# Hyperspectral imaging technology is used to identify damage in pears.

**Dr. SUNITA BAL,NALANDA INSTITUTE OF TECHNOLOGY,BHUBANESWAR**
E-MAIL-sunitabal@thenalanda.com

Abstract- Crown pears are a significant crop for the economy, but the various levels of damage have a negative impact on both their quality and economy. Sorting crown pears with various degrees of damage is necessary to improve the overall quality of crown pears. Traditional detection techniques do have significant drawbacks, though, such limited efficiency and high inaccuracy. In order to distinguish between sound and the three types of damage (levels I, II, and III, respectively) to crown pears, the hyper spectral technology was employed in this study. Reflectance (R) spectra were combined with absorbance (A) and Kubelka-Munk (K-M) spectra to increase the model's discriminatory accuracy. The three spectra were processed, and two models—the support vector machine (SVM) model and the partial least squares discriminate analysis (PLS-DA) model—were developed to separate the crown pears with various degrees of damage. The discriminate model's results demonstrate that the SVM based on R, A, and K-M spectra has a higher discrimination accuracy than PLS-DA of those spectra; the A-RAW-SVM model has the best discrimination performance, with an overall discrimination accuracy of 100% for the test and 98.98% for calibration sets, respectively. Finally, the characteristic wavelengths were extracted from the spectra using the competitive adaptive reweighted sampling (CARS) and the uninformative variables elimination (UVE), and the SVM models were constructed using the filtered R, A, and K-M. The A-RAW-CARS-SVM model has the best capacity to discriminate, according to their results, and its test and calibration sets' discriminating accuracies are 96.88% and 100%, respectively. The outcomes demonstrate that the SVM model based on a spectrum is the best at differentiating between the various degrees of damage in crown pears. This work offers an experimental and theoretical foundation for the use of hyper spectral imaging to identify crown pear damage.

## 1. Introduction

Crown pears are a significant crop for the economy and are commonly cultivated in subtropical areas. It is well-liked due of its thin, crisp, sweet, and juicy skin [1]. Owing to their thin skin, crown pears are susceptible to mechanical damage during picking, shipping, and storage after ripening. Depending on the severity of the damage, these three types of damage may be distinguished: surface damage, epidermal abrasion, and impact damage [2]. Damaged crown pears quickly rot and degrade due to the simple growth of germs, and they also spread infection to healthy crown pears, reducing their shelf life [3].The crown pears after damage can be sorted and takendifferent treatment according to the different levels of damage to obtain higher economic benefits. The crown pearswith mild damage can still be eaten normally in the early stage. The crown pears with severe damage can be con- sidered to remove the damaged part after making processed food. At present, the damaged crown pears mainly rely on manual sorting. This sorting method has a large error and low efficiency. Therefore, it is necessary to explore a precise, nondestructive, and efficient inspection method to realize the sorting of crown pears with diferent damaged levels.

In order to discriminate the diferent levels of damage of crown pears more efficiently, researchers have proposed various detection techniques to investigate. Kim et al. [4] used infrared-locked thermography to assess the damage of pears by feeding periodic thermal energy to the pear, and then, thermal radiation from the pear was collected to identify the size of the damage and the depth of the damage. The results showed that the phase information generated by thermography could be used to detect the damaged fruit. However, thermal radiation can cause some damage to the quality of fruit. Luo et al. [5] used spatial frequency domain imaging (SFDI) to detect diferent levels of damage (sound, slight, and severe damage) of pears. The results showed that the accuracy of SFDI was 100% at 527 nm; this demonstrated that the SFDI technique could be used to detect the damage level of fruit. However, SFDI is strongly influenced by the color of the peel, and the diference in peel color poses a great difficulty in characterizing the optical properties. Zhou et al. [6] used optical coherence tomography (OCT) to detect the damage of pears. The results indicated that the OCTcould beused for damage detection in pears, but the resolution and detection depth of OCT technique needed to be improved. The insignificant diference between the early damaged and sound areas of crown pears increases the difficulty of sorting by visual techniques [7]. To overcome the limitations of the above detection techniques, the hyperspectral technique is proposed to detect the damaged level of pears. The hyper- spectral technique is seen as an emerging nondestructive testing technique, and it provides a large amount of in- formation about the characteristics of sample [8].

Many scholars have studied the damage of fruit using the hyperspectral technique and have obtained many satisfactory results. Li et al. [9] used hyperspectral to achieve the detection of early damage in peaches. They compared the ability of short-wave NIR and long-wave NIR to distinguish bruises, and it concluded that the short-wave NIR had better discriminatory ability with an accuracy of 96.5% for bruised peaches. Tan et al. [10] used hyperspectral to detect early damage and determine the damaged level of apples with the accuracy of 97.5%. Sun et al. [11] used hyperspectral to identify diferent levels of damage in tomatoes with the accuracy of 90.93%. The above studies show the feasibility of using hyperspectral detection damage of fruit. Most of the studies based on hyperspectral detection quality of fruit used the reflectance spectra to establish analytical models [12].

The physical and chemical properties of fruit may changeinspection, and the samples with equatorial diameter be- tween 70 and 80 mm and weight between 250 and 280 g were stored at the room temperature of 25°C.

*Quantitative Damage Experiments.* The samples subjected to impact damage experiments using the pendulum device are shown in Figure 1(a). The pendulum consists of a pendulum arm and an impact surface, and the sample is fixed at the end of the pendulum arm. The pendulum arm is released by diferent angles, which can cause the diferent levels of damage samples. The samples caused by 30°, 40°, and 50° release are defined as level I, II, and III damage, respectively. Samples with diferent levels of damage are shown in Figure 1(b). Thesamples are divided into 4 groups, one of which served as thecontrol group, and the remaining 3 groups are subjected to 3 levels of impact damage experiments. 64 samples each of sound, I, II, and III damage are obtained by the impact damage experiments.

after being damaged, and these physical and chemical properties can be connected with the spectra [13]. The combination of Ryperspectral and chemometrics can be used to detect the change in the content of compounds in fruits [14]. However, there are a lot of background information, noise information, and useless information in the spectra when judging based on the reflectance spectra directly obtained by the spectral instrument, those lead to the low-discrimination accuracy (90.93%) of fruit with various degrees of damage. Therefore, the reflectance (*R*) spectra, absorbance (*A*) spectra, and Kubelka–Munk (*K-M*) spectra combining with chemometrics was proposed to identify the various degrees of damage of fruit to improve the identification accuracy.

## 2. Materials and Methods

*Crown Pear Sample.* The crown pear samples used for the experiment were obtained from local orchards in Jiangxi. Samples without external damage were selected by visual

*Acquisition of Hyperspectral Images.* Images of all samples are acquired by the Gaia Sorter hyperspectral instrument. The composition of the hyperspectral image acquisition system is shown in Figure 2, and the in- strument is manufactured by Dualix Spectral Imaging Ltd. Crown pear samples are sequentially placed on the transport platform, and the spectral acquisition is per- formed sequentially to generate a 3D data body con- taining image information and spectral information. The hyperspectral acquisition instrument is preheated for 30 min before acquisition to reduce the acquisition error caused by the baseline drift.

*Spectral Calibration.* The dark current and uneven distribution of light sources exist in the hyperspectral instrument when the acquisition is carried out directly, and they may cause large experimental errors, so the spectral correction process is carried out by

$$R = \frac{R_0 - R_y}{R_Z - R_y}, \quad (1)$$

where $R$ is the calibrated sample image acquisition data, $R_y$ is the all-black image acquisition data, $R_z$ is the all-white image acquisition data, and $R_0$ is the original sample image acquisition data.

*Spectral Extraction.* Spectral extraction is performed on the acquired hyperspectral images to obtain the spectral in- formation that it can be used to characterize the sample. The VIS-NIR spectra cover the wavelength range of 397.5~1014.0 nm with a resolution of 3.5 nm. The region of interest (ROI) is selected in the equatorial region of sound samples and damaged samples to reduce the uneven degree of light. The average value of the reflectance within the ROI is calculated as the reflectance spectra ($R$ of this sample. The $R$ spectra are calculated by (2) and (3) to obtain the absorbance ($A$) spectra and Kubelka–Munk ($K$-$M$) spectra, respectively.

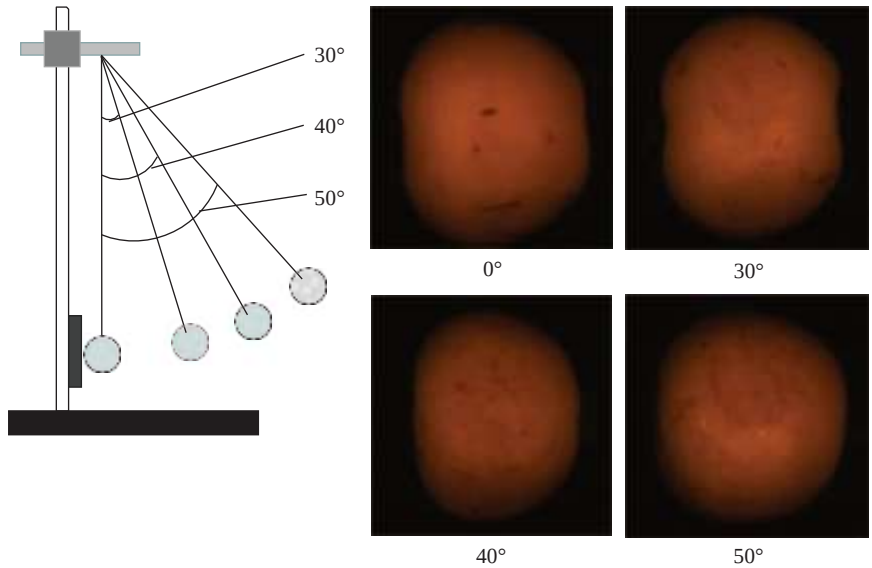



Figure 1: Pendulum quantitative damage of the experimental setup: (a) pendulum quantitative damage device, schematic; (b) sound and diferent levels of damage of the crown pear samples.

