RESEARCH ARTICLE
doi: 10.2306/scienceasia1513-1874.2026.033

ScienceAsia 52 (2): 2026: ID 2026033: **1–10**

Rapid type identification of plastic waste using valley-side slope of near-infrared spectrum

Chaoyi Shi[a,*], **Senlin Zhao**[a], **Zhenyi Xu**[b], **Fengjiao Shen**[a], **Cuiping Lu**[a], **Xianhe Gao**[a]

[a] School of Electronic Information and Automation, Hefei University, Hefei 230601 China
[b] Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230088 China

*Corresponding author, e-mail: cyshi@hfuu.edu.cn

Received 27 Jul 2025, Accepted 27 Mar 2026
Available online 5 May 2026

ABSTRACT: The identification and sorting of different types of plastic waste are of great significance for improving the efficiency and quality of recycling, which in turn avoids plastic waste pollution, reduces greenhouse gas emissions, and saves production resources. The existing identification method based on spectral analysis faces the problems of complex spectral processing, high data computation, and a slow prediction speed of the prediction model, which in turn leads to low identification efficiency. To solve these problems, a rapid plastic waste type identification method using the valley-side slope of the near-infrared spectrum was developed. Three wavelengths at the left side, center, and right side of the valley and their corresponding spectral reflectance were used to calculate the left and right valley-side slopes, respectively, to construct a valley-side slope feature. Based on the constructed valley-side slope feature, identification models realizing two-category, three-category, four-category, five-category, six-category, and seven-category identification modes were established using the Classification and Regression Tree (CART) algorithm. The ten-fold cross-validation results showed that the identification accuracies of the above identification modes reached 98.4%–100%, and the prediction speeds reached 31000–39000 observations/second. The proposed method can simplify spectral data processing, reduce model complexity, and improve prediction speed, while ensuring high identification accuracy.

KEYWORDS: plastic waste, rapid type identification, near-infrared spectrum, valley-side slope, classification and regression tree

INTRODUCTION

Plastic products have been widely used due to their light weight, excellent chemical stability, low cost, and other significant advantages, but at the same time, a large amount of plastic waste has been generated, causing substantial environmental challenges [1, 2]. According to data released by the United Nations Environment Programme, more than 430 million tons of plastics are produced annually worldwide, and the cumulative global plastic production is expected to reach 34 billion tons by 2050. According to a report by the United Nations International Atomic Energy Agency in 2024, approximately 70% of all plastics produced end up as plastic waste, and only 9% of the plastic waste is recycled. Plastic waste that is not recycled is burned, deposited in landfills, or leaks into the natural environment. The leakage of plastic waste into the natural environment can lead to reduced stability of soil ecosystems, enrichment of microplastics in aquatic organisms, and the release of harmful gases and greenhouse gases under conditions of light and high temperatures, harming human health and affecting climate change [3, 4]. In addition, the production of virgin plastics, including the sourcing and refining of raw materials, as well as the synthesis and processing of plastics, results in substantial greenhouse gas emissions, thereby contributing to climate change [5, 6]. As a result, plastic waste appears in the list of the top 10 most dangerous waste items worldwide [7]. Recycling of plastic waste can not only prevent the leakage of plastic waste into the natural environment, but also reduce the production of virgin plastics, thereby reducing environmental pollution, decreasing greenhouse gas emissions, and saving production resources [8, 9]. The United Nations and many other countries have made significant efforts in plastic waste recycling [10–12].

Plastic waste usually contains different types of plastics with different chemical compositions and physical properties, and mixing different plastics together during recycling will reduce the quality and properties of the recycled product; therefore, the most important step in the plastic recycling process is the identification and sorting of different types of plastics [13, 14]. The commonly used plastic identification and sorting methods include flotation sorting, electrostatic sorting, wind sorting, chemical analysis, and spectral analysis [15, 16]. Spectral analysis has the advantages of being non-destructive, accurate, fast, and automated [17, 18], studies have been carried out on the identification of plastic waste using near-infrared spectroscopy [19–22]. In the existing conventional spectral analysis of the plastic type, a spectrometer is first used to acquire spectral data, and then algorithms are usually employed for feature wavelength selection or data dimensionality reduction. Finally, based on the features selected in the previous step, a plastic type

identification model is established using a machine learning identification algorithm. Although the identification accuracy of the conventional spectral analysis has reached 97%–100%, the following problems still exist: (1) a sufficient number of feature wavelengths (typically nearly 10 or more) need to be retained to ensure identification accuracy; (2) although the feature wavelengths can be reduced to less than 10 through dimensionality reduction (such as Principal Component Analysis, PCA), the data processing remains complex. Additionally, the number and positions of feature wavelengths are related to the positions of the maxima and minima of the principal component loadings, and the determination of these maxima and minima is susceptible to subjective factors; (3) in order to achieve satisfactory identification results, especially for multi-category identification, complex identification algorithms (such as Support Vector Machine (SVM), Convolutional Neural Network (CNN), Artificial Neural Network (ANN), etc.) or a combination of the complex identification algorithms are required for model establishment. These problems lead to complex spectral processing, high data computation, and a slow prediction speed of the established model, which in turn affects the efficiency of industrial production.

Based on the above problems, this paper proposes a rapid type identification method for plastic waste based on the valley-side slope of the near-infrared spectrum. This method uses only three fixed feature wavelengths and their corresponding spectral reflectance to construct the identification feature and allows the establishment of an identification model using a simple identification algorithm, thus reducing the complexity of the data processing and identification model and improving the prediction speed of the model.

MATERIALS AND METHODS

Experimental system

As shown in Fig. S1, the experimental system mainly consists of four parts: light source, spectrometer, fiber optic probe, and computer. A tungsten halogen lamp with a wavelength range of 360–2400 nm (HL-2000, Ocean Optics, Inc., Shanghai, China) was used as the light source. A fiber-optic spectrometer (NIRQUEST512-2.5, Ocean Optics Inc.) with an InGaAs linear array detector was used to record the reflectance spectrum. The wavelength range of the spectrometer was 900–2500 nm with an optical resolution of 2.8 nm FWHM. A 400-μm diameter integrated reflectance probe (QR400-7-VIS-NIR, Ocean Optics Inc.) was used to conduct the source light to the sample for irradiation while simultaneously coupling the reflected light to the fiber-optic spectrometer. The probe was designed as a 6-around-1 fiber bundle with six fibers connected to the light source and a single fiber connected to the fiber-optic spectrometer. The computer (CPU: 12th Gen Intel(R) Core (TM) i7-12700 2.10 GHz, RAM: 16.0 GB, Dell Inc., Shanghai, China) obtained the spectral data from the fiber-optic spectrometer and performed the valley-side slope feature calculation and plastic type identification.

