

Hyperspectral image classification of wolfberry with different geographical origins based on three-dimensional convolutional neural network

Qingshuang Mu, Zhilong Kang, Yanju Guo, Lei Chen, Shenyi Wang & Yuchen Zhao

To cite this article: Qingshuang Mu, Zhilong Kang, Yanju Guo, Lei Chen, Shenyi Wang & Yuchen Zhao (2021) Hyperspectral image classification of wolfberry with different geographical origins based on three-dimensional convolutional neural network, International Journal of Food Properties, 24:1, 1705-1721, DOI: [10.1080/10942912.2021.1987457](https://doi.org/10.1080/10942912.2021.1987457)

To link to this article: <https://doi.org/10.1080/10942912.2021.1987457>



Published with license by Taylor & Francis Group, LLC. © 2021 Qingshuang Mu, Zhilong Kang, Yanju Guo, Lei Chen, Shenyi Wang and Yuchen Zhao



Published online: 31 Oct 2021.



Submit your article to this journal



Article views: 1485



View related articles



View Crossmark data



Citing articles: 5 View citing articles

Hyperspectral image classification of wolfberry with different geographical origins based on three-dimensional convolutional neural network

Qingshuang Mu^a, Zhilong Kang^a, Yanju Guo^a, Lei Chen^b, Shenyi Wang^a, and Yuchen Zhao^a

^aSchool of Electronic Information Engineering, Hebei University of Technology, Tianjin, China; ^bSchool of Information Engineering, Tianjin University of Commerce, Tianjin, China

ABSTRACT

The hyperspectral image is a three-dimensional (3D) hypercube with spectral and spatial continuity. Traditional hyperspectral imaging (HSI) processing mainly focuses on spectral information. However, this paper proposed a new hybrid convolutional neural network (New-Hybrid-CNN) algorithm using HSI spectral-spatial joint information. We used the algorithm combined with HSI processing to classify the origin of Chinese wolfberry from Ningxia, Qinghai, Gansu, and Xinjiang. (1) Selecting the region of interest (ROI) over the raw HSI data as input; (2) Extracting spectral-spatial joint information from the hyperspectral stack information using homogeneous 3D convolution architecture with $3 \times 3 \times 3$ convolution kernels; (3) Then the depth separable convolution (DSC) was used to learn spatial information. This algorithm combined the advantages of 3D convolution and DSC, and it effectively extracted deep spectral-spatial joint information and made the architecture more lightweight. 3D convolutional neural network (3D-CNN), hybrid spectral convolutional neural network (HybridSN), and support vector machine (SVM) were established to compare with the proposed method. The proposed algorithm made full use of the HSI information while reducing the number of parameters and training time involved in the network, and improved the classification accuracy. The classification accuracy of wolfberry origin reached more than 99%. Therefore, the New-Hybrid-CNN classifier combined with HSI had the potential to classify wolfberry origin and food detection.

ARTICLE HISTORY

Received 11 June 2021

Revised 24 September 2021

Accepted 24 September 2021

KEYWORDS

Hyperspectral image; Three-dimensional convolutional neural network; Depth separable convolution; Support vector machine; Classification of wolfberry origin

INTRODUCTION

Wolfberry contains polysaccharides, carotenoids, flavonoids, and other ingredients. Moreover, it has a wide array of biological and pharmacological activities.^[1] Wolfberry is a kind of Chinese food and famous Chinese medicine, and it has been considered to have anti-aging, anti-oxidative, and cancer prevention properties for thousands of years.^[2] The medicinal function and price of wolfberry are closely related to its geographical origin, cultivation methods, and environmental factors^[3,4] (including sunshine intensity, temperature, precipitation, and soil).

According to most traditional methods, the geographical origins of wolfberries can be identified by their color, shape, and taste. However, this subjective detection method is easily affected by personal emotions.^[5] In addition, some researchers use chemical analysis techniques to classify and identify the origin of wolfberry. These research methods take a long time and require skilled analysts to do

experiments.^[6] At present, some nondestructive technologies have been developed to evaluate the intrinsic characteristics of Horti-food products.^[7] Therefore, establishing a fast, nondestructive, and high-precision classification system to trace the origin of wolfberry is necessary.

