not captured by the LIBS-UNet network. This complementarity and diversity among the weak ensemble classifiers were also reflected in the accuracy results. Compared with the individual weak classifiers, whose accuracies ranged from 74.06% to 98.69%, the multi-model ensemble achieved an accuracy of 99.99%, representing improvements at different levels over any single model. In addition, multiple weak classifiers provide multiple perspectives for ensemble discrimination. When the data quality changes due to changes in the environment, ensemble learning to a certain extent avoids the negative impact of the limitations of a single model on the task.

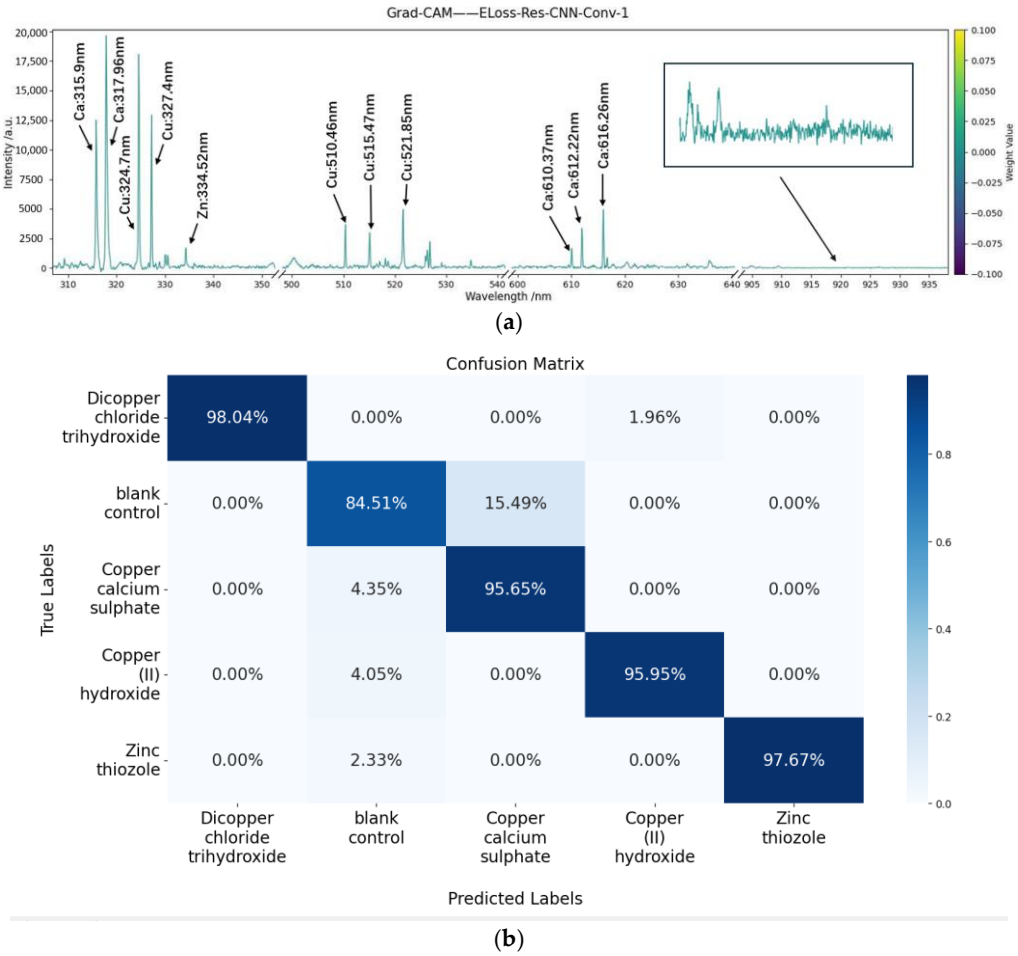


Figure 10. (**a**) ELoss-Res-CNN model Grad-CAM region of interest map; (**b**) ELoss-Res-CNN model confusion matrix diagram.

