Detection of early collision and compression bruises for pears based on hyperspectral imaging technology

Guanglai Wang, Congcong Wang, Dayang Liu

Publisher's Disclaimer

Detection of early collision and compression bruises for pears based on hyperspectral imaging technology

Guanglai Wang,[1] Congcong Wang,[1] Dayang Liu[2]

[1]College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin, Heilongjiang; [2]College of Computer and Control Engineering, Northeast Forestry University, Harbin, Heilongjiang, China

Correspondence: Dayang Liu, College of Computer and Control Engineering, Northeast Forestry University, Harbin, Heilongjiang 150040, China.
E-mail: ldy333ldy@163.com

Key words: pear; early bruise; bruised time; hyperspectral imaging technology; STOA-GA-SVM.

Conflict of interest: the authors declare no potential conflict of interest.

Availability of data and material: the data used in this study are available upon request.

Abstract

Early detection of bruising is one of the major challenges in postharvest quality sorting processes for pears. In this study, visible/near infrared (VIS/NIR) hyperspectral imaging (400–1000 nm) was utilized for early detection of pear bruise type and timing (1, 12, and 24 h post-bruise). Spectral images of nonbruised and mechanically bruised pears (collision and compression) were captured at these intervals for modeling. Spectral data was processed using principal component analysis (PCA) and uninformative variable elimination (UVE) to select optimum wavelengths. Classification models were then built using an extreme learning machine (ELM) and support vector machine (SVM), and compared with a model combining genetic algorithm, sooty tern optimization algorithm, and SVM (STOA-GA-SVM). For PCA-ELM, UVE-ELM, PCA-SVM, and UVE-SVM models, the calibration set accuracies were 98.99%, 98.98%, 96.94%, and 99.23% respectively. And the validation set accuracies were 89.29%, 87.97%, 88.78%, and 88.78% respectively. The STOA-GA-SVM model shows the best performance, and the accuracy of the calibration set and validation set is determined to be 97.19% and 92.86%, respectively. This study shows that the use of the VIS/NIR hyperspectral imaging technique combined with the STOA-GA-SVM algorithm is feasible for the rapid and nondestructive identification of the bruise type and time for pears.

Introduction

Pear (*Pyrus spp.*) is a favorite fruit of consumers due to its rich vitamins and minerals with high nutritional value. However, pears are easily bruised by various activities during harvesting, transportation, grading, and sorting, such as collision and compression (Arango *et al.*, 2021; Stropek and Gołacki, 2015). Bruising not only injures the tissue structure of the pear, leading to loss of nutritional value and increased risk of microbial infestation, but also degrades the appearance quality of pears, causing significant economic loss (Li *et al.*, 2016; Opara and Pathare, 2014). Hence, detecting and distinguishing the type and timing of pear bruising is crucial as it provides a reference for quickly identifying the process causing the bruise and taking prompt measures to reduce economic losses. Furthermore, insights into the specific types and timings of the bruises can guide the design enhancements and optimization of equipment used in the harvest, grading, packaging, storage, and transportation of fruits and vegetables (D. Y. Liu *et al.*, 2024). Consequently, accurately identifying the type and timing of pear bruises plays a crucial role. It holds significant importance for the pear

industry (Opara and Pathare, 2014; Yuan *et al.*, 2021).

Near-infrared spectroscopy (NIRS), as an advanced technique, has been successfully applied to the nondestructive detection of the internal quality of various fruits and vegetables (Nicolaï *et al.*, 2014), such as persimmons (Hasnah Ar *et al.*, 2019), kiwifruits (Liu *et al.*, 2017), and peaches (Guo *et al.*, 2016). However, NIR spectra are usually obtained at a single point for a sample. Hyperspectral imaging technology has emerged as a relatively rapid and accurate detection technique for horticultural products over the past decade (Zhang *et al.*, 2014) because it can be used to provide more detailed and complete information, including internal structure characteristics, morphological information, and chemical composition for a sample (Rahman *et al.*, 2017). Recently, hyperspectral imaging technology has been widely used for the detection of bruising in fruits, such as apple (Huang *et al.*, 2015; Xing and De Baerdemaeker, 2005), peach (Li *et al.*, 2022b; Li *et al.*, 2018; Li *et al.*, 2021), blueberry (Fan *et al.*, 2017), kiwifruit (Gao *et al.*, 2021), jujube (Thien Pham and Liou, 2022; Yuan *et al.*, 2021), etc.

Recently, hyperspectral imaging systems have been widely used to detect early bruises on pears. However, most current studies focus only on identifying early bruises (Fang *et al.*, 2019; Jiang *et al.*, 2016; Lee *et al.*, 2014; Li *et al.*, 2024), with fewer investigations simultaneously addressing multiple factors such as bruising type, and timing (Fu and Wang, 2022; Liu *et al.*, 2023; Liu *et al.*, 2024; Su *et al.*, 2022). Despite some advancements, these studies still have limitations. Currently, there are two primary methods for detecting pear bruises using hyperspectral imaging. The first method involves stepwise feature spectral extraction followed by target identification model building, which still relies on manual expertise to determine whether the selected features and classification models perform well. The second approach employs deep learning for an end-to-end methodology that autonomously extracts features from hyperspectral images to detect fruit damage. While this method shows promise, it requires extensive datasets and robust hardware.

Consequently, this paper proposes a novel approach combining the genetic algorithm (GA) and the sooty tern optimization algorithm (STOA) to optimize both the support vector machine (SVM) hyperparameters and feature selection simultaneously. This improved STOA-GA-SVM method enables the automatic discovery of abstract features in the spectrum and obtains optimal feature classification results. The feasibility of the proposed STOA-GA-SVM model will be analyzed in detecting pear bruises and the associated bruise timing.

Materials and Methods

In this section, the details for each part of the experimental process are presented, including sample preparation, experimental setup, spectral data acquisition, data preprocessing, feature selection, model building and model evaluation, and the overall flow is shown in Figure 1.

Sample preparation

All pears (variety 'Crystal') in this study were purchased from the local market in Harbin, Heilongjiang, China. A total of 210 pears without any visual defects (such as scars, cuts, shrivel, etc.) were selected, washed and numbered individually. To keep the samples at approximately the same temperature as room temperature, they were stored in an environment with a temperature of 22 °C and relative humidity of 60% for 24 h. The samples were randomly divided into three batches, each containing 70 pears. One batch was used for preparing collision bruised samples, a second batch was used for obtaining compression bruised samples, and a third batch without any treatment was used as a control and named nonbruised samples.