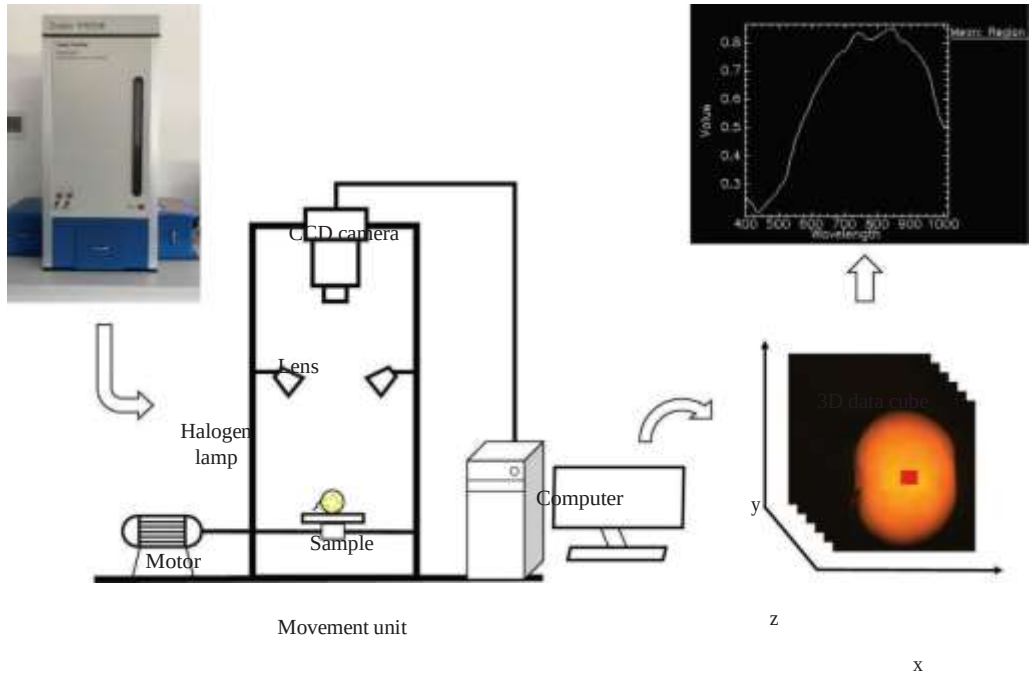



Figure 2: Schematic diagram of a hyperspectral acquisition system.

$$A = -\lg R, \tag{2}$$

$$K - M = \frac{(1-R)^2}{2R}. \tag{3}$$

*Spectral Data Pretreatment.* The *R*, *A*, and *K-M* spectral data may contain information about anomalies caused by some unrelated characteristics of the sample in addition to the information about the sample being tested. Therefore, the raw data need to be processed to eliminate or attenuate the influence of anomalous information and retain valid information to improve the applicability and robustness of the discrimination model. The pretreatment methods of Gaussian filter (GF), standard normal variation (SNV), Savitzky–Golay derivative (SGD), multiplicative scatter correction (MSC), and correlation optimized warping (COW) are commonly used.

*Characteristic Wavelength Selection.* There are multiple correlations between *R*, *A*, and *K-M* spectra at diferent wavelengths, resulting in redundant information in the

spectra. This redundant information reduces the discrimination speed and accuracy of the model, and therefore, the redundant information needs to be eliminated [15]. In practice, a good performing classification model requires not only high discrimination accuracy but also has fast discrimination speed; therefore, the original spectra can be downscaled by the competitive adaptive reweighted sampling (CARS) and the uninformative variables elimination (UVE) [16]. CARS is a sampling method for wavelength selection based on regression coefficients, and UVE is a sampling method for rejecting uninformative variables based on the value of the stability of the model variables. These two methods are widely used for the selection of characteristic wavelengths of spectral data [17].

*Model Building.* The Kennard–Stone (KS) algorithm achieves stratified sampling by uniformly selecting samples in the characteristic space based on the Euclidean distance between variables, reducing the impact on the final results due to the introduction of additional bias in the partitioning process [18]. The KS algorithm is used to divide the spectral data of the samples into two mutually exclusive calibration sets and test set (calibration set: test set 3 : 1). The partial least squares discriminant analysis (PLS-DA) and the sup- port vector machine (SVM) models are based on *R*, *A*, and *K-M* spectra, respectively. The PLS-DA is used to find a linear regression model by projecting the predictor and observed variables into a new space [19]. SVM can be used to perform nonlinear classification by kernel methods with a decision boundary of the maximum margin hyperplane solved for the learned samples [20]. The precision and accuracy can be used to evaluate the performance of the classification model. The precision based on the confusion matrix is usually evaluated by the *F*1 value, which is a har- monic mean based on the accuracy and recall. The closer the *F*1 value is closer to 1, the higher the precision of the model [21]. The *F*1 value is calculated by the following equation:

$$F1 = \frac{\mathrm{TP}}{2 \times \mathrm{TP} + \mathrm{FN} + \mathrm{FP}},$$

where TP is the true case, FN is the false counter case, and FP is the false positive case.

The *F*1 value is applicable to the binary classification model, for the evaluation of the accuracy rate of the multiclassification model needs to calculate the *F*1 value of each category, and the average of it is taken to get the macro-*F* value of the model, and the macro-*F* value is calculated by

$$\mathrm{Macro} - F = \frac{1}{n}\sum_{i=1}^{n} \mathrm{Fi},$$

plotted. The mean values of the damage levels of the *R* spectra are shown in Figure 3(a), with the sound samples exhibiting the highest spectral reflectance, followed by the level I damage and level II damage and finally the level III damage. There is a significant decrease in the spectral reflectance of the damaged crown pear, and the reason for the spectral diference is mainly due to the rupture of the cell membrane of the crown pear after damage, which led to water loss and easy oxidation, which is consistent with the results of research on apple damage [22]. Figures 3(b) and 3(c) show that the curve characteristics of the *A* spectral and *K-M* spectral are diferent, and the *A* spectral and *K-M* spectral curves show that the spectral values of level III damage, level I damage, level II damage, and sound samples become lower in sequence. The diferent trends and characteristics presented by the three spectral curves are the basis for model discrimination.

*Spectral Pretreatment.* The spectra are usually pretreated before building a discriminant model. The purpose of pretreatment is to extract valid information and eliminate background information and noise. The *R*, *A*, and *K-M* spectra of crown pears were pretreated based on GF, MSC, SNV, COW, and SGD, and the changes in the spectra after pretreatment were described as an example of reflectance spectra. Figure 4 shows that the spectral characteristics of the *R* spectra change significantly after diferent pretreatments, and these changes are more pronounced at the wave peaks and troughs. Gaussian filtering is calculated by Gaussian function to reduce the self-noise of raw spectra. MSC and SNV methods are mainly used to eliminate the scattering efect from the inhomogeneous distribution of particles and particle size. COW aligns two signals by segmenting the spectra with linear stretching and compression. SGD can be used to efectively eliminate the baseline and other background interference by performing 1st derivative on the spectral curve.

*Model Building and Analysis.* The accuracy of the dis- crimination results of PLS-DA based on *R* spectra, *A* spectra, and *K-M* spectra is shown in Table 1. The best model based on the *R* spectra is the model pretreated by SNV with the discrimination accuracies of 64.06% and 65.63% for the test and calibration sets. The best model based on *A* spectra is the model pretreated by COW with discrimination accuracies of 60.94% and 54.17% for the test and calibration sets. The best model based on *K-M* spectra is the model based on RAW

## 3. Results and Discussion

*Spectral Analysis of Crown Pear.* The *R*, *A*, and *K-M* spectral data for sound and samples with diferent levels of damage were averaged, and their spectral curves were

the test and calibration sets, respectively. From the dis- criminant results, it is clear that the *K-M* spectra exhibits the best modeling performance in the PLS-DA discriminant model.

The accuracy of the discrimination results of the SVM model based on *R* spectra, *A* spectra, and *K-M* spectra is shown in Table 2. The best model based on *R* spectra is the model based on RAW with discrimination accuracy of 96.88% and 98.96% for the test and calibration sets. The best model based on *A* spectra is the model based on RAW with