5. Conclusions

The qualitative analysis of pesticide residues in traditional Chinese medicines is of great significance for drug safety and even the protection of human life. With the introduction of LIBS technology and its integration with deep learning techniques, both identification speed and identification cost have been substantially improved compared with those of traditional analytical methods. In this study, serial and residual CNN architectures were combined with the LIBS domain-specific model LIBS-UNet to analyze pesticide application scenarios under laboratory conditions, simulating the cultivation of Xinhui dried tangerine peel. By adopting soft independent modeling, an accuracy of 99.99% was achieved, representing an improvement of 11.01% compared with hard independent modeling, whereas weak single classifiers achieved accuracy improvements ranging from 1.3% to 25.93%. The horizontally distributed nature of ensemble learning reduces the requirements for hardware computational capacity and GPU memory, thereby improving the portability of the model.

Author Contributions: Conceptualization, J.S. and C.J.; methodology, J.S., W.B. and D.S.; validation, W.B., D.S. and Y.S.; formal analysis, W.B. and F.W.; investigation, D.S.; resources, J.S. and C.J.; data curation, W.B., D.S., Y.S. and F.W.; writing—original draft preparation, W.B. and D.S.; writing—review and editing, Y.S.; visualization, F.W.; supervision, J.S. and C.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key Research and Development Program (NKRDP) (2022YFC2406204), the National Key Research, Shandong Province Key Research and Development Plan (2023CXPT041 and 2024CXPT054), the Taishan Industrial Experts Program (tscx202306125), Social welfare projects of Jinan City and Science and Technology for the Benefit of the People Project in Qingdao City, Shandong Province (25-1-5-smjk-5-nsh).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: Author Feifei Wang and Yuxiao Song was employed by the company Jinan Guoke Medical Technology Development Co. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LIBS	Laser-induced breakdown spectroscopy
HPLC-MS/MS	High-performance liquid chromatography–tandem mass spectrometry
GC-MS/MS	Gas chromatography–tandem mass spectrometry
LC/MS-MS	Liquid chromatography–tandem mass spectrometry
SPE	Solid-phase extraction
CNN	Convolutional neural network
LSTM	Long short-term memory
RS	Random subspace
LDA	Linear discriminant analysis
SVM	Support vector machine
PLS-DA	Partial least squares discriminant analysis
RF	Random forest
ET	Extra trees
GBDT	Gradient boosting decision tree
NIST	National Institute of Standards and Technology
MSE	Mean squared error
TP	True positive
TN	True negative
FP	False positive
FN	False positive
CCD	Charge-coupled device
ALS	Alternating least squares
SGD	Stochastic gradient descent
Eloss-CNN	EnsembleLoss-Convolutional Neural Network
Eloss-Res-CNN	EnsembleLoss-Residual-Convolutional Neural Network
Eloss-LIBS-UNet	EnsembleLoss-LIBS-UNet
Grad-CAM	Gradient-weighted class activation mapping

References

1. Zhang, R.; Zhang, M.; Chen, Y.; Wang, C.; Zhang, C.; Heuberger, H.; Li, H.; Li, M. Future development of Good Agricultural Practice in China under globalization of traditional herbal medicine trade. *Chin. Herb. Med.* **2021**, *13*, 472–479. [CrossRef] [PubMed]
2. Ali, A.H. High-performance liquid chromatography (HPLC): A review. *Ann. Adv. Chem.* **2022**, *6*, 010–020. [CrossRef]
3. Teonata, N.; Wijaya, V.A.; Vitaloka, V.S.; Attamimi, M.T.; Kartikawati, M. An introduction to different types of gas chromatography. *J. Sains Dan Terap. Kim.* **2021**, *15*, 8–17. [CrossRef]

