



OPEN ACCESS

EDITED BY

John-Lewis Zinia Zaukuu,
Kwame Nkrumah University of Science
and Technology, Ghana

REVIEWED BY

Guantao Xuan,
Shandong Agricultural University,
China
Sicong Liu,
Tongji University, China

*CORRESPONDENCE

Long Xue
ultimata@163.com

SPECIALTY SECTION

This article was submitted to
Nutrition and Food Science
Technology,
a section of the journal
Frontiers in Nutrition

RECEIVED 14 July 2022

ACCEPTED 28 September 2022

PUBLISHED 19 October 2022

CITATION

Li J, He L, Liu M, Chen J and Xue L
(2022) Hyperspectral dimension
reduction and navel orange surface
disease defect classification using
independent component
analysis-genetic algorithm.
Front. Nutr. 9:993737.
doi: 10.3389/fnut.2022.993737

COPYRIGHT

© 2022 Li, He, Liu, Chen and Xue. This
is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided
the original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Hyperspectral dimension reduction and navel orange surface disease defect classification using independent component analysis-genetic algorithm

Jing Li¹, Liang He¹, Muhua Liu¹, Jinyin Chen² and Long Xue^{1,3*}

¹Jiangxi Key Laboratory of Modern Agricultural Equipment, College of Engineering, Jiangxi Agricultural University, Nanchang, China, ²Collaborative Innovation Center of Postharvest Key Technology and Quality Safety of Fruits and Vegetables in Jiangxi Province, Nanchang, China, ³Key Laboratory of Optics-Electrics Application of Biomaterials of Jiangxi Province, Jiangxi Agricultural University, Nanchang, China

Canker is a common disease of navel oranges that is visible before harvest, and penicilliosis is a common disease occurring after harvest and storage. In this research, the typical fruit surface, canker spots, penicillium spore, and hypha of navel oranges were, respectively, identified by hyperspectral imaging. First, the light intensity on the edge of samples in hyperspectral images was improved by spherical correction. Then, independent component images and weight coefficients were obtained using independent component analysis. This approach, combined with use of a genetic algorithm, was used to select six characteristic wavelengths. The method achieved dimension reduction of hyperspectral data, and the testing time was reduced from 46.21 to 1.26 s for a self-developed online detection system. Finally, a deep learning neural network model was established, and the four kinds of surface pixels were identified accurately.

KEYWORDS

canker, penicilliosis, navel orange, hyperspectral image, ICA-GA

Introduction

Navel orange is an important fruit crop grown in Jiangxi Province, China and many other agricultural areas in subtropical regions. The fruit is generally round, oblate, or elliptic in shape and usually orange-yellow or orange-red in color (1). The calyx of the fruit has a few immature pericardial groups, which form the characteristic navel, and

the core is solid or semi-solid. In the process of transportation and packaging of navel oranges after harvest, a spore from a single navel orange can cause an entire batch to start spoiling within a short time (2). Therefore, the post-harvest sorting of navel orange is conducive to improving their value and increasing the preservation period, and technology to accomplish this objective is urgently needed for fruit production (3).

In present-day fruit sorting operations, machine vision technology is mainly used to assess fruit appearance quality, including fruit shape and visible defects, for example, brown spots, spores, granulation, edema, etc. However, for some defects, such as decay caused by fungi, machine vision methods seem ineffective. The main reason is that the skin color of a decayed area can be close to that of the normal skin areas, so it is hard to distinguished using images from a color camera. However, the skin damage is related to the characteristics of the fruit skin that can be detected by near infrared spectroscopy (NIRS). Moreover, NIRS can also be used for non-destructive testing of the internal components of fruits, such as soluble solids (4), acidity, and other traits (5, 6). Although NIRS technology can reflect the internal quality of fruit, only a small part of the fruit surface can be collected, and the fruit appearance quality cannot be perfectly determined.

Hyperspectral image technology combines the advantages of spectrum- and image-based approaches (7). It not only can achieve fruit appearance image inspection, but can also collect the surface spectra which can be used to detect the invisible defect (8, 9). The hyperspectral image is a three-dimensional image. Through the selection of characteristic wavelengths of hyperspectral images, simultaneous external and internal fruit quality testing can be realized; this is a major development trend in the field of fruit detection (10). Hyperspectral image technology can be used to detect meat quality (11), orange spores, cucumber frostbite (12), guava maturity (13), strawberry ripeness (14), Moisture of Okra (15), and many other food defects (16–18). However, hyperspectral images contain large amounts of wavelength information, thus requiring a long time to collect information; accordingly, its online detection applications are limited. Therefore, the study of hyperspectral data dimension reduction technology has become the key to improving the speed of online detection. Current dimension reduction methods mainly include genetic algorithm (19), principal component analysis (20, 21), independent component analysis (22), and some deep learning (23) methods. However, most approaches have only been applied to hyperspectral test platforms, and online detection has not yet been realized.

In this study, the hyperspectral images of navel orange samples were dynamically collected using a self-designed online detection system. The spectra ranged from 975.18 to 2,196.2 nm.

The method employed combined independent component analysis (ICA) with genetic algorithm (GA) to select the characteristic spectra. The dimensions of the hyperspectral images were reduced, and then the cankered and spored fruit were separated. This study had the following specific objectives: (1) spectral preprocessing of the collected hyperspectral images; (2) use of an ICA method to obtain independent component images and weight coefficients; (3) combining these data with GA after characteristic selection of wavelengths to achieve dimensionality reduction of hyperspectral data; (4) establishment of a deep neural network method of classified of navel orange surface defects and subsequent automated classification of fruits with a normal versus defective surface. Six characteristic wavelengths were selected by ICA-GA, and the surface defect detection model of navel orange was established based on LSTM. Finally, the online detection of navel orange was achieved.

Materials and methods

Experimental materials

The navel orange samples used in this experiment were collected in November 2019 from Ganzhou City, Jiangxi Province, China. After harvest, the navel oranges were first cleaned in the laboratory. Thirty normal navel orange samples without defects (group 1) and thirty navel oranges with canker defects on their surfaces (group 2) were selected. Then, the hyperspectral image of the samples in categories 1 and 2 were acquired. Meantime, one navel orange inoculated with penicilliosis was stored with some normal navel oranges in the laboratory. After 30 days, 30 infected navel orange were selected as samples with penicilliosis on their surfaces (group 3), and the hyperspectral image data of them were also acquired. Images of samples from each category are shown in **Figures 1A–C**, respectively.

In order to distinguish navel oranges with surface defects online, it is necessary to identify the above defects on surfaces of navel orange samples. In this study, navel orange surfaces were mainly divided into four categories: normal navel orange surface without defects, cankered surface, penicillium spore surface, and hypha surface. **Figure 1A** shows a navel orange without defects on its surface, and the category is defined as Surface. Since normal navel oranges have a large surface range, only rectangular areas A and B were selected. **Figure 1B** shows the surface of a navel orange with cankers, as indicated by the arrow in area C. This type of sample is referred to as Canker. **Figure 1C** is the image of a penicilliosis navel orange sample. Areas D and E were, respectively, the surface of penicillium spores and penicillium hypha, which are referred to as Spore and Hypha, respectively.