In preparing the samples with compression bruises, the samples were placed on the platform of a computer-controlled universal testing machine (WDW-100, Changchun Kexin Test Equipment Co Ltd., Changchun, China). The movement speed of the lower indenter of the universal testing machine was set to 10 mm/min. When the deformation of each pear sample reached 5 mm, the movement of the indenter was stopped and held for 20 s, and then the sample was taken out. Collision bruise samples were prepared by using steel balls with a weight of 225 g placed on a slope inclined at 20°. The balls were rolled down the slope from a height of 10 cm above the ground to hit the equatorial zone of the pear, with one impact made on each sample. Figure 2 shows selected images of pear samples. There is no significant difference in appearance between pear samples immediately post-bruising and the nonbruised samples. Three days post-compression bruising, it remains challenging to observe the bruise even after peeling off the skin. In comparison, three days post-collision bruising, the bruised areas become slightly brown and visible once the skin is removed.

Acquisition of hyperspectral images and spectral extraction

Figure 3 shows the structural information for the hyperspectral acquisition system. This system mainly consists of the following components: a charged couple device (CCD) camera, an imaging

spectrograph (FX10E, SPECIM Ltd., Oulu, Finland) that covers the spectral range of 400-1000 nm coupled with a zoom lens (f/1.7, SPECIM Ltd., Oulu, Finland), an illumination unit consisting of three pairs of 50 W quartz tungsten halogen lamps adjusted at an angle of approximately 45° to illuminate the camera's field of view, an electric-driving displacement platform, and a computer (Inter(R) Core(TM)i5-5200U CPU @2.2 GHz, Acer, Taiwan, China). The entire collection system (except the computer) was placed in a dark box to avoid stray light that might affect the measured reflectance from the samples.

In this study, the distance between the lens and conveyer platform was 40 cm, and the bruised area of the pear was placed horizontally facing upward on the platform. To obtain high-quality images without saturation and distortion, the stepper motor was moved at a speed of 5.5 mm/s. NIR hyperspectral image acquisition was carried out at room temperature (22 ± 2 °C). Each hyperspectral image of the pear samples was recorded as a three-dimensional image (X, Y, Z) including 1198·512 pixels in the spatial dimension (X, Y) and 224 spectral bands ranging from 400 to 1000 nm in 2.68 nm intervals between contiguous bands in the spectral dimension (Z).

After bruising, hyperspectral images of the control group and experimental group samples were collected multiple times at 1, 12, and 24 h, respectively. In this way, seven sets of pear hyperspectral images were obtained. The acquired hyperspectral images needed to be corrected due to the adverse effects caused by the dark current and noise of the instrument. The resulting spectral reflectance obtained after correction is shown by Eq. (1):

$$R = \frac{R_o - D}{W - D} \tag{1}$$

where R is the corrected image, R_o is the original image, W is the white standard reference image, and D is the dark image.

To extract the spectral data for each sample, ENVI 5.1 (The Environment for Visualizing Images, Research System Inc., USA) software was used to select the regions of interest (ROIs) of nonbruised and bruised dates. ROIs 4×4 pixels in size were manually plotted on the data, and then the average spectrum within the ROI was calculated and recorded. A total of 490 spectral samples were collected, comprising 70 non-bruised samples and 420 bruised samples. The bruised samples include 210 compression bruises and 210 collision bruises, with 70 samples for each type at three different post-injury intervals (1 hour, 12 hours, and 24 hours)

Spectral data preprocessing

The spectral data obtained by the hyperspectral imaging system contain not only sample information but also background information and noise (Yuan *et al.*, 2022). Therefore, it is necessary to preprocess the spectral data. Multiple scattering correction (MSC) is a transformation method used to compensate for additive and/or multiplicative effects in spectral data (Maleki *et al.*, 2007).

Wavelength selection methods

During hyperspectral imaging acquisition, a large amount of spectral data is generated. In the full spectra, some spectral bands are highly correlated, and some can contain redundant information (Dong *et al.*, 2015). Optimal wavelength selection enables one to eliminate redundant spectra data to improve the prediction accuracy and facilitate model interpretation (Jie *et al.*, 2013). Hence, it is necessary to find the most influential effective wavelengths (EWs) for quality assessment. In this study, principal component analysis (PCA) and uninformative variable elimination (UVE) were applied to select EWs from the full spectra (FS).

Principal component analysis

PCA is an excellent dimensionality reduction method with the advantages of enhancing the information content of hyperspectral data, isolating noisy signals and reducing the dimensionality of the data, which can be used to transform mutually redundant data information into mutually uncorrelated data and replace the original spectral information with less data information (Ji *et al.*, 2019; Li *et al.*, 2018; Li *et al.*, 2021).

Uninformative variable elimination

UVE is a method built upon a PLS model with a central goal of eliminating uninformative variables, thereby enhancing model precision and reducing bias in the eigenvalues. The optimization process consists of determining the optimal model complexity using RMSEP as the metric. Then, a random variable matrix R is combined with X to form matrix XR. From there, PLS models are calculated, regression coefficients are determined, and variables are evaluated based on the criterion *cj = bj / s(bj)*. Those variables below the maximum absolute value are discarded from X and incorporated into a new X matrix, Xnew. The final PLS models are constructed and predictions are made, with the new

model's predictive capacity gauged through cross-validated RMSEPnew. Depending on the comparison of RMSEPnew and RMSEP, the optimization process may be terminated or repeated with adjusted parameters. The UVE method, by dismissing uninformative variables and integrating a feedback loop for continual model enhancement, provides an effective tool in improving model performance and predictability (Abbott *et al.*, 1997).

Model building methods

The pear spectral data were divided into two sets: A total of 392 spectral data were selected randomly from 490 spectral data as the calibration set for building the calibration model, and the remaining 98 spectral data were used as the validation set to verify the predictive ability of the calibration model. None of the samples were used for both the calibration and validation sets.

Extreme learning machine

The Extreme Learning Machine (ELM) is a single hidden layer machine learning model that can be used to generate the first layer parameters through a stochastic policy. It has been rapidly developed in various fields because it does not require backpropagation to correct the parameters, greatly increases the speed of model operation, and has a simple structure (Xiao *et al.*, 2022).

Support vector machine

SVM is a supervised classification method based on statisticlial learning theory and structural risk minimization. It achieves data classification by finding the hyperplane that maximizes the data interval. SVM has evolved to include various kernel functions for different tasks, such as linear kernels, sigmoid kernels, radial basis function (RBF) kernels, and polynomial kernels. SVM-RBF, formed by the combination of SVM and RBF, is adept at handling nonlinear classification problems. In addition, a grid search procedure is used to optimize the SVM parameters c and g (Yang *et al.*, 2021; Zhang *et al.*, 2021).