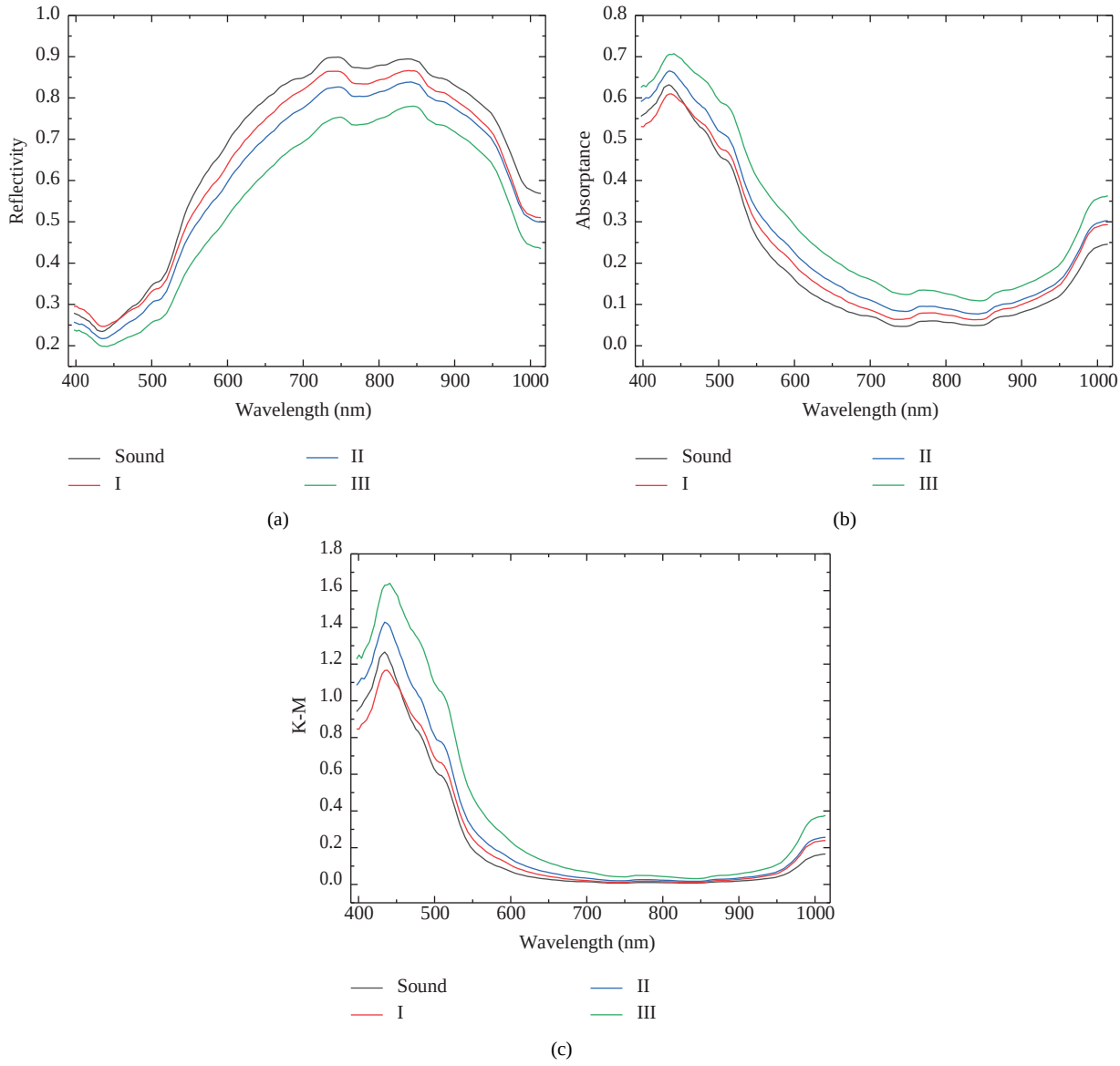



Figure 3: Raw spectral curves of(a) *R*, (b) *A* and, (c) *K-M*.

discrimination accuracies of 100% and 98.98% for the test and calibration sets. The best model based on *K-M* spectra is the model based on COW with discrimination accuracy of 95.31% and 98.44% for the test and calibration sets. From the discriminant results, it is clear that the SVM model has the best discriminant performance based on *A* spectra.

From the discriminant results in Tables 1 and 2, it is clear that the performance of the SVM model by GF is best among the models based on *R* spectra. Among the models based on the *A* spectra, the SVM model when the *A* spectra is not pretreated has the best performance. Among the models based on *K-M* spectra, the SVM model by COW pre-treatment has the best performance. The discrimination results of the models show that the SVM model has a higher discrimination accuracy than the PLS-DA model. The reason for the poor discriminatory ability of PLS-DA may be due to the fact that the spectral information does not have a linear correlation with the level of cellular water loss of the sample under test reducing the predictive ability. In contrast, SVM is a nonlinear classification by the kernel method, so it obtains high discrimination accuracy in the detection diferent damage levels of crown pears. Cao et al. [23] used hyper-spectral combined with SVM and PLS models to detect the damage level of pears, and the results showed that the SVM model was more accurate in predicting the damage level of pears. The results of the experimental discrimination show that SVM can be used to efectively discriminate the damage level of pears.

The discriminant results of the optimal models in *R*, *A*, and *K-M* spectra are analyzed separately. The optimal model based on *R* spectra is R-GF-SVM, and the confusion matrix of the model discrimination results is shown in Figure 5(a). 2 level II damage samples in the test set are misclassified as level I damage sample: 1 sound sample in the calibration set

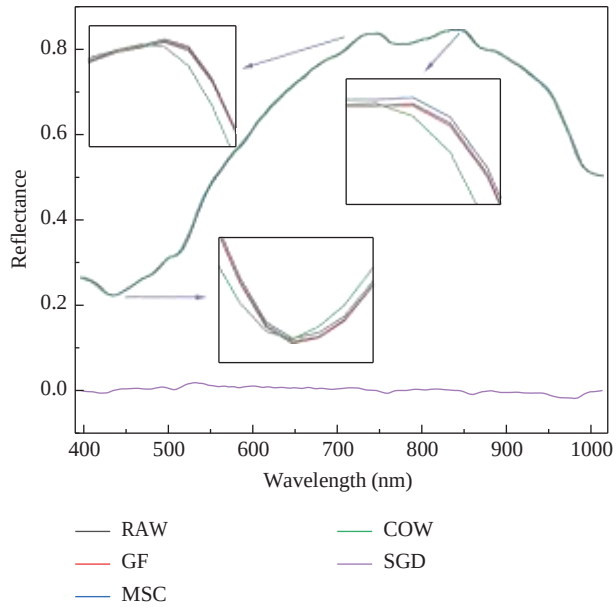



Figure 4: Curves of raw and pretreated *R* spectra.

Table 1: Discriminant accuracy of the PLS-DA model based on *R*, *A*, and *K-M* spectra.

| Spectra | Pretreatment | Accuracy (%), test set/calibration set | | | | |
|---|---|---|---|---|---|---|
| | | Sound | I | II | III | Total |
| *R* | RAW | 50.00/66.67 | 87.50/62.5 | 31.25/60.42 | 68.75/70.83 | 63.33/65.10 |
| | GF | 50.00/66.67 | 87.50/62.5 | 31.25/60.42 | 68.75/70.83 | 63.33/65.10 |
| | SNV | 62.50/75 | 87.50/64.58 | 43.75/60.42 | 62.50/62.5 | 64.06/65.63 |
| | MSC | 62.50/75 | 87.50/64.58 | 43.75/60.42 | 62.50/62.5 | 64.06/65.63 |
| | COW | 43.75/66.67 | 100/70.83 | 37.5/85.42 | 68.75/77.08 | 62.5/75 |
| | SGD | 50/72.92 | 81.25/81.25 | 43.75/70.83 | 62.5/77.08 | 59.38/74.48 |
| *A* | RAW | 43.75/68.75 | 87.50/75 | 37.5/79.17 | 62.50/72.92 | 57.81/73.96 |
| | GF | 50.00/68.75 | 93.75/75 | 37.5/79.17 | 62.50/70.83 | 57.81/73.44 |
| | SNV | 68.75/72.92 | 68.75/68.75 | 37.5/60.42 | 56.25/62.5 | 57.81/66.15 |
| | MSC | 68.75/72.92 | 68.75/68.75 | 37.5/60.42 | 56.25/62.5 | 57.81/66.15 |
| | COW | 62.5/62.5 | 35.42/93.75 | 18.75/54.17 | 68.75/64.58 | 60.94/54.17 |
| | SGD | 50/79.17 | 62.5/39.58 | 25/43.75 | 62.5/54.17 | 50/48.96 |
| *K-M* | RAW | 50/75 | 87.5/77.08 | 50/81.25 | 43.75/81.25 | 64.06/75.52 |
| | GF | 50/70.83 | 87.5/77.08 | 31.25/81.25 | 56.25/81.25 | 56.25/74.48 |
| | SNV | 43.75/81.25 | 75/70.83 | 37.5/62.5 | 56.25/60.42 | 53.13/68.75 |
| | MSC | 43.75/79.17 | 81.25/68.75 | 37.5/58.33 | 56.25/62.5 | 52.08/67.19 |
| | COW | 43.75/75 | 93.75/81.25 | 43.75/62.5 | 68.75/68.75 | 62.5/71.88 |
| | SGD | 37.5/58.33 | 93.75/64.58 | 31.25/54.17 | 50/56.25 | 53.13/58.33 |

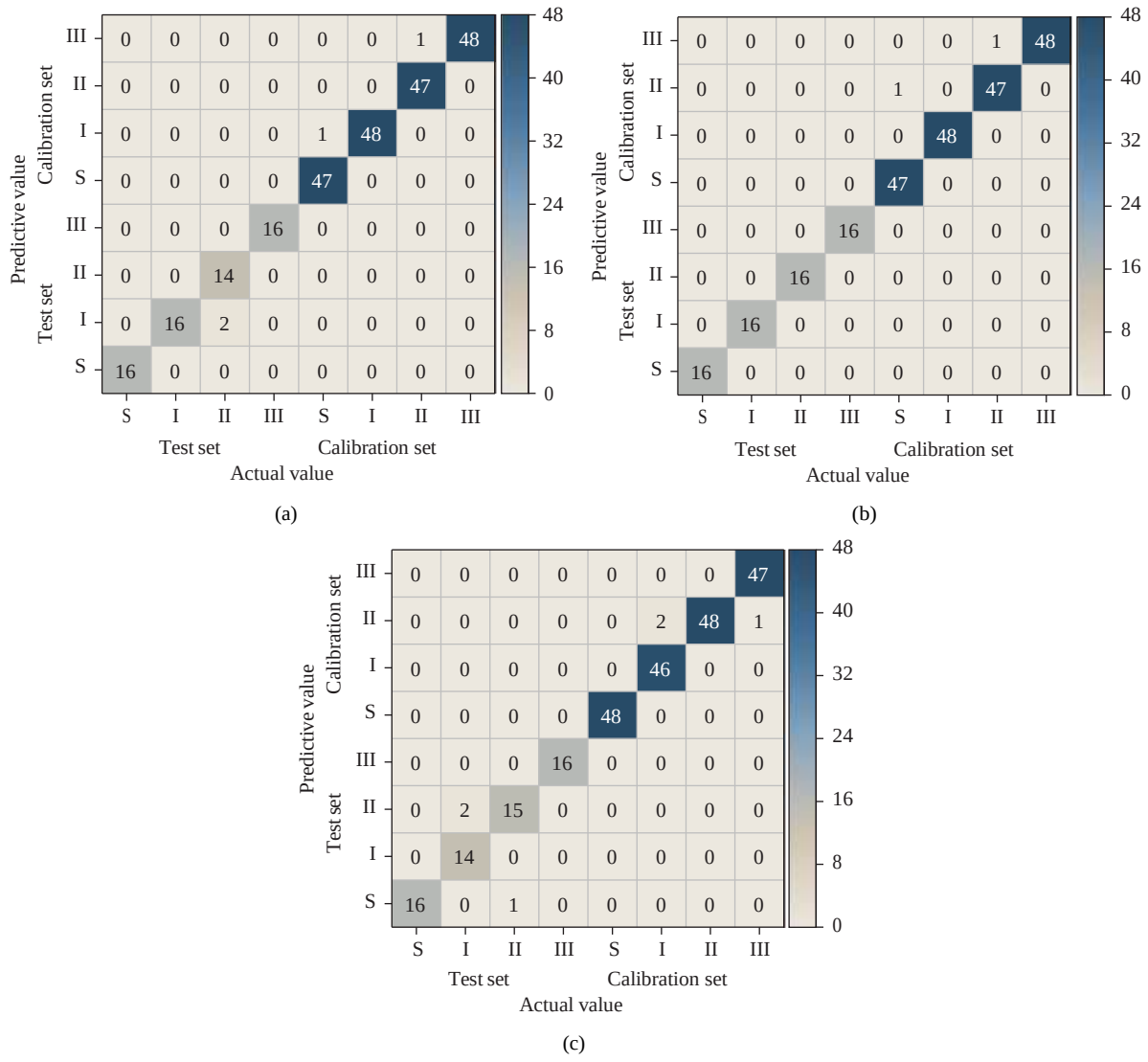
is misclassified as level I damage sample, and 1 level II damage sample is misclassified as level III damage sample. The optimal model based on *A* spectra is A-RAW-SVM, and the confusion matrix of model discrimination results is shown in Figure 5(b). There are no misclassified samples in the test set; 1 sound sample in the calibration set is misclassified as level II damage sample, and 1 level II damage sample is misclassified as level III damage sample. The optimal model based on *K-M* spectra is K-M-COW-SVM, and the confusion matrix of the model discrimination results is shown in Figure 5(c). 2 level I damage samples in the test set are misclassified as level II damage sample, and 1 level II damage sample is misclassified as 1 sound sample. 2 level I damage samples in the calibration set are misclassified as level II damage samples, and 1 level III damage sample is misclassified as level II damage sample.

