Development of deep learning method for predicting firmness and soluble solid content of postharvest Korla fragrant pear using Vis/NIR hyperspectral reflectance imaging

Xinjie Yu⁎, Huanda Lu, Di Wu

⁎ Corresponding author.
E-mail address: xjyu1979@nit.zju.edu.cn (X. Yu).

Received 29 November 2017; Received in revised form 22 February 2018; Accepted 23 February 2018

ARTICLE INFO

Keywords:
Stacked auto-encoders
Fully-connected neural network
Pixel-level spectral features
Fruit quality
Non-destructive detection

ABSTRACT

The objective of this research was to develop a deep learning method which consisted of stacked auto-encoders (SAE) and fully-connected neural network (FNN) for predicting firmness and soluble solid content (SSC) of postharvest Korla fragrant pear (Pyrus brestschneideri Rehd). Firstly, deep spectral features in visible and near-infrared (380–1030 nm) hyperspectral reflectance image data of pear were extracted by SAE, and then these features were used as input data to predict firmness and SSC by FNN. The SAE-FNN model achieved reasonable prediction performance with $R^2 = 0.890$, RMSEP = 1.81 N and RPD = 3.05 for firmness, and $R^2 = 0.921$, RMSEP = 0.22% and RPD = 3.68 for SSC. This research demonstrated that deep learning method coupled with hyperspectral imaging technique can be used for rapid and nondestructive detecting firmness and SSC in Korla fragrant pear, which would be useful for postharvest fruit quality inspections.

1. Introduction

Korla fragrant pear (Pyrus brestschneideri Rehd) is one of the most popular fruit in China (Wang et al., 2014). This pear is attracting an increasing number of Chinese consumers and is exported to many countries around the world because of its sweet and fragrant taste. Firmness and soluble solid content (SSC) are the most important edible quality attributes of Korla fragrant pear fruit, and they directly influence on consumer satisfaction (Fu et al., 2007; Chen et al., 2006; Wang et al., 2017). Thus, they must be taken into account when determining the effect of the postharvest storage period and conditions on fruit quality control. Currently, the industry standard for the determination of fruit firmness is penetrometer test carried out using a Magness-Taylor to penetrate fruit flesh to a depth (Guo et al., 2015). The traditional method used to measure fruit SSC requires juice extracted from the fruit pulp, and is carried out by using digital refractometer. These destructive methods cannot meet the requirements of rapid and automatic monitoring firmness and SSC changes of fruit during postharvest storage. Therefore, there is a considerable need for nondestructive method for the assessment of firmness and SSC in Korla fragrant pear to instruct postharvest handling.

Visible and near-infrared (Vis/NIR) hyperspectral imaging has been one of the most successful techniques for nondestructive detection of firmness and/or SSC of fruit such as apple (Mendoza et al., 2011; Fan et al., 2016), blueberries (Leiva-Valenzuela et al., 2013), banana (Rajkumar et al., 2012), peach (Lu and Peng, 2006) and grapes (Baiano et al., 2012). Recently, Vis/NIR hyperspectral image has also been applied to predict firmness and SSC of pear fruit (Li et al., 2016). The principle for such detections is based on measuring the remission of spectrum from fruit surface in the form of reflectance, interactance or transmission. Since the measured spectrum is related to the composition and structure of the fruit, the relevant wavelength changes in spectrum can thus be utilized to be correlated with the firmness and/or SSC by using some chemometric methods (Huang et al., 2017). However, as Vis/NIR hyperspectral image integrates imaging and spectroscopy to obtain both spatial and spectral information simultaneously from a sample to be measured, the huge amounts of data in hyperspectral image increase the data processing load, resulting in the difficulty of quantitative modeling for the prediction of composition and structure properties of fruit (Liu et al., 2014). In order to reduce the data processing load, the majority of previous studies concerning the application of the hyperspectral imaging to the evaluation of fruit properties used an averaged spectrum or a limited number of representative spectra extracted from region of interest (ROI), which make it difficult to reveal the global information in hyperspectral image relating the differences in composition and structure properties of a...
sample (Wang et al., 2016; Roggo et al., 2005). Aiming to deal with large datasets composed of a high number of hyperspectral image data for the early bruise detection on apples, Ferrari et al. (2015) recently proposed a hyperspectrogram-based approach, which can reduce each hyperspectral image (both spatial and spectral information) into a single signal. The reduced signals can then be analyzed simply using multivariate analysis methods such as partial least squares regression (PLSR), least-squares support vector machine (LS-SVM) and multilinear regression (MLR). However, in the hyperspectrogram-based approach, multivariate analysis of the reduced signals is made at the image level, and not a pixel level. So far, very little work has been carried out on the big data analysis of hyperspectral image based on a large number of pixel-level spectra for the fruit quality detection.

Deep learning is a new area of machine learning research (LeCun et al., 2015), which has a deep structure of artificial neural networks with capable of processing very large-scale data, and has dramatically improved the modeling performance in many data analysis tasks (Zhao and Du, 2016; Farias et al., 2016; Cole et al., 2017). In the previous study, our research group had confirmed that deep learning classification algorithm coupled with pixel-level spectra in Vis/NIR hyperspectral image (450–1010 nm) achieved satisfactory total classification accuracy of 98.28% for discriminating freshness of shrimp product (Yu et al., 2017). In this work, a novel deep learning regression method is further developed for the prediction of firmness and SSC of postharvest pear fruit. This deep learning method composed of a stacked auto-encoders (SAE) (Suk et al., 2015) and a feed-forward fully-connected neural network (FNN) (Biganzoli et al., 1998), in which SAE is trained by using a large number of pixel spectra to automatically learn spectral features of Vis/NIR hyperspectral image, and the spectral features are fed to FNN to quantitatively predict the corresponding firmness and SSC. To the best of our knowledge, this is the first time that deep learning is applied to the detection of quality attributes of pear fruit. The main object of this study is to investigate the potential of SAE-FNN method together with Vis/NIR hyperspectral imaging technique for detecting firmness and SSC of postharvest Korla fragrant pear fruit.