Materials and spectrum acquisition

The experimental samples were plastic waste provided by Hefei Taiho Intelligent Technology Group Co., Ltd., Hefei, China, including seven commonly used plastics, namely Polyethylene Terephthalate (PET), Polyethylene (PE), High-Density Polyethylene (HDPE), Polypropylene (PP), Polystyrene (PS), Polyvinyl Chloride (PVC), and Acrylonitrile Butadiene Styrene (ABS). The samples have been cleaned and shredded into irregularly shaped fragments measuring approximately 3 cm × 3 cm, and the thicknesses of the samples ranged from 0.2 mm to 1.5 mm. The samples cover five colors (white, transparent, green, blue, and yellow) in total, with each type of sample containing a subset of these colors. Each type of plastic comprises 45 samples. Four spectra were measured at four different positions on each sample, resulting in 180 spectra for each plastic type. The color and thickness of the corresponding samples were recorded for each spectrum. A total of 1260 spectra were obtained for the seven types of plastics. The average spectrum for each type of plastic is shown in Fig. 1a.

Feature construction method

Owing to the different elemental and molecular compositions of different types of plastics, there are differences in the number, position, width, and relative intensity of the valleys in the near-infrared spectrum [23], as shown in Fig. 1a. Therefore, valley information can be used for plastic identification. Thus, in this study, a feature construction method using the valley-side slope is proposed for the type identification of plastics. As shown in Fig. 1b, for each valley, three wavelengths at the left side, center, and right side of the valley, which are respectively named left wavelength (λ_l), center wavelength (λ_c), and right wavelength (λ_r), are selected. As shown in Eqs. (1)–(3), the three wavelengths λ_l, λ_c, λ_r and their corresponding reflectance (R_l, R_c, and R_r) are employed to calculate the left valley-side slope (S_L) and the right valley-side slope (S_R). Finally, valley-side slope feature F is constructed.

$$S_L = \frac{R_c - R_l}{\lambda_c - \lambda_l} \tag{1}$$

$$S_R = \frac{R_r - R_c}{\lambda_r - \lambda_c} \tag{2}$$

$$F = [S_L, S_R] \tag{3}$$

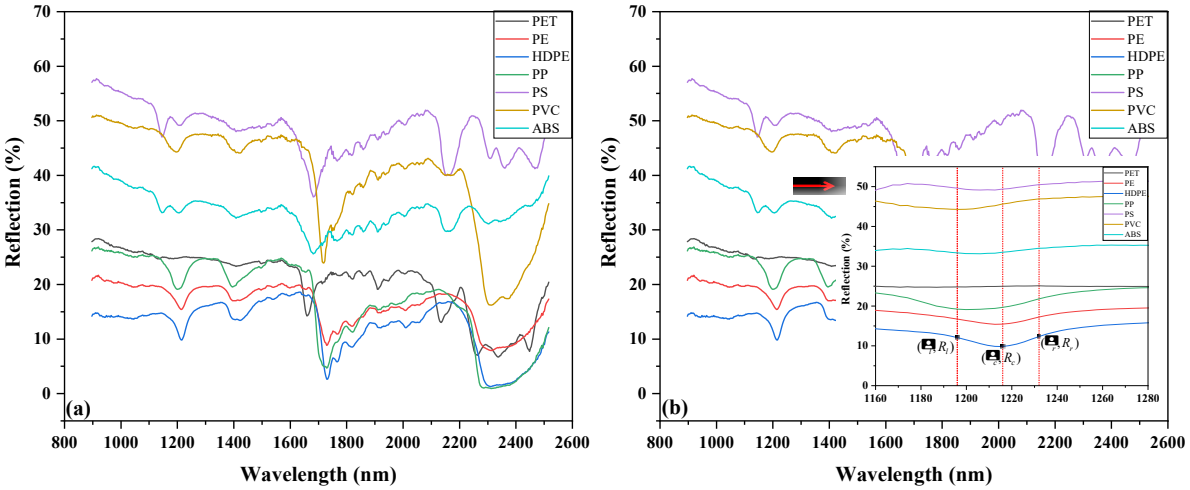

Fig. 1 (a) Average spectra of PET, PE, HDPE, PP, PS, PVC, and ABS plastics; (b) The three wavelengths and the corresponding reflectance at a certain valley.

RESULTS AND DISCUSSION

Valley-side slope feature construction and selection

Owing to the wide application and high recycling value of PET plastic [24], valleys were selected based on the criterion of distinguishing PET from other types of plastic. Six valleys were selected and analyzed, and the selected wavelengths at the six valleys were (1196 nm, 1215 nm, 1231 nm), (1385 nm, 1391 nm, 1401 nm), (1652 nm, 1659 nm, 1668 nm), (1713 nm, 1719 nm, 1725 nm), (2116 nm, 2129 nm, 2145 nm), and (2279 nm, 2292 nm, 2304 nm). Each spectrum had six valley-side slope features, namely F_1, F_2, F_3, F_4, F_5, and F_6. As there were 180 spectra for PET, PE, HDPE, PP, PS, PVC, and ABS plastics, each type of plastic possessed 180 F_1 features, as well as 180 F_2, F_3, F_4, F_5, and F_6 features; the average values of each feature are shown in Table 1. The spectrometer-measured spectrum in this study, from which the above wavelengths were selected, was not preprocessed (such as smoothing), because, in practical applications, the three selected wavelengths used for the calculation of a valley-side slope feature will be measured by three filter-equipped single-point detectors, rather than being selected from the spectrometer-measured spectrum, and no spectral preprocessing is required for the three discrete wavelengths. Therefore, in order to keep consistent with the practical applications, the spectrometer-measured spectrum in this study was not preprocessed.

