Infrared Physics & Technology 120 (2022) 104003

Contents lists available at ScienceDirect

Infrared Physics and Technology

journal homepage: www.elsevier.com/locate/infrared

Developing deep learning based regression approaches for prediction of firmness and pH in Kyoho grape using Vis/NIR hyperspectral imaging

Min Xu [a,b], Jun Sun [a,*], Kunshan Yao [a], Qiang Cai [a], Jifeng Shen [a], Yan Tian [a], Xin Zhou [a]

[a] School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, Jiangsu, China
[b] School of Electronic Engineering, Changzhou College of Information Technology, Changzhou 213164, Jiangsu, China

ARTICLE INFO

Keywords:
Deep learning
Fruit quality
Hyperspectral imaging
Kyoho grape
Nondestructive detection
Stacked auto-encoders

ABSTRACT

Firmness and pH, the most important quality attributes of grapes, are directly associated with their quality and price. This study aimed to predict firmness and pH of Kyoho grape using hyperspectral imaging (HSI) via a deep learning approach. Stacked auto-encoders (SAE) were applied to extract deep spectral features based on the pixel-level spectra of each sample over the wavelength range of 400.68–1001.61 nm. Subsequently, these features were used as input data to construct deep learning models for assessing firmness and pH. Additionally, the successive projections algorithm and competitive adaptive reweighed sampling (as wavelength selection algorithms) as well as partial least squares (PLS) and least squares support vector machine (LSSVM) (as modeling approaches) were investigated as conventional spectra analysis approaches for comparison. The results showed that the SAE-LSSVM model achieved the optimal performance, with $R_p^2 = 0.9232$, $RMSEP = 0.4422 \quad N$, and $RPD = 3.26$ for firmness, and the SAE-PLS model yielded satisfactory accuracy, with $R_p^2 = 0.9005$, $RMSEP = 0.0781$, and $RPD = 2.82$ for pH. The overall results revealed that SAE could be used as an alternative to deal with high-dimensional hyperspectral image data. Combined with HSI, it could non-destructively and rapidly detect firmness and pH in grapes; this significantly facilitates post-harvest management and may provide a valuable reference for evaluating other internal quality attributes of fruit.

1. Introduction

Kyoho grape (Vitis labruscana cv. Kyoho) is one of the most popular fruits among consumers and contains various nutrients (melatonin, vitamins, anthocyanins, hydroxytyrosol, etc.) that are beneficial to human health [1,2]. Table grapes, which can be processed into high value-added products (raisins, grape juice, wine, grape seed oil, etc.), provide more options for consumers and generate more profits for merchants. Firmness and pH, the important quality attributes of grapes, are directly associated with their taste and affect the purchase decision and satisfaction of consumers [3,4]. Accordingly, the firmness and pH of grapes should be considered during their shelf life. Currently, fruit firmness is measured using the Magness-Taylor (MT) penetrometer test based on the industry standard [5]. The determination of fruit pH requires the extraction of juice from the pulp and detection using a pH meter [6]. Although the conventional methods are objective and reliable, they are destructive, time-consuming, laborious, and cannot satisfy the requirements for rapid and nondestructive determination. Consequently, a nondestructive method is required for evaluating the firmness and pH of grapes to facilitate post-harvest management.

Recently, hyperspectral imaging (HSI) has been broadly recognized as a promising method to nondestructively assess the firmness and/or pH of fruits such as pears [7], bananas [8], kiwifruit [9], persimmon [10], lychee [11], and apple [12]. In addition, HSI was used to detect the firmness and pH of grape [13]. Its detection principle is based on the object interaction with light, where the spectral properties of the measured object change with wavelength. Because the variation in each spectrum is determined by the chemical composition and microstructure of an object, the firmness and/or pH can be reflected by the change in the relevant wavelength using chemometric approaches [14]. Hyperspectral images acquired using modern high-resolution spectroscopy devices contain abundant spectral and spatial data, which increase the processing load and modeling difficulty [15]. To reduce computational cost, previous algorithms such as the successive projections algorithm (SPA) and competitive adaptive reweighed sampling (CARS) have been proposed to reduce the dimensionality of large datasets [16].

* Corresponding author.
E-mail address: sun2000jun@sina.com (J. Sun).

https://doi.org/10.1016/j.infrared.2021.104003
Received 21 September 2021; Received in revised form 24 November 2021; Accepted 17 December 2021
Available online 21 December 2021

Nevertheless, most existing relevant studies use the average spectrum of all pixels or a limited number of representative spectra from each sample region of interest (ROI) to manage hyperspectral data and establish models for predicting fruit quality attributes. However, the application of mean or typical spectra renders it difficult to reveal the global information of hyperspectral data [17]. Additionally, the big data analysis of hyperspectral image data using numerous pixel-level spectra to predict the firmness and pH of grapes is scarce.

Deep learning is a new research area in machine learning that is an alternative for managing high-dimensional hyperspectral image data and has been applied to regression issues in previous studies [18–22]. These studies suggested that deep learning for prediction can perform equivalently to or better than conventional machine learning methods. The aim of this study was to predict the firmness and pH of grapes using HSI and a deep learning method. To the best of our knowledge, this is the first study that uses deep learning to evaluate the internal attributes of grapes. Specifically, the deep learning approach included two main stages, in which stacked auto-encoders (SAE) were trained using pixel-level spectra to extract deep features from numerous hyperspectral image data. Subsequently, these deep spectral features were input to partial least squares (PLS) or least squares support vector machine (LSSVM) models to predict the firmness and pH of grape. Conventional spectral analysis methods, including PLS and LSSVM (as modeling approaches) as well as CARS and SPA (as wavelength selection algorithms), have been investigated and compared.

2. Materials and methods

2.1. Sample preparation

Kyoho grapes at commercial maturity were harvested manually from an orchard located in Jurong, Zhenjiang, Jiangsu Province, China. To obtain samples that reflect a wide range of growth conditions, the grapes were harvested from different trees and canopy levels, and then wrapped in grape-specific bags (YT-16A, 'Aiguo'). Then, they were placed in a fresh-keeping box and transported to a laboratory at Jiangsu University in Zhenjiang within 4 h. Grapes (240 in total) with undamaged surface were selected from the upper, middle, and bottom of each cluster after they arrived at the laboratory, and then placed in the chamber, where the temperature and humidity of the laboratory were maintained at 20 ℃±2 ℃ and 60 %±5 %, respectively. To account for the effect of environmental elements, experimental procedures including hyperspectral image acquisition as well as reference firmness and pH measurements were performed on the same day. Afterwards, the samples were segregated into calibration and prediction sets at a ratio 2:1 via sample set partitioning based on joint x-y distances algorithm [23]. To evaluate the performances of the different models objectively, the samples in the calibration and prediction sets were maintained consistent throughout the prediction analysis.

2.2. HSI system, images acquisition, and correction

Hyperspectral images of grapes were acquired using an HSI device (Model: VNIR-HIS-sCMOS, Isuzu Optics, Co, Ltd., Taiwan) equipped with a spectrograph (ImSpectorV10, Spectral Imaging Ltd., Oulu, Finland) with a spectral resolution of 2.8 nm, a charge-coupled device camera (Zyla 4.2 Plus, Andor Technology, Inc., UK) with a resolution of 2048 pixels in the spatial dimension and 478 bands in the spectral dimension, a light source unit (3900-ER, Illumination Technology, Inc., USA), lens (F/2.4. f = 23 mm, OLE23, Specim Ltd., Finland), and a displacement platform (MSI300, Isuzu Optics, Taiwan, China) (Fig. 1a). In this experiment, the light source was first preheated for approximately 30 min, then the samples were placed on the sample holder and the distance between the samples and the lens was adjusted to 0.45 m. Next, the exposure time of the samples, and the references for the whiteboard and blackboard were set as 46.2, 11.2 and 11.2 ms, respectively. Finally, the computer-controlled displacement platform was moved at a speed of 2.46 mm/s until the samples were completely scanned to obtain the hyperspectral images of grapes in the band range of 400.68–1001.61 nm. Considering the effect of uneven light intensity distribution and the interference of the imaging system, the obtained hyperspectral images had to be corrected for further processing. The correction steps are available in a paper published previously by our research group [24].