The precision of the model discrimination is calculated, and the discrimination results are shown in Figure 5. The accuracy rates of the R-GF-SVM model, A-RAW-SVM model, and K-M-COW-SVM model are calculated by (4), and the results are shown in Table 3.

The macro-*F* value of the R-GF-SVM model, A- RAW-SVM model, and A-RAW-SVM is 0.9844, 0.9922, and 0.9767, respectively. The closer the macro-*F* value is to 1, the

Table 2: Discriminant accuracy of SVM models based on *R*, *A*, and *K-M* spectra.

| Spectra | Pretreatment | Accuracy (%), test set/calibration set | | | | |
|---|---|---|---|---|---|---|
| | | Sound | I | II | III | Total |
| *R* | RAW | 100/97.92 | 100/100 | 87.5/97.92 | 93.75/100 | 95.31/98.96 |
| | GF | 100/97.92 | 100/100 | 87.5/97.92 | 100/100 | 96.88/98.96 |
| | SNV | 98.44/100 | 81.25/95.83 | 81.25/100 | 81.25/97.92 | 84.38/98.44 |
| | MSC | 93.75/100 | 81.25/89.58 | 81.25/93.75 | 87.50/95.83 | 85.94/94.79 |
| | COW | 87.5/95.83 | 93.75/83.33 | 75/87.5 | 100/93.75 | 89.06/90.10 |
| | SGD | 93.75/95.83 | 81.25/87.5 | 68.75/85.42 | 87.5/95.83 | 81.25/91.15 |
| *A* | RAW | 100/98.44 | 100/100 | 100/98.44 | 100/100 | 100/98.98 |
| | GF | 100/98.44 | 93.75/100 | 93.75/100 | 100/100 | 96.88/99.48 |
| | SNV | 93.75/100 | 75/93.75 | 87.5/95.83 | 81.25/100 | 84.38/96.88 |
| | MSC | 93.75/100 | 75/93.75 | 87.5/95.83 | 81.25/98.44 | 84.38/96.88 |
| | COW | 100/97.92 | 87.5/100 | 87.5/100 | 93.75/95.83 | 92.19/98.44 |
| | SGD | 100/97.92 | 62.5/87.5 | 62.5/91.67 | 75/89.58 | 75/91.67 |
| *K-M* | RAW | 100/100 | 93.75/97.92 | 87.5/100 | 93.75/100 | 93.75/99.48 |
| | GF | 100/100 | 93.75/97.92 | 87.5/100 | 93.75/100 | 93.75/99.48 |
| | SNV | 87.5/100 | 85.71/91.67 | 68.75/89.58 | 87.5/95.83 | 79.69/94.27 |
| | MSC | 87.5/100 | 85.71/91.67 | 68.75/89.58 | 87.5/95.83 | 79.69/94.27 |
| | COW | 100/100 | 87.5/95.83 | 93.75/100 | 100/97.92 | 95.31/98.44 |
| | SGD | 75/100 | 87.5/100 | 68.75/100 | 93.75/97.92 | 79.69/99.48 |




(a) (b)

(c)

Figure 5: Confusion matrix of the discriminant results of the optimal SVM model based on *R*, *A*, and *K-M* spectra. Confusion matrix of discriminant results of the (a) R-GF-SVM model, (b) A-RAW-SVM model, and (c) K-M-COW-SVM model.

higher the precision of the model is indicated [24]. From Tables 2 and 3, it can be found that the performance model based on *A* spectra is best, and its accuracy and precision are higher than *R* spectra and *K-M* spectra.

*Selection of Characteristic Wavelength.* To simplify the spectral dataset and enhance the speed of the model run, CARS and UVE were used to filter the spectra with better modeling results for the characteristic wavelengths, respectively. A total of nine spectra were selected, and their discrimination accuracies were all greater than 89.06%, which were R-RAW, R-GF, R-COW, A-RAW, A-GF, A-COW, K-M-RAW, K-M-GF, and K- M-COW. The subset of wavelengths with the smallest RMSECV was obtained by first building a partial least squares model in CARS, and then, the wavelengths in the model were selected that correspond to the larger regression coefficients [25]. The A-RAW spectra are used as an example to introduce the characteristic wavelength selection of CARS, and its se-lection process is shown in Figure 6. CARS is performed 100 times sampling by sampling, and the number of sample variables gradually decreases as the number of samples sampled increase. RMSECV decreases and then increases to a minimum value of 0.4902 at 20 runs, with the characteristic wavelength determined by the minimum RMSEP during the sampling run [26].

The PLS model was established in UVE by cross-validation, and then, the ratio of the mean and standard deviation of the regression coefficients of the PLS model was calculated to select valid spectral information [27]. The wavelength selection process of UVE is introduced by using A-RAW spectra as an example, and it its selection process is shown in Figure 7. Yellow curve is the original variable and red curve is the filtered-free variable. The dotted line in Figure 7 shows the threshold splitting line with upper and lower limits of ±37.9147. According to the UVE selection principle, wavelengths within the threshold split line should be discarded, and wavelengths outside the threshold split line region are selected as characteristic wavelengths [28].

The results of the *R* spectra characteristic wavelength se-lection are shown in Figure 8. The R-RAW spectra have 10 characteristic wavelengths selected by the CARS algorithm, accounting for 5.68% of the total number of spectra. The 21 characteristic wavelengths selected by UVE for the R-RAW spectra are shown in Figure 8(b), accounting for 11.93% of the total number of spectra. The 10 characteristic wavelengths selected by CARS for R-GF spectra are shown in Figure 8(c), accounting for 5.68% of the total number of spectra. The 57 characteristic wavelengths selected by the R-GF spectra by the UVE algorithm are shown in Figure 8(d), accounting for 32.39% of the total number of spectra. The 16 characteristic wavelengths selected by CARS for the R-COW spectra are shown in Figure 8(e), accounting for 9.09% of the total number of spectra. The 111 characteristic wavelengths of the R-COW spectra selected by the UVE algorithm are shown in Figure 8(f), accounting for 63.07% of the total number of spectra.

The results of characteristic wavelength selection for the *A* spectra are shown in Figure 9. The 75 characteristic wavelengths selected by CARS for the A-RAW spectra are shown in Figure 9(a), accounting for 42.61% of the number of wavelengths in the full spectra. The 98 characteristic wavelength points selected by CARS for the A-RAW spectra are shown in Figure 9(b), accounting for 55.68% of the number of wavelengths in the full spectra. The 85 characteristic wavelengths selected by CARS for the A-GF spectra are shown in Figure 9(c), accounting for 48.3% of the number of wavelengths in the full spectra. The 97 characteristic wavelengths of the A-GF spectra selected by UVE are shown in Figure 9(d), accounting for 55.11% of the number of wavelengths in the full spectra. The 6 characteristic wavelengths selected by CARS for the A-COW spectra are shown in Figure 9(e), accounting for 3.41% of the number of wavelengths in the full spectra. The 71 characteristic wavelengths of the A-COW spectra selected by UVE are shown in Figure 9(f ), accounting for 40.34% of the number of wavelengths in the full spectra.

The results of characteristic wavelength selection for *K-M* spectra are shown in Figure 10. 75 characteristic wavelengths selected by CARS for K-M-RAW spectra are shown in Figure 10(a), accounting for 42.61% of the number of wave-lengths in the full spectra. The 98 characteristic wavelengths selected by CARS for K-M-RAW spectra are shown in Figure 10(b), accounting for 55.68% of the number of wave-lengths in the full spectra. The 43 characteristic wavelengths selected by CARS for K-M-GF spectra are shown in Figure 10(c), accounting for 24.43% of the number of wavelengths in the full spectra. The 103 characteristic wavelengths of K-M-GF spectra selected by UVE are shown in Figure 10(d), accounting for 58.52% of the number of wavelengths in the full spectra. The 33 characteristic wavelengths selected by CARS for K-M-COW spectra are shown in Figure 10(e), accounting for 18.75% of the number of wavelengths in the full spectra. The 60 characteristic wavelengths of A-GF spectra selected by UVE are shown in Figure 10(f), accounting for 34.09% of the number of wave-lengths in the full spectra.