To determine the most effective valley-side slope feature, a scatterplot analysis was performed for F_1, F_2, F_3, F_4, F_5 and F_6. Fifty-four F_1 features were randomly selected for each type of plastic, and there were 378 F_1 features in total for the seven types of plastic. An axis-centered scatter plot of these 378 F_1 features was drawn (Fig. 2a), with S_L as the horizontal coordinate, S_R as the vertical coordinate, and the coordinate origin at the center of the graph. Similarly, axis-centered scatter plots were drawn for F_2, F_3, F_4, F_5, and F_6 (Fig. 2b–f). It can be seen from Fig. 2a that the F_1 features of the seven types of plastics clustered separately according to the type of plastic, with no overlap between different types. As for F_2, F_3, F_4, F_5, and F_6, there were different degrees of overlap between different types of plastics, leading to the inability to distinguish certain plastic types. Therefore, F_1 was selected as the feature for plastic type dentification.

Color and thickness influence on the valley-side slope feature

The influence of the color and thickness of the plastic on the valley-side slope feature was analyzed. To analyze the influence of color, 20 spectra and their corresponding F_1 features were randomly selected for transparent, blue, and green PET plastic with a thickness of 0.2 mm. The spectra of the transparent, blue, and green PET plastics are shown in Fig. 3a, from which the spectra of the three different color PET plastics have the same shape and differ only in the magnitude of spectral reflectance. Furthermore, the F_1 feature was analyzed with an axis-centered scatter plot (Fig. 3c), and the 20 F_1 features were still clustered together and did not overlap with those of other types of plastics, so they could still be distinguished from the F_1 features of other plastic types. The influence of color was also analyzed for other types of plastics. The results indicate that color does not affect the clustering of the valley-side slope feature F_1 and, therefore, does not affect identification. It should be noted that the near-infrared spectroscopy cannot be used to identify dark-colored plastics, such as black plastics, because most of the light is absorbed and little light is reflected, making

Table 1 Valley-side slope features of seven types of plastics. Six valleys were selected and analyzed, and each spectrum had six valley-side slope features: F_1, F_2, F_3, F_4, F_5, and F_6. As there were 180 spectra for PET, PE, HDPE, PP, PS, PVC, and ABS plastics, each type of plastic possessed 180 F_1 features, as well as 180 F_2, F_3, F_4, F_5, and F_6 features; the average values of each feature are shown.

Plastic type	F_1	F_2	F_3	F_4	F_5	F_6
PET	[0.0146, 0.0119]	[−0.0156, 0.0118]	[−0.2812, 0.1683]	[0.1839, 0.2236]	[−0.4065, 0.0880]	[0.0657, −0.0331]
PE	[−0.1595, 0.2778]	[−0.1564, 0.0014]	[0.0649, 0.0180]	[−0.3771, −0.2215]	[0.0406, −0.0108]	[−0.1129, −0.0570]
HDPE	[−0.2021, 0.1953]	[−0.1471, −0.0210]	[0.0718, 0.0527]	[−0.3579, −0.2348]	[0.0491, 0.0134]	[−0.2881, −0.1504]
PP	[0.0403, 0.2096]	[−0.1236, 0.0585]	[0.0844, 0.0354]	[0.0067, 0.0787]	[0.0080, −0.0358]	[−0.0109, −0.0005]
PS	[−0.0248, 0.0789]	[−0.0433, 0.0052]	[−0.1815, −0.3506]	[0.3203, 0.3317]	[−0.2991, −0.5771]	[−0.3250, −0.2032]
PVC	[0.1307, 0.1536]	[−0.0756, −0.0033]	[−0.0487, −0.1922]	[0.0128, 0.6114]	[−0.1034, −0.1005]	[−0.4299, −0.1252]
ABS	[0.0097, 0.0755]	[−0.0368, −0.0008]	[−0.1412, −0.1987]	[0.2132, 0.2582]	[−0.1773, −0.2414]	[−0.1194, −0.0197]

it impossible to obtain effective spectral information.

The influence of the thickness was then analyzed. Twenty spectra and their corresponding F_1 features were randomly selected for blue PET plastic with thicknesses of 0.2, 1.1, and 1.5 mm. The spectra of these 0.2, 1.1, and 1.5 mm blue PET plastics are shown in Fig. 3b, from which the spectra of the three different thickness blue PET plastics also have the same shape and differ only in the magnitude of spectral reflectance. Furthermore, the F_1 feature was analyzed with an axis-centered scatter plot (Fig. 3d), and the F_1 features of the PET plastics were still clustered together and did not overlap with those of other types of plastics, so they could still be distinguished from the F_1 features of other plastic types. The influence of thickness on other types of plastics was also analyzed. The results indicate that thickness does not affect the clustering of the valley-side slope feature F_1 and, therefore, does not affect identification.

Model establishment and performance evaluation

The Classification and Regression Trees (CART) algorithm is a basic classification and regression method, and the basic principle is to establish a binary tree-structured model by recursively dividing the training dataset [25]. The classification process is interpretable and easy to understand, and the training and classification of the model are simple and fast. Based on the CART algorithm, six identification models that realized the following six identification modes were established: two-category (PET versus others), three-category (PET, PE versus others), four-category (PET, PE, HDPE versus others), five-category (PET, PE, HDPE, PP versus others), six-category (PET, PE, HDPE, PP, PS versus others), and seven-category (PET, PE, HDPE, PP, PS, PVC, and ABS). As described in previously, there were 180 F_1 features for each type of plastic; thus, there were 1260 F_1 features in total. In each model, Arabic numerals were used $(1, 2, 3, \ldots)$ to label the different categories of plastics. Subsequently, a stratified sampling method was used to divide the 1260 F_1 features into training set and testing set according to a set ratio. Finally, the identification model based on the CART algorithm (hereafter referred to as F_1-CART model) was trained using the training set and evaluated using the testing set.

For comparison, using the modeling steps described above, six identification models were also established using the SVM algorithm with a Linear Kernel (hereafter referred to as F_1-SVM model). The SVM algorithm is a widely used classification algorithm with excellent performance [26]; however, its complexity is higher than that of the CART algorithm.