2. Materials and methods

2.1. Pear samples

Korla fragrant pear fruit at commercial maturity were hand-harvested from an orchard located in Korla, Xinjiang, China (86.10’N, 41.69’E) on 4 September 2016. The fruit were picked in different trees and in different canopy layers to get representative samples which contain a wider range of growth conditions, then wrapped in EPE foam fruit nets, put into boxes, and transported to Ningbo city, Zhejiang province of China within 72 h. Upon arrival at the laboratory, a total of 180 pear fruit with uniform size (diameter of 55-65 mm) and weight (120–130 g) were chosen out, then placed in chambers where the temperature (20 ± 1 °C) and relative humidity (90–95%) were controlled to produce a wide distribution in firmness and SSC as the fruit were ripened. Fifteen pear fruit were randomly taken out every other day (from 8 September to 30 September 2016) to be scanned by a Vis/NIR hyperspectral imaging system, and to measure their reference firmness and SSC by traditional destructive methods. The reference measured firmness and SSC of pear fruit were calibrated with the hyperspectral image information using PLSR, LS-SVM and SAE-FNN models, respectively. In the experiments, 135 samples were randomly chosen out and used as calibration set, whereas, the remaining 45 samples were used as prediction set for estimating the performance of the calibration models.

2.2. Reference measurements of firmness and SSC

Firmness and SSC were measured from one side of the pear fruit where spectral acquisition had been carried out. Firstly, a skin of 2 cm² was removed around the equatorial section of pear fruit (Fig. 1a). The firmness of the peeled tissue was analyzed destructively using a 3.5 mm diameter Magness-Taylor (M-T) probe, which was attached to a fruit sclerometer (Model: GY-1, Zhejiang Tuopu Instrument, Co. Ltd., Hangzhou, China). A downward pressure with velocity of 1 mm s⁻¹ was forced until the plunger has penetrated 10 mm into the peeled tissue. The maximum force was recorded and used as the measurement of pear fruit M-T firmness (Hertog et al., 2004) which expressed in newton (N) with an accuracy of 0.1 N. Then, the juice of flesh whose position was in the same peeled tissue was extracted by using a manual fruit squeezer. The juice was dropped to a digital refractometer (Model: PR-101α, Atago Co., Ltd., Tokyo, Japan) for SSC measurement. The refractive index accuracy is ± 0.1% and the Brix (%) range is 0.0–45.0%.

2.3. Hyperspectral imaging system

Before reference measurements of firmness and SSC, an in-house developed line-scan Vis/NIR hyperspectral imaging system was used to acquire hyperspectral reflectance images from Korla fragrant pears. The hyperspectral imaging system consisted of four components: a spectral imaging system, a lighting system, a translation stage and a computer (Fig. 2). The spectral imaging system was composed of a spectrograph (ImSpectorV10, Spectral Imaging Ltd., Finland) with spectral range of 380–1030 nm, which was connected to the CCD camera (B1621M, Imperx Inc., USA) with max resolution of 1632 pixels in the spatial dimension and 1232 bands in the spectral dimension, and a standard C-
mount zoom lens (F/1.4, f = 23 mm, 21–1001917, Schneider Optics Inc., Germany). The lighting system was composed of a 150 W quartz tungsten-halogen light source (3900-ER, Illumination Technology, Inc., USA), which shines on sample through two bifurcated line light guides. The translation stage consisted of a sample holder integrated to a liner slider which was driven by a stepper motor (IRCP-0076-1COMB, Isuzu Optics Corp., Taiwan, China). The computer (Dell Inc., USA) installed with system control software (Spectral Image System, Isuzu Optics Corp., Taiwan, China).

2.4. Image acquisition and correction

Before scanning samples, several system parameters were set to be fixed values for acquiring accurate hyperspectral data. The distance between the upper surface of fruit samples and the CCD camera lens was set to 450 mm. The speed of translation stage was 1.5 mm s⁻¹, and the exposure time of the CCD camera was 21 ms. The CCD camera used a two-dimensional detector array with 816 × 512 (spatial × spectral) configuration of Vis/NIR hyperspectral imaging system. The acquired hyperspectral image was further corrected according to the Eq. (1):

\[ I_c = \frac{I_o - B}{W - B} \times 100 \]  

(1)

where \( I_c \) is the corrected hyperspectral image, \( I_o \) is the original hyperspectral image, B is the black image recorded by closing completely the aperture of the camera, and W is the white image obtained by a white surface board with uniform, stable and high reflectance standard. The corrected hyperspectral images were the basis for the subsequent image analysis to extract the spectra, predict and visualize the firmness and SSC of Korla fragrant pear.

2.5. Spectra extraction

To know if firmness and SSC of Korla fragrant pear could be predicted using the hyperspectral imaging data, a representative ROI for each pear fruit was chosen manually from the corrected hyperspectral sample. In our experiments, the mean spectra were used as original data for spectral data preprocessing, optimum wavelengths selection, as well as building traditional regression models of PLSR and LS-SVM. Meanwhile, for the purpose of predicting firmness and SSC of pear fruit using deep learning method, 1000 pixel spectra in the ROI of each pear sample were randomly selected out to create a big data set. These randomly selected pixel spectra were firstly used to train SAE in the SAE-FNN model, and then the trained SAE was applied on the mean spectra to extract informative spectral features, which finally used as inputs of FNN in the SAE-FNN to predict the firmness and SSC of pear fruit.