From the characteristic wavelength selection results of CARS and UVE, it can be seen that the number of char-acteristic wavelengths selected by UVE is significantly more than the number of characteristic wavelengths selected by CARS. The reason for this may be that both CARS and UVE are characteristic wavelength selection methods based on the regression coefficients of the PLS model, but UVE uses the regression coefficients as a measure of wavelength impor-tance, while the CARS uses the absolute value of the re-gression coefficients as a measure of wavelength importance [22, 23]. Zhou et al. [6] used hyperspectral to detect sugar content in pears, and the characteristic wavelength selection of spectral wavelengths was performed by CARS and UVE, and their results showed that the number of characteristic wavelengths selected by UVE (390) was more than the number of characteristic wavelengths selected by CARS (42).

*Results and Analysis of Spectral Models of Characteristic Wavelengths.* The SVM model based on the characteristic spectra selected by CARS and UVE screening was selected with the number of wavelengths from 6 (A-COW-CARS) to 111 (K-M-RAW-UVE). The discrimination results of the

Table 3: Results of macro-*F* for SVM models based on R-GF, A-RAW, and K-M-COW spectra.

| Discriminant model | Fi value | | | | Macro-*F* value |
|---|---|---|---|---|---|
| | Sound | I | II | III | |
| R-GF-SVM | 0.9921 | 0.9771 | 0.9760 | 0.9922 | 0.9844 |
| A-RAW-SVM | 0.9921 | 1 | 0.9844 | 0.9922 | 0.9922 |
| K-M-COW-SVM | 0.9922 | 0.9677 | 0.9546 | 0.9921 | 0.9767 |

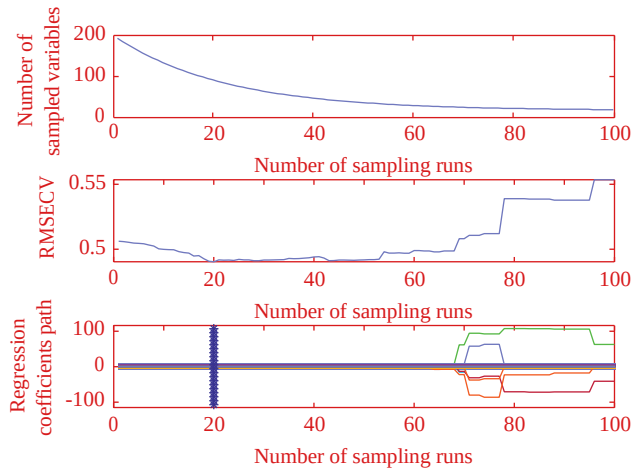



Figure 6: Process and results of A-RAW-CARS characteristic wavelength selection.

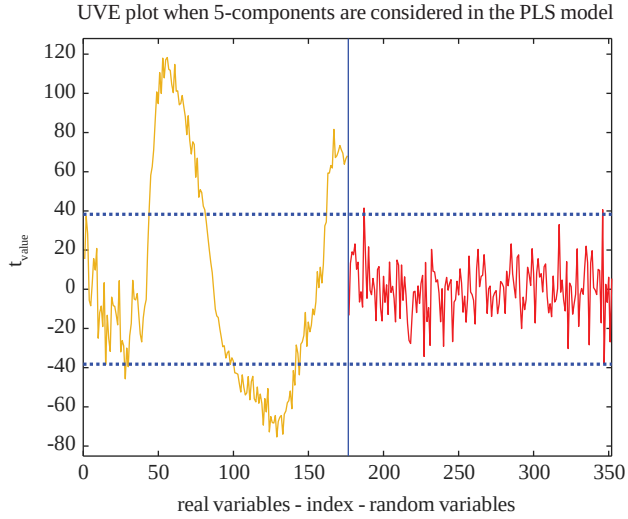



Figure 7: Process and results of A-RAW-UVE characteristic wavelength selection.

SVM model are shown in Table 4. The 58 characteristic wavelengths of the R-GF spectra selected by UVE in the *R* spectra have the highest discrimination accuracy, with 95.32% and 100% for the test and calibration sets. The 75 characteristic wavelengths of the A-RAW spectra selected by CARS in the *A* spectra have the highest discrimination accuracy, with 96.88% and 100% for the test and calibration sets. The 43 characteristic wavelengths of K-M-GF spectra selected by CARS in *K-M* spectra have the highest discrimination accuracy, with 93.75% and 96.88% for the test and calibration sets.

The optimal model based on the *R* spectra after characteristic wavelength selection is the R-RAW-UVE-SVM model, and the confusion matrix of model discrimination results are shown in Figure 11(a). 1 sound sample in the test set is misclassified as level I damage sample, and 1 level II damage sample is misclassified as level I damage sample. 2 sound samples in the calibration set are misclassified as level II damage samples, 2 level I damage samples are misclassified as level II damage samples, 2 level II damage samples are misclassified as level I damage samples and level III damage, and 2 level III damage samples are misclassified

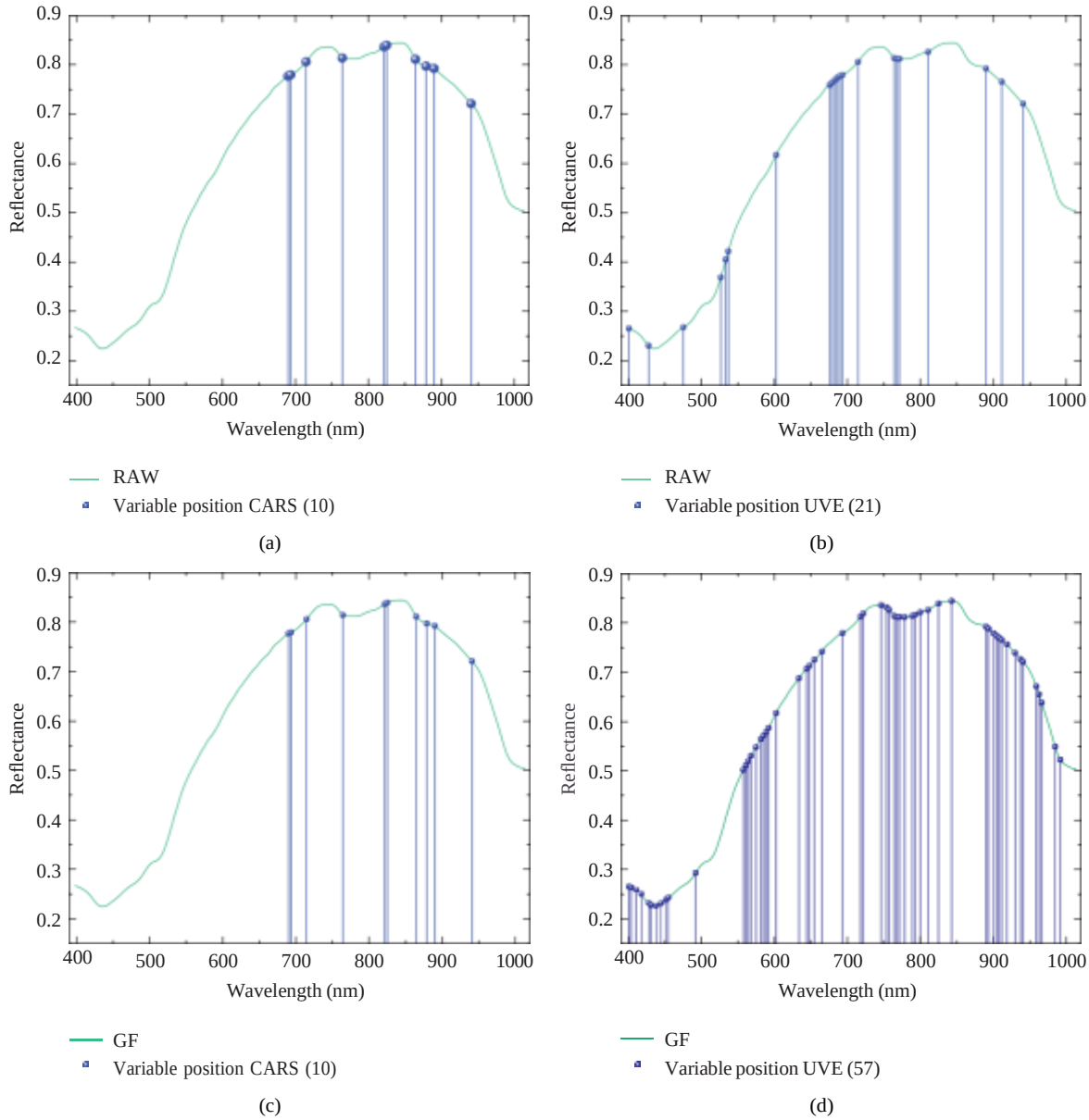



Figure 8: Continued.

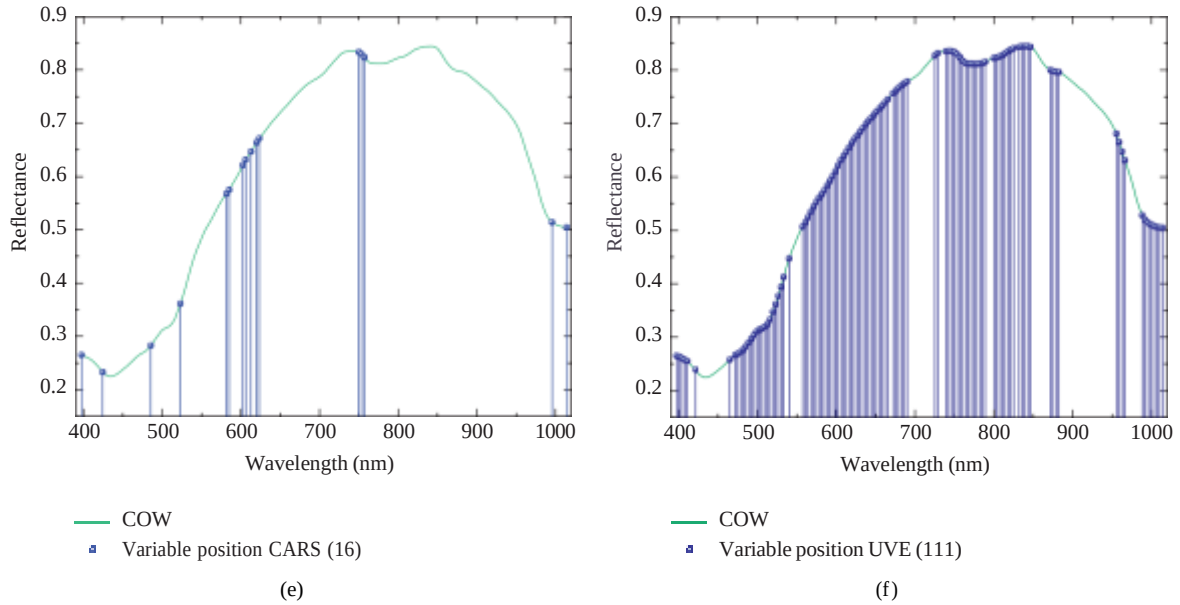



Figure 8: Results of *R* spectra characteristic wavelength selection based on CARS and UVE. Characteristic wavelength of R-RAW spectra selected based on (a) CARS and (b) UVE. Characteristic wavelength of R-GF spectra selected based on (c) CARS and (d) UVE. Characteristic wavelength of R-COW spectra selected based on (e) CARS and (f) UVE.