Ten-fold cross-validation was performed for the F_1-CART models and F_1-SVM models; the average identification accuracy and average prediction speed are listed in Table 2. Overall, both the F_1-CART model and the F_1-SVM model demonstrated good accuracy in the six identification modes, but the F_1-CART model was slightly better than the F_1-SVM model. The accuracy of F_1-CART model was 100% in the two-category, three-category, and four-category identification modes, while the accuracy of F_1-SVM model reached 100% only in the two-category and three-category identification modes, with an accuracy of 99.6% in the four-category identification mode. In the five-category, six-category, and seven-category identification modes, the accuracies of F_1-CART model and F_1-SVM model were no less than 98.4%.

Regarding the prediction speed, which is described by the number of samples that the model can process and provide prediction results per second (observations per second, obs/s), the F_1-CART model significantly outperformed F_1-SVM model. In the two-category and three-category identification modes, the two models possessed comparable prediction speeds; however, in the four-category, five-category, six-category, and seven-category identification modes, the prediction speed of the F_1-SVM model decreased significantly as the number of categories increased, with the prediction speed of the seven-category identification mode decreasing to 36.6% of the prediction speed of the two-category identification mode. The F_1-CART model, on the other hand, showed only a slight decrease in prediction speed, and the prediction speed of the seven-category identification mode was still 80% of that of the two-category identification mode, which was 2 times faster than that of the F_1-SVM model.

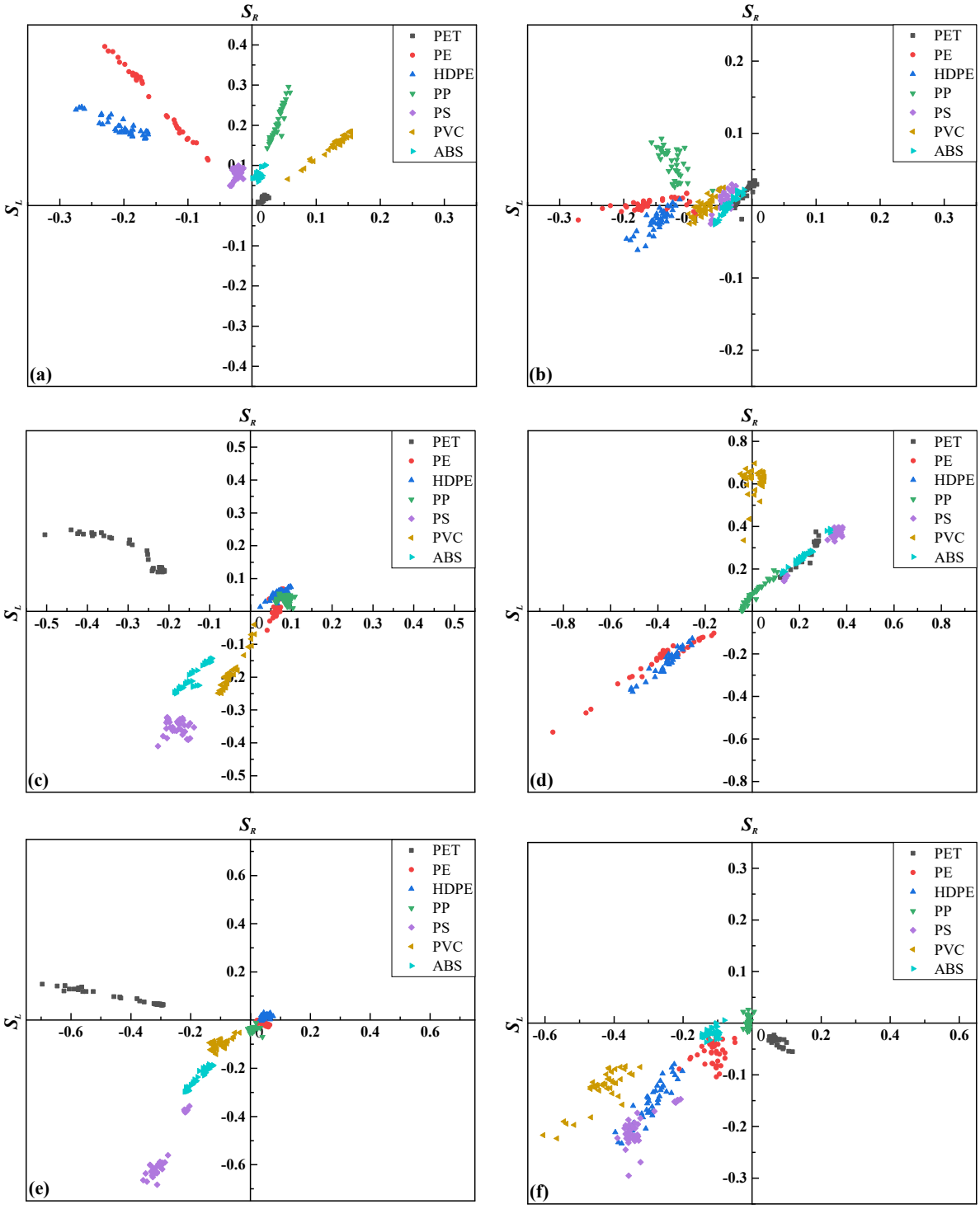

Fig. 2 Axis-centered scatter plots of F_1 (a), F_2 (b), F_3 (c), F_4 (d), F_5 (e), and F_6 (f).

The above analysis shows that, using the valley-side slope feature F_1, the simple classification algorithm CART can be used to establish the plastic identification model with high accuracy; its prediction speed was significantly faster, and the identification performance was slightly better than that of the model established using the complex classification algorithm SVM.