2.6. Data preprocessing and wavelength selection

It is well-known that preprocessing of the spectra data may help to improve the performance of regression models like PLSR and LS-SVM. In this study, the original mean spectra data were preprocessed by multiplicative signal correction (MSC) method (Rady et al., 2017) to minimize noise resulting from various electronic sources and variation in sample conditions. The MSC preprocessing was carried out in the Unscrambler 9.7 software (CAMO Technologies Inc., Woodbridge, New Jersey).

On the other hand, wavelength selection can eliminate irrelevant variables to improve performance for regression models. Therefore, it is important to select optimum wavelengths that may perform equally well or even better than using full spectra for identifying the firmness and SSC in pear. In this study, successive projections algorithm (SPA) (Fan et al., 2017) was applied to select optimum wavelengths. SPA method was performed on the MSC preprocessed mean spectra in the calibration set by a SPA toolbox (available at http://www.ele.ita.br/~kawakami/spa/) in Matlab 8.1 (The Math Works, Natick, USA) and the parameters used in the toolbox were set to default values.

2.7. Calibration models for firmness and SSC prediction

2.7.1. PLSR

PLSR is a widely used chemometric method for building calibration models (Wold et al., 2001). The PLSR algorithm has inferential capability, which can be used to model a possible linear relationship between dependent variables Y (target composition and structure properties) and independent variables X (spectra data). During modeling, latent variables (LVs) in PLSR are selected in order to maximise the covariance between the projected objects in X space and corresponding projections in Y space. In this study, the PLSR algorithm was applied to build calibration models for firmness and SSC. The PLSR modeling was carried out in the Unscrambler 9.7 software (CAMO Technologies Inc., Woodbridge, New Jersey).

2.7.2. LS-SVM

LS-SVM is an efficient approach that can be used to establish linear and nonlinear calibration models by respectively using nonlinear radial basis function (RBF) kernel and linear kernel (Suykens et al., 2002). In our work, aiming to find out the best calibration model for firmness and SSC, both linear LS-SVM (LS-SVM(L)) and nonlinear RBF LS-SVM (LS-SVM(R)) were applied to establish calibration models. During the leave one out cross validation in the grid-search step for LS-SVM training, two parameters including \( \gamma \) and \( \sigma^2 \) were optimized for RBF kernel, and one parameter of \( \gamma \) was optimized for linear kernel. The LS-SVM was executed in the MATLAB 8.1 (The Math Works, Natick, USA) using a LS-SVM toolbox (available at http://www.esat.kuleuven.be/sista/lssvmlab/).
2.7.3. SAE-FNN

SAE-FNN was further developed for the prediction of firmness and SSC, and its performance was compared with that of PLSR and LS-SVM. The SAE-FNN is a deep neural network architecture, in which SAE is used as an unsupervised manner to extract spectral features from hyperspectral image by training pixel-level spectra, and the extracted spectral features are input into FNN to predict firmness and SSC.

The basic structure of the SAE is an auto-encoder, which aims to extract features from input data (Zhu et al., 2016) by non-linear modeling of deep neural networks. As shown in Fig. 3a, the auto-encoder has one input layer of d units, one hidden layer of h units, one output layer of d units, and an activation function \( f(\cdot) \). During the learning of auto-encoder, it first transfers the original input \( x \in \mathbb{R}^d \) to the hidden layer \( y \in \mathbb{R}^h \), this process is called “encoding”. Then, the hidden layer \( y \) is transferred back to the output layer \( z \in \mathbb{R}^d \), this process is called “decoding”. These two steps can be formulated as Eqs. (2) and (3):

\[
y = f(w_y x + b_y),
\]
\[
z = f(w_z y + b_z),
\]

where \( w_y \) is the input-to-hidden weight matrix, \( w_z \) is the hidden-to-output weight matrix, \( b_y \) and \( b_z \) denote the bias of hidden and output units, and \( f(\cdot) \) denotes the activation function. In the experiments, “SoftPlus” non-linearity activation: \( \text{softplus}(x) = \log(1 + e^x) \), a smooth version of the rectifying non-linearity, was used in the current study for the activation function \( f(\cdot) \), since “SoftPlus” can help to gain easier supervised and unsupervised training (Huang and Wang, 2017). The constraint between \( w_y \) and \( w_z \) is denoted as Eq. (4):

\[
w_y = w_z^T = W.
\]

Finally, the auto-encoder is trained by minimize the “error” between input data and output data, which is denoted as:

\[
\arg\min_{W, b_y, b_z} \sum_i [c(x_i, z_i)],
\]

where \( z \) is reconstructed output data which has the same size as the given input data \( x \). The auto-encoder can be trained in an unsupervised way to reproduce input data \( x \) at reconstructed data \( z \) by hidden representation \( y \), this means that the hidden representation \( y \) can argmin \( c(x, z) \), be employed as reduced feature variables representing the original input \( x \).