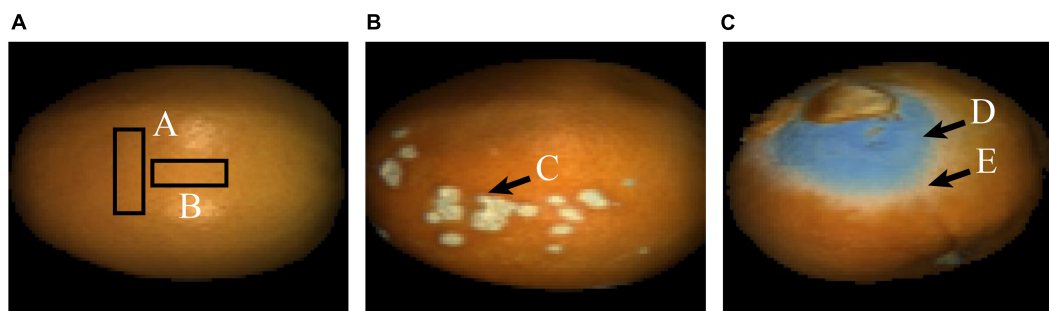


FIGURE 1 False-color images of the typical navel orange. (A) Navel orange sample with an intact surface; (B) navel orange sample with canker; (C) navel orange sample with penicilliosis.

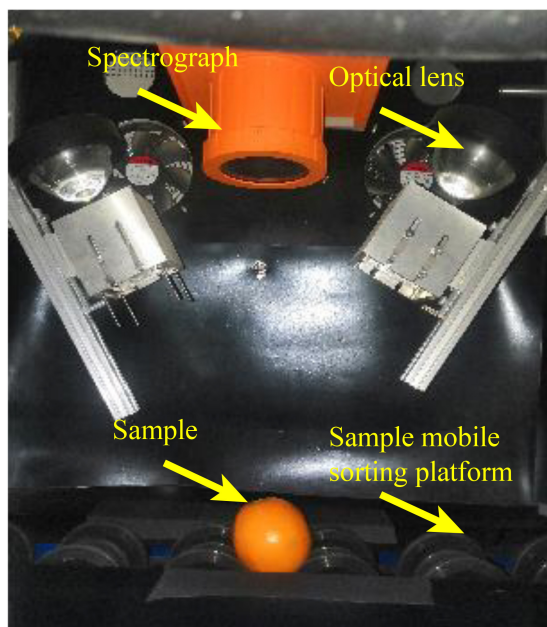


FIGURE 2 Hyperspectral imaging sorting system.

Hyperspectral imaging sorting system

The hyperspectral image system mainly includes a hyperspectral camera (SWIR-CL-400-N25E, SPECIM, Finland), two 150-Watt halogen lamps, two focusing lenses, and a mobile sorting platform setup, as shown in **Figure 2**. Before collecting hyperspectral images, the lamp was turned on for about 15 min, and the white and black calibration images of the hyperspectral camera were captured. The white calibration image was acquired from a 99.9% reflectance white board (Spectralon SRT-99-100, Labsphere Inc., North Sutton, NH, USA) 30 cm below the camera. The black calibration image was acquired when the lens was completely shielded.

The hyperspectral images of samples were corrected using Equation (1):

$$I = \frac{I_{data} - I_{black}}{I_{white} - I_{black}} \quad (1)$$

Here, I_{data} represents the acquired hyperspectral image, I_{black} represents the black calibration image, I_{white} represents the white calibration image, and I represents the corrected image. The exposure time and the frame rate of the hyperspectral camera were 1.8 ms and 120 f/s, respectively.

Spectral preprocessing

The average spectrum of rectangular areas A (in **Figure 1**) corrected by the black and white calibration images is shown in **Figure 3**. It can be seen that there is considerable noise at both ends of the spectrum, at ranges of 946.43–975.18 nm and 2,196.2–2,256.82 nm, respectively. Therefore, in the subsequent processing, these two parts of spectra were eliminated, and the spectral range within 975.18–2,196.2 nm between the two black rectangular boxes in **Figure 3** was used. The red asterisk in **Figure 3** represents the three wave peaks of the surface spectral curve of navel oranges, which are respectively, 1,078.31, 1,266.07, and 1,655.72 nm. The hyperspectral images corresponding to these three bands were selected to produce false color images of navel oranges, like the one shown in **Figure 1**.

As navel oranges are spherical fruits, the reflected light on the upper surface of the navel oranges collected by the hyperspectral camera is the strongest. In contrast, the reflected light at the edges of the navel oranges is very weak. In order to improve the reflected light intensity on the surface of the spherical fruits and increase the light intensity at the edges, the hyperspectral images of each spherical fruit were corrected (24). This is achieved through the following correction process: (1) the binary-value image of the sample is built using a 1,078.31 nm wavelength; (2) the center of gravity of the binarization region is found, and the image is then segmented into 16 equal parts; (3)

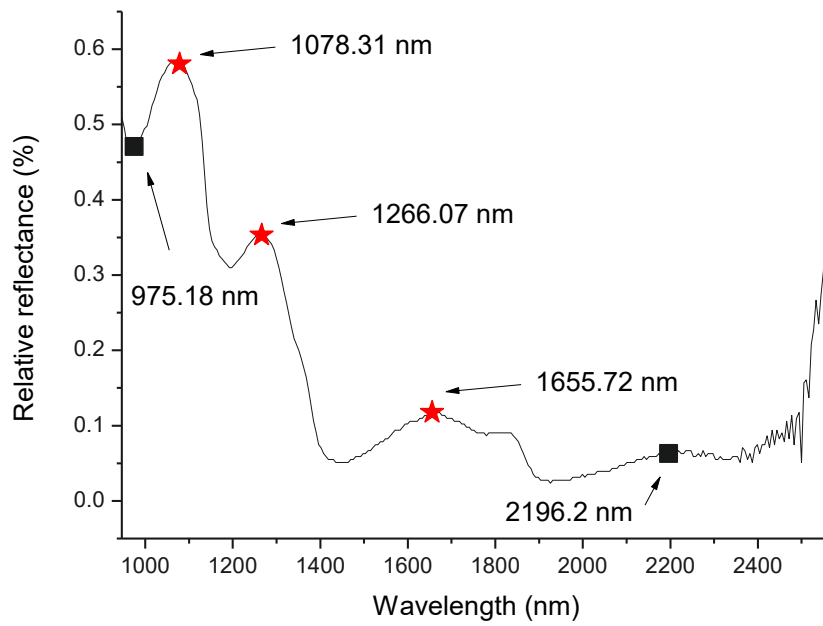


FIGURE 3
Reflection spectrum of a normal navel orange surface.