Hybrid algorithm

The STOA-GA-SVM hybrid algorithm is a method for simultaneous feature selection and parameter optimization (Jia *et al.*, 2022). The flowchart for the STOA-GA-SVM algorithm is shown in Figure

4, and the detailed process is described as follows: (1) the data within the dataset is normalized so that all the data range between [0,1], and then each feature is binned so that the solution of the feature is limited to {0,1}; (2) the initialized populations are generated based on the maximum and minimum values of population size and parameters; (3) the parameters c, g for the generated SVM and the corresponding feature subsets are input to the SVM to complete the training and testing, and the individual fitness value f_1 and the population average fitness value f_2 are calculated; (4) if $f_1 < f_2$, the individual position according to the selection, crossover and mutation operations of GA are updated, otherwise the current individual position is updated according to the global optimization and local optimization of STOA; (5) binary features with the solution "1" are selected from the dataset, and the features c and g selected from the dataset are input into the SVM together to construct the STOA-GA-SVM classifier; (6) the fitness value is calculated using cross-validation, and the optimal solution is updated if there is a better solution than the current one; (7) it is determined whether the maximum number of iterations is reached, and, if so, the optimal value is output, otherwise one skips to step 3 and continues with the process. The processing of the spectral data was carried out using MATLAB R 2016b.

Model evaluation

Accuracy, Sensitivity (also known as True Positive Rate, TPR), and Specificity (also known as True Negative Rate, TNR) are key metrics used to evaluate classification models. Accuracy describes the overall correctness of the model, expressing the ratio of correctly classified samples out of all samples. Sensitivity, on the other hand, measures the model's performance on positive samples by showing the proportion of actual positives which are correctly identified. Specificity complements Sensitivity by measuring the proportion of actual negatives that are correctly identified, demonstrating the model's performance on negative samples. These metrics are calculated using the following formulas:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$TPR = \frac{TP}{TP + FN} \tag{3}$$

$$TNR = \frac{TN}{TN + FP} \tag{4}$$

In these equations, *TP* represents the number of True Positives, which are positive samples that have been correctly identified. *FN* stands for False Negatives, which are positive samples that are erroneously identified as negative. *TN* denotes True Negatives negative samples that are correctly classified, and *FP* represents False Positives negative samples that are incorrectly classified as positive. These metrics holistically provide a comprehensive understanding of the model's capability in predicting different classes.

Results

Hyperspectral reflectance spectra

As depicted in Figure 5, the reflection spectra curves for nonbruised pears and pears with various bruises exhibit a similar trend, with significant noise interference observed in the wavelength range of 400-450 nm. Consequently, a total of 204 wavelengths ranging from 450-1000 nm were selected for subsequent analysis to avoid this noise. Additionally, two distinct absorption valleys are observed in the S1 region (650-800 nm) and S2 region (950-1000 nm). The spectral absorption valleys around 680nm are predominantly absorbed by the carotenoids and chlorophyll pigments present on the surface of the pear (Abbott *et al.*, 1997). Similarly, the absorption valleys appearing near 960nm are indicative of the pear's moisture absorption, effectively reflecting the water content information within the fruit (ElMasry *et al.*, 2008; Huang *et al.*, 2015).

In general, the reflectance of bruised areas is lower than that of healthy areas across the visible to near-infrared spectrum (450-1000 nm), which is a trend in agreement with Lee et al (Lee *et al.*, 2014). This discrepancy in reflectance likely stems from the altered cellular structure in bruised areas, variations in water content, and shifts in biochemical components like sugars, acids, and phenolic compounds, culminating in pronounced differences in spectral absorption characteristics between bruised and unbruised pears.

Selection of the Effective Wavelengths

PCA

By using PCA to reduce the dimensionality of the spectral data, the original 204 dimensions were reduced to 75 dimensions with a cumulative contribution of over 99.99%, fully characterizing the original spectral information. The contribution rates and cumulative contribution rates for the first

five principal components are shown in Table 1, with a cumulative contribution rate of 99.12%. And the Figure 6 depicted the score plots of the first three principal components of the PCA performed, which highlights clusters with different bruising type and bruise time for pears. Among the bruised samples, the majority of clusters are relatively dense, and there is contact and overlap between them, making them more challenging to distinguish.

UVE

UVE is performed by introducing random variables in the spectral matrix and then building a PLS cross-validation model to use the variable with the lowest RMSECV value as the best variable. As shown in Figure 7, the left side of the vertical line represents the stable value of the spectral variable, the right side of the vertical line represents the stable value of the noise variable, and the two horizontal dashed lines denote the selection thresholds for the variables. The inner dashed line denotes the rejected useless information, and the outer side denotes the useful information. In this study, using UVE allowed us to focus on pertinent spectral features for the detection of fruit bruises. By eliminating uninformative variables, we could build a more effective model with improved predictive capability. The influence of this process on our results is thus one of enhancement of prediction accuracy and model efficiency. A total of 50 features are retained.

Bruise identification results

The parameters used in this study are detailed in Table 2. For PCA-ELM and UVE-ELM, the number of hidden layer nodes is set at 77 and 96, respectively. The c and g values for PCA-SVM, UVE-SVM, and STOA-GA-SVM are respectively set at 256 and 16, 256 and 147.03, and 400 and 42.71.

Five different methods were applied to identify the bruise type and time for the pears. The specific results are presented in Table 3. The accuracies of the PCA-ELM calibration and validation sets were 98.99% and 89.29%, respectively. For UVE-ELM, the calibration and validation sets showed accuracies of 98.98% and 87.97%, respectively. The PCA-SVM calibration and validation sets achieved accuracies of 96.94% and 88.78%, respectively, while those of the UVE-SVM were 99.23% and 88.78%, showing varying degrees of overfitting. The STOA-GA-SVM model exhibited accuracies of 97.19% and 92.86% for the calibration and validation sets, respectively.

The STOA-GA-SVM model exhibits the highest overall accuracy, as well as boasting the highest

sensitivity, achieving a rate of 91.85%. This indicates the model's effectiveness in predicting the type and timing of pear bruising. In food quality control and inspection, the ability to identify all negative samples—often at the expense of over-detection—is generally prioritized. Therefore, high specificity is often considered more important than high sensitivity. The STOA-GA-SVM model excels in this regard, with an impressive specificity of 98.82%, outperforming all other models. This underscores its superiority and applicability in the field.

The number of features retained by the STOA-GA-SVM model falls between that of PCA and UVE (59, compared to 75 for PCA and 50 for UVE); its prediction time of 0.0833 seconds was longer than that of PCA-ELM, UVE-ELM and UVE-SVM, but shorter than that of PCA-SVM. This timing makes it practically viable, and the STOA-GA-SVM model emerged as the most effective classification model.

Discussion

According to Table 4, the PCA-SVM model achieves an accuracy of 95.23% for identifying nonbruised pears, while the UVE-SVM model attains a recognition accuracy of 96.67% for pears with 1 h bruise. Additionally, other models such as PCA-ELM, UVE-ELM, and STOA-GA-SVM demonstrate perfect accuracy, reaching 100% for both nonbruised and 1 h bruised pears. As illustrated in Figure 5, the average spectral reflectance of nonbruised pears is significantly higher than that of other bruised types, making them easily distinguishable. For 1 h bruised pears, chemical changes within the first hour post-injury result in distinct spectral differences compared to nonbruised, 12 h bruised, and 24 h bruised pears, facilitating their easier identification (Li *et al.*, 2022a).