On the other hand, it should be stressed that stacking multilayers of auto-encoder network namely SAE can learn better high-level deep features compared to the single layer auto-encoder. Therefore, in this study, SAE that was established by stacking multilayers of auto-encoder network was conducted to extract deep spectral features from hyperspectral image of pear fruit (illustrated in Fig. 3b). In the experiments, we used large number of spectra (1000 random pixel spectra for each pear fruit sample in the calibration set, a total of 135,000 spectra) to train the SAE in deep learning way (Xing et al., 2016). After training of SAE, the decoding part of SAE is removed and the encoding part of SAE is retained, then a FNN is added to the last encoding layer of SAE to produce a SAE-FNN regression network for prediction. The FNN is a fully-connected back propagation neural network, and the output of FNN is a single unit denoted as:

\[
\hat{Y} = f(w_y y_i + b_i),
\]

where \( \hat{Y} \) is the output value (predicted firmness or SSC value), \( f(\cdot) \) denotes the “SoftPlus” non-linearity activation function, \( w_y \) is the weight matrix, \( b_i \) denote the bias of output unit, and \( y_i \) is the outputs of the last encoding layer which are also the deep spectral features \( \hat{Y} = f(w_y y_i + b_i) \) pre-trained by the SAE. We can use the error between the predicted firmness (or SSC) results and the reference measured firmness (or SSC) to fine-tune the whole SAE-FNN model in a supervised manner. The error function for fine-tuning the SAE-FNN is defined as the following mean squared error (MSE) function:

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{Y} - Y_i)^2,
\]

where \( n \) is the total number of training pear fruit sample, \( Y_i \) is the reference measured firmness (or SSC) of sample \( i \), and \( \hat{Y} \) is the corresponding predicted firmness or SSC value. Here, we used an “adam” optimizer to solve the optimization problem of minimizing the MSE function (Gopalarakrishnan et al., 2017).

The key steps of the developed SAE-FNN model can be described as two stages that mainly including (1) SAE pre-training: a SAE is used to extract deep spectral features from hyperspectral image, and (2) FNN fine-tuning: a FNN layer is added to the last encoding layer of the pre-trained SAE, so that, a SAE-FNN regression model is established based on the deep spectral features. SAE pre-trained the network weights by using large number of randomly selected pixel spectra as input data, and the learned weights (deep spectral features) acted as the initial weights for the SAE-FNN network. For the purpose of predicting the firmness and SSC of pear fruit, the whole SAE-FNN network weights were finally fine-tuned by using 135 samples (in the calibration set) with mean spectra and their corresponding firmness (or SSC) values (illustrated in Fig. 3c).

Our SAE-FNN algorithm was carried out on a windows 7 system, which has an Intel I7-7700K processor. A python language-based Keras framework (Chollet, 2015) was used for the deep learning of SAE-FNN algorithm. A cross_validation function in the machine learning framework of Scikit-learn (Pedregosa et al., 2011) was used to perform cross-validation in SAE-FNN models. In the experiments, a GPU of NVIDIA GT750M was also used to make the learning procedure faster.
2.8. Model evaluation

During the modeling, the samples in calibration set were used to establish calibration models based upon PLSR, LS-SVM and SAE-FNN. The established calibration models were then used to predict the values of firmness and SSC in the prediction set. The performances of these models were compared and evaluated according to their capability of prediction. To evaluate the capability of prediction of the models, some parameters were calculated, including coefficients of determination by calibration set ($R^2_C$) and prediction set ($R^2_P$), root mean square errors in calibration set (RMSEC) and prediction set (RMSEP), and ratio of prediction to deviation of calibration set (RPDC) and prediction set (RPDP) that were calculated by dividing the standard deviation (SD) of the reference measured values by RMSEC and RMSEP, respectively. Generally, a good model should have high values of $R^2_C$, $R^2_P$, RPDC and RPDP, while low values of RMSEC and RMSEP. Specifically, the value of $R^2$ in the range of 0.82–0.90 and the RPDP value higher than 2.0 indicate a good performance of the calibration model, while the value of $R^2$ above 0.90 and the RPDP value larger than 3.0 are considered sufficient for a particular prediction purpose (Viscarra Rossel et al., 2006).

3. Results and discussion

3.1. Reference measurement results of firmness and SSC

Reference measurements of firmness and SSC in Korla fragrant pear fruit were performed according to the destructive methods described in “Reference measurements of firmness and SSC” section. Minimum, mean, maximum and standard deviation values of firmness and SSC of the fruit are analyzed and presented in Table 1. For all 180 measured Korla fragrant pear samples, the firmness changed from 87.2 N to 60.8 N with a mean value of 72.3 N and a standard deviation of 5.35 N; the SSC changed from 15.5% to 12.2% with a mean value of 13.9% and a standard deviation of 0.74%. This relatively high variability appeared in firmness and SSC is because the pear fruit were harvested at different trees and different canopy layer positions, and stored for various periods. The mean values (averaged from 15 fruit samples for every other day) of firmness and SSC of the fruit samples stored during a twenty days postharvest period with standard deviation bars are illustrated in Fig. 4a and b respectively. A decreasing tendency for both firmness and SSC in samples was appeared throughout the storage of 23 d. The mean value of firmness of Korla fragrant pear changed from about 81.5 N to 64.3 N, and the SSC changed from about 15.0% to 12.9%. The changing tendency of the firmness and SSC in this study appears to be similar to that reported by the authors of Li et al. (2010). Their results showed that the firmness of Korla fragrant pear changed from about 82.3 N to 64.7 N, and the SSC changed from about 15.6% to 13.5%, during a twenty seven-days postharvest storage at refrigerator room with 20 °C, relative humidity 90–95%.

For the calibration modeling purpose, all pear samples were randomly divided into calibration set of 135 samples and prediction set of 45 samples, and the statistical results of firmness and SSC were displayed in the last two rows in Table 1. The firmness ranged from 60.8 N to 85.8 N for the calibration set and 61.8 N to 87.2 N for the prediction set, respectively; while the corresponding SSC ranged from 15.3% to 12.4% and from 15.5% to 12.2%. During modeling, the relatively wide range of data variability in the reference measurement of firmness and SSC in both calibration and prediction sets might help to the establishing of robust calibration models.