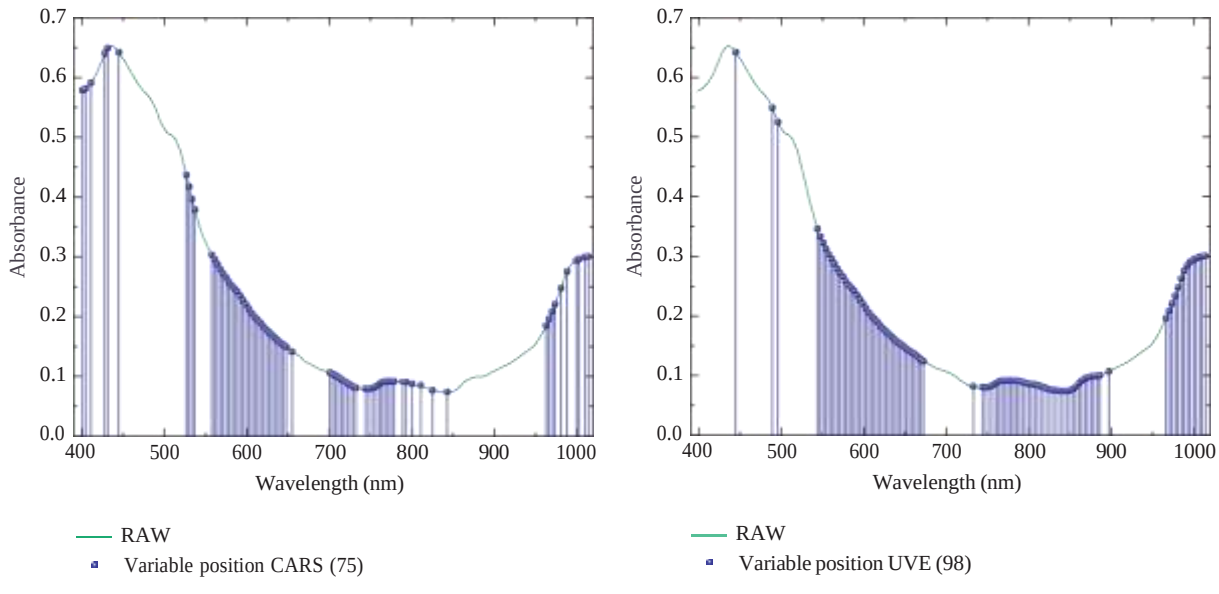



Figure 9: Continued.

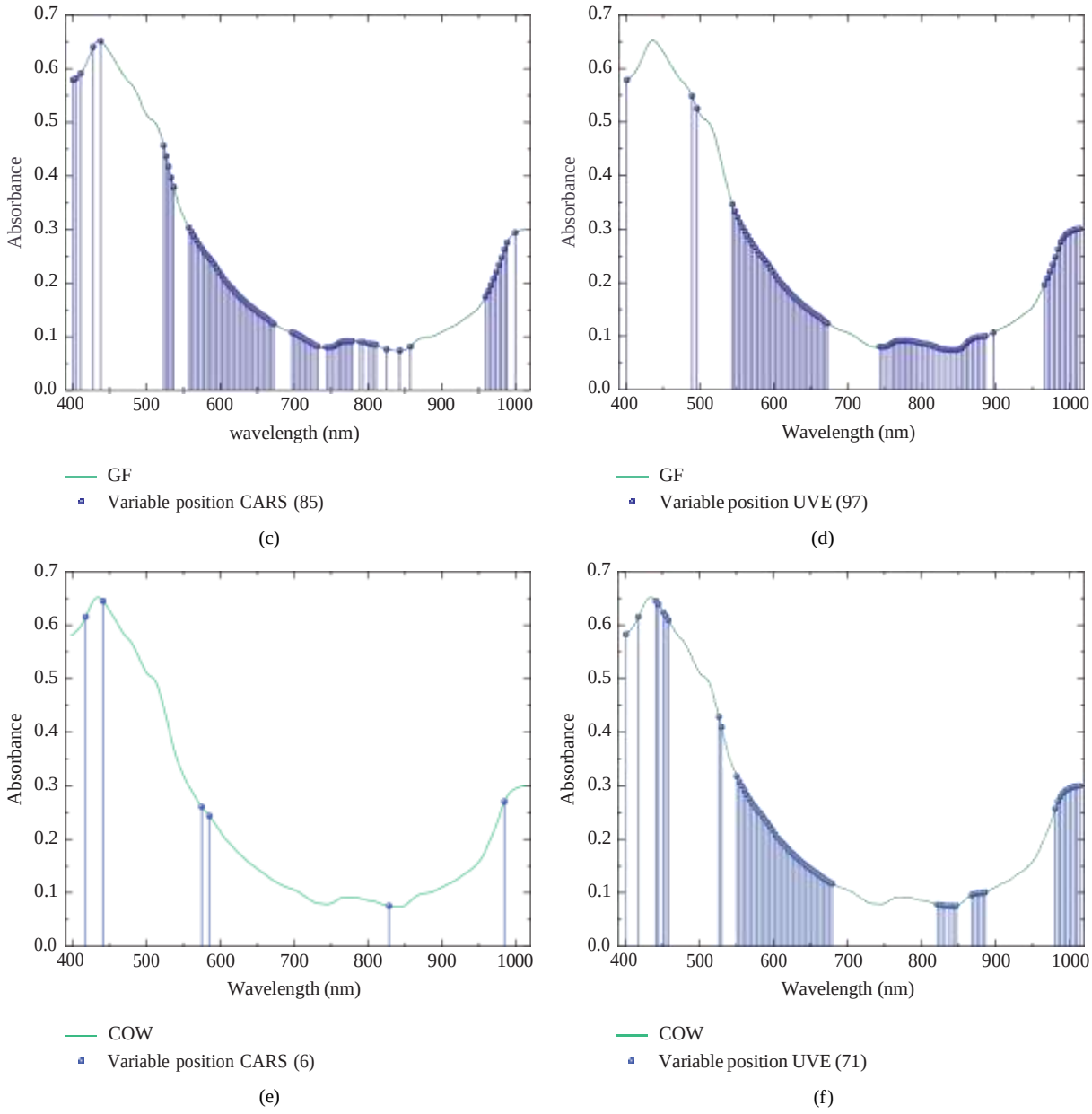



Figure 9: Results of the selection of *A* spectra characteristic wavelengths based on CARS and UVE. Characteristic wavelength of the A-RAW spectra selected based on (a) CARS and (b) UVE. Characteristic wavelength of the A-GF spectra selected based on (c) CARS and (d) UVE. Characteristic wavelength of the A-COW spectra selected based on (e) CARS and (f) UVE.

as level II damage samples. The optimal model based on the *A* spectra after characteristic wavelength selection is A-RAW-CARS-SVM, and the confusion matrix of the model discrimination results is shown in Figure 11(b). 1 sound sample in the test set is misclassified as the level II damage sample, and 1 level II damage sample was misclassified as the level I damage sample. There are no misclassified samples in the calibration set. The optimal model based on the *A* spectra after characteristic wavelength selection is A-RAW-CARS-SVM. The confusion matrix of the model discrimination results is shown in Figure 11(b), where 1 sound sample in the test set is misclassified as level II damage sample and 1 level II damage sample is misclassified as the level I damage sample. There are no misclassified samples in the calibration set. The optimal model based on *K-M* spectra is the K-M-GF-UVE-SVM model. The confusion matrix of model discrimination results is shown in Figure 11(c). 1 sound sample in the test set is misclassified as level I damage sample; 2 level II damage samples is misclassified as level I damage samples. 2 sound samples in the calibration set are misclassified as level II damage samples, 1 sound sample is misclassified as level I damage sample, 2 level I damage samples are misclassified as level II damage samples, 1 level I damage sample is misclassified as level III damage sample, 2 level III damage samples are misclassified as level I damage, 2 level III damage samples are misclassified as level II damage