TABLE 1 Spectra number of each category.

Category	Spectra number	Selected spectra number	Spectra number of training dataset	Spectra number of validation dataset	Spectra number of test dataset
Spore	1,954	1,300	1,065	123	112
Hypha	1,374	1,300	1,044	135	121
Canker	3,302	1,300	1,016	161	123
Surface	3,775	1,300	1,035	121	144
Total	10,405	5,200	4,160	540	500

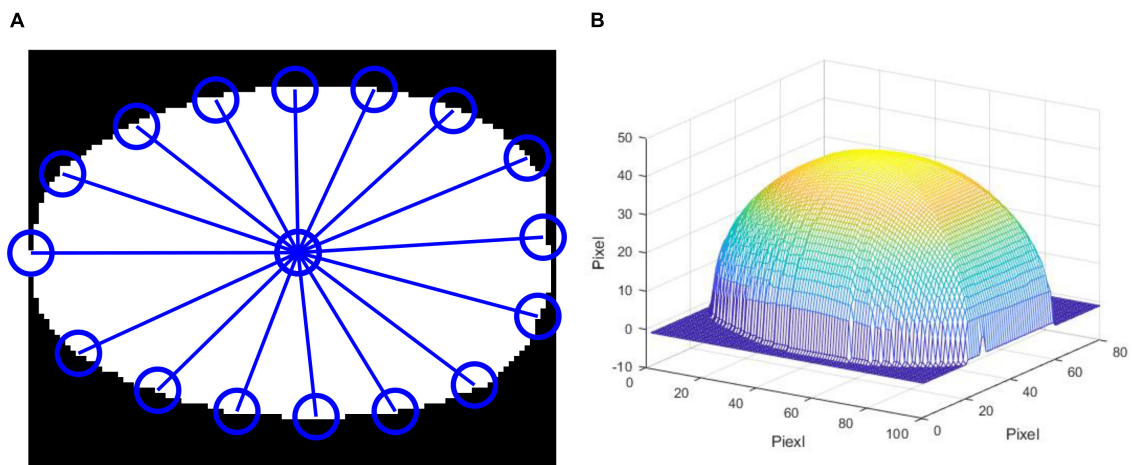


FIGURE 4
Correction of the spherical fruit. (A) segmentation of the mask image, (B) 3D reconstruction of spherical fruit.



FIGURE 5

False-color images after spherical correction. (A) Navel orange sample with an intact surface; (B) navel orange sample with cankers; (C) navel orange sample with penicilliosis.

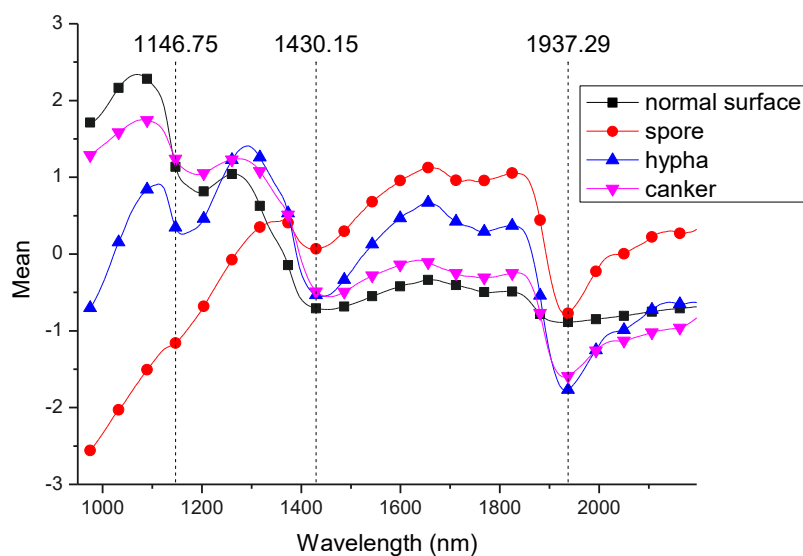


FIGURE 6

Spectral curves of the four categories of surface after spherical correction and standard normal variate (SNV) transformation.

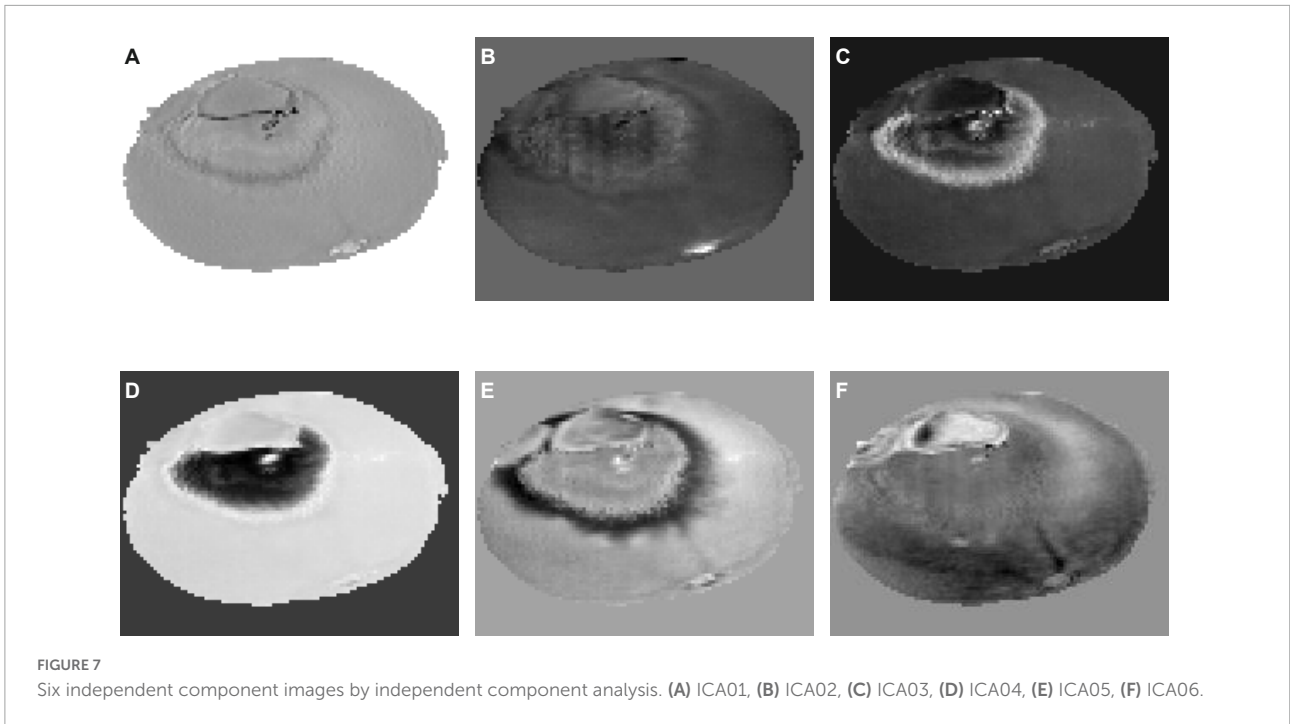
the height of the sample based on the distance from the center of gravity to points in each of the 16 equal parts is calculated; (4) the height of each pixel in the segmented binary image according to four quadrants is calculated, and the reflected light intensity is then corrected. Then, standard normal variate (SNV) transformation is used to eliminate baseline drift.