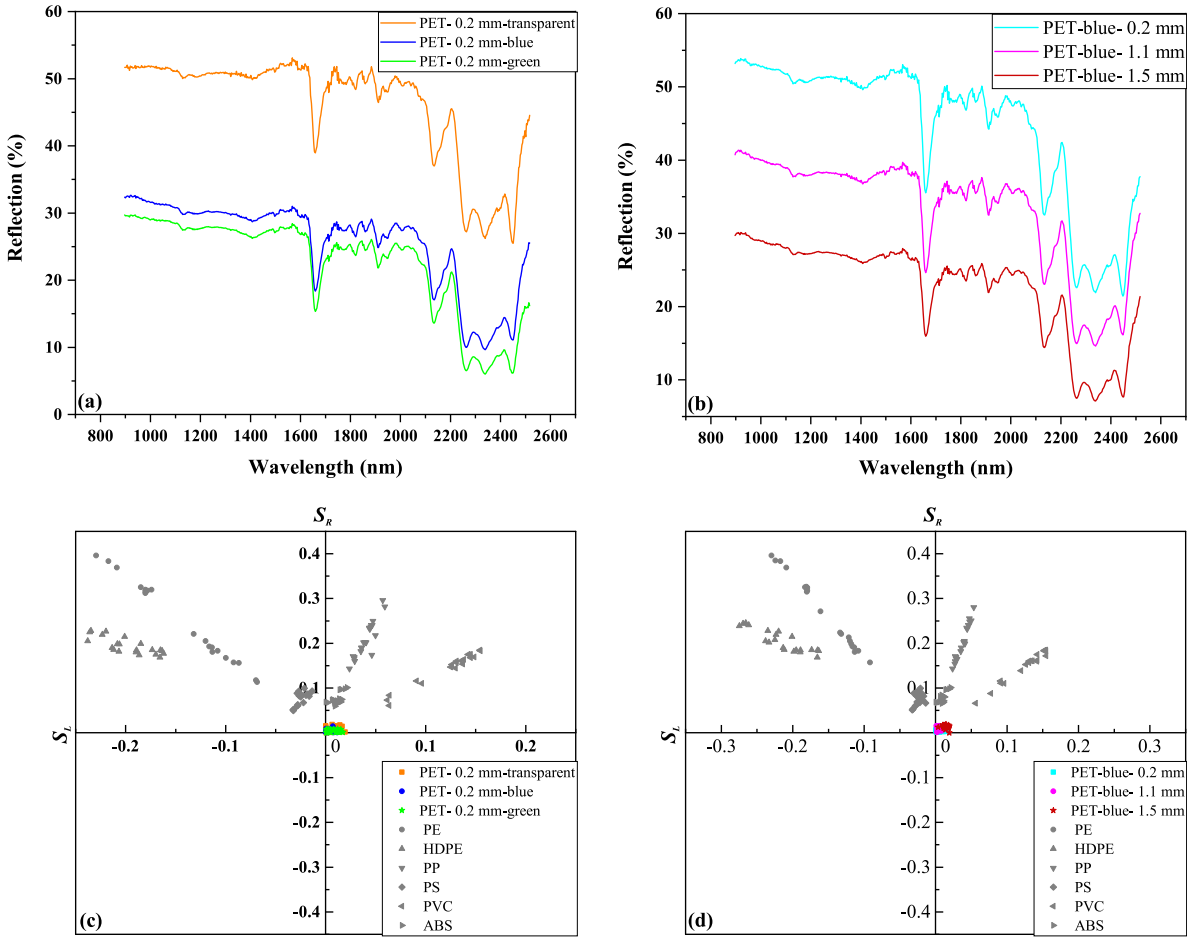

Fig. 3 (a) Spectra of the transparent, blue, and green PET plastics with a thickness of 0.2 mm; (b) Spectra of the blue PET plastics with thicknesses of 0.2, 1.1, and 1.5 mm; (c) Axis-centered scatter plot of the F_1 features of the spectra of the transparent, blue, and green PET plastics with a thickness of 0.2 mm; (d) Axis-centered scatter plot of the F_1 features of the blue PET plastics with thicknesses of 0.2, 1.1, and 1.5 mm.

Table 2 Average accuracies and prediction speed of the F_1-CART and F_1-SVM identification models. Using the F_1 feature, the F_1-CART model and the F_1-SVM model were established. Ten-fold cross-validation was performed for both models, and the performance of the two types of models was analyzed.

Identification mode	Average accuracy (%)		Prediction speed (obs/s)	
	F_1-CART	F_1-SVM	F_1-CART	F_1-SVM
two-category (PET versus others)	100	100	39000	41000
three-category (PET, PE versus others)	100	100	38000	36000
four-category (PET, PE, HDPE versus others)	100	99.6	37000	29000
five-category (PET, PE, HDPE, PP versus others)	99.3	99.3	34000	26000
six-category (PET, PE, HDPE, PP, PS versus others)	98.6	98.7	32000	19000
seven-category (PET, PE, HDPE, PP, PS, PVC, ABS)	98.4	98.4	31000	15000

Comparison with the conventional spectral analysis method

To further verify the validity of the F_1-CART model, it was compared with the conventional spectral analysis method, which includes spectral preprocessing, feature wavelength selection, and model training steps.

Through appropriate spectral preprocessing, negative factors such as instrument noise, light scatter, spectral data drift, as well as spectral interference can be eliminated or reduced, thereby the performance of the established model can be improved [27]. An appropriate feature wavelength selection method can effectively reduce the data dimensionality and reduce

Table 3 Average accuracies and prediction speed of the SPA-CART, SPA-SVM, and F_1-CART identification models. To further verify the validity of the F_1-CART model, it was compared with the conventional spectral analysis method, which includes spectral preprocessing, feature wavelength selection, and model training steps.

Identification mode	Average accuracy (%)			Prediction speed (obs/s)		
	SPA-CART	SPA-SVM	F_1-CART	SPA-CART	SPA-SVM	F_1-CART
two-category (PET versus others)	99.8	100	100	46000	43000	39000
three-category (PET, PE versus others)	99.5	99.9	100	41000	32000	38000
four-category (PET, PE, HDPE versus others)	99.2	99.7	100	35000	18000	37000
five-category (PET, PE, HDPE, PP versus others)	99.0	99.6	99.3	29000	17000	34000
six-category (PET, PE, HDPE, PP, PS versus others)	98.7	99.3	98.6	17000	8100	32000
seven-category (PET, PE, HDPE, PP, PS, PVC, ABS)	98.5	99.1	98.4	9500	5600	31000

Table 4 Average accuracies and prediction speed of the PCA-CART, PCA-SVM, and F_1-CART identification models. Using the six feature wavelengths selected through PCA, the PCA-CART model and the PCA-SVM model were respectively established. The performance of the models was compared with the F_1-CART model.