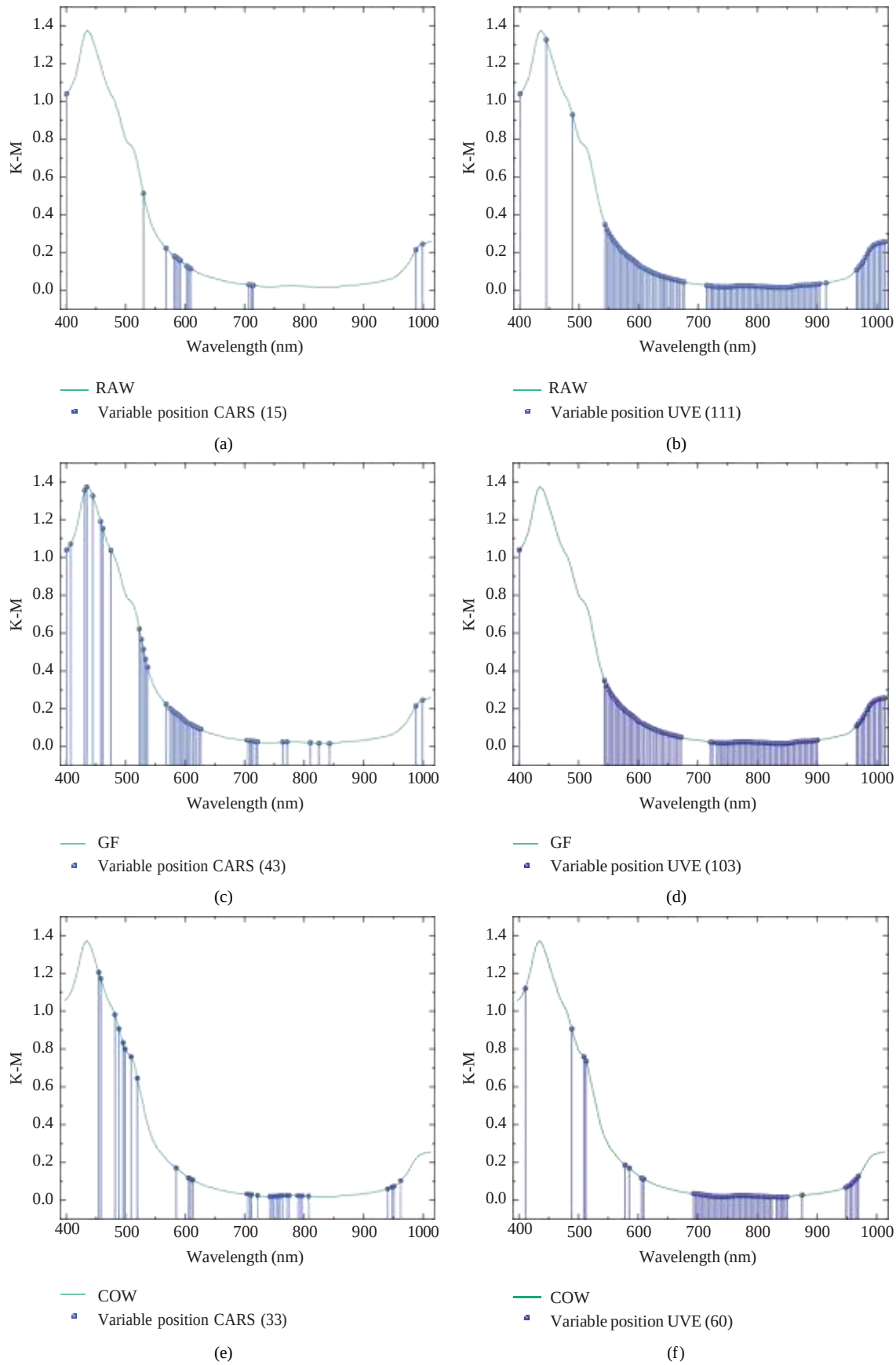



Figure 10: Results of *K-M* spectral characteristic wavelength selection based on CARS and UVE. Characteristic wavelength of the K-M-RAW spectra selected based on (a) CARS and (b) UVE. Characteristic wavelength of the K-M-GF spectra selected based on (c) CARS and (d) UVE. Characteristic wavelength of the K-M-COW spectra selected based on (e) CARS and (f) UVE.

Table 4: Discriminant results of SVM based on the characteristic wavelengths of *R*, *A*, and *K-M* spectra.

| Spectra | Pretreatment | Wavelength selection | Number of characteristic wavelengths | Discriminant accuracy (%) | |
|---|---|---|---|---|---|
| | | | | Test set | Calibration set |
| R | RAW | CARS | 10 | 84.38 | 81.81 |
| | GF | | 10 | 85.94 | 83.33 |
| | COW | | 16 | 93.75 | 92.19 |
| | RAW | UVE | 21 | 96.88 | 95.83 |
| | GF | | 58 | 95.31 | 100 |
| | COW | | 111 | 95.31 | 93.75 |
| *A* | RAW | CARS | 75 | 96.88 | 100 |
| | GF | | 85 | 93.75 | 97.4 |
| | COW | | 6 | 82.81 | 90.63 |
| | RAW | UVE | 98 | 90.63 | 97.4 |
| | GF | | 97 | 89.06 | 91.67 |
| | COW | | 71 | 92.19 | 99.48 |
| *K-M* | RAW | CARS | 15 | 87.5 | 91.67 |
| | GF | | 43 | 93.75 | 96.88 |
| | COW | | 33 | 78.13 | 86.46 |
| | RAW | UVE | 111 | 93.75 | 94.79 |
| | GF | | 103 | 95.31 | 94.27 |
| | COW | | 60 | 93.75 | 94.27 |

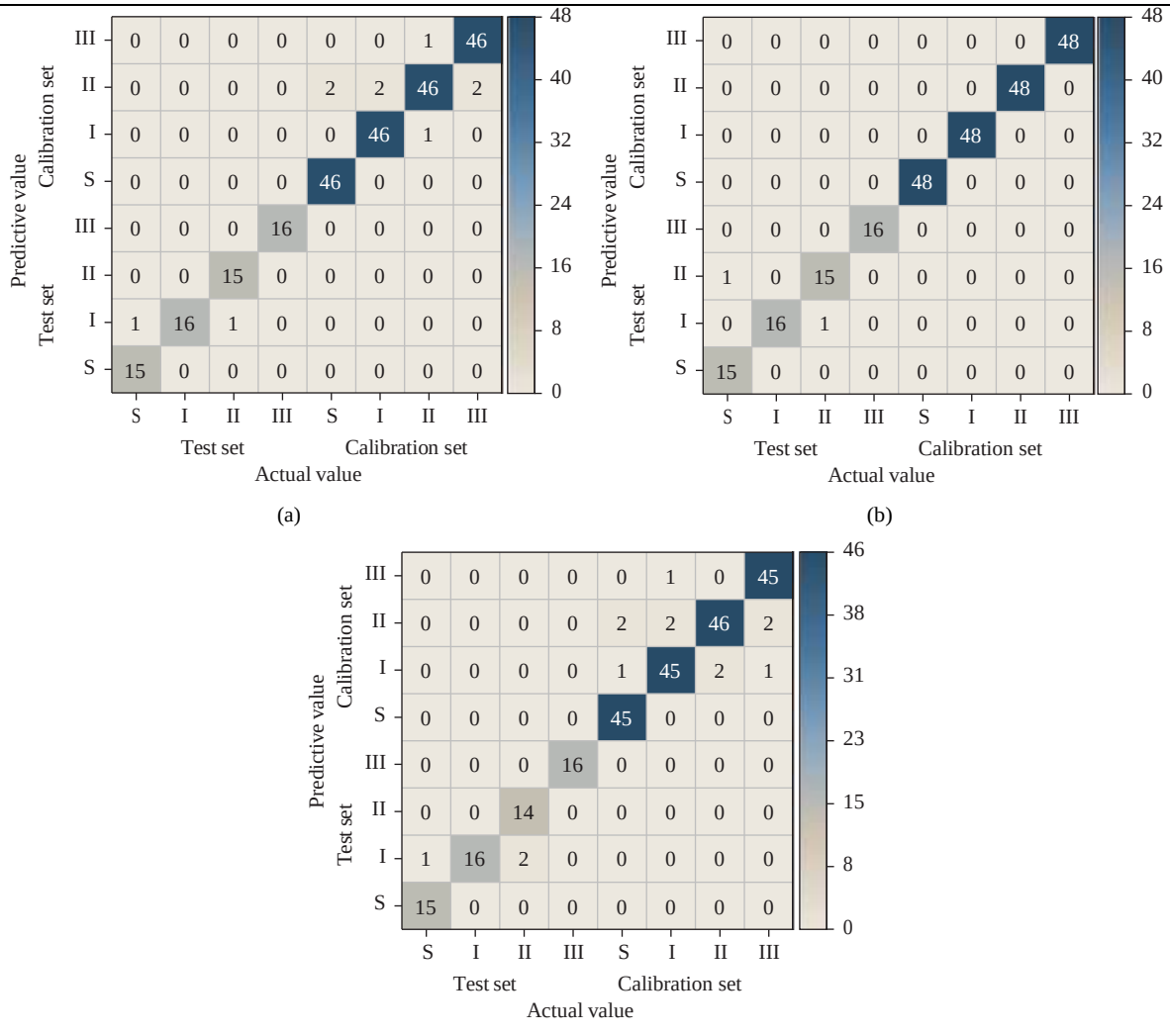



Figure 11: Confusion matrix of the optimal SVM model discrimination results based on *R*, *A*, and *K-M* spectra after wavelength selection. Confusion matrix for the discriminant results of the (a) R-RAW-UVE-SVM model, (b) A-RAW-CARS-SVM model, and (c) K-M-GF-UVE-SVM model.

Table 5: Macro-*F*1 results of SVM model discrimination based on R-RAW-UVE, A-RAW-CARS, and K-M-GF-UVE spectra.

| Spectral models | Fi value | | | | Macro-*F* value |
|---|---|---|---|---|---|
| | Sound | I | II | III | |
| R-RAW-UVE | 0.9760 | 0.9582 | 0.9313 | 0.9764 | 0.9605 |
| A-RAW-CARS | 0.9921 | 0.9922 | 0.9844 | 1 | 0.9922 |
| F-UVE | 0.9677 | 0.9243 | 0.9231 | 0.9683 | 0.9481 |

samples, and 1 level III damage sample is misclassified as level I damage sample.

The discriminant results are shown in the confusion matrix of Figure 11, the macro-*F* values of the R-RAW-UVE-SVM model, A-RAW-CARS-SVM model, and K-M-GF-UVE-SVM model are calculated by (4), and the results are shown in Table 5.