As inferred from Figure 8, the PCA-ELM model exhibits relatively low recognition accuracy for compression bruising for 24 h and collision bruising for 12 h, at 70% and 73.33%, respectively. Similarly, the UVE-ELM model shows low recognition accuracy for these conditions, both at 80%. Additionally, the UVE-ELM model achieves only 77.78% accuracy for collision bruising for 24 h. The PCA-SVM model has recognition accuracies of 80% for compression bruising for 24 h and 73.33% for collision bruising for 12 h. The UVE-SVM model demonstrates accuracies of 70% and 80% for compression bruising for 24 h and collision bruising for 12 h, respectively. Moreover, the UVE-SVM model achieves only 76.92% accuracy for compression bruising for 12 h. These four methods show low recognition accuracy for compression bruising for 24 h and collision bruising for 12 h. This is

mainly due to the misclassification of samples with compression bruising for 24 h as having compression bruising for 1 h, collision bruising for 1 h, and collision bruising for 24 h, and samples with collision bruising for 12 h as having collision bruising for 24 h. Additionally, samples with compression bruising for 1 h are often misclassified as having collision bruising for 1 h, and samples with compression bruising for 12 h are misclassified as having compression bruising for 1 h, compression bruising for 24 h, and collision bruising for 12 h. Samples with collision bruising for 1 h are often misclassified as having compression bruising for 1 h and 24 h. For the recognition accuracy of collision bruising for 24 h, except for the UVE-ELM model, all other models achieve 100% accuracy.

From the above description and Table 4, it can be seen that the identification accuracy for pears with bruising at 12 h and 24 h post-injury ranges between 78.57%-85.71% and 84.21%-89.47%, respectively. It is noteworthy that within this context, the STOA-GA-SVM model utilized in this study is equal to or surpasses other models in the accuracy of both the timing and the type of bruise recognition. When excluding pear samples with 1-hour bruising, the detection accuracy for bruise timing incrementally improves. This implies a correlation between the progression of time and the increase of detection accuracy. This is because, after pear bruise, the bruised area begins to brown, reaching a peak around 12 h. Thus, pear samples with 1 h bruise are more distinguishable from intact, 12 h bruised, and 24 h bruised pears. Notably, the models exhibited higher accuracy in recognizing types of bruising wherein they were able to identify collision bruising with a range of 97.62%-100% accuracy, as compared to compression bruising ranging from 91.43%-94.29%. The main reason is the differing mechanisms of pear bruise by compression and collision (Guo *et al.*, 2021). Compression bruising typically results in a more extensive bruised area on the pear's surface, leading to a more uniform disruption of the cellular structure and shallower bruise depth. In contrast, collision bruising is likely to cause localized, deeper internal injuries, with the bruised areas being more concentrated and profound. The results presented in this paper bear similarity to those of other comparable studies (Fan *et al.*, 2017; Huang *et al.*, 2015; Li *et al.*, 2021; Liu *et al.*, 2023; Liu *et al.*, 2024; Yuan *et al.*, 2021). However, the scenarios in these studies are not as complex as those addressed in this paper, with most focusing only on the degree of bruise or the time of bruise for identification. This reinforces that STOA-GA-SVM can effectively automate the selection of appropriate spectral features and parameters for the model, holding its ground not only in bruise detection but also in estimating the

timing of the bruise occurrence.

While the method employed in this study enhances the accuracy of bruise recognition to a certain degree, its ability to discern compression bruising and accurately determine the timing of bruising occurrence remains an area of improvement. There's room for improvement in the work ahead, with future research centered on identifying more effective ways to further exploit the spectral information of compression bruising and bruising timing to improve recognition accuracy. Additionally, this study only conducted bruising experiments on the equatorial region of pears and collected hyperspectral data with the hyperspectral lens vertically aligned to the bruised area. Future experiments should randomly bruise various parts of the pear and collect hyperspectral images from multiple angles to cover all possible real-world scenarios in pear bruise detection.

Conclusions

This study proposes that VIS/NIR hyperspectral imaging technology in the spectral range of 400–1000 nm shows good potential for the effective detection of different bruise types and times for pears. Spectral images for nonbruised pears and pears subjected to mechanical collision and compression bruises were taken at three different time intervals (1, 12, and 24 h). Spectral data from the ROI was then extracted from the hyperspectral images using ENVI. After analyzing and processing the spectral data obtained for the samples, PCA-ELM, UVE-ELM, PCA-SVM, UVE-SVM and STOA-GA-SVM classification models were constructed to achieve accurate classification of bruise type and bruise time for pears. The results demonstrate that the STOA-GA-SVM is an optimal model for detecting pear bruises and bruise timing. In the STOA-GA-SVM model, the accuracy for the calibration set and validation set is 97.19% and 92.86%, respectively. Additionally, the model achieves the highest sensitivity and specificity among the comparison models, with values of 91.85% and 98.82% on the validation set. This model can be used to accurately identify pears with and without bruising and can be used to better identify pears with collision bruising and the bruising time. For compression bruising and bruising time recognition accuracy, the model can meet the actual use requirements. This study shows that the use of the VIS/NIR hyperspectral technique combined with the STOA-GA-SVM algorithm is feasible for the rapid and nondestructive identification of the bruise type and time for pears.

References

Abbott, J. A., Lu RenFu, L. R., Upchurch, B. L., Stroshine, R. L. 1997. Technologies for nondestructive quality evaluation of fruits and vegetables.

Arango, J. D., Staar, B., Baig, A. M., Freitag, M. 2021. Quality control of apples by means of convolutional neural networks - Comparison of bruise detection by color images and near-infrared images. Procedia CIRP, 99: 290-294.

Dong, J., Guo, W., Wang, Z., Liu, D., Zhao, F. 2015. Nondestructive Determination of Soluble Solids Content of 'Fuji' Apples Produced in Different Areas and Bagged with Different Materials During Ripening. Food Analytical Methods, 9(5): 1087-1095.

ElMasry, G., Wang, N., Vigneault, C., Qiao, J., ElSayed, A. 2008. Early detection of apple bruises on different background colors using hyperspectral imaging. LWT - Food Science and Technology, 41(2): 337-345.

Fan, S., Li, C., Huang, W., Chen, L. 2017. Detection of blueberry internal bruising over time using NIR hyperspectral reflectance imaging with optimum wavelengths. Postharvest Biology and Technology, 134: 55-66.

Fang, Y., Yang, F., Zhou, Z., Lin, L., Li, X. 2019. Hyperspectral Wavelength Selection and Integration for Bruise Detection of Korla Pears. Journal of Spectroscopy, 2019.

Fu, X., Wang, M. 2022. Detection of Early Bruises on Pears Using Fluorescence Hyperspectral Imaging Technique. Food Analytical Methods, 15(1): 115-123.