3.2. Spectral characteristics of pears

The Vis/NIR mean spectra averaged from all pixels in ROI of two Korla fragrant pears with firmness value of 81.0 N, SSC value of 14.4% (sample A), and firmness value of 66.4 N, SSC value of 13.4% (sample B) were shown in Fig. 5. Both spectra have similar tendency displayed

![Fig. 4. Mean values with indicated standard deviations of firmness (a) and SSC (b) of Korla fragrant pear on different storage days.](image)

### Table 1
Reference measurement results of firmness and SSC.

| Sample seta | Number of samples | Firmness (N) | | | | SSC (%) | | |
|-------------|-------------------|--------------|---|---|---|-------------|---|
|             |                   | Range        | Meanb | SD  | Range        | Mean | SD   |
| Sep 8       | 15                | 77.4–78.2    | 81.5 | ± 2.91 | 14.4–15.5   | 15.0 | ± 0.29  |
| Sep 10      | 15                | 74.2–81.2    | 77.9 | ± 2.05 | 14.1–15.1   | 14.6 | ± 0.30  |
| Sep 12      | 15                | 71.0–78.0    | 75.4 | ± 2.02 | 13.5–15.0   | 14.4 | ± 0.37  |
| Sep 14      | 15                | 70.5–78.2    | 74.8 | ± 2.18 | 13.7–14.8   | 14.3 | ± 0.36  |
| Sep 16      | 15                | 70.3–76.7    | 74.5 | ± 1.73 | 13.6–14.7   | 14.3 | ± 0.39  |
| Sep 18      | 15                | 68.2–75.6    | 72.6 | ± 2.04 | 13.1–14.7   | 14.2 | ± 0.41  |
| Sep 20      | 15                | 70.5–75.6    | 73.3 | ± 1.70 | 13.2–14.6   | 13.9 | ± 0.45  |
| Sep 22      | 15                | 64.2–75.0    | 71.1 | ± 2.55 | 12.9–14.3   | 13.9 | ± 0.37  |
| Sep 24      | 15                | 67.0–72.6    | 69.5 | ± 1.35 | 12.7–14.2   | 13.6 | ± 0.40  |
| Sep 26      | 15                | 61.8–72.4    | 67.9 | ± 2.93 | 12.4–13.9   | 13.3 | ± 0.47  |
| Sep 28      | 15                | 61.1–71.9    | 65.0 | ± 2.46 | 12.2–13.6   | 13.0 | ± 0.42  |
| Sep 30      | 15                | 60.8–67.7    | 64.3 | ± 1.89 | 12.4–13.4   | 12.9 | ± 0.29  |
| All samples | 180               | 60.8–87.2    | 72.3 | ± 3.55 | 12.2–15.5   | 13.9 | ± 0.74  |
| Calibration set | 135     | 60.8–85.8    | 72.3 | ± 5.30 | 12.4–15.3   | 14.0 | ± 0.71  |
| Prediction set | 45            | 61.8–87.2    | 72.5 | ± 5.52 | 12.2–15.5   | 13.9 | ± 0.81  |

*a 15 pear samples were examined for every other day. 135 samples were randomly selected out to form the calibration set. The remaining 45 samples formed the prediction set.  
b Firmness and SSC of sample set are expressed in units of N and % respectively, and given as Mean ± Standard Deviation (SD).
throughout the Vis/NIR spectral region (380–1030 nm). The reflectance of sample A was lower than that of sample B in the Vis region (380–750 nm), but was higher than sample B in the NIR region (750–1030 nm). According to the Vis wavelength region curves displayed in Fig. 5, the first absorption bands presented high absorptions on the green (500–550 nm) and red (625–740 nm) spectral regions. The Vis wavelength region of the spectra was dominated by a large number of fruit showing high absorptions on green (500–550 nm) spectral regions (dos Santos Neto et al., 2017). The region of 625–740 nm with rapid changed reflectance is well known as the “red edge” in the electromagnetic spectrum, which can be used as a potential index for detecting chlorophyll in fruit with different ripeness (Zude-Sasse et al., 2002). Meanwhile, the trend of the green and red spectral regions can also reflect the Korla fragrant pear color which is characterized by a green and yellow color with a slight red blush, but might also vary from green to yellow depending on postharvest storage time (Jia et al., 2016). For the NIR wavelength region, some small weak peaks observed in the range of 750–900 nm are mostly attributed to the third overtone stretch of O–H functional groups related to water in pear fruit (Martinsen and Schaeare, 1998). An absorptions region is observed in the range of 900–980 nm, which correspond to the third overtones of C–H functional groups (910 nm) related to sugar, and the second O–H overtone (960 nm) related to moisture in fruit (Guthrie et al., 2005). The variation of spectral reflectance in Vis and NIR regions might have the potential to discriminate physiochemical properties between Korla fragrant pear samples with different firmness and SSC.