Identification mode	Average accuracy (%)			Prediction speed (obs/s)		
	PCA-CART	PCA-SVM	F_1-CART	PCA-CART	PCA-SVM	F_1-CART
two-category (PET versus others)	99.9	99.9	100	45000	40000	39000
three-category (PET, PE versus others)	98.6	99.8	100	42000	28000	38000
four-category (PET, PE, HDPE versus others)	98.4	99.4	100	36000	21000	37000
five-category (PET, PE, HDPE, PP versus others)	98.0	98.9	99.3	28000	18000	34000
six-category (PET, PE, HDPE, PP, PS versus others)	97.6	98.1	98.6	16000	10000	32000
seven-category (PET, PE, HDPE, PP, PS, PVC, ABS)	97.3	97.9	98.4	11000	7000	31000

multi-collinearity, thus reducing the modeling complicacy and improving the prediction accuracy and efficiency [28]. Therefore, pre-experiments were conducted to select appropriate spectral preprocessing method and feature wavelength selection method. Three preprocessing methods, Multiplicative Scatter Correction (MSC), Savitzky-Golay (SG), and Standard Normal Variate Transformation (SNV), were evaluated, and four feature wavelength selection methods, Successive Projections Algorithm (SPA), Genetic Algorithm (GA), Variable Combination Population Analysis (VCPA), and Competitive Adaptive Reweighted Sampling (CARS), were analyzed. The CART and SVM algorithms were employed to establish models for performance evaluation of the above methods.

After a comprehensive evaluation of accuracy and prediction speed, the combination of SG and SNV methods was selected for preprocessing of the full spectrum, and the SPA method was selected for feature wavelength selection. The SG and SNV methods were sequentially used for spectral preprocessing, and then the SPA method was implemented for feature wavelength selection. Eight feature wavelengths were selected from the preprocessed full spectrum. Finally, using the eight feature wavelengths, six identification models (hereafter referred to as the SPA-CART model) were established based on the CART algorithm, realizing six identification modes. For a more comprehensive comparative analysis, six identification models based on the SVM algorithm (hereafter referred to as the SPA-SVM model) were established using the same modeling steps.

The ten-fold cross-validation results of the SPA-CART models and SPA-SVM models are listed in Table 3. The results indicated that the SPA-CART model, although faster than the F_1-CART model in two-category and three-category modes, was significantly slower than the F_1-CART model in

Table 5 Average accuracies and prediction speed of the SPEM-CART, SPEM-SVM, and F_1-CART identification models.

Identification mode	Average accuracy (%)			Prediction speed (obs/s)		
	SPEM-CART	SPEM-SVM	F_1-CART	SPEM-CART	SPEM-SVM	F_1-CART
two-category (PET versus others)	98.7	99.8	100	5500	2600	39000
three-category (PET, PE versus others)	98.3	99.2	100	5400	2300	38000
four-category (PET, PE, HDPE versus others)	97.8	98.9	100	4900	2100	37000
five-category (PET, PE, HDPE, PP versus others)	97.5	98.4	99.3	4600	1800	34000
six-category (PET, PE, HDPE, PP, PS versus others)	96.6	97.9	98.6	3300	1700	32000
seven-category (PET, PE, HDPE, PP, PS, PVC, ABS)	96.1	97.6	98.4	2100	1600	31000

four-category, five-category, six-category, and seven-category modes. Overall, the accuracy was lower than that of the F_1-CART model. Thus, the SPA-CART model did not perform as well as the F_1-CART model did. For the SPA-SVM model, the identification accuracies for the six identification modes were equivalent to those of the F_1-CART model, with accuracies in the three-category and four-category modes slightly lower than those of the F_1-CART model, and the accuracies in the five-category and six-category and seven-category modes were slightly higher than those of the F_1-CART model. However, the prediction speed of the SPA-SVM model was significantly lower than that of the F_1-CART model, with the prediction speed of the SPA-SVM model being only 18% of that of the F_1-CART model in the seven-category mode. Overall, the F_1-CART model was able to achieve the identification accuracy of the SPA-SVM model that was established using the complex classification algorithm SVM and eight feature wavelengths selected from the preprocessed spectrum; however, the F_1-CART model possessed a significant advantage in the prediction speed.

PCA is a classical method that is widely used for data dimensionality reduction and feature extraction [29]. In contrast to SPA, GA, VCPA, and CARS, the feature wavelength selection method based on PCA takes the loadings of principal components as the indication to guide the selection of feature wavelengths [30]. The F_1-CART model was compared with the model established using the feature wavelengths selected based on PCA. The PCA results of the spectral data showed that the cumulative contribution rate of the first two principal components reaches 93.7%. Therefore, the number of principal components is set to 2. The loadings of the first two components were then plotted against wavelengths, and six wavelengths (1145 nm, 1212 nm, 1672 nm, 1732 nm, 1902 nm, and 2157 nm) situated at the maxima or minima of the plots were determined as feature wavelengths. Using the six feature wavelengths selected from the preprocessed full spectrum, identification models using CART algorithm (hereafter referred to as the PCA-CART model) and SVM algorithm (hereafter referred to as the PCA-SVM model) were respectively established.

The ten-fold cross-validation results of the PCA-CART models and the PCA-SVM models are listed in Table 4. The results indicated that the accuracy of the F_1-CART model was better than those of the PCA-CART model and PCA-SVM model. Regarding prediction speed, despite being slightly faster than the F_1-CART model in certain modes (PCA-CART in two-category and three-category modes, and PCA-SVM in two-category mode), the PCA-CART model and PCA-SVM model exhibited a significant decrease in prediction speed as the number of categories increased, which were notably lower than that of the F_1-CART model. In summary, the F_1-CART model outperformed both the PCA-CART model and PCA-SVM model in terms of both accuracy and prediction speed.