2.3. Spectra extraction

The corrected hyperspectral images of the samples were further analyzed using the ENVI 5.0 (ITT Visual Information Solutions, Boulder, CO, USA) software. The entire grape region was selected as the ROI, and the spectral extraction process is shown in Fig. 1b. Specifically, the grayscale was obtained using an image of 823.18 nm owing of the prominent difference in spectral reflectance between the target and background at this band. Next, the threshold segmentation method was applied to process the grayscale. The threshold should be higher than the maximum reflectance of the background and lower than the minimum reflectance of the target. Hence, the threshold was set as 0.04, which resulted in a binary mask that included the target (white, all pixels are 1) and background (black, all pixels are 0). Then, the background was removed by multiplying the binary mask with the original image to obtain the ROI of the samples. Finally, the average spectrum of all pixels of each sample in the ROI was calculated and regarded as the raw

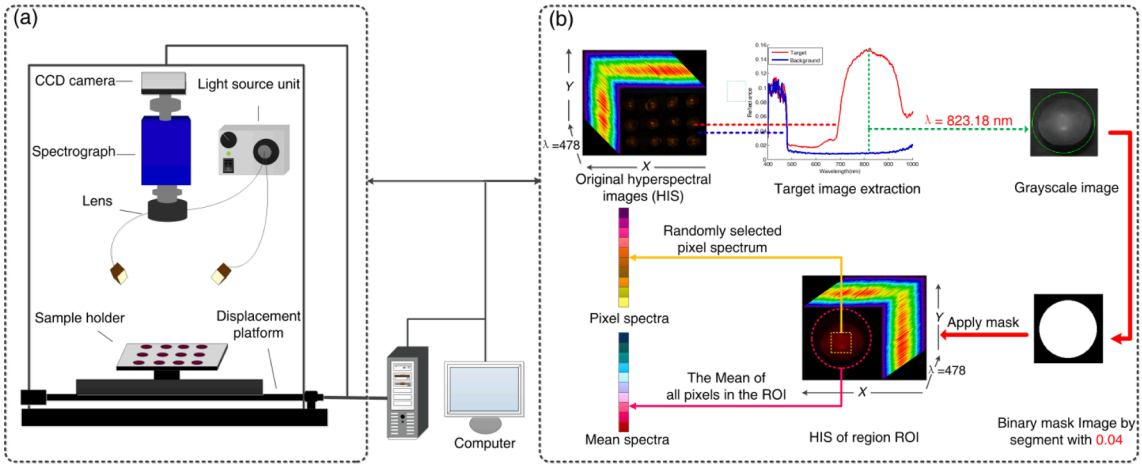

Fig. 1. Hyperspectral images acquisition system (a), and spectra extraction (b).

spectrum for conventional wavelength selection and modeling analysis. In addition, a total of 600 average spectra of 20 pixels × 20 pixels were randomly selected from the ROI of each sample to create a large dataset for extracting deep spectral features and establishing deep learning regression models.

2.4. Reference measurement of firmness and pH

Upon hyperspectral image acquisition, the firmness was first measured destructively using a 3.5 mm diameter MT probe attached to a fruit sclerometer (Model: GY-4, Le qing Aidebao Instrument, Co. Ltd., LeQing, Zhejiang, China). A download pressure with a velocity of 1 mm s^{-1} was forced until the plunger penetrated 10 mm into the samples. The maximum force was recorded and used to measure the firmness, expressed in newton (N), with an accuracy of 0.1 N. Then, the samples were peeled and juiced using a blender (Model: JYL-B065, Joyoung), and their pH was determined using a pH meter with 0.01 accuracy (Model: FE28, Mettler Toledo Instrument, Co. Ltd., Shanghai, China). To reduce the testing errors and obtain more reliable pH reference values, each grape sample was measured thrice, and the corresponding average of three measurements was set as the final pH reference measured value in this experiment.

2.5. Spectra preprocessing and wavelength selection

It is well known that spectral preprocessing can facilitate the development of reliable models. In this study, the response of the detector at the beginning of the spectra was insignificant, and the spectra in this region were noisy. Hence, only the spectral range of 493.79–1001.61 nm (400 bands) was used for further analysis [25]. Multiplicative scatter correction (MSC) was adopted to relieve the effect of scattering on the spectra due to sample heterogeneity, and the denoising algorithm ensemble empirical mode decomposition coupled with discrete wavelet transform (EEMD-DWT) proposed by Xu et al. was used to minimize the noise intensity [26]. Furthermore, a Monte Carlo (MC) outlier detection algorithm was introduced to identify potential outliers in the samples to further improve the performances of the models [27].

In addition, wavelength selection is crucial in ensuring faster and more cost-effective variables for decreasing computational cost; furthermore, it contributes positively to eliminating redundant variables and promoting the interpretability of the models. Hence, conventional wavelength algorithms, i.e., the SPA and CARS were used to select the feature wavelengths of the spectra as input to the PLS (or LSSVM) for predicting the firmness and pH of grapes and comparing them with deep learning models.

2.6. Calibration models for firmness and pH prediction

2.6.1. PLS

As a multivariate statistical method, PLS can adequately consider the interpretation of spectra on response variables (firmness and pH) [28]. It is particularly suitable for cases in which the number of variables is higher than the sample size. In PLS modeling, the spectral data are projected onto latent variables (LVs) used to explain the variance of the response variables and are used as a new eigenvector for dimensionality reduction and spectral data compression. The optimal number of LVs was obtained using the lowest root mean square error of cross-validation (*RMSECV*) using the full cross-validation (CV) method. PLS modeling was implemented in the Unscrambler X 10.4 software (CAMO Technologies Inc., Woodbridge, New Jersey).

2.6.2. LSSVM

The LSSVM was developed based on a support vector machine (SVM) [29]. It can transform the quadratic convex programming problem into a linear equation solving problem, thereby reducing the computational cost and increasing the convergence speed. The radial basis function (RBF) was used as the kernel function of the LSSVM, and the regularization parameter γ and kernel parameter σ^2 affect the precision of the RBF kernel-based LSSVM. To avoid setting the parameter blindly, the grid search (GS) method was employed to optimize the parameters (γ, σ^2). When implementing the GS algorithm, a crude search with a large step size was first performed, followed by a specified search with a small step size. For each combination of γ and σ^2, the *RMSECV* was calculated, and the optimal parameters (γ, σ^2) for obtaining the lowest *RMSECV* were determined. The LSSVM was operated in MATLAB 2014a (The Math Works, Natick, USA) using an LSSVM toolbox (available at http://www.esat.kuleuven.be/sista/lssvmlab/).

2.6.3. SAE-PLS and SAE-LSSVM

SAE-PLS (or SAE-LSSVM), a deep-learning-based regression model, was developed to predict the firmness and pH of grapes. Concretely, the SAE was utilized as an unsupervised manner to extract deep features from pixel-level spectra, whereas PLS or LSSVM was used in a supervised manner to predict firmness and pH based on these deep spectral features.

An SAE exhibits a hierarchical deep network architecture that comprises multiple autoencoders (AEs). Each AE comprises a three-layer network that includes a d-unit input layer, an h-unit hidden layer, and a d-unit output layer; the purpose of AE is to extract representative features from the original input as well as possible via self-learning [30]. The AE can be categorized into two components: an encoder and a decoder. When performing encoding, $x \in R^d$ is mapped to the hidden layer $y \in R^h$ by an activation function $f(\cdot)$, whereas the hidden layer $y \in R^h$ is re-mapped to the output layer $z \in R^d$ via a function $f(\cdot)$ in the decoding stage. These two stages can be defined as follows.

$$y = f(w_y x + b_y) \tag{1}$$

$$z = f(w_z y + b_z) \tag{2}$$

Where w_y is an $(h \times d)$ input-to-hidden weight matrix; w_z is a $(d \times h)$ hidden-to-output weight matrix; $b_y \in R^h$ and $b_z \in R^d$ are the basis vectors of the hidden layer and output layer, respectively. A nonlinear function $sigmoid(x) = 1/1 + e^{-x}$ is selected as the activation function $f(\cdot)$. The purpose of AE training is to minimize the error between the reconstructed output data and input data, as follows.

$$\begin{aligned} &\underset{W,b_y,b_z}{argmin}[c(x,z)] \\ &s.t. \quad w_y = w_z^{'} = W. \end{aligned} \tag{3}$$

Where z is the reconstructed output data with the same size as the original input data x. The AE can train the model in an unsupervised manner, allowing the output data z to be identical to the input data x via the deep features y in the hidden layer. It is noteworthy that the AE can be stacked to construct the SAE, which can yield abstract high-level deep features compared with a single layer AE. Therefore, in this study, the SAE was used to obtain deep spectral features from the hyperspectral data of grapes.