### 3.3. PLSR calibration models of firmness and SSC

PLSR calibration models for the prediction of firmness and SSC were developed using the original mean spectra (without preprocessing) in the 380–1030 nm wavelength range with 512 bands. Before the development of calibration models, the number of LVs was optimized under full cross validation by PLSR models with 6 for firmness and 7 for SSC. No outliers were found in the calibration set during the development of models. The results of the optimal PLSR models for firmness and SSC are presented in Tables 2 and 3, respectively. The PLSR models obtained the values of $R^2$, RMSEP and RPD were 0.842, 2.17 N and 2.54 for firmness, and $R^2 = 0.832$, RMSEP = 0.33% and RPD = 2.45 for SSC. According to the standard for evaluating the calibration models, the PLSR models yielded good prediction performance with $R^2 > 0.82$ and RPD > 2.0 for both firmness and SSC in Korla fragrant pear, showing that the models can be used for practical applications. The mean spectra preprocessed by MSC method were also used to develop PLSR calibration models for the prediction of firmness and SSC. The number of LVs was optimized with 11 for both firmness and SSC. As shown in Tables 2 and 3, the PLSR models based on the MSC-preprocessed mean spectra achieved the values of $R^2$, RMSEP and RPD were 0.792, 2.49 N and 2.13 for firmness, and $R^2 = 0.756$, RMSEP = 0.39% and RPD = 2.08 for SSC. The results indicated that the performances of the PLSR models with the MSC-preprocessed mean spectra were even worse than the PLSR models using original mean spectra.

In order to identify the effective wavebands for firmness (or SSC), SPA was further carried out on the MSC-preprocessed mean spectra of 135 samples in calibration set. The optimal variable number of wavebands was selected by comparing the RMSEC values of different variable numbers from 3 to 20. As a result, twelve wavebands (as displayed in Fig. 6a) were chosen as effective wavebands with RMSEC = 2.10 N for the firmness, which were centered at around 380, 403, 451, 453, 517, 526, 675, 692, 712, 749, 896 and 982 nm. Fig. 6b showed that fourteen wavebands including 380, 512, 577, 629, 666, 674, 687, 692, 695, 723, 737, 830, 838 and 963 nm were most effective for SSC with RMSEC = 0.32%. The above wavebands are well consistent with the spectral characteristics of Korla fragrant pear that have been analyzed in the “Spectral characteristics of pears” section. Therefore, according to the effective wavelengths analysis, it can be safely confirmed that hyperspectral image on the spectral range (380–1030 nm) is feasible to rapidly determine firmness and SSC of Korla fragrant pear.

After effective wavelengths selection by SPA, the selected wavebands in calibration and prediction set were then respectively used to establish PLSR calibration models as well as evaluate the performance of the models. The number of LVs was optimized under full cross validation by PLSR models with 11 for both firmness and SSC. The prediction results of PLSR calibration models for firmness and SSC were listed in Tables 2 and 3. In Table 2, when twelve SPA selected wavebands were used to establish PLSR model for firmness, the results of $R^2$, RMSEP and RPD were 0.796, 2.47 N and 2.23. Comparing with the performance of PLSR model using MSC-preprocessed mean spectra, the prediction performance of PLSR model was slightly improved by applying SPA selected wavebands, but the difference between the prediction results obtained by MSC-preprocessed mean spectra and SPA selected wavebands was not so significant. In Table 3, when using fourteen SPA selected wavebands to predict SSC, the PLSR model obtained a prediction result with $R^2 = 0.684$, RMSEP = 0.45% and RPD = 1.80. This result was worse than the result obtained by MSC-preprocessed mean spectra.

From the above results, when PLSR was used to establish calibration models, the best performance for firmness and SSC prediction was obtained by using non-preprocessed original mean spectra as input data. Previous research results indicated that PLSR model coupled with Vis/NIR spectra achieved $R^2 = 0.916$ for predicting SSC in ‘Cuiguian’ pear (Sun et al., 2009) and $R^2 = 0.912$ for predicting firmness in “Fengshui” pear (Liu et al., 2008). In our work, the best PLSR model yielded prediction performance with $R^2 = 0.842$ for firmness and $R^2 = 0.832$ for SSC, which was worse than the former prediction result of “Cuiguian” pear and “Fengshui” pear. Although the difference between the prediction results of Korla fragrant pear, “Cuiguian” pear and “Fengshui” pear might be caused by some uncertain factors, such as the variety of fruit to be detected, training dataset and parameters in the PLSR model, we believe that there is still highly necessary to further investigate if it is possible to increase prediction accuracy by applying the other calibration models such as LS-SVM and SAE-FNN.

### 3.4. LS-SVM calibration models of firmness and SSC

LS-SVM calibration models were established for the prediction of firmness and SSC by using original mean spectra as input data, and their performance was compared with that of PLSR. When LS-SVM was used to establish calibration models for the firmness, the optimal parameters of $C$ and $\gamma$ were achieved with $C = 6.760 \times 10^3$ and $\gamma = 4.188 \times 10^3$.
in LS-SVM model with RBF kernel (LS-SVM(R)), and $\gamma = 0.156$ in LS-SVM model with linear kernel (LS-SVM(L)); meanwhile, the optimal parameters of $\gamma$ and $\sigma^2$ were achieved with $\gamma = 6.515 \times 10^3$ and $\sigma^2 = 5.683 \times 10^2$ in the LS-SVM(R) model for SSC, and $\gamma = 0.142$ in the LS-SVM(L) model for SSC. The results of the optimal LS-SVM models for firmness and SSC are presented in Tables 2 and 3, respectively. From the results, we find that although all LS-SVM models with original mean spectra superior to the PLSR models in calibration set for both firmness and SSC, the accuracy produced by the LS-SVM models for unknown samples in prediction set is lower than the PLSR models with original mean spectra. Here, the result was achieved by LS-SVM(R) with $R^2 = 0.939$, $\text{RMSEP} = 2.11$ N and $\text{RPDP} = 2.45$ for the SSC prediction, and the LS-SVM(L) obtained a result with $R^2 = 0.938$, $\text{RMSEP} = 2.46$ N and $\text{RPDP} = 2.24$ for the firmness prediction, and $R^2 = 0.784$, $\text{RMSEP} = 0.37\%$ and $\text{RPDP} = 2.19$ for the SSC prediction. These results were worse than the result obtained by the PLSR models with original mean spectra. The worse prediction performance of LS-SVM models with original mean spectra can also be observed in the plots of measured vs. predicted firmness and SSC for 45 samples in prediction set that are displayed in Fig. 7. By comparing the compactness of samples around the regression line, LS-SVM models with original mean spectra (Fig. 7b,c,f and g) are found to be with slightly lower correlation coefficient of measured values and predicted values than that of PLSR models with original mean spectra (Fig. 7a and e).