Gao, M., Guo, W., Huang, X., Du, R., Zhu, X. 2021. Effect of pressing and impacting bruises on optical properties of kiwifruit flesh. Postharvest Biology and Technology, 172: 111385.

Guo, W., Gao, M., Cheng, J., Zhou, Y., Zhu, X. 2021. Effect of mechanical bruises on optical properties of mature peaches in the near-infrared wavelength range. Biosystems Engineering, 211: 114-124.

Guo, W., Gu, J., Liu, D., Shang, L. 2016. Peach variety identification using near-infrared diffuse reflectance spectroscopy. Computers and Electronics in Agriculture, 123: 297-303.

Hasnah Ar, N., Aris Purwanto, Y., Budiastra, I. W., Sobir. 2019. Prediction of soluble solid content, vitamin C, total acid and firmness in astringent persimmon (Diospyros kaki L.) cv. Rendeu using NIR spectroscopy. IOP Conference Series: Materials Science and Engineering, 557: 012086.

Huang, W., Li, J., Wang, Q., Chen, L. 2015. Development of a multispectral imaging system for

online detection of bruises on apples. *Journal of Food Engineering, 146*: 62-71.

Ji, Y., Sun, L., Li, Y., Ye, D. 2019. Detection of bruised potatoes using hyperspectral imaging technique based on discrete wavelet transform. *Infrared Physics & Technology, 103*: 103054.

Jia, H., Li, Y., Sun, K. 2022. Simultaneous Feature Selection Optimization Based on Hybrid Sooty Tern Optimization Algorithm and Genetic Algorithm. *Acta Automatica Sinica, 48*(6): 1601-1615.

Jiang, H., Zhang, C., He, Y., Chen, X., Liu, F., Liu, Y. 2016. Wavelength Selection for Detection of Slight Bruises on Pears Based on Hyperspectral Imaging. *Applied Sciences-Basel, 6*(12).

Jie, D., Xie, L., Fu, X., Rao, X., Ying, Y. 2013. Variable selection for partial least squares analysis of soluble solids content in watermelon using near-infrared diffuse transmission technique. *Journal Of Food Engineering, 118*(4): 387-392.

Lee, W.-H., Kim, M. S., Lee, H., Delwiche, S. R., Bae, H., Kim, D.-Y., Cho, B.-K. 2014. Hyperspectral near-infrared imaging for the detection of physical damages of pear. *Journal of Food Engineering, 130*: 1-7.

Li, B., Yin, H., Liu, Y.-d., Zhang, F., Yang, A. k., Su, C.-t., Ou-yang, A.-g. 2022a. Detection storage time of mild bruise's yellow peaches using the combined hyperspectral imaging and machine learning method. *Journal of Analytical Science and Technology, 13*(1): 24.

Li, B., Yin, H., Liu, Y.-d., Zhang, F., Yang, A. k., Su, C.-t., Ou-yang, A.-g. 2022b. Study on qualitative impact damage of yellow peaches using the combined hyperspectral and physicochemical indicators method. *Journal of Molecular Structure, 1265*: 133407.

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

Li, J., Yan, J., Ritenour, M. A., Wang, J., Cao, J., Jiang, W. 2016. Effects of 1-methylcyclopropene on the physiological response of Yali pears to bruise damage. *Scientia Horticulturae, 200*: 137-142.

Li, X., Liu, Y., Jiang, X., Wang, G. 2021. Supervised classification of slightly bruised peaches with respect to the time after bruising by using hyperspectral imaging technology. *Infrared Physics & Technology, 113*: 103557.

Li, Y., You, S., Wu, S., Wang, M., Song, J., Lan, W., . . . Pan, L. 2024. Exploring the limit of detection on early implicit bruised 'Korla' fragrant pears using hyperspectral imaging features and spectral variables. *Postharvest Biology and Technology, 208*.

Liu, D., Li, Q., Li, W., Yang, B., Guo, W. 2017. Discriminating forchlorfenuron-treated kiwifruits using a portable spectrometer and Vis/NIR diffuse transmittance spectroscopy technology. Analytical Methods, 9(28): 4207-4214.

Liu, D., Lv, F., Wang, C., Wang, G., Zhang, H., Guo, J. 2023. Classification of early mechanical damage over time in pears based on hyperspectral imaging and transfer learning. Journal of Food Science, 88(7): 3022-3035.

Liu, D. Y., Zhang, H. T., Lv, F., Tao, Y. R., Zhu, L. K. 2024. Combining transfer learning and hyperspectral imaging to identify bruises of pears across different bruise types. Journal of Food Science, 89(5): 2597-2610.

Maleki, M. R., Mouazen, A. M., Ramon, H., De Baerdemaeker, J. 2007. Multiplicative Scatter Correction during On-line Measurement with Near Infrared Spectroscopy. Biosystems Engineering, 96(3): 427-433.

Nicolaï, B. M., Defraeye, T., De Ketelaere, B., Herremans, E., Hertog, M. L. A. T. M., Saeys, W., . . . Verboven, P. 2014. Nondestructive measurement of fruit and vegetable quality. Annual review of food science and technology, 5: 285-312.

Opara, U. L., Pathare, P. B. 2014. Bruise damage measurement and analysis of fresh horticultural produce—A review. Postharvest Biology and Technology, 91: 9-24.

Rahman, A., Kandpal, L., Lohumi, S., Kim, M., Lee, H., Mo, C., Cho, B.-K. 2017. Nondestructive Estimation of Moisture Content, pH and Soluble Solid Contents in Intact Tomatoes Using Hyperspectral Imaging. Applied Sciences, 7(1): 109.

Stropek, Z., Gołacki, K. 2015. A new method for measuring impact related bruises in fruits. Postharvest Biology and Technology, 110: 131-139.

Su, C.-T., Li, B., Yin, H., Zou, J.-P., Zhang, F., Liu, Y.-D. 2022. Identification of Damage in Pear Using Hyperspectral Imaging Technology. Journal of Spectroscopy, 2022.

Thien Pham, Q., Liou, N.-S. 2022. The development of on-line surface defect detection system for jujubes based on hyperspectral images. Computers and Electronics in Agriculture, 194: 106743.

Xiao, D., Huang, J., Li, J., Fu, Y., Li, Z. 2022. Inversion study of cadmium content in soil based on reflection spectroscopy and MSC-ELM model. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 283: 121696.

Xing, J., De Baerdemaeker, J. 2005. Bruise detection on 'Jonagold' apples using hyperspectral

imaging. *Postharvest Biology and Technology, 37*(2): 152-162.

Yang, J., Sun, L., Xing, W., Feng, G., Bai, H., Wang, J. 2021. Hyperspectral prediction of sugarbeet seed germination based on gauss kernel SVM. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 253:* 119585.