In the prediction phase, the decoding component of the SAE was removed, whereas the encoding component was retained. Subsequently, PLS (or LSSVM) was incorporated into the last encoding layer of the SAE to construct the SAE-PLS (or SAE-LSSVM) model. The developed models included two main phases: training and fine-tuning. The network weights were obtained by training the pixel spectra in the SAE (Fig. 2a). Afterwards, the learned weights were regarded as initial weights for extracting the deep spectral features based on the input average spectra using the SAE network. Finally, these deep features were used to establish deep learning models to predict the firmness and pH of the grapes. The coefficient of determination (R^2) between the predicted value and the reference measured firmness (or pH) was used to fine-tune the entire SAE-PLS (or SAE-LSSVM) network in a supervised manner

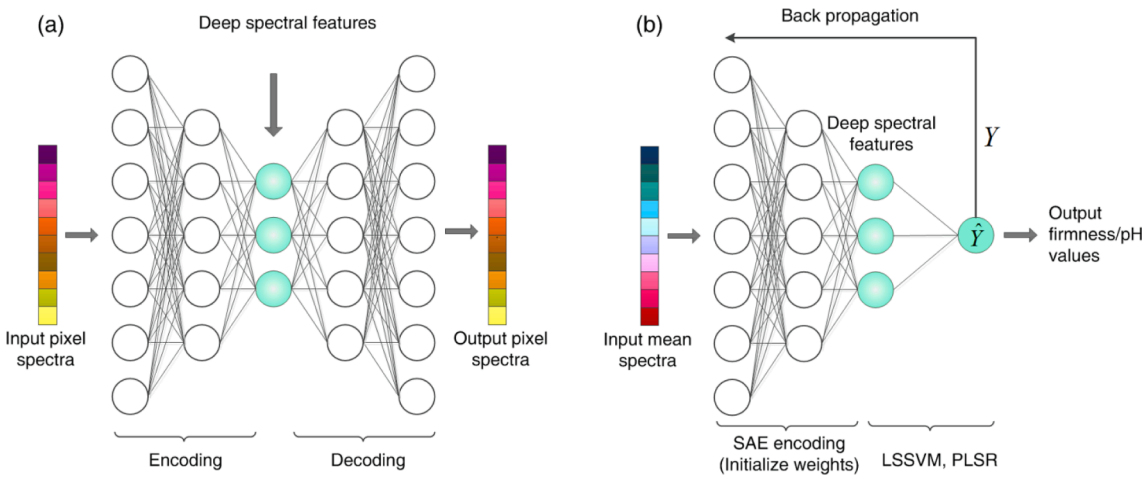

Fig. 2. The SAE-PLS (or SAE-LSSVM) method. unsupervised training of SAE using random pixel spectra (a), supervised fine-tuning of SAE-PLS (or SAE-LSSVM) using average spectra and firmness (or pH).

(Fig. 2b). Specifically, based on R^2, the approximate range of the last hidden layer unit (deep feature) in the encoding component of the SAE was obtained, and the optimal number of units in the range above was determined based on the highest R^2 of the calibration set for fine-tuning the deep learning models. The R^2 is expressed as follows.

$$R^2 = 1 - \left[\sum_{i=1}^{n} (Y_i - \widehat{Y}_i)^2 / \sum_{i=1}^{n} (Y_i - \overline{Y}_i)^2 \right] \tag{4}$$

Where n is the number of training set samples; Y_i and $\widehat{Y}_i$ are the measured and predicted values of firmness or pH in the sample i, respectively; $\overline{Y}_i$ represents the mean values of measured firmness or pH in the calibration set. The SAE was executed in a Windows 7 system, which comprised an Intel i7 8700 K processer with eight 3.7 GHz cores; it was developed based on a Keras framework (available at http://github.com/fchollet/keras), which is a Python library exploiting deep learning models. In this experiment, two NIVDIA GT 1080Ti GPUs were used, which improved the learning rate of the SAE.

2.7. Model evaluation

The performance of the model was evaluated based on coefficient of determination of calibration (R_c^2), root mean square error of calibration (*RMSEC*), coefficient of determination of prediction (R_p^2), root mean square error of prediction (*RMSEP*), and residual prediction deviation (*RPD*), which are expressed as follows.

$$R_c^2, R_p^2 = 1 - \left[\sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 / \sum_{i=1}^{n} (y_i - \overline{y}_i)^2 \right] \tag{5}$$

$$RMSEC, \quad RMSEP = \sqrt{\sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 / n} \tag{6}$$

$$RPD = \frac{SD}{RMSEP} \tag{7}$$

Where, y_i and $\widehat{y}_i$ are the measured and predicted values of firmness (or pH) in the *ith* sample calibration and prediction set, respectively; $\overline{y}_i$ represents the mean values of the reference value in the calibration and prediction sets; *SD* is the standard deviation of the measured value in the prediction set; n is the number of samples in the corresponding sample set. Generally, higher R_c^2, R_p^2, and *RPD* values correspond to lower *RMSEC* and *RMSEP* values, signifying the better performance of the model [31].

3. Results and discussion

3.1. Spectra analysis and preprocessing

Fig. 3a shows the mean spectra curve of the samples with 400 bands. The spectra exhibit similar tendencies, with two distinct peaks (about 640 and 820 nm) and valleys (about 670 and 980 nm). In the visible region, the distinct absorption peaks at approximately 640 and 670 nm correspond to the reflection bands of anthocyanin pigments and chlorophyll, respectively. These two characteristic peaks represent the red and green-yellow colors of the ripe grape, respectively [32]. In the near-infrared region, the relative reflectance spectra show different peak values at around 820 nm, which is primarily attributed to the third overtone stretch of C-H associated with sugar in grapes [33], whereas the reverse absorption band near 980 nm corresponds to the second overtone of O-H of water in the fruit [34,35]. The variation in the spectral reflectance among the samples allows the physiochemical properties of grapes with different firmness and pH values to be determined.

The acquired mean spectra indicate clear scattering and noise (Fig. 3a). Therefore, MSC and EEMD-DWT were applied to preprocess the raw spectra. After preprocessing, the spectral features are more pronounced and the curves smoother (Fig. 3b). Moreover, outlier detection prior to modeling is important because the outliers in the samples will contaminate the dataset and deteriorate the performance of the models. Hence, an MC outlier detection method that can simultaneously identify outliers in both the spectra and reference values (firmness or pH) was used to detect potential outliers in the samples. The algorithm provides the prediction error of each sample, and the distribution of the standard deviation and mean values of the prediction error can be used to identify the outliers. Samples with large mean values and standard deviations can be regarded as potential outliers. As shown in Fig. 3c-d, one sample numbered 161 for firmness and two samples numbered 157 and 161 for pH are detected as latent outliers. Then, PLS models with full CV for firmness and pH evaluation were established based on datasets before and after eliminating each potential outlier (Table 1). It is observed that the PLS model with samples after outlier removal shows better accuracy for firmness and pH assessment compared with the full samples. Thus, the outliers selected via the MC outlier detection approach were removed from the samples for subsequent analysis.