Theoretically, the full spectrum contains more information; therefore, the identification model established using the full spectrum (i.e., without feature wavelength selection) may exhibit better performance, but it must also be noted that the redundant information in the full spectrum may also affect the identification performance. In this study, the spectrum in the range of 900 nm to 2500 nm was acquired. Because of the instability at both ends of the spectrum, the spectra in the 900–1000 nm and 2400–2500 nm bands were discarded, and the spectrum in the 1000–2400 nm range was used as the full spectrum. Based on the full spectrum preprocessed by the SG and SNV methods, six identification models realizing six identification modes were established using the CART algorithm (hereafter referred to as the SPEM-CART model), and similarly, six identification models were also established using the SVM algorithm (hereafter referred to as the SPEM-SVM model). The performance of the models was compared with the F_1-CART model. The ten-fold cross-validation results of the SPEM-CART models and SPEM-SVM models are listed in Table 5. The results showed that the F_1-CART model outperformed the SPEM-CART model and SPEM-SVM model in terms of the accuracy and prediction speed. In particular, F_1-CART model possessed a significant advantage in terms of prediction speed.

CONCLUSION

A rapid type identification method for plastic waste based on the valley-side slope of the near-infrared spectrum was proposed. Compared with existing spectral analysis methods, the valley-side slope feature of the proposed method requires only three fixed feature wavelengths and is computationally simple, thereby reducing the complexity of spectral data processing. Using the valley-side slope feature, identification models were established for six identification modes based on the low-complexity CART algorithm, and the type identification accuracies reached 98.4%–100%. Compared with the more complex SPA-SVM model, which demonstrated the best comprehensive performance in terms of accuracy and prediction speed among all models established based on the CART and SVM algorithms using either the full spectrum or selected feature wavelengths, the F_1-CART model demonstrated equivalent identification accuracy, whereas the prediction speed was significantly higher. Thus, the type identification method proposed in this paper can not only reduce the complexity of spectral data processing but also allow the use of a simpler identification algorithm for model establishment, which simplifies the identification model and improves the prediction speed. This study validates the effectiveness of the

valley-side slope feature in the type identification of plastic waste and provides theoretical and technical support for the development of a simple and efficient identification technique for plastic waste. In future work, an identification system that can be applied to industrial production will be developed using single-point detectors to detect the three selected feature wavelengths (i.e., 1196 nm, 1215 nm, and 1231 nm) instead of detecting the full spectrum, and the identification model will be optimized to further improve the accuracy of the simultaneous identification of multiple types of plastics.

Appendix A. Supplementary data

Supplementary data associated with this article can be found at https://dx.doi.org/10.2306/scienceasia1513-1874.2026.033.

Acknowledgements: This work was supported by the National Natural Science Foundation of China under Grant No. 42405132, Scientific Research Program of Higher Education Institutions of Anhui Province under Grant No. 2023AH052188, and Academic Degree Program of Anhui Province under Grant No. 2022tsxwd048.

REFERENCES

1. Borrelle SB, Ringma J, Law KL, Monnahan CC, Lebreton L, McGivern A, Murphy E, Jambeck J, et al (2020) Predicted growth in plastic waste exceeds efforts to mitigate plastic pollution. *Science* **369**, 1515–1518.
2. Suzuki G, Uchida N, Tanaka K, Matsukami H, Kunisue T, Takahashi S, S Takahashi S, Viet PH, et al (2022) Mechanical recycling of plastic waste as a point source of microplastic pollution. *Environ Pollut* **303**, 119114.
3. Wojnowska-Baryła I, Bernat K, Zaborowska M (2022) Plastic waste degradation in landfill conditions: the problem with microplastics, and their direct and indirect environmental effects. *Int J Environ Res Public Health* **19**, 13223.
4. Sooksawat T, Wattanakornsiri A, Kohkaew R, Page LM, Tongnunui S (2023) Microplastic accumulation in local dominant shellfish from the Khwae Noi basin in western Thailand and its environmental factors. *ScienceAsia* **49**, 445–453.
5. Ford HV, Jones NH, Davies AJ, Godley BJ, Jambeck JR, Napper IE, Suckling CC, Williams GJ, et al (2022) The fundamental links between climate change and marine plastic pollution. *Sci Total Environ* **806**, 150392.
6. Drewniok MP, Gao Y, Cullen JM, Cabrera Serrenho A (2023) What to do about plastics? Lessons from a study of United Kingdom plastics flows. *Environ Sci Technol* **57**, 4513–4521.
7. Rangel-Buitrago N, Arroyo-Olarte H, Trilleras J, Arana VA, Mantilla-Barbosa E, Gracia A, Mendoza AV, Neal WJ, et al (2021) Microplastics pollution on Colombian Central Caribbean beaches. *Mar Pollut Bull* **170**, 112685.
8. Pottinger AS, Geyer R, Biyani N, Martinez CC, Nathan N, Morse MR, Liu C, Hu S, et al (2024) Pathways to reduce global plastic waste mismanagement and greenhouse gas emissions by 2050. *Science* **386**, 1168–1173.
9. Ahmed N (2023) Utilizing plastic waste in the building and construction industry: A pathway towards the circular economy. *Constr Build Mater* **383**, 131311.
10. Islam MS, Lee Z, Shaleh A, Soo HS (2024) The United Nations Environment Assembly resolution to end plastic pollution: Challenges to effective policy interventions. *Environ Dev Sustain* **26**, 10927–10944.
11. Liu X, Lu X, Feng Y, Zhang L, Yuan Z (2022) Recycled WEEE plastics in China: Generation trend and environmental impacts. *Resour Conserv Recycl* **177**, 105978.
12. Kasznik D, Łapniewska Z (2023) The end of plastic? The EU's directive on single-use plastics and its implementation in Poland. *Environ Sci Policy* **145**, 151–163.
13. Adarsh UK, Kartha VB, Santhosh C, Unnikrishnan VK (2022) Spectroscopy: A promising tool for plastic waste management. *TrAC Trends Anal Chem* **149**, 116534.
14. Lim J, Ahn Y, Cho H, Kim J (2022) Optimal strategy to sort plastic waste considering economic feasibility to increase recycling efficiency. *Process Saf Environ Prot* **165**, 420–430.
15. Lim J, Ahn Y, Kim J (2023) Optimal sorting and recycling of plastic waste as a renewable energy resource considering economic feasibility and environmental pollution. *Process Saf Environ Prot* **169**, 685–696.
16. Liu Q, Martinez-Villarreal S, Wang S, Tien NNT, Kammoun M, Roover QD, Len A, Richel A (2024) The role of plastic chemical recycling processes in a circular economy context. *Chem Eng J* **498**, 155227.
17. Howard IA, Busko D, Gao G, Wendler P, Madirov E, Turshatov A, Moesslein J, Richards BS (2024) Sorting plastics waste for a circular economy: Perspectives for lanthanide luminescent markers. *Resour Conserv Recycl* **205**, 107557.
18. Wu X, Li J, Yao L, Xu Z (2020) Auto-sorting commonly recovered plastics from waste household appliances and electronics using near-infrared spectroscopy. *J Cleaner Prod* **246**, 118732.
19. Neo ERK, Yeo Z, Low JSC, Goodship V, Debattista K (2022) A review on chemometric techniques with infrared, Raman and laser-induced breakdown spectroscopy for sorting plastic waste in the recycling industry. *Resour Conserv Recycl* **180**, 106217.
20. Munz M, Kreiß, Krüger L, Schmidt LK, Bochow M, Bednarz M, Bannick CG, Oswald SE (2023) Application of high-resolution near-infrared imaging spectroscopy to detect microplastic particles in different environmental compartments. *Water Air Soil Pollut* **234**, 286.
21. Adarsh UK, Gowd EB, Bankapur A, Kartha VB, Chidangil S, Unnikrishnan VK (2022) Development of an inter-confirmatory plastic characterization system using spectroscopic techniques for waste management. *Waste Manage* **150**, 339–351.
22. Ozawa K (2023) Identification of overlapping plastic sheets using short-wavelength infrared hyperspectral imaging. *Opt Express* **31**, 12328–12338.
23. Pilapitiya PNT, Ratnayake AS (2024) The world of plastic waste: A review. *Cleaner Mater* **11**, 100220.
24. Jang JY, Sadeghi K, Seo J (2022) Chain-extending modification for value-added recycled PET: A review. *Polym Rev* **62**, 860–889.
25. Costa VG, Salcedo-Sanz S, Pedreira CE (2024) Efficient evolution of decision trees via fully matrix-based fitness evaluation. *Appl Soft Comput* **150**, 111045.