Yuan, R., Guo, M., Li, C., Chen, S., Liu, G., He, J., . . . Fan, N. 2022. Detection of early bruises in jujubes based on reflectance, absorbance and Kubelka-Munk spectral data. *Postharvest Biology and Technology, 185:* 111810.

Yuan, R., Liu, G., He, J., Wan, G., Fan, N., Li, Y., Sun, Y. 2021. Classification of Lingwu long jujube internal bruise over time based on visible near-infrared hyperspectral imaging combined with partial least squares-discriminant analysis. *Computers and Electronics in Agriculture, 182:* 106043.

Zhang, B., Huang, W., Li, J., Zhao, C., Fan, S., Wu, J., Liu, C. 2014. Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. *Food Research International, 62:* 326-343.

Zhang, L., Sun, H., Li, H., Rao, Z., Ji, H. 2021. Identification of rice-weevil (Sitophilus oryzae L.) damaged wheat kernels using multi-angle NIR hyperspectral data. *Journal of Cereal Science, 101:* 103313.

Table 1. Contribution rate and cumulative contribution rate of the first 5 principal components.

PC	PC1	PC2	PC3	PC4	PC5
contribution rates	80.62	12.68	4.36	1.10	0.36
cumulative contribution rates	80.62	93.30	97.65	98.76	99.12

Table 2. Parameters for each model.

Methods	PCA-ELM	UVE-ELM	PCA-SVM	UVE-SVM	STOA-GA-SVM
	hidden layer nodes number		c, g		
Value	77	96	256, 16	256, 147.03	400, 42.71

Table 3. Comprehensive identification results for different models: accuracy, sensitivity, and specificity on both calibration and validation sets.

Model	Accuracy		Sensitivity(TPR)		Specificity(TNR)		Time (s)
	Cal.*	Val.*	Cal.	Val.	Cal.	Val.	Val.
PCA-ELM	98.99%	89.29%	98.52%	90.01%	99.74%	98.52%	0.0009
UVE-ELM	98.98%	87.97%	99.25%	87.72%	99.87%	98.33%	0.0009
PCA- SVM	96.94%	88.78%	96.96%	88.47%	99.49%	98.16%	0.1409
UVE-SVM	99.23%	88.78%	99.24%	87.58%	99.87%	98.16%	0.0482
STOA-GA-SVM	97.19%	**92.86%**	95.98%	**91.85%**	99.32%	**98.82%**	0.0833

*Cal. means calibration set, Val. means validation set.

Table 4. Validation accuracy of recognizing the bruise time and bruise type.

Methods	Nonbruised	Bruising time			Bruising type	
		Bruise 1 h	Bruise 12 h	Bruise 24 h	Compression	Collision
PCA-ELM	**100%**	**100%**	82.14%	**89.47%**	**94.29%**	97.62%
UVE-ELM	**100%**	**100%**	**85.71%**	84.21%	91.43%	97.62%
PCA-SVM	95.23%	**100%**	78.57%	**89.47%**	**94.29%**	97.62%
UVE-SVM	**100%**	96.67%	78.57%	**89.47%**	91.43%	97.62%
STOA-GA-SVM	**100%**	**100%**	**85.71%**	89.47%	94.29%	**100%**
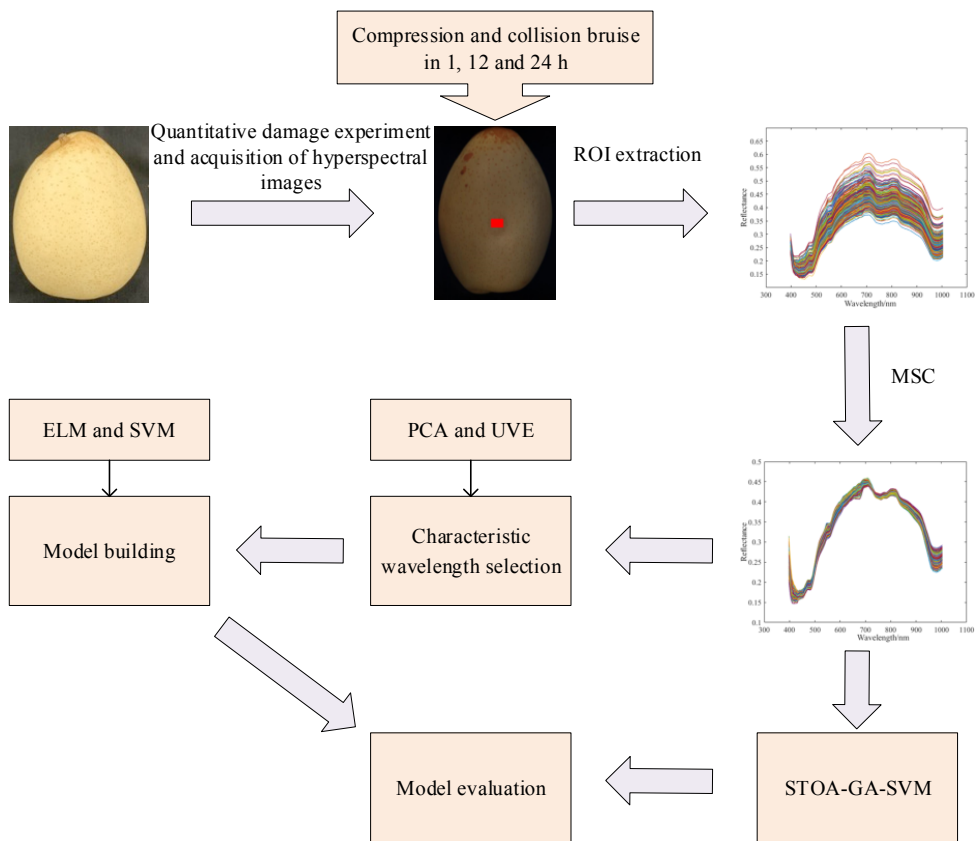
Figure 1. Processing methods for spectral data.