3.2. Reference measurement results of firmness and pH

After removing the outliers, the samples were divided into

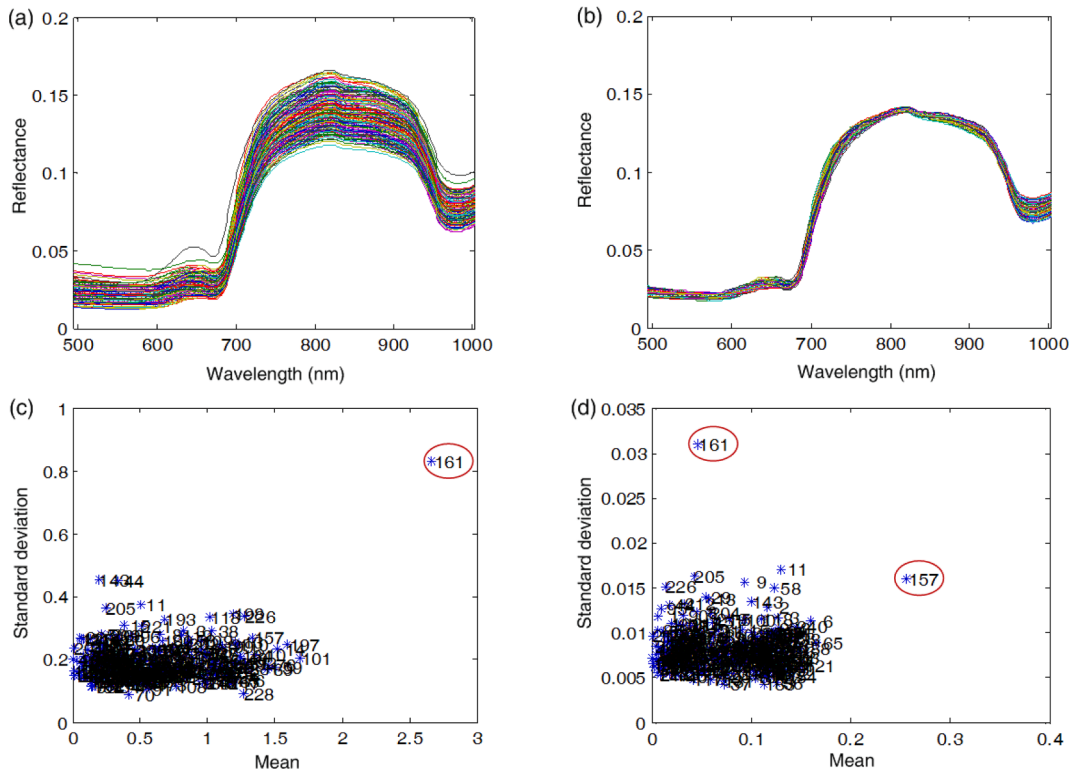

Fig. 3. the original average spectra of samples (a), the preprocessed spectra by using MSC and EEMD-DWT (b), identification results of outliers of firmness (c) and pH (d) using MC outlier detection method.

Table 1
Prediction results of Firmness and pH in grape using PLS model before and after removing the outliers.

Parameters	No. of samples	Outlier number	LVs[a]	R^2_{cal}	*RMSEC*	R^2_{cv}	*RMSECV*
Firmness (N)	240	None	10	0.8577	0.6674	0.8221	0.7989
	239	161	12	0.9004	0.5529	0.8598	0.6621
pH	240	None	10	0.9327	0.0748	0.8328	0.0911
	239	157	11	0.937	0.0724	0.8454	0.0869
	239	161	10	0.9277	0.0752	0.8557	0.0831

calibration and prediction sets, and the statistical references of firmness and pH of all samples are presented in Table 2. The ranges of firmness and pH reference were 6.52–13.98 N and 3.21–4.47, respectively. In detail, for the calibration set, the ranges of firmness and pH were 6.52–13.98 N and 3.21–4.47, respectively. For the prediction set, the ranges of firmness and pH were 7.8–13.49 N and 3.43–4.44, respectively. The firmness and pH ranges of the calibration set encompass those of the prediction set, which facilitates the development of regression models.

Table 2
Statistics of measurement of samples in both calibration and prediction sets, respectively.

Parameters	Sample set	No. of samples	Min.	Max.	Mean	Standard Deviation
Firmness (N)	Total samples	239	6.52	13.98	10.48	1.76
	Calibration set	159	6.52	13.98	10.15	1.82
	Prediction set	80	7.8	13.49	11.14	1.44
pH	Total samples	238	3.21	4.47	3.86	0.27
	Calibration set	159	3.21	4.47	3.83	0.28
	Prediction set	79	3.43	4.44	3.93	0.22

3.3. *Feature wavelength selection and modeling analysis using CARS and SPA*

CARS and SPA were applied to screen the optimal wavelength subset from the average spectra to predict the grape firmness (or pH). An example of the wavelengths selected via CARS for the prediction of grape firmness is shown in Fig. 4. The *RMSECV* of five-fold CV was obtained based on 50 MC samples (Fig. 4b). As shown, the *RMSECV* first decreases, which implies that wavelengths unrelated to the firmness prediction are removed; subsequently, the *RMSECV* value increases gradually, which is attributed to the removal of some valid variables. When the minimum *RMSECV* (0.8792) was obtained in the 28th sampling run, the corresponding optimal characteristic wavelength subset is located on the vertical line denoted by a blue asterisk (Fig. 4c), and the number of selected wavelengths is 22 (Fig. 4a). In addition, Fig. 5 shows the wavelength selection results for the determination of firmness using the SPA. Fig. 5a shows the change tendency of the root square error of validation (*RMSEV*) with the wavelength number by the model; when the selected key wavelength number is 18 (red square symbol), the *RMSEV* of the model is the optimal value, and the wavelength indices (red solid circular symbol) corresponding to the 18 selected wavelengths

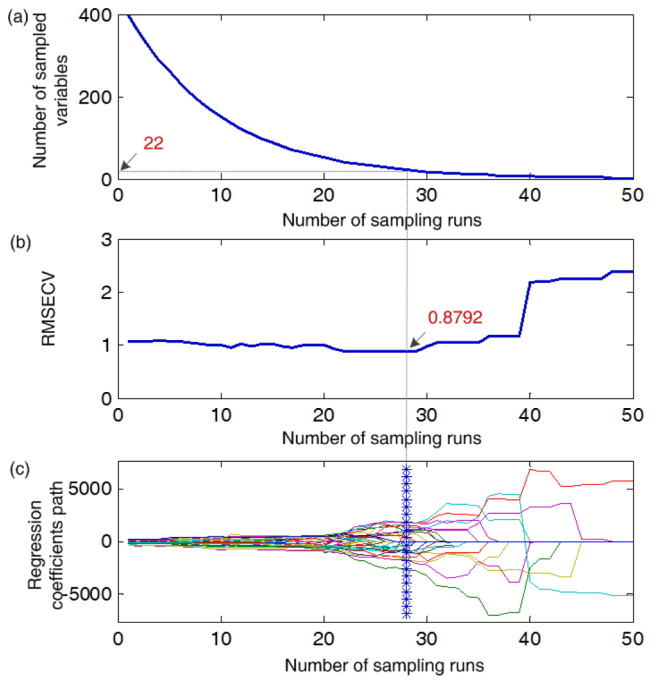

Fig. 4. Feature wavelength selection by CARS for firmness. (a), (b) and (c) show the change trend of wavelengths number, five-fold RMSECV values and the regression coefficients path of each wavelength with the increase of MC sampling runs, respectively.

are shown in Fig. 5b. Similarly, the feature wavelengths identified by CARS and SPA for the prediction of grapes pH were 19 and 14, respectively (Fig. 6). As shown in Fig. 6, the wavelengths selected by CARS and SPA are primarily distributed around features of the mean spectrum, which is consistent with the spectral profiles of grapes presented in Section 3.1. Therefore, based on the effective wavelength analysis, it was proven that HSI is feasible for determining the firmness and pH of grapes.