Establishment of deep neural network model

Deep neural networks are widely used in deep learning and image recognition (25–27). In this study, a long short-term memory (LSTM) deep neural network was applied to 217 spectra in the range of 975.18–1,146.75 nm to establish a classification model. Before establishing a LSTM deep neural network, training and validation datasets for training model

should be provided. From Figure 3, it can be seen the reflection is the strongest at a wavelength of 1,078.31 nm. Based on the 1,078.31 nm hyperspectral images, the tool “Image Labeler” in MATLAB (MathWorks Ltd., Natick, MA, USA) was used to select the most representative pixels on the surface of navel oranges from the prediction dataset. For this dataset, 15 penicilliosis navel oranges, six rotten navel oranges, 30 cankered navel oranges, and 16 normal navel oranges were selected. After sample selection, the image “label” and file “gTruth” were generated for each navel orange. The image “label” contains the features selected from each navel orange. The file “gTruth” contains the file storage location of the corresponding file and the storage location and classification name of the corresponding image “label.”

Thus, 10,405 data points were obtained for training. As shown in Table 1, 1,300 data points for each category were randomly selected as the modeling data, thus obtaining a total of



5,200 modeling data points. These data were randomly divided into a training dataset, validation dataset, and test dataset according to the approximate ratio of 0.8:0.1:0.1.

The LSTM model is mainly composed of five layers: (1) sequenceInputLayer, (2) bilstmLayer, (3) fullyConnectedLayer, (4) softmaxLayer, and (5) classificationLayer. The second layer has 200 hidden layer nodes, and the fifth layer has four output categories. The numbers of hidden layer nodes are selected by comparing the results using 50, 100, 150, 200, and 250 hidden layer nodes. The LSTM cell with a Forget Gate can be mathematically expressed as follows:

$$f_t = \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f),$$

$$i_t = \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i),$$

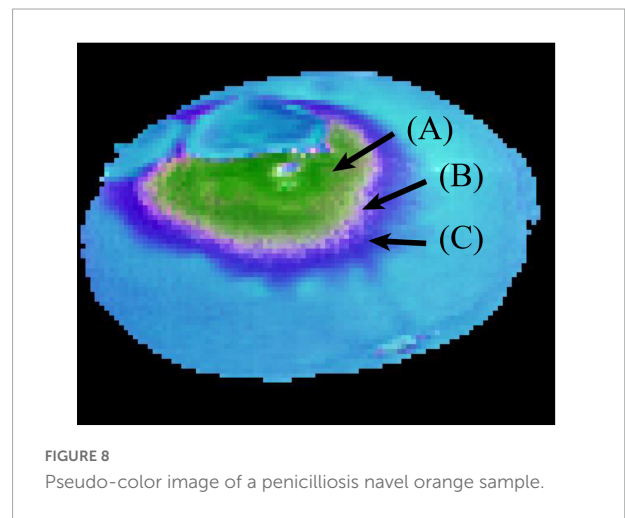
$$\tilde{c}_t = \tanh(W_{\tilde{c}h}h_{t-1} + W_{\tilde{c}x}x_t + b_{\tilde{c}}),$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t,$$

$$o_t = \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o),$$

$$h_t = o_t \cdot \tanh(c_t) \tag{2}$$

Where f_t is the forget gate, and the value of f_t can decide what information will be thrown away from the cell state. Where c_t denotes the cell state of LSTM. W_i , W_c , and W_o are the weights.



Selection of characteristic wavelengths

The data collected by the hyperspectral imaging system are usually very large. Therefore, a method of model establishment that uses a small number of characteristic wavelengths cannot only reduce the time spent in modeling, but also effectively avoid the problem of information similarity between adjacent wavelengths while reducing the over-fitting of the model. Therefore, using characteristic wavelengths to establish prediction models can reduce data calculation and meet the efficiency requirements of online detection.

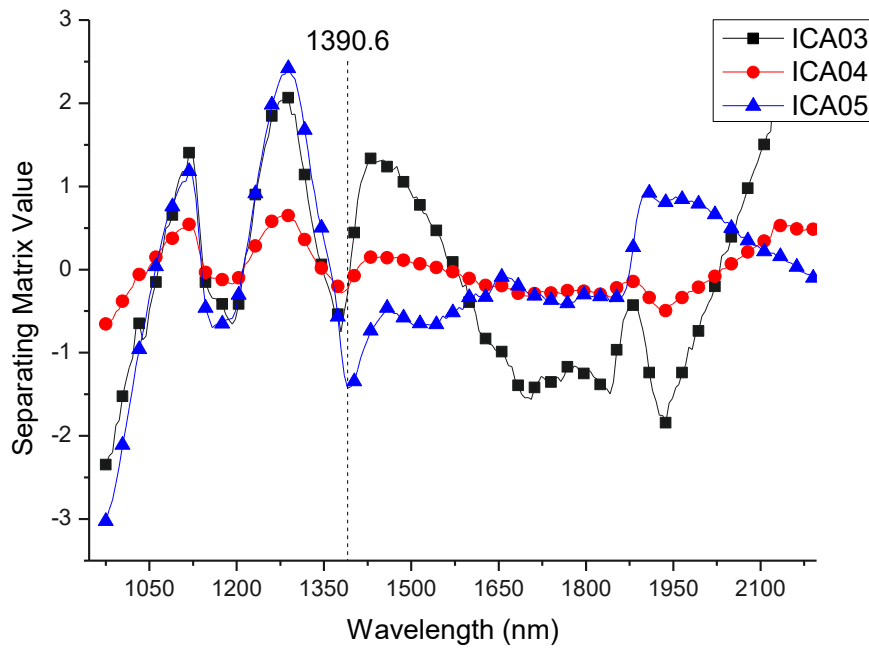


FIGURE 9
Separating matrix of ICA03, ICA04, ICA05.