Table 5 shows that the macro-*F* value of the R-RAW-UVE-SVM model, RAW-CARS-SVM model, and K-M-GF-SVM model is 0.9605, 0.9922, and 0.9481, respectively. The closer the macro-*F* value is to 1, the higher the precision of the model is indicated. From Tables 4 and 5, it can be found that the accuracy and precision of the SVM model based on the *A* spectra after characteristic wavelength selection are higher than those of the *R* spectra and *K-M* spectra after characteristic wavelength selection. The results demonstrate that the *A* spectra by CARS has the best model discrimination performance. The CARS, which is widely used for characteristic wavelength selection of spectral data, is an efective characteristic wavelength selection method. Xuan et al. [29] used the characteristic wavelengths of *R* spectra selected by CARS to build MLR models to assess soluble solids (SSC) and hardness of peaches, and the results showed that hyperspectral combined with MLR had best discrimination results. Zhan et al. [30] selected the characteristic wavelengths of hyperspectral by pairing methods such as CARS and UVE. The results showed that the nonlinear model partial least squares support vector machine (LS-SVM) by CARS-selected spectra had the best results for pear fragrance identification.

## 4. Conclusions

In this study, the *R* spectra of sound as well as level I, II, and III damaged crown pears were acquired separately using a hyperspectral acquisition system, and then, the *A* spectra and *K-M* spectra were obtained from the *R* spectra by transformation of the equations. Linear PLS-DA models and nonlinear SVM models were developed for each of the three spectra and the discriminant results were explained. It was found that as the level of damage of the crown pears increases, the spectral reflectance decreased and the spectral absorbance increased. In the *R*, *A*, and *K-M* spectra models, the nonlinear SVM model was found to outperform the linear PLS-DA model by discriminating the results. The SVM discriminant model based on the *A* spectra outperformed the *R* and *K-M* spectra, and the discriminant accuracy of the test and correction sets of the SVM model based on A-RAW spectra was 100% and 98.98%, respectively. Moreover, the macro-*F* value of the model was 0.9922, and the A-RAW-SVM model showed high discrimination accuracy. Two false samples were found in the SVM discriminant model based on the characteristic wavelengths of the A-RAW spectra after CARS selection, and the number of false samples was the same as before CARS selection. The macro-*F* value of the A-RAW-CARS-SVM model after CARS selection was 0.9922, and the accuracy of the model was the same as that of the full spectra model, while the discriminant time of the model time was greatly reduced. In conclusion, the hyperspectral technology can be used to discriminate diferent damage levels of crown pears, and the nonlinear SVM model based on the *A* spectra has better discriminative efect on diferent damage levels of crown pears.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (12103019) and National Science and Technology Award Backup Project Cultivation Plan (20192AEI91007).

## References

[1] H. Zhan, C. Qin, and F. Guo, “Study on oligosaccharides treatment alleviating peel browning of Huangguan pear,” *Journal of Chinese Institute of Food Science and Technology*, vol. 22, no. 4, pp. 276–284, 2022.

[2] Z. Ren, Y. Cheng, and Y. Guan, “Efects of mechanical damage on browning in ‘Yali’ pear,” *Food Science and Technology*, vol. 46, no. 6, pp. 35–41, 2021.

[3] S. Zhang, X. Wu, S. Zhang, Q. Cheng, and Z. Tan, “An effective method to inspect and classify the bruising degree of apples based on the optical properties,” *Postharvest Biology and Technology*, vol. 127, pp. 44–52, 2017.

[4] G. Kim, G. H. Kim, J. Park, D. Y. Kim, and B. K. Cho, “Application of infrared lock-in thermography for the quantitative evaluation of bruises on pears,” *Infrared Physics & Technology*, vol. 63, pp. 133–139, 2014.

[5] Y. Luo, X. Jiang, and X. Fu, “Spatial frequency domain imaging system calibration, correction and application for pear surface damage detection,” *Foods*, vol. 10, no. 9, p. 2151, 2021.

[6] Y. Zhou, J. Mao, D. Wu et al., “Nondestructive early detection of bruising in pear fruit using optical coherence tomography,” *Korean journal of horticultural science and technology*, 2019.

[7] E. Mavridou, E. Vrochidou, G. A. Papakostas, T. Pachidis, and V. G. Kaburlasos, “Machine vision systems in precision agriculture for crop farming,” *Journal of Imaging*, vol. 5, no. 12, p. 89, 2019.

[8] M. R. Mogollon, A. F. Jara, C. Contreras, and J. Zofoli, “Quantitative and qualitative VIS-NIR models for early determination of internal browning in ‘Cripps Pink’ apples

during cold storage," *Postharvest Biology and Technology*, vol. 161, Article ID 111060, 2020.

[9] J. Li, L. Chen, and W. Huang, "Detection of early bruises on peaches (Amygdalus persica L.) using hyperspectral imaging coupled with improved watershed segmentation algorithm," *Postharvest Biology and Technology*, vol. 135, pp. 104–113, 2018.

[10] W. Tan, L. Sun, F. Yang et al., "The feasibility of early detection and grading of apple bruises using hyperspectral imaging," *Journal of Chemometrics*, vol. 32, no. 10, Article ID e3067, 2018.

[11] Y. Sun, I. Pessane, L. Pan, and X. Wang, "Hyperspectral characteristics of bruised tomatoes as afected by drop height and fruit size," *Lebensmittel-Wissenschaft & Technologie*, vol. 141, Article ID 110863, 2021.

[12] N. T. Vetrekar, R. S. Gad, I. Fernandes et al., "Non-invasive hyperspectral imaging approach for fruit quality control application and classification: case study of apple, chikoo, guava fruits," *Journal of Food Science & Technology*, vol. 52, no. 11, pp. 6978–6989, 2015.

[13] I. A. d Moraes, L. J. P. Cruz-Tirado, and D. F. Barbin, "Online measurement of carambola (Averrhoa carambola L.) physicochemical properties and estimation of maturity stages using a portable NIR spectrometer," *Scientia Horticulturae*, vol. 304, Article ID 111263, 2022.

[14] J. P. Cruz-Tirado, P. R. Lopes de França, and D. Fernandes Barbin, "Chia (Salvia hispanica) seeds degradation studied by fuzzy-c mean (FCM) and hyperspectral imaging and chemometrics - fatty acids quantification," *Scientia Agropecuaria*, vol. 13, no. 2, pp. 167–174, 2022.

[15] Y. Fang, F. Yang, Z. Zhou, L. Lin, and X. Li, "Hyperspectral wavelength selection and integration for bruise detection of korla pears," *Journal of Spectroscopy*, vol. 2019, Article ID 6715247, 8 pages, 2019.

[16] T. Wang, G. Li, and C. Dai, "Soluble Solids Content prediction for Korla fragrant pears using hyperspectral imaging and GsMIA," *Infrared Physics & Technology*, vol. 123, Article ID 104119, 2022.

[17] J. Li, S. Fan, and W. Huang, "Assessment of multiregion local models for detection of SSC of whole peach (*Amygdalus persica* L.) by combining both hyperspectral imaging and wavelength optimization methods," *Journal of Food Process Engineering*, vol. 41, no. 8, Article ID e12914, 2018.

[18] X. Wei, J. He, S. Zheng, and D. Ye, "Modeling for SSC and firmness detection of persimmon based on NIR hyperspectral imaging by sample partitioning and variables selection," *Infrared Physics & Technology*, vol. 105, Article ID 103099, 2020.

[19] J. M. Amigo, C. Ravn, N. B. Gallagher, and R. Bro, "A comparison of a common approach to partial least squares-discriminant analysis and classical least squares in hyperspectral imaging," *International Journal of Pharmaceutics*, vol. 373, no. 1-2, pp. 179–182, 2009.

[20] K. Tan and P. J. Du, "Hyperspectral remote sensing image classification based on support vector machine," *Journal of Infrared and Millimeter Waves*, vol. 27, no. 2, pp. 123–128, 2008.

[21] X. Li, H. Jiang, X. Jiang, and M. Shi, "Identification of geographical origin of Chinese chestnuts using hyperspectral imaging with 1D-CNN algorithm," *Agriculture*, vol. 11, no. 12, p. 1274, 2021.

[22] X. Luo, T. Takahashi, K. Kyo, and S Zhang, "Wavelength selection in vis/NIR spectra for detection of bruises on apples by ROC analysis," *Journal of Food Engineering*, vol. 109, no. 3, pp. 457–466, 2012.

[23] F. Cao, D. Wu, J. T. Zheng, Y. D. Bao, Z. Y. Wang, and Y He, "[Detection of pear injury based on visible-near infrared spectroscopy and multispectral image]," *Guang pu xue yu Guang pu fen xi*[illegible]*Guang pu*, vol. 31, no. 4, pp. 920–923, 2011.

[24] M. H. Hu, Y. Zhao, and G. T. Zhai, "Active learning algorithm can establish classifier of blueberry damage with very small training dataset using hyperspectral transmittance data," *Chemometrics and Intelligent Laboratory Systems*, vol. 172, pp. 52–57, 2018.

[25] S. Gao and J. Xu, "Discrimination and distribution visualization of total acidity and hardness of red globe grapes based on hyperspectral imaging," *Food Science*, pp. 1–5, 2022.

[26] W. H. Su, C. Yang, Y. Dong et al., "Hyperspectral imaging and improved feature variable selection for automated determination of deoxynivalenol in various genetic lines of barley kernels for resistance screening," *Food Chemistry*, vol. 343, Article ID 128507, 2021.

[27] S. Wang, M. Huang, and Q. Zhu, "Model fusion for prediction of apple firmness using hyperspectral scattering image," *Computers and Electronics in Agriculture*, vol. 80, pp. 1–7, 2012.

[28] M. H. Hu, Q. L. Dong, B. L. Liu, and U. L. Opara, "Prediction of mechanical properties of blueberry using hyperspectral interactance imaging," *Postharvest Biology and Technology*, vol. 115, pp. 122–131, 2016.

[29] G. Xuan, C. Gao, and Y. Shao, "Spectral and image analysis of hyperspectral data for internal and external quality assessment of peach fruit," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 272, Article ID 121016, 2022.

[30] B. S. Zhan, J. H. Ni, and J. Li, "[Hyperspectral technology combined with CARS algorithm to quantitatively determine the SSC in Korla fragrant pear]," *Spectroscopy and Spectral Analysis*, vol. 34, no. 10, pp. 2752–2757, 2014.