PLS and LSSVM models were constructed using the full or feature bands selected using CARS and SPA to detect the firmness (or pH) of the grapes. The prediction results of the PLS and LSSVM models based on different input datasets are listed in Table 3. For firmness (or pH) prediction using PLS, when the optimal LVs numbers were determined to be 12 for firmness and 10 for pH under full mean spectra based on the lowest *RMSECV*, the optimal accuracy was achieved, where $R_p^2 = 0.8573$,*RMSEP* $= 0.674$ N and $RPD = 2.13$ for firmness; and $R_p^2 = 0.8688$, $RMSEP = 0.0818$, and $RPD = 2.69$ for pH. For firmness (or pH) prediction using the LSSVM, when the combination of (γ,σ^2) was $(5.9 \times 10^1,\ 3.4 \times 10^3)$ for firmness and $(2.3 \times 10^2,\ 2.2 \times 10^4)$ for pH, the best accuracy was obtained, where $R_p^2 = 0.8726$, $RMSEP = 0.6112$ N, and $RPD = 2.36$ for firmness; and $R_p^2 = 0.8645$,*RMSEP* $= 0.0822$, and $RPD = 2.68$ for pH. The scatter plots of measured vs. predicted firmness and pH of samples the PLS and LSSVM models based on the full wavelengths are shown in Fig. 7a-b and e-f, respectively. As shown in Table 3, the LSSVM models outperform the PLS models for firmness prediction, whereas the PLS models are more suitable for detecting the pH. This is because the relationship between the spectra and firmness is primarily nonlinear, whereas the relationship between the spectra and pH is approximately linear. The RPD values of all models exceed 2.0, which indicates that the developed models are feasible for predicting the firmness (or pH) of grapes [36]. Among them, the models developed using full wavelengths demonstrate the best performance. Nevertheless, the full-band-based models result in a higher modeling cost and prevent the implementation of online detection. In fact, the models constructed using fewer effective wavelengths selected using CARS and SPA can reduce the modeling time and be more useful for practical applications; however, the accuracy of feature-wavelength-based models would deteriorate by varying degrees. Accordingly, in the next section, a deep learning approach for extracting deep features and establishing prediction models is introduced; the approach is analyzed to determine whether it can further improve the firmness and pH prediction of grapes.

3.4. Deep features extraction and modeling analysis using SAE-PLS and SAE-LSSVM

SAE-PLS (or SAE-LSSVM) was developed to predict the firmness and pH of grapes. The method comprises two stages: training and fine-tuning. The SAE was first trained in an unsupervised manner using 600 randomly selected pixel spectra of each sample (96.000 spectra in total) in the calibration set. Additionally, an SAE network with a symmetric structure (400, 200, 100, 50, h, 50, 100, 200, 400) was designed, where h represents the unit of the last hidden layer of the encoding part, denoted as SAE (h). Finally, the last hidden layer units of the encoder were extracted as deep features using the SAE. An example of the training result of SAE (24) is shown in Fig. 8. After executing the SAE with 100 iterating epochs, the mean validation error between the input average spectra and output reconstructed average spectra indicates a low value of 3.725×10^{-3}(Fig. 8a), suggesting that the original spectra can be well reproduced in the reconstructed spectra by the SAE (Fig. 8b). The deep spectral feature curves of grapes with different hidden units (200, 100, 50, and 24) are shown in Fig. 8c-f. Compared Fig. 8f and 8c, it is observed that the spectral features become more abstract and differentiable as the number of units in the last encoding layer of SAE-encoding reducing from 200 to 24. Based on the principle of the SAE, the deep spectral features extracted in the last hidden layer provide the most useful information regarding the original spectra of the grapes.

After training the SAE, the decoding part of the SAE was removed, the encoding part of the SAE was retained, and PLS (or LSSVM) was

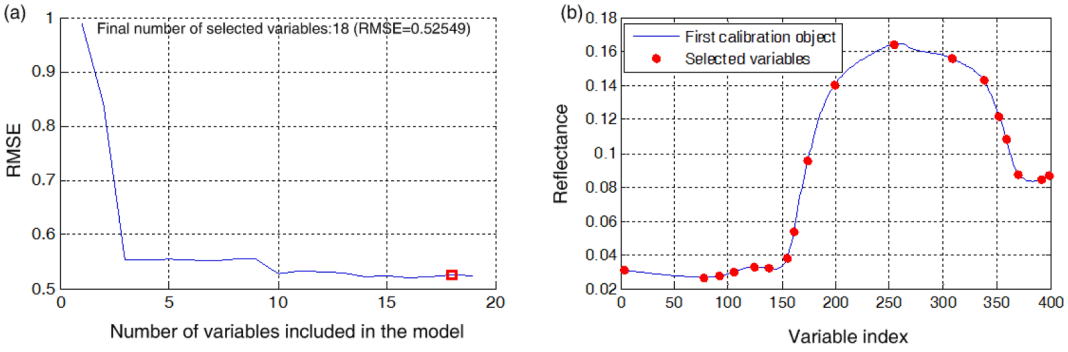

Fig. 5. Feature wavelength selection by SPA for firmness. The RMSE change with wavelength (variable) number in the model (a), and the Variable index of selected 18 effective wavelengths (b).

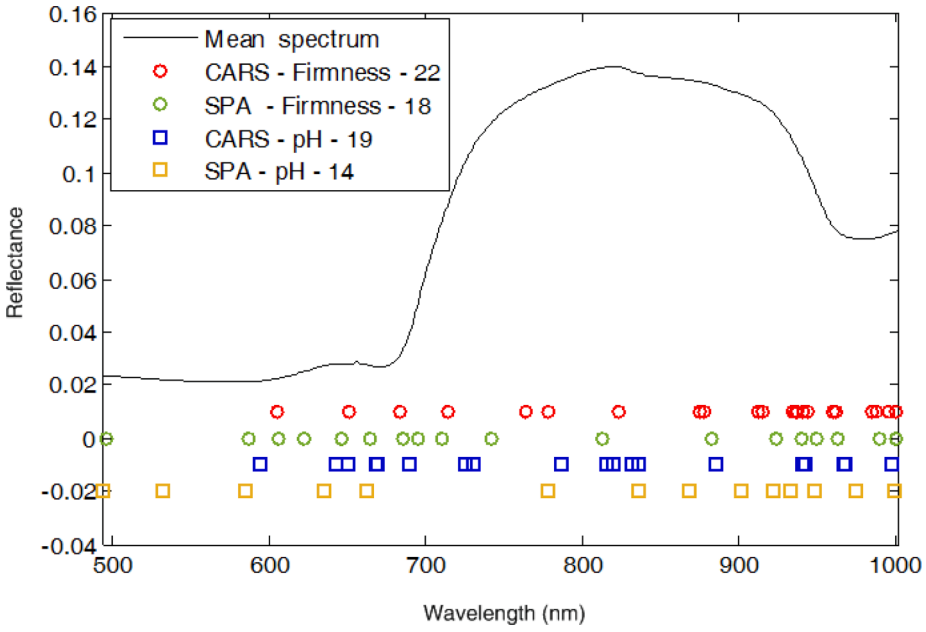

Fig. 6. The distribution of firmness and pH feature wavelength selected by CARS and SPA. Note: a, b and c in a-b-c represent wavelength selection algorithm, firmness (or pH) and number of feature wavelengths.

Table 3
Results of Firmness and pH in grape using PLSR and LSSVM with full or feature wavelength.

Parameters	Model types	Data type	N.W.[b]	LVs or (γ, σ^2)[c]	Calibration set		Prediction set		
					R_c^2	RMSEC	R_p^2	RMSEP	RPD
Firmness (N)	PLS	Full	400	12	0.8966	0.5767	0.8573	0.674	2.13
		CARS	22	10	0.8882	0.595	0.8489	0.7036	2.05
		SPA	18	11	0.8918	0.5795	0.8509	0.6966	2.07
	LSSVM	Full	400	$(5.9 \times 10^1, 3.4 \times 10^3)$	0.9112	0.5211	0.8726	0.6112	2.36
		CARS	22	$(8.4 \times 10^1, 6.3 \times 10^3)$	0.9018	0.5401	0.8607	0.6521	2.17
		SPA	18	$(2.2 \times 10^1, 1.1 \times 10^3)$	0.9041	0.5389	0.8642	0.6505	2.21
pH	PLS	Full	400	10	0.916	0.0812	0.8688	0.0818	2.69
		CARS	19	10	0.9034	0.0871	0.8402	0.0899	2.45
		SPA	14	9	0.8971	0.0899	0.824	0.0945	2.33
	LSSVM	Full	400	$(2.3 \times 10^2, 2.2 \times 10^4)$	0.9296	0.0759	0.8645	0.0822	2.68
		CARS	19	$(7.2 \times 10^2, 3.7 \times 10^4)$	0.8863	0.0945	0.8365	0.0911	2.41
		SPA	14	$(6.6 \times 10^2, 3.2 \times 10^4)$	0.9019	0.0878	0.8194	0.0958	2.30

Note: b: N.W. is the number of wavelengths; c:(γ, σ^2) is the optimized parameter of LSSVM modeling based on RBF kernel.

incorporated into the last layer to construct the regression model. The trained weights from the SAE-encoding stage were used as the initial weights of the models. Subsequently, the entire regression model was fine-tuned using the R^2 of the calibration set between the predicted value and the reference measured firmness (or pH) in a supervised manner. Finally, the predicted set was used to assess the performance of the model. The units of the last hidden layer (h) of the encoding part in the SAE were invariably affected the prediction accuracy of the regression models. Therefore, h was set as the values (10, 20, 30, and 40) to train the SAE, and the range of best units was estimated. Next, the optimal unit number in the range above was determined based on the highest R^2 of the average spectra and firmness (or pH) in the calibration set.