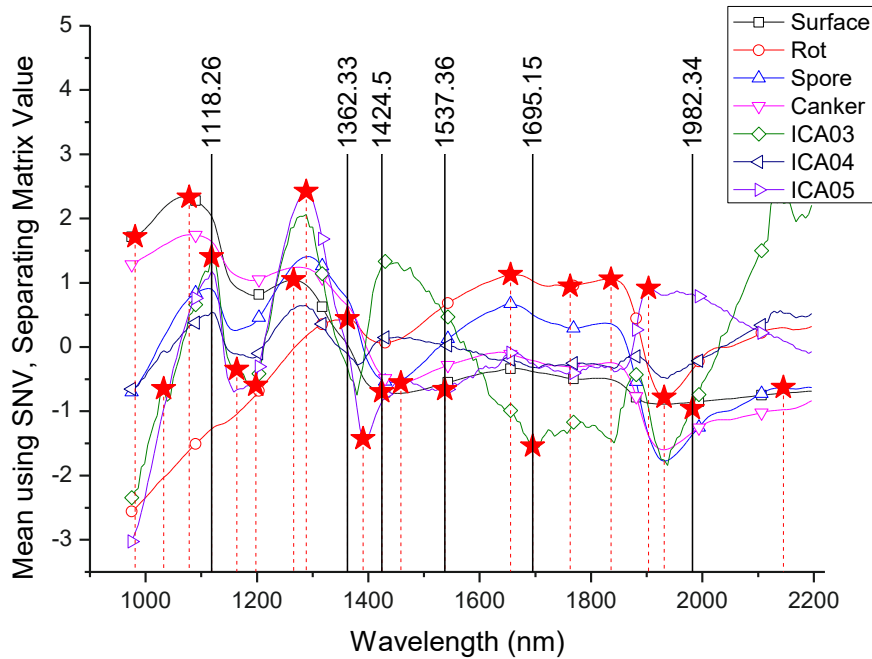


FIGURE 10
Selection of characteristic wavelengths.

The extraction methods of characteristic wavelengths include principal component analysis (PCA) and independent principal component analysis (ICA). ICA has a superior applicability to solve the problem of blind source separation in

spectral analysis (28). Accordingly, this study adopted the ICA method to extract characteristic spectra. The main steps are as follows. First, the binarization extraction of the hyperspectral image is conducted to obtain a mask, which is used to

TABLE 2 List of 21 characteristic wavelengths.

Number of wavelengths	Characteristic wavelengths (nm)
21	980.930, 1,032.55, 1,078.31, <i>1,118.26</i> , 1,163.83, 1,197.95, 1,266.07, 1,288.74, <i>1,362.33</i> , 1,390.60, <i>1,424.50</i> , 1,458.38, <i>1,537.36</i> , 1,655.72, <i>1,695.15</i> , 1,762.73, 1,835.94, 1,903.51, 1,931.66, <i>1,982.34</i> , 2,145.61

Italic values are the result of wavelength selection.

separate the background and the fruit area. According to the position of fruit pixels in the mask, the hyperspectral image is subjected to two-dimensional decomposition, thus obtaining a two-dimensional array. The array included the spectra of the fruit surface, but not the background part in the mask. Then, the first six independent components and the corresponding weight coefficients were obtained by ICA analysis of spectral data. Finally, six independent components were reconstructed according to the positions of the mask pixel points, and the values of independent components. Then the characteristic wavelengths can be selected by the ICA analysis based on the corresponding weight coefficients.

Genetic algorithm

For online detection, the number of characteristic wavelengths should be reduced as much as possible. Therefore, on the basis of the selected wavelengths using ICA analysis, a genetic algorithm (GA) was applied to select characteristic wavelengths once more (29, 30). A genetic algorithm is a kind of efficient global search optimization algorithm that can find the optimal solution within a large solution space, and its global search ability can greatly improve search efficiency while avoiding local minimal solutions. In this paper, the main genetic algorithm program was written using the software MATLAB. The characteristic wavelengths were selected by GA based on the accuracy of the test dataset as the fitness function. The following main parameters were used for the genetic algorithm: six characteristic wavelengths were selected without repetition each time. That is, the number of individuals was six, the population contained six individuals, the maximum genetic algebra was 30, and the generation gap was 0.9.

Results and discussion

Spherical fruit correction

The segmented and reconstructed navel orange 3D image is shown in [Figure 4](#).

For the corrected hyperspectral images, the 1,078.31, 1,266.07, and 1,655.72 nm wavelength images were extracted and fused into false-color images, as shown in [Figure 5](#). Compared with [Figure 1](#), the light intensity of the edge was significantly improved.

After spherical correction, the hyperspectral images of navel oranges were decomposed into two dimensions, and the spectra were preprocessed using standard normal variate (SNV) transformation to eliminate baseline drift. Corrected characteristic spectra of normal Surface, Canker, Spore, and Hypha samples are shown in [Figure 6](#). In the wavelength range of 975.18–1,146.75 nm, the spectral value of Surface is the highest, while the spectral value of Spore is the lowest. In the wavelength range of 1,146.75–1,430.15 nm, the spectral curves of Canker, Spore, and Hypha show some differences. At the wavelength of 1,430.15 nm and 1,937.29 nm, the spectral values of Surface and Hypha were the lowest. The four types of spectra all had a double-peak structure, and at 1,430.15 nm, there was a significant trough. At 1,937.29–2,196.26 nm, the spectral curves of Canker, Spore, and Hypha samples all increased, while the Surface spectral curve was relatively stable.

Model using full bands

The LSTM01 model was obtained by using the whole spectra, which included 217 bands. After training, the accuracy of the training, validation, and test dataset were 91.80, 90.56, and 92.81%, respectively. The modeling time was about 3,265.691 s, and the average test time for each orange was about 46.21 s.

Characteristic wavelengths using independent component analysis

Six independent component images were obtained, as shown in [Figure 7](#). ICA01–ICA06 represent the first through sixth independent component images, respectively. [Figure 7](#) shows the ICA result of a navel orange sample with penicilliosis. It can be seen that ICA03, ICA04, and ICA05 are sensitive to penicilliosis navel orange images. ICA03 and ICA05 are sensitive to penicillium hypha images, and ICA04 is sensitive to penicillium spore category images.

[Figure 8](#) is a pseudo-color image of a penicilliosis navel orange sample synthesized by ICA03, ICA04, and ICA05. The figure shows that the spore region (A) and hypha region (C) are relatively obvious, and the region (B) between regions (A) and (C) is the transitional region between spores and hyphae. On this basis, the characteristic wavelength can be selected through the weight coefficients of the wavelength variates of the three independent components.