26. Yang C, Oh SK, Yang B, Pedrycz W, Wang L (2022) Hybrid fuzzy multiple SVM classifier through feature fusion based on convolution neural networks and its practical applications. *Expert Syst Appl* **202**, 117392.
27. Yang W, Xiong Y, Xu Z, Li L, Du Y (2022) Piecewise preprocessing of near-infrared spectra for improving prediction ability of a PLS model. *Infrared Phys Technol* **126**, 104359.
28. Wang T, Zheng Y, Xu L, Yun YH (2025) Comprehensive comparison on different wavelength selection methods using several near-infrared spectral datasets with different dimensionalities. *Spectrochim Acta Part A* **331**, 125767.
29. Maouardi ME, Braekeleer KD, Bouklouze A, Heyden YV (2024) Comparison of near-infrared and mid-infrared spectroscopy for the identification and quantification of argan oil adulteration through PCA, PLS-DA and PLS. *Food Control* **165**, 110671.
30. Serranti S, Gargiulo A, Bonifazi G (2012) Classification of polyolefins from building and construction waste using NIR hyperspectral imaging system. *Resour Conserv Recycl* **61**, 52–58.

Appendix A. Supplementary data

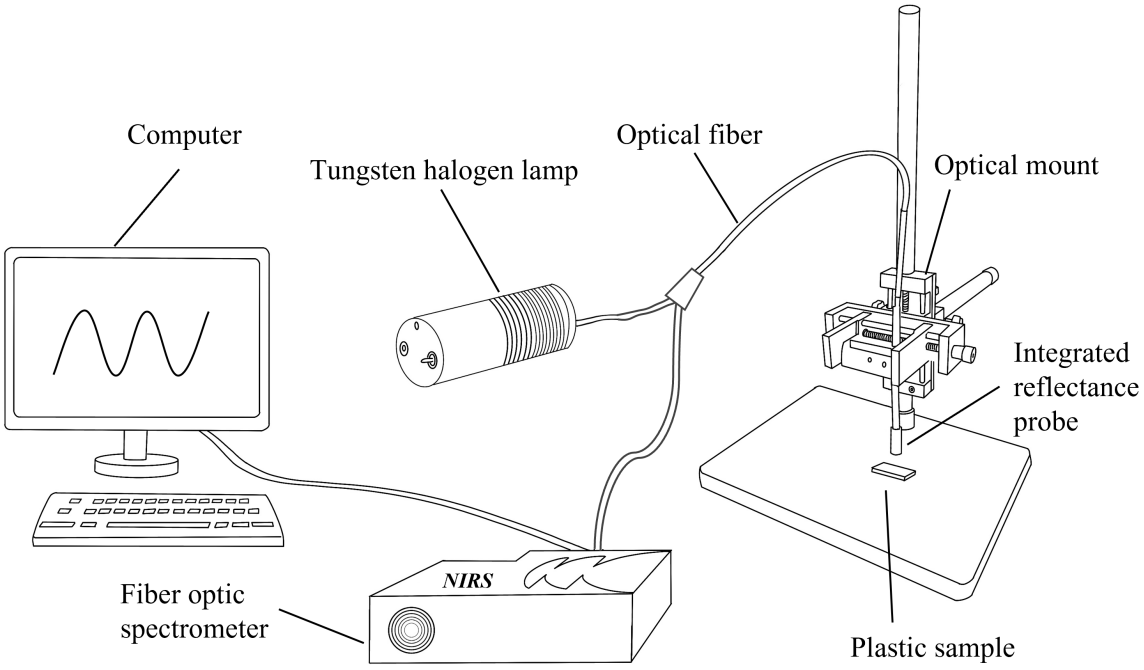

Fig. S1 Schematic diagram of the plastic waste identification system using valley-side slope feature of the near-infrared spectrum.