In our experiments, a batch of size 100 and an epoch value of 100 were used to train the SAE and fine-tune the entire network. The prediction of firmness and pH using regression models developed with full pixel spectra and different deep features are listed in Table 4. For firmness (or pH) prediction using F-PLS (or SAE-PLS), the optimal result for firmness was achieved by SAE (30), where $R_p^2 = 0.9009$,$RMSEP = 0.4663 \quad N$, and $RPD = 3.09$, whereas the best accuracy for pH was obtained by SAE (19), where $R_p^2 = 0.9005$, $RMSEP = 0.0781$, and $RPD = 2.82$. For firmness (or pH) prediction using F-LSSVM (or SAE-LSSVM), the optimal result for firmness was achieved by SAE (24), where $R_p^2 = 0.9232$,$RMSEP = 0.4422 \quad N$, and $RPD = 3.26$, whereas the best accuracy for pH was obtained by SAE (22), where $R_p^2 = 0.8943$,$RMSEP = 0.0792$, and $RPD = 2.78$. Table 4 shows that the performances of the SAE (24)-LSSVM and SAE (19)-PLS models are optimal for firmness and pH, respectively. The scatter plots of the measured vs. predicted firmness and pH of the samples based on the best performances of SAE-PLS and SAE- LSSVM are shown in Fig. 7c-d and g-h, respectively. Comparing Table 4 with Table 3, it is observed that the models established based on deep spectral features yield better performance than traditional modeling method such as PLS or the LSSVM. These results are attributable to the following two reasons: the SAE can extract deep abstract features from hyperspectral data that contain significant amounts of multicollinearity, which enables global information in the hyperspectral image data to be revealed. Second, the deep learning models were developed using many pixel spectra; as such, overfitting was prevented and prediction accuracy improved [37]. In

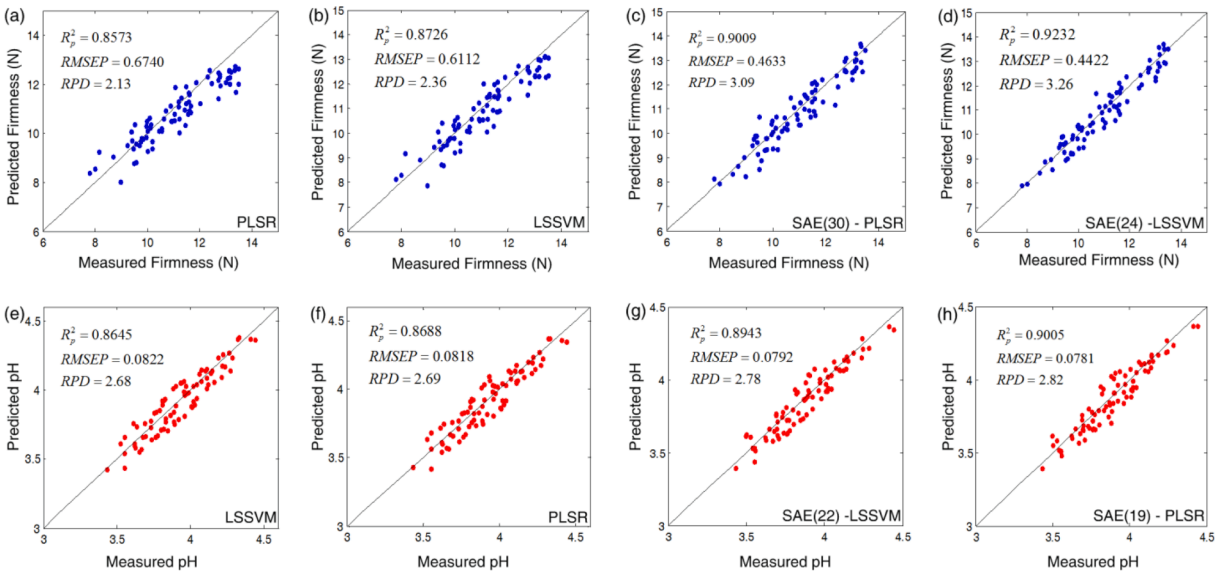

Fig. 7. Scatter plots of predicted vs. measured the firmness and pH of grapes using the models of PLS, LSSVM, SAE-PLS and SAE-LSSVM. Firmness (a)-(d), and pH (e)-(h).

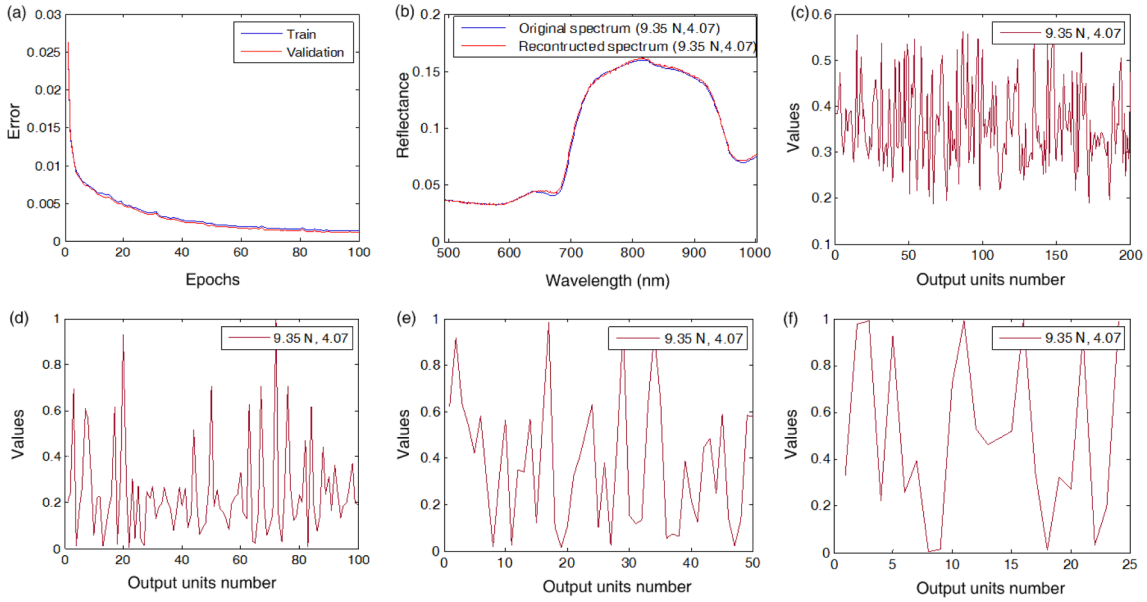

Fig. 8. The training result of a SAE. Reconstruction error of training (a), original input spectra and reconstructed spectra curves(b), deep spectral features of encoding part with 200, 10,050 and 24 units (c)-(f).

addition, the predictive accuracy of the best SAE-based models is superior to that of the full-pixel-spectra-based models listed in Table 4. In our opinion, this shows that the SAE-based deep learning method can effectively capture useful and important features from HSI data and improve the predictive accuracy to estimate the firmness (or pH) of grapes. Consequently, the overall results show that HSI coupled with the deep learning method can non-destructively and rapidly detect internal quality attributes in fruit.