HSI technology is a combination of spectral technology and imaging technology. It can obtain internal and external information of the measured sample in the wavelength range of 780–2500 nm.^[8] It has the advantages of being rapid, accurate, noninvasive, easy to use, and is widely used in remote sensing images processing.^[9] Visible and NIR ranges are the most spectra the researchers have focused on for agricultural applications.^[10] The near-infrared hyperspectral imaging (NIR-HSI) technology has been widely used in food quality analysis, variety, and origin of identification.^[11] For example: identifying six different commercial teas,^[12] identifying the origin of Longjing tea,^[13] distinguishing coffee beans with varying levels of roasting,^[14] determining the polysaccharide content in wolfberry,^[15] estimating the ripeness of strawberries in the field and the laboratory,^[16] etc. Therefore, NIR-HSI technology has great potential in determining the geographic origin of wolfberry.

HSI information is presented in the format of 3D hypercubes. The first two dimensions represent spatial information, and the third dimension represents spectral information.^[12] Extracting data information from hypercubes requires advanced pattern recognition. The traditional HSI processing method only extracts the relevant spectral information from the ROI, and most food quality evaluation studies only use the one-dimensional spectral data information.^[13,17,18] However, studies have shown that combining one-dimensional spectral information and 2D spatial information is better than only using spectral information or texture information. For example, the classification accuracy of green tea types was improved by fusion of spectral and texture information^[19]; the spectral and texture feature information were combined to distinguish fresh pork and frozen-thawed pork.^[20] The disadvantage of the above methods is that the spectral data and spatial data are extracted separately. It cannot make full use of the spectral-spatial joint information, which affects the classification performance.

To solve the above problems, the researchers introduced 3D-CNN into the HSI classification. For example, researchers proposed a new deep 3D-CNN model to extract combined features and classify the meat in the HSI image.^[21] Applying the 3D convolution kernel operation to the 3D-HSI, we can learn the spectral-spatial joint information of the data cube simultaneously and fully use the structural characteristics of 3D-HSI data.^[21–24] However, when training 3D-CNN, we often meet various optimization difficulties due to increased parameters, such as over-fitting, gradient disappearance, explosion, etc. These problems limit the training and application of 3D-CNN. Other researchers^[25,26] proposed to connect 3D convolution and 2D convolution operations in series to obtain spatial and spectral information, reducing the complexity of the 3D-CNN model. Compared with ordinary convolution, deep separable neural networks can significantly reduce model parameters, calculation time, and model complexity under almost the same recognition accuracy.^[27]

A lot of research work has been carried out to identify the origin of wolfberry. Most of the research is based on wolfberry samples' internal chemical composition to determine the geographical origin. Besides, some research only focuses on the spectral information in wolfberry's hyperspectral data and seldom uses its spatial data. According to our knowledge, there are no published studies about the classification of wolfberry origin using hybrid convolutional neural networks and HSI technology. Inspired by these studies, this study designed a New-Hybrid-CNN algorithm combined with HSI technology to classify wolfberry from different origins. The New-Hybrid-CNN algorithm has the advantages of short training time, high classification accuracy, and lightweight structure, which will help future research on detection systems for food production areas and types. Specifically, the objectives of this study were: (i) Obtain the hyperspectral data information of four different origins of wolfberry, connect the 3D convolution operation and DSC in series to establish a New-Hybrid-CNN algorithm, and apply it to the classification of the origin of wolfberry. (ii) Compare and evaluate the New-Hybrid-CNN with three other hyperspectral image classification methods: SVM, 3D-CNN, and HybridSN. (ii) Analyze the impact of samples from different input windows on the performance of the New-Hybrid-CNN algorithm.



MATERIALS AND METHODS

Sample preparation

The wolfberry samples were purchased from local farmers in four main producing areas, including Zhongning County (approximately 37.48°N and 105.66°E, Zhongwei, Ningxia, China), Tsaidam Basin (approximately 35–39.2°N and 90.16°–99.16°E, Qinghai, China), Jingtai County (approximately 37.18° N and 104.06°E, Gansu, China), Jinghe County (approximately 44.60°N and 82.89°E, Bortala Mongolia, Xinjiang, China). All of them meet the implementation standards of “Wolfberry”: GB/T18672. The wolfberry was stored in plastic bags at room temperature. As shown in Figure 1, about 150 g of wolfberry with uniform appearance and size from each producing area were collected as samples.