(a) Nonbruised pear   (b) Pears after collision bruise   (c) Pears peeled after 3 days of collision bruise

(d) Nonbruised pear   (e) Pears after compression bruise   (f) Pears peeled after 3 days of compression bruise

Figure 2. Pears samples: nonbruised (a, d), immediately post-bruise (b, e), and peeled after 3 days of bruising (c, f).
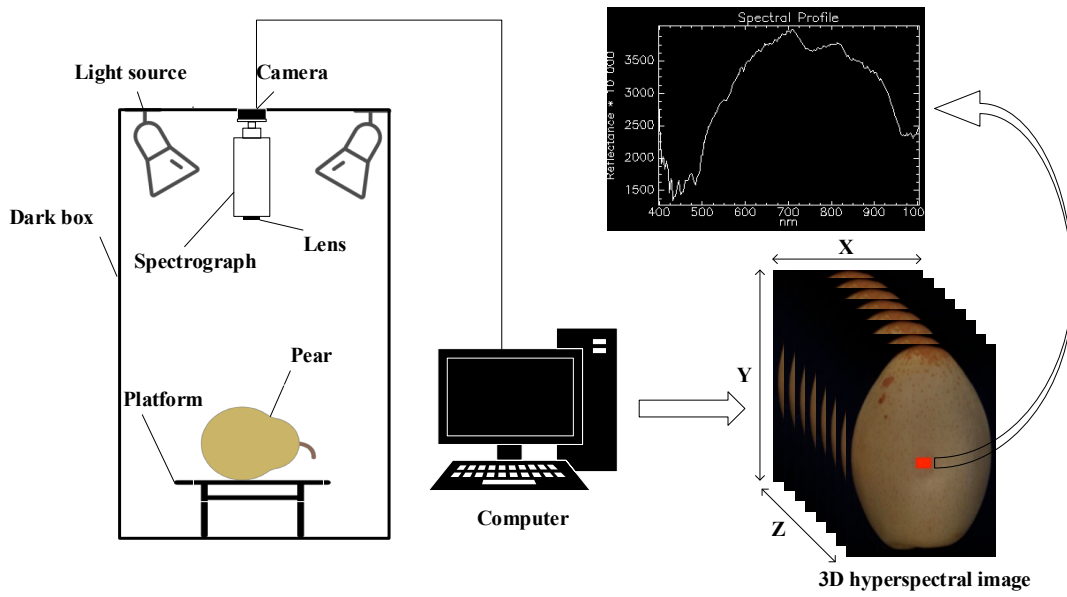
Figure 3. Data acquisition system and hyperspectral 3D spectral image.
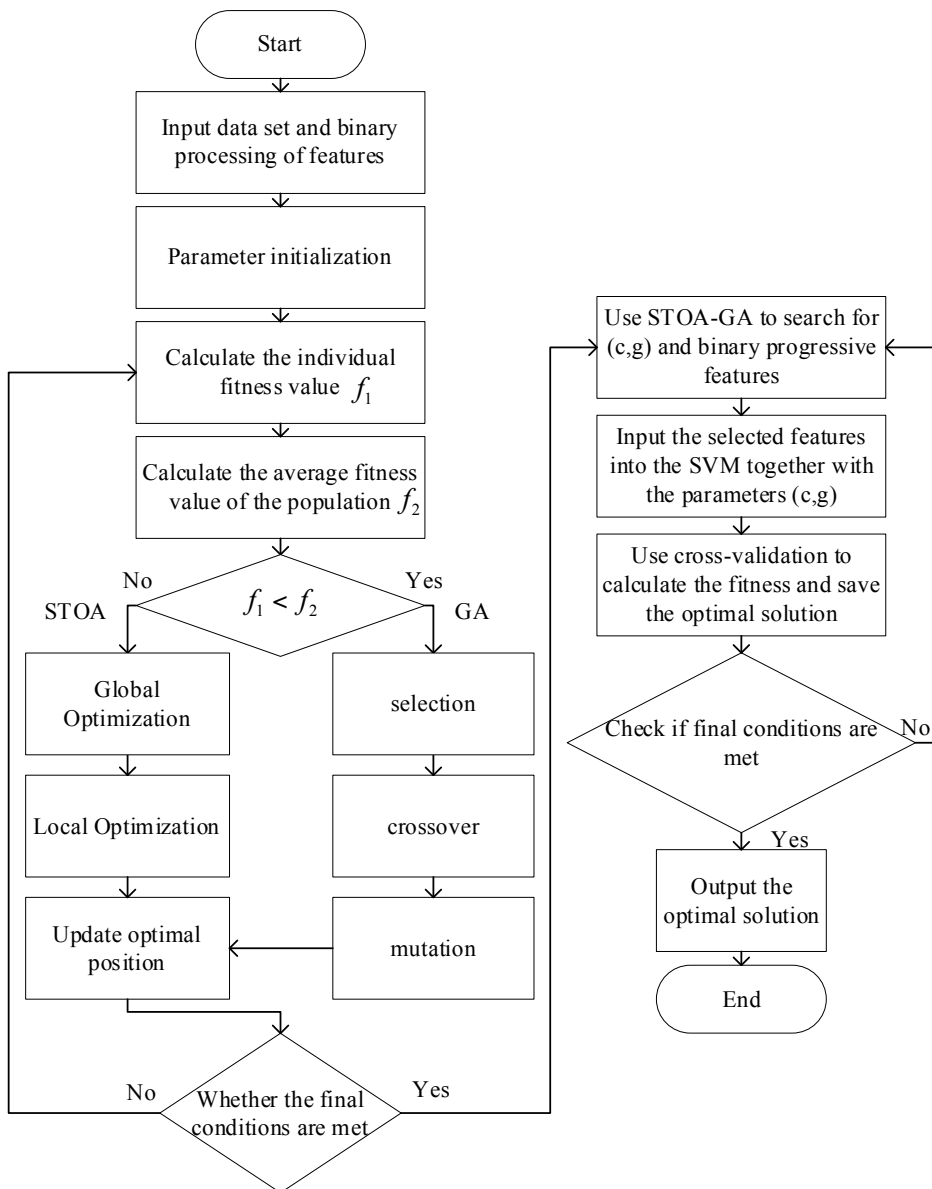
Figure 4. STOA-GA-SVM hybrid algorithm flowchart.
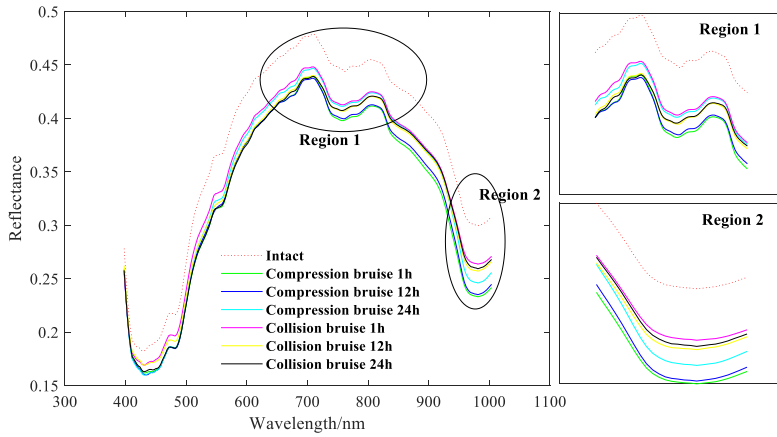
Figure 5. Comparison of the average spectra measured for pear samples in each sample group in the experiment.