4. Conclusions

This study suggests that it is feasible to apply HSI coupled with deep learning to detect the firmness and pH of Kyoho grapes non-destructively. Because deep learning involves the training of significant amounts of data to achieve the best network, the training time of SAE-based models under a GPU is slightly longer than that of conventional models. However, the computational time required for feature extraction and prediction analysis is not affected by online inspection. This is because the training process is typically implemented offline before the prediction stage. In addition, the merit of automatic features learned by an SAE can decrease significantly the computing time in comparison with other conventional algorithms that must extract features manually. It is noteworthy that the SAE-based deep learning methods employ many pixel level spectra to develop deep learning models with high precision, unlike other traditional algorithms, thereby alleviating the problem of insufficient samples. To the best of our knowledge, this is the first study where a deep learning method is used for the internal quality assessment of grapes. The satisfied results obtained using deep learning models (SAE-PLS or SAE-LSSVM) support the development of portable devices and promote the use of deep learning schemes for further predicting of other internal quality attributes of fruit. In future studies, enabled by progress in hardware

Table 4
Results of Firmness and pH in grape using PLS and LSSVM models developed with full pixels spectra and different deep features extracted by SAE.

Parameters	Model types	N.D.F.[d]	LVs or (γ, σ^2)	Calibration set		Prediction set		
				R_c^2	RMSEC	R_p^2	RMSEP	RPD
Firmness (N)	F-PLS[e]	400	11	0.9201	0.5026	0.8704	0.6113	2.36
	SAE-PLS	40	12	0.9106	0.5373	0.8648	0.6359	2.26
		30	13	0.9329	0.4799	0.9009	0.4663	3.09
		20	12	0.9289	0.4933	0.8831	0.5720	2.52
		10	11	0.9012	0.5529	0.8518	0.6798	2.12
	F-LSSVM[f]	400	$(3.8 \times 10^2, 5.4 \times 10^3)$	0.9289	0.4858	0.8921	0.5597	2.57
	SAE-LSSVM	40	$(5.2 \times 10^1, 2.2 \times 10^2)$	0.9196	0.5146	0.8668	0.6147	2.34
		30	$(4.3 \times 10^2, 1.9 \times 10^3)$	0.9321	0.4803	0.8989	0.5323	2.71
		24	$(4.2 \times 10^3, 7.6 \times 10^3)$	0.9486	0.4121	0.9232	0.4422	3.26
		20	$(6.6 \times 10^3, 9.5 \times 10^4)$	0.9265	0.4992	0.8797	0.5964	2.41
		10	$(9.6 \times 10^2, 1.5 \times 10^3)$	0.9153	0.5187	0.8658	0.6201	2.32
pH	F-PLS	400	11	0.9338	0.0748	0.8801	0.0805	2.73
	SAE-PLS	40	10	0.9328	0.0759	0.8691	0.0815	2.7
		30	11	0.9341	0.0741	0.8812	0.0809	2.72
		20	11	0.9356	0.0726	0.8905	0.0801	2.75
		19	12	0.9418	0.0697	0.9005	0.0781	2.82
		10	10	0.9202	0.0791	0.8624	0.0828	2.66
	F-LSSVM	400	$(4.5 \times 10^4, 2.4 \times 10^3)$	0.9298	0.0756	0.8708	0.0834	2.64
	SAE-LSSVM	40	$(3.7 \times 10^3, 7.6 \times 10^2)$	0.9278	0.0761	0.8676	0.082	2.68
		30	$(1.8 \times 10^4, 1.3 \times 10^3)$	0.9334	0.0744	0.8857	0.0804	2.74
		22	$(7.5 \times 10^4, 9.1 \times 10^2)$	0.9405	0.0704	0.8943	0.0792	2.78
		20	$(2.9 \times 10^4, 7.3 \times 10^2)$	0.9305	0.0752	0.8729	0.0813	2.71
		10	$(5.5 \times 10^4, 3.3 \times 10^2)$	0.9229	0.0776	0.8539	0.0839	2.62

Note: d: N.D.F. is the number of deep spectral features (i.e. hidden layer nodes) extracted by SAE; e: F-PLS is PLS based on the full pixels spectra; f: F-LSSVM is LSSVM based on the full pixels spectra.

computing ability, training and prediction tasks will be completed in a much shorter time, thereby further facilitating the practical application of deep learning in terms of speed and cost-effectiveness. In addition, additional grape samples and varieties will be investigated to obtain more accurate and robust models that can be transferred to portable HSI devices to achieve industrial detection in multiple environments.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is partially supported by National natural science funds projects (31971788), Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD-2018-87), Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX21_3386), Project funded by China Postdoctoral Science Foundation (2021M701479), Jiangsu University Student Research project (20AB0003), Jiangsu University Student Innovation Project (130).

References

[1] R.C. Colombo, S.R. Roberto, M.A. da Cruz, D.U. de Carvalho, L.Y. Yamamoto, S. L. Nixdorf, J. Pérez-Navarro, S. Gómez-Alonso, M. Shahab, S. Ahmed, L.S. A. Gonçalves, R.T. de Souza, I. Hermosín-Gutiérrez, Characterization of the phenolic ripening development of 'BRS Vitoria' seedless table grapes using HPLC–DAD–ESI-MS/MS, J. Food Compos. Anal. 95 (2021) 103693, https://doi.org/10.1016/j.jfca.2020.103693.

[2] D. Solairaj, Q. Yang, N.N. Guillaume Legrand, M.N. Routledge, H. Zhang, Molecular explication of grape berry-fungal infections and their potential application in recent postharvest infection control strategies, Trends Food Sci. Technol. 116 (2021) 903–917.

[3] I. Balic, T. Ejsmentewicz, D. Sanhueza, C. Silva, T. Peredo, P. Olmedo, M. Barros, J. C. Verdonk, R. Paredes, C. Meneses, H. Prieto, A. Orellana, B.G. Defilippi, R. Campos-Vargas, Biochemical and physiological study of the firmness of table grape berries, Postharvest Biol. Technol. 93 (2014) 15–23, https://doi.org/10.1016/j.postharvbio.2014.02.001.

[4] A. Benelli, C. Cevoli, L. Ragni, A. Fabbri, In-field and non-destructive monitoring of grapes maturity by hyperspectral imaging, Biosyst. Eng. 207 (2021) 59–67, https://doi.org/10.1016/j.biosystemseng.2021.04.006.

[5] M.L.A.T. Hertog, R. Ben-Arie, E. Róth, B.M. Nicolaï, Humidity and temperature effects on invasive and non-invasive firmness measures, Postharvest Biology and Technology 33 (1) (2004) 79–91, https://doi.org/10.1016/j.postharvbio.2004.01.005.

[6] A.M. Fernandes, C. Franco, A. Mendes-Ferreira, A. Mendes-Faia, P.L.D. Costa, P. Melo-Pinto, Brix, pH and anthocyanin content determination in whole Port wine grape berries by hyperspectral imaging and neural networks, Comput. Electron. Agric. 115 (2015) 88–96, https://doi.org/10.1016/j.compag.2015.05.013.

[7] J. Li, W. Huang, C. Zhao, B. Zhang, A comparative study for the quantitative determination of soluble solids content, pH and firmness of pears by Vis/NIR spectroscopy, J. Food Eng. 116 (2) (2013) 324–332, https://doi.org/10.1016/j.jfoodeng.2012.11.007.

[8] C. Xie, B. Chu, Y. He, Prediction of banana color and firmness using a novel wavelengths selection method of hyperspectral imaging, Food Chem. 245 (2018) 132–140, https://doi.org/10.1016/j.foodchem.2017.10.079.

[9] T. Ma, Y. Xia, T. Inagaki, S. Tsuchikawa, Non-destructive and fast method of mapping the distribution of the soluble solids content and pH in kiwifruit using object rotation near-infrared hyperspectral imaging approach, Postharvest Biol. Technol. 174 (2021) 111440, https://doi.org/10.1016/j.postharvbio.2020.111440.

[10] X. Wei, J. He, S. Zheng, D. Ye, Modeling for SSC and firmness detection of persimmon based on NIR hyperspectral imaging by sample partitioning and variables selection, Infrared Phys. Technol. 105 (2020) 103099, https://doi.org/10.1016/j.infrared.2019.103099.