[Figure 9](#) shows the separating matrix values of the wavelengths for independent components ICA03, ICA04, and

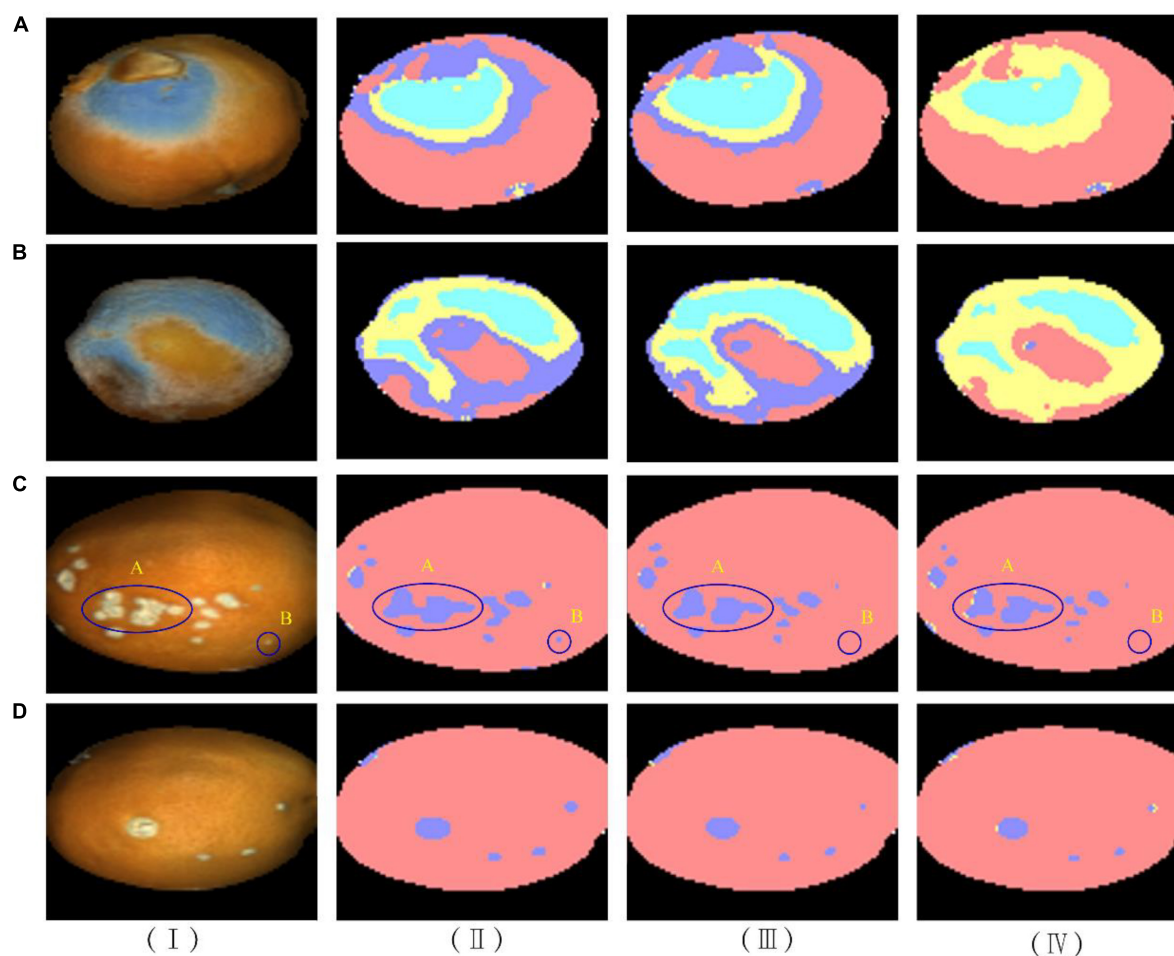


FIGURE 11

Result of the three LSTM models for four typical samples. (I) Sample hyperspectral images and corresponding images for models (II) LSTM01, (III) LSTM02, (IV) LSTM03.

ICA05. It can be seen that ICA03 and ICA05 have similar curve trends in the range of 975.18–1,390.6 nm. The curves both contain two wave peaks. In the range of 1,390.6–2,196.26 nm, the three independent component curves showed obvious differences. According to the curve trend of the independent component, the characteristic wavelengths were selected based on the peaks and troughs of the curves. In the selection process, in order to contain as much information about the classified features as possible, the selection of feature wavelengths also included the peaks and troughs of the spectral curves of the four categories in Figure 6. Thus, a total of 21 characteristic wavelengths were selected, which are marked with asterisks in Figure 10 and Table 2.

The LSTM02 model was obtained by training the LSTM model with 21 characteristic wavelengths. The accuracies of the training dataset, validation dataset, and test dataset were 94.87, 94.07, and 95.01%, respectively. The modeling time was about 328.4015 s, and the average test time for each orange was about

3.71 s. Compared with the LSTM01 model trained with full spectra, it can be seen that the accuracy has been improved, the modeling time has been reduced by about 10 times, and the number of wavelengths used has been reduced from 217 to 21.

Characteristic wavelengths using genetic algorithm

Based on the 21 characteristic wavelengths selected by ICA, GA was used for further dimension reduction, and the six characteristic wavelengths are 1,118.26, 1,362.33, 1,424.50, 1,537.36, 1,695.15, and 1,982.34 nm. In Table 2, the values specified in italic font are the six characteristic wavelengths. Then, the LSTM03 model was established with six characteristic wavelengths. The accuracies of the modeling set, prediction set, and verification set were 94.03, 93.148, and 93.41%. The modeling time was about 105.223 s, and the average test time

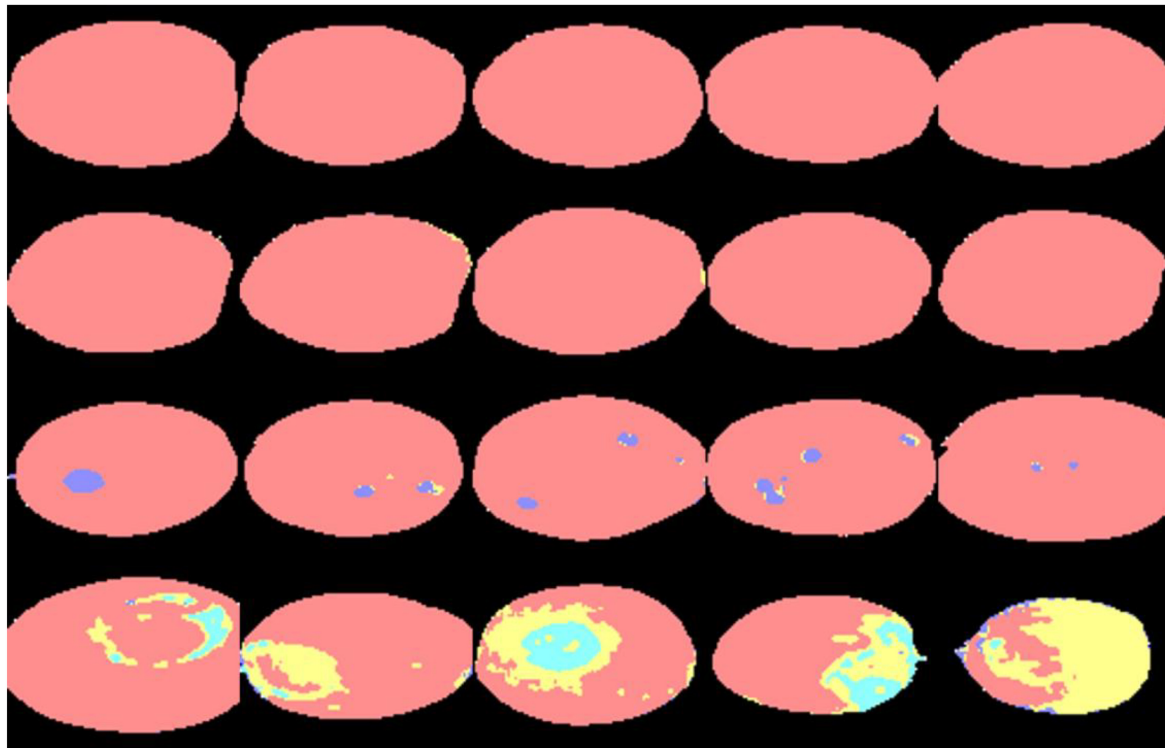


FIGURE 12
Discrimination results of 20 navel orange samples.

for each orange was about 1.26 s. It can be seen that through the selection of wavelengths by GA, the number of wavelengths is reduced from 21 to 6, and the time required for modeling is also reduced by about three-fold compared with the LSTM02 model.