4. Ma, D.; Wang, S.; Shi, Y.; Ni, S.; Tang, M.; Xu, A. The development of traditional Chinese medicine. *J. Tradit. Chin. Med. Sci.* **2021**, *8*, S1–S9. [CrossRef]
5. Thomas, J.; Joshi, H.C. Review on Laser Induced Breakdown Spectroscopy: Methodology and Technical Developments. *Appl. Spectrosc. Rev.* **2024**, *59*, 124–155. [CrossRef]
6. Castorena, J.; Oyen, D.; Ollila, A.; Legget, C.; Lanza, N. Deep spectral CNN for laser induced breakdown spectroscopy. *Spectrochim. Acta Part B At. Spectrosc.* **2021**, *178*, 106125. [CrossRef]
7. Yang, F.; Xu, W.; Cui, Z.; Liu, X.; Xu, X.; Jia, L.; Chen, Y.; Shu, R.; Li, L. Convolutional neural network chemometrics for rock identification based on laser-induced breakdown spectroscopy data in Tianwen-1 pre-flight experiments. *Remote Sens.* **2022**, *14*, 5343. [CrossRef]
8. Ma, S.; Xu, S.; Pan, C.; Fang, J.; Han, F.; Xia, Y.; Ding, W.; Chen, Y.; Hu, Y. TrCSL: A Transferred CNN-SE-LSTM Model for High-accuracy Quantitative Analysis of Laser-induced Breakdown Spectroscopy with Small Samples. *J. Anal. At. Spectrom.* **2025**, *40*, 1810–1820. [CrossRef]
9. Hu, Z.; Cheng, C.; Xu, H.; Huang, L. Automatic identification of modal parameters of random subspaces based on spectral clustering. *Eng. Mech.* **2026**, *43*, 39–49. [CrossRef]
10. Chu, Y.; Chen, F.; Sheng, Z.; Zhang, D.; Zhang, S.; Wang, W.; Jin, H.; Qi, J.; Guo, L. Blood cancer diagnosis using ensemble learning based on a random subspace method in laser-induced breakdown spectroscopy. *Biomed. Opt. Express* **2020**, *11*, 4191–4202. [CrossRef] [PubMed]
11. Shih, M.; Yuan, Y.; Shi, G. Comparative analysis of LDA, PLS-DA, SVM, RF, and voting ensemble for discrimination origin in greenish-white to white nephrites using LIBS. *J. Anal. At. Spectrom.* **2024**, *39*, 1560–1570. [CrossRef]
12. Wang, W.; Shi, S.; Liu, Y.; Hou, Z.; Qi, J.; Guo, L. Staging classification of omicron variant SARS-CoV-2 infection based on dual-spectrometer LIBS (DS-LIBS) combined with machine learning. *Opt. Express* **2023**, *31*, 42413–42427. [CrossRef] [PubMed]
13. Pokrajac, D.D.; Sivakumar, P.; Markushin, Y.; Milovic, D.; Holness, G.; Liu, J.; Melikechi, N.; Rana, M. Modeling of Laser-Induced Breakdown Spectroscopic Data Analysis by an Automatic Classifier. *Int. J. Data Sci. Anal.* **2019**, *8*, 213–220. [CrossRef] [PubMed]
14. Ma, J. Laser-Induced Breakdown Spectroscopy Signal Enhancement Technology and Its Application in Quantitative Analysis of Cd Elements in Soil. Master's Thesis, Henan Agricultural University, Zhengzhou, China, 2024. [CrossRef]
15. Fu, X. Study on the Influencing Factors and Methods of Rapid Measurement of Soil Elements Using LIBS. Ph.D. Thesis, Southwest University, Chongqing, China, 2020. [CrossRef]
16. Fu, X.; Peng, Y.; Liu, Y.; Lin, Y.; Gui, G.; Gacanin, H.; Adachi, F. Semi-supervised specific emitter identification method using metric-adversarial training. *IEEE Internet Things J.* **2023**, *10*, 10778–10789. [CrossRef]
17. *DB44/T 604-2009*; Product of geographical indication—Xinhui mandarin orange peel. Guangdong Provincial Quality and Technical Supervision Bureau: Jiangmen Guangdong, China.
18. Kukreti, A.; Reddy, N.K. Safeguarding Citrus: Exploring State-of-the-Art Management. In *Challenges in Plant Disease Detection and Recent Advancements*; IntechOpen: London, UK, 2024; Volume 157. [CrossRef]
19. Pan, G.; Shi, J.; Qiu, R.; Wang, H.; Wam, Q.; Zhang, Z.; Wang, K. Baseline correction algorithm for laser-induced breakdown spectroscopy. *Appl. Opt.* **2022**, *43*, 538–543. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.