[11] H. Pu, D. Liu, L. Wang, D. Sun, Soluble Solids Content and pH Prediction and Maturity Discrimination of Lychee Fruits Using Visible and Near Infrared Hyperspectral Imaging, Food Anal. Methods 9 (1) (2016) 235–244, https://doi.org/10.1007/s12161-015-0186-7.

[12] R. Pourdarbani, S. Sabzi, D. Kalantari, J.I. Arribas, Non-destructive visible and short-wave near-infrared spectroscopic data estimation of various physicochemical properties of Fuji apple (Malus pumila) fruits at different maturation stages, Chemometrics and Intelligent Laboratory Systems 206 (2020) 104147, https://doi.org/10.1016/j.chemolab.2020.104147.

[13] C. Kanchanomai, S. Ohashi, D. Naphrom, W. Nemoto, P. Maniwara, K. Nakano, Non-destructive analysis of Japanese table grape qualities using near-infrared spectroscopy, Hortic. Environ. Biotechnol. 61 (4) (2020) 725–733, https://doi.org/10.1007/s13580-020-00256-4.

[14] W. Zheng, Y. Bai, H. Luo, Y. Li, X.i. Yang, B. Zhang, Self-adaptive models for predicting soluble solid content of blueberries with biological variability by using near-infrared spectroscopy and chemometrics, Postharvest Biol. Technol. 169 (2020) 111286, https://doi.org/10.1016/j.postharvbio.2020.111286.

[15] Y. Lu, W. Saeys, M. Kim, Y. Peng, R. Lu, Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress, Postharvest Biol. Technol. 170 (2020) 111318, https://doi.org/10.1016/j.postharvbio.2020.111318.
[16] Y. Yun, H. Li, B. Deng, D. Cao, An overview of variable selection methods in multivariate analysis of near-infrared spectra, TrAC, Trends Anal. Chem. 113 (2019) 102–115, https://doi.org/10.1016/j.trac.2019.01.018.
[17] C. Wang, B. Liu, L. Liu, Y. Zhu, J. Hou, P. Liu, X. Li, A review of deep learning used in the hyperspectral image analysis for agriculture, Artif. Intell. Rev. 54 (7) (2021) 5205–5253, https://doi.org/10.1007/s10462-021-10018-y.
[18] C. Zhang, W. Wu, L. Zhou, H. Cheng, X. Ye, Y. He, Developing deep learning based regression approaches for determination of chemical compositions in dry black goji berries (Lycium ruthenicum Murr.) using near-infrared hyperspectral imaging, Food Chem. 319 (2020) 126536, https://doi.org/10.1016/j.foodchem.2020.126536.
[19] Y. Fan, C. Zhang, Z. Liu, Z. Qiu, Y. He, Cost-sensitive stacked sparse auto-encoder models to detect striped stem borer infestation on rice based on hyperspectral imaging, Knowl.-Based Syst. 168 (2019) 49–58, https://doi.org/10.1016/j.knosys.2019.01.003.
[20] D. Ahn, J. Choi, H. Kim, J. Cho, K. Moon, T. Park, Estimating the Composition of Food Nutrients from Hyperspectral Signals Based on Deep Neural Networks, Estimating the Composition of Food Nutrients from Hyperspectral Signals Based on Deep Neural Networks Sensors 19 (7) (2019) 1560, https://doi.org/10.3390/s19071560.
[21] X. Yu, H. Lu, D. Wu, Development of deep learning method for predicting firmness and soluble solid content of postharvest Korla fragrant pear using Vis/NIR hyperspectral reflectance imaging, Postharvest Biol. Technol. 141 (2018) 39–49, https://doi.org/10.1016/j.postharvbio.2018.02.013.
[22] X. Zhou, J. Sun, Y. Tian, B. Lu, Y. Hang, Q. Chen, Hyperspectral technique combined with deep learning algorithm for detection of compound heavy metals in lettuce, Food Chem. 321 (2020) 126503, https://doi.org/10.1016/j.foodchem.2020.126503.
[23] R.K.H. Galvão, M.C.U. Araujo, G.E. José, M.J.C. Pontes, E.C. Silva, T.C.B. Saldanha, A method for calibration and validation subset partitioning, Talanta 67 (4) (2005) 736–740, https://doi.org/10.1016/j.talanta.2005.03.025.
[24] X. Zhou, J. Sun, Y. Tian, Q. Chen, X. Wu, Y. Hang, A deep learning based regression method on hyperspectral data for rapid prediction of cadmium residue in lettuce leaves, Chemometrics and Intelligent Laboratory Systems 200 (2020) 103996, https://doi.org/10.1016/j.chemolab.2020.103996.
[25] J. Li, R. Zhang, J. Li, Z. Wang, H. Zhang, B. Zhan, Y. Jiang, Detection of early decayed oranges based on multispectral principal component image combining both bi-dimensional empirical mode decomposition and watershed segmentation method, Postharvest Biol. Technol. 158 (2019) 110986, https://doi.org/10.1016/j.postharvbio.2019.110986.
[26] M. Xu, J. Sun, X. Zhou, N. Tang, J. Shen, X. Wu, Research on nondestructive identification of grape varieties based on EEMD-DWT and hyperspectral image, J. Food Sci. 86 (5) (2021) 2011–2023, https://doi.org/10.1111/1750-3841.15715.
[27] D. Cao, Y. Liang, Q. Xu, H. Li, X. Chen, A new strategy of outlier detection for QSAR/QSPR, J. Comput. Chem. 31 (3) (2010) 592–602, https://doi.org/10.1002/jcc.21351.
[28] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometrics and Intelligent Laboratory Systems 58 (2) (2001) 109–130, https://doi.org/10.1016/S0169-7439(01)00155-1.
[29] J. Suykens, J. Vandewalle, Least squares support vector machine classifiers, Neural Process. Lett. 9 (1999) 293–300, https://doi.org/10.1023/A:1018628609742.
[30] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323 (6088) (1986) 533–536, https://doi.org/10.1038/323533a0.
[31] V.M. Gomes, A.M. Fernandes, A. Faia, P. Melo-Pinto, Comparison of different approaches for the prediction of sugar content in new vintages of whole Port wine grape berries using hyperspectral imaging, Computers and Electronics in Agriculture 140 (2017) 244–254, https://doi.org/10.1016/j.compag.2017.06.009.
[32] A. Baiano, C. Terracone, G. Peri, R. Romaniello, Application of hyperspectral imaging for prediction of physico-chemical and sensory characteristics of table grapes, Comput. Electron. Agric. 87 (2012) 142–151, https://doi.org/10.1016/j.compag.2012.06.002.
[33] Y. Shao, K. Wang, G. Xuan, C. Gao, Z. Hu, Soluble solids content monitoring for shelf-life assessment of table grapes coated with chitosan using hyperspectral imaging, Infrared Phys. Technol. 115 (2021) 103725, https://doi.org/10.1016/j.infrared.2021.103725.
[34] R. Beghi, V. Giovenzana, L. Brancadoro, R. Guidetti, Rapid evaluation of grape phytosanitary status directly at the check point station entering the winery by using visible/near infrared spectroscopy, J. Food Eng. 204 (2017) 46–54, https://doi.org/10.1016/j.jfoodeng.2017.02.012.
[35] G. Özdoğan, X. Lin, D. Sun, Rapid and noninvasive sensory analyses of food products by hyperspectral imaging: Recent application developments, Trends Food Sci. Technol. 111 (2021) 151–165, https://doi.org/10.1016/j.tifs.2021.02.044.
[36] B.M. Nicolaï, K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K.I. Theron, J. Lammertyn, Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review, Postharvest Biol. Technol. 46 (2) (2007) 99–118, https://doi.org/10.1016/j.postharvbio.2007.06.024.
[37] X. Yu, H. Lu, Q. Liu, Deep-learning-based regression model and hyperspectral imaging for rapid detection of nitrogen concentration in oilseed rape (Brassica napus L.) leaf, Chemometrics and Intelligent Laboratory Systems 172 (2018) 188–193, https://doi.org/10.1016/j.chemolab.2017.12.010.