Prediction results

The prediction results show that the prediction accuracy of LSTM01, LSTM02, and LSTM03 models were 92.81, 95.01, and 93.41%, respectively. Although the prediction accuracy of LSTM03 model was slightly lower than that of LSTM02, the prediction time of LSTM03 model for each navel orange was the shortest, only 1.26 s, the prediction time of LSTM02 was 3.7 s, and the prediction time of LSTM01 is the longest, 46.21 s. Thus, in order to ensure the accuracy and efficiency at the same time, the LSTM03 model was better than LSTM02, which is conducive to the online monitoring of fruits by accurately predicting the location of penicillium spore defects.

Figure 11 shows the classification results of four navel oranges in the application of the three LSTM models. In the column (I), (A) and (B) are pseudo-color images of penicilliosis navel oranges, while (C) and (D) are pseudo-color images of navel oranges with canker. The proportions and positions of penicillium disease and canker on the navel orange surface

differ. In the column (II), the prediction results of the LSTM01 model are shown. It can be seen that part of the hypha area was misjudged as canker in (A) and (B), but the location of penicillium disease was fairly accurate. The predictions for (C) and (D) were relatively accurate. A smaller canker part was detected in circular region B. However, two canker regions were connected in elliptic region A. In the column (III), the prediction results of the LSTM02 model are shown; (A) and (B) show that the prediction results of penicilliosis were good, but some hypha area were still misjudged as canker. The canker locations in (C) and (D) also had high recognition rates. Compared with the predicted results of the LSTM01 model, two canker parts in the elliptic region A were identified separately, while the smaller canker was not successfully identified and was instead misjudged as normal surface, as shown in circular region B. Among the results of LSTM03 in the column (IV), the prediction results showed that hypha areas in (A) and (B) are accurately identified. The discriminant results of canker in (C) and (D) are similar to that of the LSTM02 model, and the overall area of the discriminant result is slightly smaller than that of the LSTM02 model.

Figure 12 shows the classification results for 20 navel oranges using the LSTM03 model. The first and second rows are the discriminant results of hyperspectral data from navel orange samples collected on the first and fifth days after

picking, respectively. It can be seen that the hyperspectral image data collected on the fifth day included misjudged hyphae, especially the edge part of navel oranges. The third row is the discrimination result of hyperspectral image data with canker, which shows that this model can accurately identify the canker location and size on navel orange surfaces. The fourth row shows the discrimination results of hyperspectral image data collected after 30 days in the laboratory. This model can accurately identify penicilliosis on navel oranges.

Conclusion

Based on ICA and wavelength optimization by GA, a total of six characteristic wavelengths were selected to establish a deep learning neural network model. The model was used to classify and detect surface defects of navel oranges. The categories included penicillium spore, penicillium hypha, canker, and normal navel orange surfaces. The characteristic wavelengths were, respectively, 1,118.26, 1,362.33, 1,424.50, 1,537.36, 1,695.15, and 1,982.34 nm. Through the selection of characteristic wavelengths, the test time of the LSTM03 model for each navel orange sample was reduced from 46.21 s (with the full spectrum model) to 1.26 s. It was conducive to the hyperspectral online detection of fruit, and its prediction accuracy was also improved. It can be seen that this method can be used to detect the surface defects of navel oranges online. However, the hyperspectral images were collected in this experiment through translation movement of navel orange samples. Thus, only images from the upper surface of the navel orange were collected, while lower surface images were not. Therefore, a major research focus includes capturing and analyzing images of the entire surface of navel oranges in subsequent online detection work.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

References

- Zhang L, Ling W, Yan Z, Liang Y, Guo C, Ouyang Z, et al. Effects of storage conditions and heat treatment on the hesperidin concentration in Newhall navel orange (*Citrus sinensis* Osbeck Cv. Newhall) juice. *J Food Compos Anal.* (2020) 85:103338. doi: 10.1016/j.jfca.2019.10.3338
- Gao Y, Liu Y, Kan C, Chen M, Chen J. Changes of peel color and fruit quality in navel orange fruits under different storage methods. *Sci Hort.* (2019) 256:108522. doi: 10.1016/j.scienta.2019.05.049
- Gao Y, Kan C, Wan C, Chen C, Chen M, Chen J. Quality and biochemical changes of navel orange fruits during storage as affected by cinnamaldehyde

Author contributions

JL: conceptualization, funding acquisition, writing – original draft review, and editing. ML: writing – original draft and visualization. LX: supervision and writing – review and editing. JC: conceptualization and funding acquisition. LH: conceptualization, investigation, software, and writing – review and editing. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Key Science and Technology Research Project in Jiangxi Province Department of Education (GJJ180169) and National Natural Science Foundation of China (31871526). This work was also partially funded by the Collaborative Innovation Center of Postharvest Key Technology and Quality Safety of Fruits and Vegetables in Jiangxi Province (JXGS-05) and supported by Jiangxi Province Science and Technology Innovation Team Building Plans (20171BCB24006).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

-chitosan coating. *Sci Hort.* (2018) 239:80–6. doi: 10.1016/j.scienta.2018.05.012

4. Song J, Li G, Yang X, Liu X, Xie L. Rapid analysis of soluble solid content in navel orange based on visible-near infrared spectroscopy combined with a swarm intelligence optimization method. *Spectrochim Acta A Mol Biomol Spectrosc.* (2020) 228:117815. doi: 10.1016/j.saa.2019.117815

5. Borba KR, Spricigo PC, Aykas DP, Mitsuyuki MC, Colnago LA, Ferreira MD. Non-invasive quantification of vitamin C, citric acid, and sugar in 'valência' oranges using infrared spectroscopies. *J Food Sci Technol.* (2020) 58:731–8. doi: 10.1007/s13197-020-04589-x