The mean spectra preprocessed by MSC were further applied to establish LS-SVM calibration models for the prediction of firmness and SSC. The LS-SVM(R) models for firmness and SSC were optimized with $\gamma = 2.989 \times 10^3$ and $\sigma^2 = 9.053 \times 10^2$, and $\gamma = 1.546 \times 10^3$ and $\sigma^2 = 5.269 \times 10^3$, respectively. The LS-SVM(L) models for firmness and SSC are presented in Tables 2 and 3, respectively. From the results, we find that although all LS-SVM models with original mean spectra superior to the PLSR models in calibration set for both firmness and SSC, the accuracy produced by the LS-SVM models for unknown samples in prediction set is lower than the PLSR models with original mean spectra. Here, the result was achieved by LS-SVM(R) with $R^2 = 0.841$, $\text{RMSEP} = 2.11$ N and $\text{RPDP} = 2.24$ for the SSC prediction, and the LS-SVM(L) obtained a result with $R^2 = 0.914$, $\text{RMSEP} = 2.38$ N and $\text{RPDP} = 2.32$ for the firmness prediction, and $R^2 = 0.795$, $\text{RMSEP} = 2.47$ N and $\text{RPDP} = 2.23$ for the firmness prediction, and $R^2 = 0.935$, $\text{RMSEP} = 2.17$ N and $\text{RPDP} = 2.17$ for the SSC prediction. These results were worse than the result obtained by the PLSR models with original mean spectra. The worse prediction performance of LS-SVM models with original mean spectra can also be observed in the plots of measured vs. predicted firmness and SSC for 45 samples in prediction set that are displayed in Fig. 7. By comparing the compactness of samples around the regression line, LS-SVM models with original mean spectra (Fig. 7b,c,f and g) are found to be with slightly lower correlation coefficient of measured values and predicted values than that of PLSR models with original mean spectra (Fig. 7a and e).
SSC were optimized with $\gamma = 1.101$ and $\gamma = 0.925$, respectively. Comparing to the LS-SVM models with original mean spectra, the performances of both LS-SVM(R) and LS-SVM(L) models were improved by using MSC spectra. The LS-SVM(R) model with MSC spectra was considered as the better model for firmness and SSC prediction. In Table 2, the LS-SVM(R) model with MSC spectra yielded a relatively high accuracy for firmness with $R^2_p = 0.841$, RMSEP = 2.17 N and RPDp = 2.54; while in Table 3, the result for SSC was achieved with $R^2_p = 0.890$, RMSEP = 0.33% and RPDp = 2.45. However, these results were still worse than that of PLSR models with original spectra. From the above results, it can be safely confirmed that the MSC spectral preprocessing can improve the performance of the prediction of firmness and SSC in LS-SVM models. However, the best performance achieved by LS-SVM(R) model with MSC spectra is not satisfied, since it is still worse than the result produced by the PLSR model with original spectra. Recent studies have reported that deep learning method can improve the generalization performance in many applications such as wind power prediction (Qureshi et al., 2017) and passenger flow prediction (Liu and Chen, 2017). In the next section, deep learning method was conducted to establish calibration models to investigate if it was possible to further increase prediction accuracy for firmness and SSC.

3.5. SAE-FNN calibration models of firmness and SSC

The SAE-FNN modeling was divided into two stages including feature extraction and regression. Additionally, as the SAE was always used as a powerful tool for noise reduction, we only used original pixel and mean spectra without any preprocessing to train the SAE-FNN models. In the feature extraction stage, 1000 randomly selected pixel spectra of each pear sample in the calibration set (a total of 135,000 spectra) were considered as a big dataset to pre-train the SAE in an unsupervised way. After pre-training, the weights of the encoding part
of the pre-trained SAE acted as the initial weights of the SAE-FNN. By this way, the pre-trained encoding part of SAE in the SAE-FNN can be used to extract deep spectral features from original input spectral data. In the regression stage, 135 mean spectrum samples with measured firmness (or SSC) in the calibration set were used as original input spectral data to fine-tune the whole SAE-FNN model in a supervised manner (the original input mean spectral data were firstly reduced to deep spectral features by the encoding part of the SAE, and then the features were used as initial inputs to build regression model for firmness and SSC by the SAE-FNN), while remain 45 mean spectrum samples with measured firmness (or SSC) in the prediction set were used as unknown samples to evaluate the performance of the fine-tuned SAE-FNN calibration model.