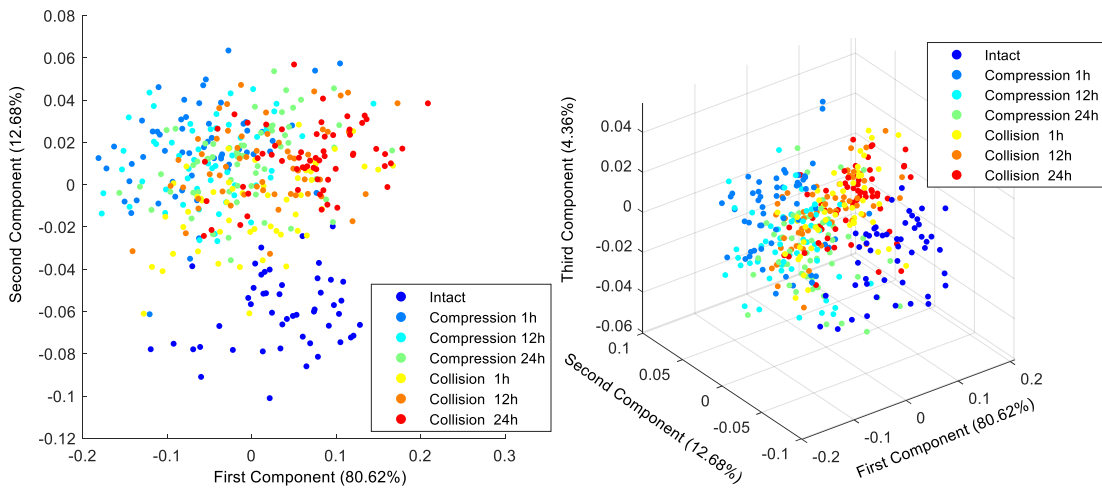
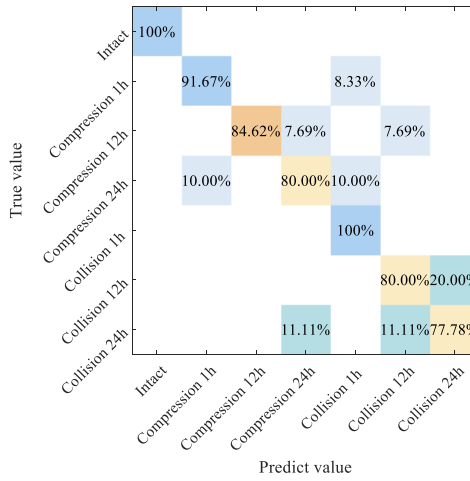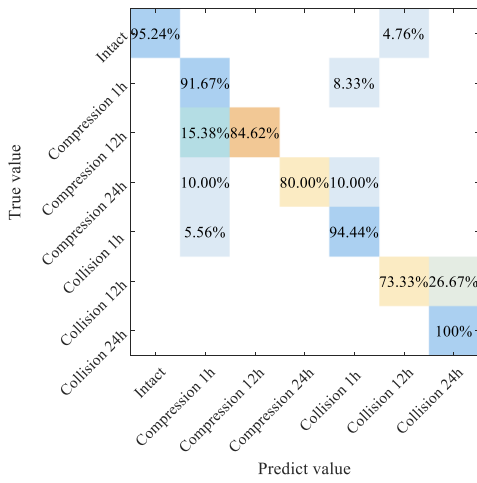**Figure 6. Score plots of the first three principal components of the PCA.**
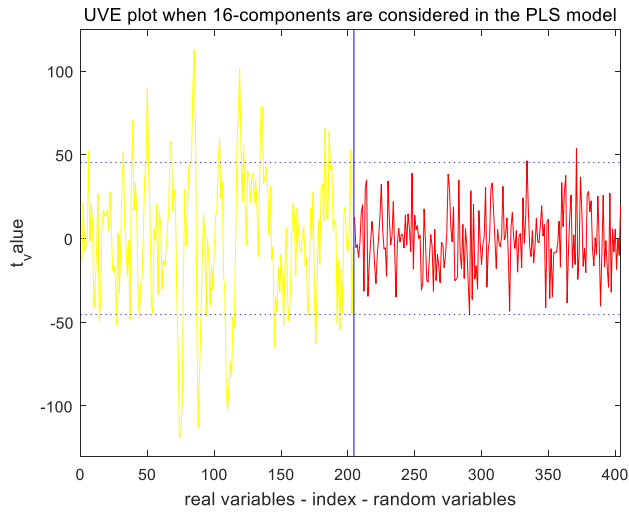
Figure 7. Stability distribution of variables used for prediction by UVE.
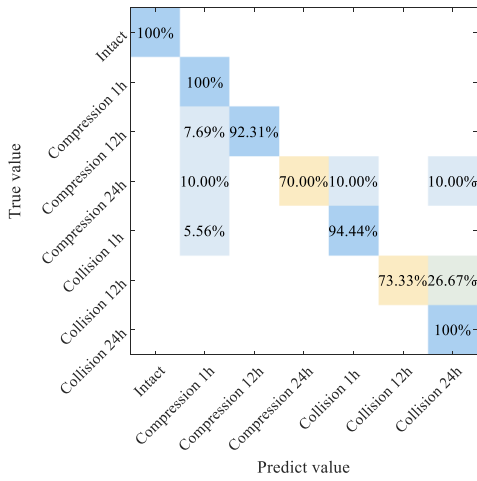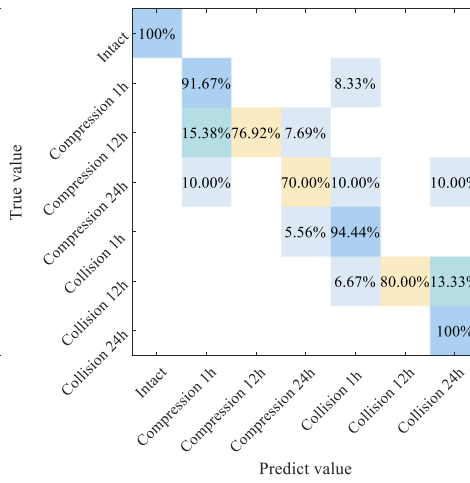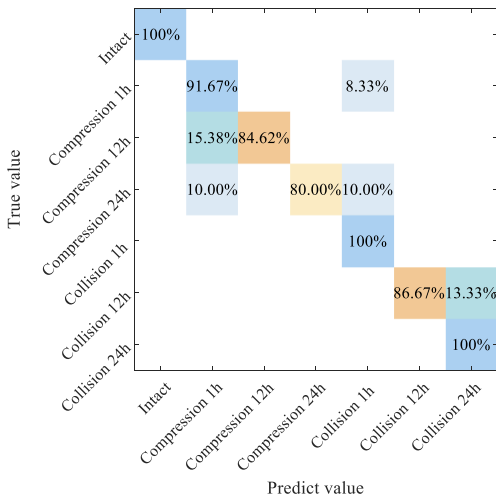
(a) Results of the PCA-ELM

(b) Results of the UVE-ELM

(c) Results of the PCA-SVM

(d) Results of the UVE-SVM

(e) Results of the STOA-GA-SVM

Figure 8. Visualized confusion matrix for five different modeling methods.