6. Singh H, Sridhar A, Saini SS. Ultra-low-cost self-referencing multispectral detector for non-destructive measurement of fruit quality. *Food Anal Methods*. (2020) 13:1879–93. doi: 10.1007/s12161-020-01810-7
7. Feng L, Wu B, Zhu S, He Y, Zhang C. Application of visible/infrared spectroscopy and hyperspectral imaging with machine learning techniques for identifying food varieties and geographical origins. *Front Nutr*. (2021) 8:680357. doi: 10.3389/fnut.2021.680357
8. Li J, Zhang Y, Liu M, Chen J, Xue L. Rapid detection and visualization of mechanical bruises on “Nanfeng” mandarin using the hyperspectral imaging combined with ICA_LSQ method. *Food Anal Methods*. (2019) 12:1936–9751. doi: 10.1007/s12161-019-01546-z
9. Zhang C, Wu W, Zhou L, Cheng H, Ye X, He Y. 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*. (2020) 319:126536. doi: 10.1016/j.foodchem.2020.126536
10. Badaró AT, Garcia-Martin JF, López-Barrera MDC, Barbin DE, Alvarez-Mateos P. Determination of pectin content in orange peels by near infrared hyperspectral imaging. *Food Chem*. (2020) 323:126861. doi: 10.1016/j.foodchem.2020.126861
11. Al-Sarayreh M, Reis MM, Yan WQ, Klette R. Potential of deep learning and snapshot hyperspectral imaging for classification of species in meat. *Food Control*. (2020) 117:107332. doi: 10.1016/j.foodcont.2020.107332
12. Cen H, Lu R, Zhu Q, Mendoza F. Nondestructive detection of chilling injury in cucumber fruit using hyperspectral imaging with feature selection and supervised classification. *Postharvest Biol Technol*. (2016) 111:352–61. doi: 10.1016/j.postharvbio.2015.09.027
13. Mayorga-Martínez AA, Olvera-Trejo D, Elías-Zúñiga A, Parra-Saldívar R, Chuck-Hernández C. Non-destructive assessment of guava (*Psidium guajava* L.) maturity and firmness based on mechanical vibration response. *Food Bioprocess Technol*. (2016) 9:1471–80. doi: 10.1007/s11947-016-1736-8
14. Gao Z, Shao Y, Xuan G, Wang Y, Liu Y, Han X. Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning. *Artif Intell Agric*. (2020) 4:31–8. doi: 10.1016/j.aiia.2020.04.003
15. Xuan G, Gao C, Shao Y, Wang X, Wang Y, Wang K. Maturity determination at harvest and spatial assessment of moisture content in Okra using vis-NIR hyperspectral imaging. *Postharvest Biol Technol*. (2021) 180:111597. doi: 10.1016/j.postharvbio.2021.111597
16. Zhao J, Ouyang Q, Chen Q, Wang J. Detection of bruise on pear by hyperspectral imaging sensor with different classification algorithms. *Sensor Lett*. (2010) 8:570–6.
17. Zareef M, Chen Q, Hassan MM, Arslan M, Hashim MM, Ahmad W, et al. An overview on the applications of typical non-linear algorithms coupled with NIR spectroscopy in food analysis. *Food Eng Rev*. (2020) 12:173–90. doi: 10.1007/s12393-020-09210-7
18. Xuan G, Gao C, Shao Y. Spectral and image analysis of hyperspectral data for internal and external quality assessment of peach fruit. *Spectrochim Acta A Mol Biomol Spectrosc*. (2022) 272:121016. doi: 10.1016/j.saa.2022.121016
19. Zhang D, Xu L, Liang D, Xu C, Jin X, Weng S. Fast prediction of sugar content in Dangshan pear (*Pyrus* Spp.) using hyperspectral imagery data. *Food Anal Methods*. (2018) 11:2336–45. doi: 10.1007/s12161-018-1212-3
20. Kang X, Xiang X, Li S, Benediktsson JA. Pca-based edge-preserving features for hyperspectral image classification. *IEEE Trans Geosci Remote Sens*. (2017) 55:7140–51. doi: 10.1109/TGRS.2017.2743102
21. Liu Q, Zhou D, Tu S, Xiao H, Zhang B, Sun Y, et al. Quantitative visualization of fungal contamination in peach fruit using hyperspectral imaging. *Food Anal Methods*. (2020) 13:1262–70. doi: 10.1007/s12161-020-01747-x
22. Shi J, Wang Y, Li Z, Huang X, Shen T, Zou X. Characterization of invisible symptoms caused by early phosphorus deficiency in cucumber plants using near-infrared hyperspectral imaging technology. *Spectrochim Acta A Mol Biomol Spectrosc*. (2022) 267:120540. doi: 10.1016/j.saa.2021.120540
23. Li S, Song W, Fang L, Chen Y, Ghamisi P, Benediktsson JA. Deep learning for hyperspectral image classification: an overview. *IEEE Trans Geosci Remote Sens*. (2019) 57:6690–709. doi: 10.1109/TGRS.2019.2907932
24. Gomez-Sanchis J, Molto E, Camps-Valls G, Gomez-Chova L, Alexios N, Blasco J. Automatic correction of the effects of the light source on spherical objects. An application to the analysis of hyperspectral images of citrus fruits. *J Food Eng*. (2008) 85:191–200. doi: 10.1016/j.jfoodeng.2007.06.036
25. Agarwal M, Al-Shuwaili T, Nugaliyadde A, Wang P, Wong KW, Ren Y. Identification and diagnosis of whole body and fragments of *Trogoderma granarium* and *Trogoderma variabile* using visible near infrared hyperspectral imaging technique coupled with deep learning. *Comput Electron Agric*. (2020) 173:105438. doi: 10.1016/j.compag.2020.105438
26. Xin Z, Jun S, Yan T, Quansheng C, Xiaohong W, Yingying HA. Deep learning based regression method on hyperspectral data for rapid prediction of cadmium residue in lettuce leaves. *Chemom Intell Lab Syst*. (2020) 200:103996. doi: 10.1016/j.chemolab.2020.103996
27. Zhou X, Sun J, Tian Y, Lu B, Hang Y, Chen Q. Hyperspectral technique combined with deep learning algorithm for detection of compound heavy metals in lettuce. *Food Chem*. (2020) 321:126503. doi: 10.1016/j.foodchem.2020.126503
28. Mishra P, Cordella CBY, Rutledge DN, Barreiro P, Roger JM, Diezma B. Application of independent components analysis with the jade algorithm and NIR hyperspectral imaging for revealing food adulteration. *J Food Eng*. (2016) 168:7–15. doi: 10.1016/j.jfoodeng.2015.07.008
29. Zehtabian A, Ghasseman H. Combining genetic algorithm with Pdes for improving the performance of statistical region merging based object extraction. *J Indian Soc Remote Sens*. (2016) 44:395–407. doi: 10.1007/s12524-015-0515-8
30. Feng Y-Z, Sun D-W. Near-Infrared hyperspectral imaging in tandem with partial least squares regression and genetic algorithm for non-destructive determination and visualization of *Pseudomonas* loads in chicken fillets. *Talanta*. (2013) 109:74–83. doi: 10.1016/j.talanta.2013.01.057