The numbers of units in different layers of the SAE were set as 512, 300, 100, h, 100, 300 and 512, where h represented the units of the last layer (deep spectral features) in the encoding part of SAE. An example of the pre-training of SAE is displayed in Fig. 8. After the SAE is trained with 100 iterating epochs, the average validation error between the original input mean spectra and the reconstructed output mean spectra tends to be a small value of $5.848 \times 10^{-5}$ (Fig. 8a), showing that the SAE can well reproduce the original input spectra (Fig. 8b) at reconstructed spectra (Fig. 8f). The spectral features of the spectra are calculated from first (300 units), second (100 units) and last (10 units) layers in the encoding part of SAE and illustrated as curves in Fig. 8c–e respectively. Specifically, when the first layer of the SAE was used to extract spectral features, the original input mean spectra with 512 bands were transferred (encoded) to an output with 300 variables using Eq. (2); when the second layer of the SAE was used to extract spectral features, the 300 variables encoded by the first layer were used as inputs for the second layer, and then further transferred to an output with 100 variables; when the last layer of the SAE was used to extract spectral features, the 100 variables encoded by the second layer were transferred to an output with 10 variables. It can be observed that the spectral features become more abstract and differentiable, as the number of units in the encoding layers reducing from 300 to 10. According to the principle of deep learning of SAE, the deep spectral features in the last layer carry the most valuable information representing the original input spectra. After the unsupervised pre-training of SAE, the original mean spectrum samples in the calibration set and prediction set were used for supervised fine-tuning and evaluating the SAE-FNN calibration model, respectively. In order to find the optimal deep spectral features for the firmness and SSC prediction, we pre-trained a series of SAE by varying the unit number of the last encoding layer (deep spectral features) chosen from 10, 20, 30, 40, 50, 60 and 70. Here, the SAE was pre-trained with a batch size of 100 and an epoch value of 100, and the SAE-FNN model was fine-tuned with a batch size of 100 and an epoch value of 10,000. Finally, seven SAE-FNN calibration models with different deep spectral features, which denoted as SAE-FNN(10), SAE-FNN(20), SAE-FNN(30), SAE-FNN(40), SAE-FNN(50), SAE-FNN(60) and SAE-FNN(70) respectively, were established to predict the firmness and SSC of Korla fragrant pear.

The firmness prediction performances of the established SAE-FNN models are listed in Table 2. From Table 2, it can be observed that all seven SAE-FNN models yielded reasonable results with $R^2 > 0.82$ and RPD$_P > 2.0$, further proving that the Vis/NIR hyperspectral imaging technology was suitable for predicting firmness in Korla fragrant pear fruit. Moreover, all of the SAE-FNN models achieved better result than that yielded by PLSR and LS-SVM models. The best result for firmness was obtained by the SAE-FNN(50) model with $R^2 = 0.931$, RMSEC = 1.39 N, RPD$_C = 3.81$ for calibration set; and $R^2 = 0.890$, RMSEP = 1.81 N, RPD$_P = 3.05$ for prediction set. On the other hand, the performances of the SAE-FNN models for the SSC of Korla fragrant pear are listed in Table 3. Here, most of the SAE-FNN models achieved a better result than that obtained by the PLSR and LS-SVM models, except for SAE-FNN(60) and SAE-FNN(70). The best prediction of SSC was obtained by the SAE-FNN(40) model, with $R^2 = 0.936$, RMSEC = 0.18%, RPD$_C = 3.94$ for calibration set; and $R^2 = 0.921$, RMSEP = 0.22%, RPD$_P = 3.68$ for prediction set. The superiority of the SAE-FNN models for the prediction of firmness and SSC in Korla fragrant pear can also be revealed by the high correlation coefficient of measured values and predicted values produced by the SAE-FNN(50) and SAE-FNN(40) models displayed in Fig. 7d and h, respectively.

According to the above results obtained by the SAE-FNN models, it was safely confirmed that deep learning method was suitable for extracting informative features from HSI data and predicting firmness and SSC of Korla fragrant pear. The reasonable performance yielded by the SAE-FNN model could be explained by the following reasons. Firstly, the collected hyperspectral data of the pear fruit is very complicated and tends to show nonlinear data, owning to various undesired effects.
caused by multi-scattering in the acquisition process, inhomogeneous samples, environmental changes, and instrumental variations (Chen et al., 2013). In this study, the SAE acted as a highly efficient non-linear data reduction method which can extract non-linear deep spectral features from the hyperspectral image data. This might result in the improved efficacy of spectral feature extraction and prediction accuracy for firmness and SSC of Korla fragrant pear. Secondly, large amounts of data trained in deep learning model were proved to be an effective way for avoiding over-fitting as well as improving generalization performance in both regression and classification task (Donahue et al., 2014; Yu et al., 2017). Our deep-learning-based SAE-FNN model was developed by pre-training SAE on a very large number of spectra of pear fruit, which is help to extract informative spectral features from the hyperspectral data. Hence, the performances of SAE-FNN models for the firmness and SSC prediction were higher than those of PLSR and LSSVM models.

4. Conclusions

This work demonstrates that it is possible to apply deep learning method together with Vis/NIR hyperspectral imaging technique for predicting postharvest firmness and SSC of Korla fragrant pear. The results obtained by the SAE-FNN would encourage more research efforts on using deep learning as a novel chemometrical method for postharvest quality detection of fruit. Considering the line scanning speed of 1.5 mm/s used in the Vis/NIR hyperspectral imaging system was not fast enough for real-time application, in the future research, hyperspectral images of pear fruit acquired through a higher scanning speed should also be investigated for the establishing of SAE-FNN models for firmness and SSC prediction.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (no. 31402352), the Ningbo Science and Technology Special Project of China (no. 2017C110002), and the Ningbo Natural Science Foundation (no. 2015A610131).

References


Yu et al. 2017). Our deep-learning-based SAE-FNN model was developed by pre-training SAE on a very large number of spectra of pear fruit, which is help to extract informative spectral features from the hyperspectral data. Hence, the performances of SAE-FNN models for the firmness and SSC prediction were higher than those of PLSR and LSSVM models.


学霸图书馆
www.xuebalib.com

本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，
提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。
图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

图书馆导航：
图书馆首页 文献云下载 图书馆入口 外文数据库大全 疑难文献辅助工具