Hyperspectral imaging system and image acquisition

A line-scan HSI system with reflection mode was used during the experiment to capture hyperspectral images of wolfberry. The system consists of a spectrograph (ImSpector N17E, Spectral Imaging Ltd., Oulu, Finland), a CCD hyperspectral camera (Zelos-258GV, Kappa optronics GmbH, Germany), a lighting unit composed of four 35 W halogen lamps, and a conveyor stage (PSA200-11-X, Zolix., Ltd., Beijing, China) and a computer (V10E, Isuzu Optical Corp., Taiwan, China) supporting spectral cube data acquisition software. The system collected a total of 25 spectral bands in the wavelength range of 900–1700 nm.



Ningxia



Qinghai



Gansu



Xinjiang

Figure 1. RGB image of the Chinese wolfberries from four different geographical origins.

The image acquisition was carried out in a dark environment to avoid light interference. We took the sample of wolfberry from one of the origins and placed it evenly on a piece of black cardboard. Then, the black cardboard was put on the moving platform at a rate of 16.8 mm/s. The distance from the lens to the sample was 380 mm, and the exposure time was adjusted to 20 ms. Images were collected with a spectral resolution of approximately 3.2 nm. The ENVI 5.3 (ITT Visual Information Solutions, Boulder, Utah, USA) and Python 3.8 were used for image processing. Before image processing, the raw hyperspectral images needed to be corrected using white and dark reference images. The white reference image (I_{white}) was acquired using a white Teflon tile with nearly 100% reflectance, and the dark reference image (I_{dark}) was obtained by completely covering the camera lens with an opaque cap. Then the calibrated image (I_c) was calculated by the original hyperspectral image (I_{raw}) following Equation (1):

$$I_c = \frac{(I_{raw} - I_{dark})}{(I_{white} - I_{dark})} \quad (1)$$

The denominator can eliminate the spatial inhomogeneity of the light source, and the subtraction in the numerator can reduce the systematic noise. As a whole, the percentage generated by division can simplify subsequent calculations. More importantly, the spectrum from the calibration image can be interpreted as molecular bonds or components that assign features.^[28]

Data collection

The original NIR-HSI obtained by the hyperspectral imaging system contained 256 dimensions information. The data information at the beginning and end were both affected by the “noise” of the instrument.^[29] Therefore, to avoid noise, we removed some bands (900–998 nm and 1622–1700 nm) and selected 190 spectral bands (998–1622 nm) as each type of wolfberry raw spectral data.

One image was acquired of the wolfberry samples from each origin. After collecting and correcting the four kinds of wolfberry images, the ENVI5.3 analysis software was used to crop manually a region of interest (ROI) from the corrected hyperspectral image. The 50 × 50 area in the middle of the hyperspectral image was selected to remove the background and insignificant pixels. So, the ROI included 50 × 50 pixels and 190 frequency bands (as shown in Figure 2, taking a sample of Gansu wolfberry as an example).

Classification methods

In this section, we adopted four different classification methods to classify the origin of wolfberry: SVM, 3D-CNN, HybridSN, and New-Hybrid-CNN. The flow of the experiment is shown in Figure 3.

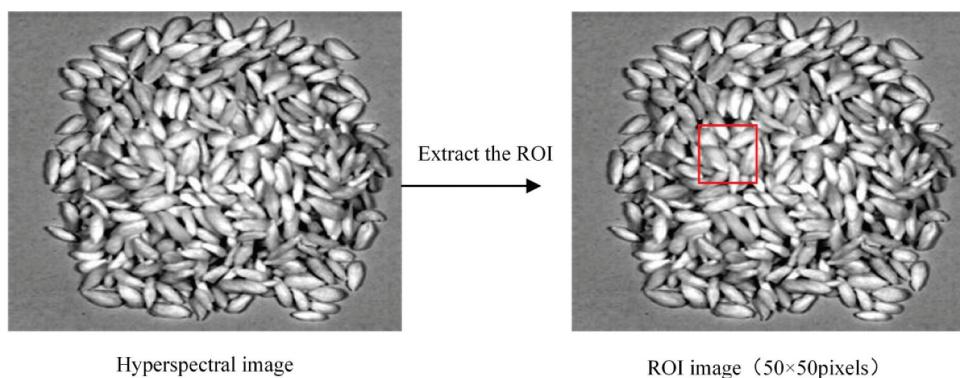


Figure 2. Flow chart of ROI extraction for NIR hyperspectral imaging of Gansu Chinese wolfberry.

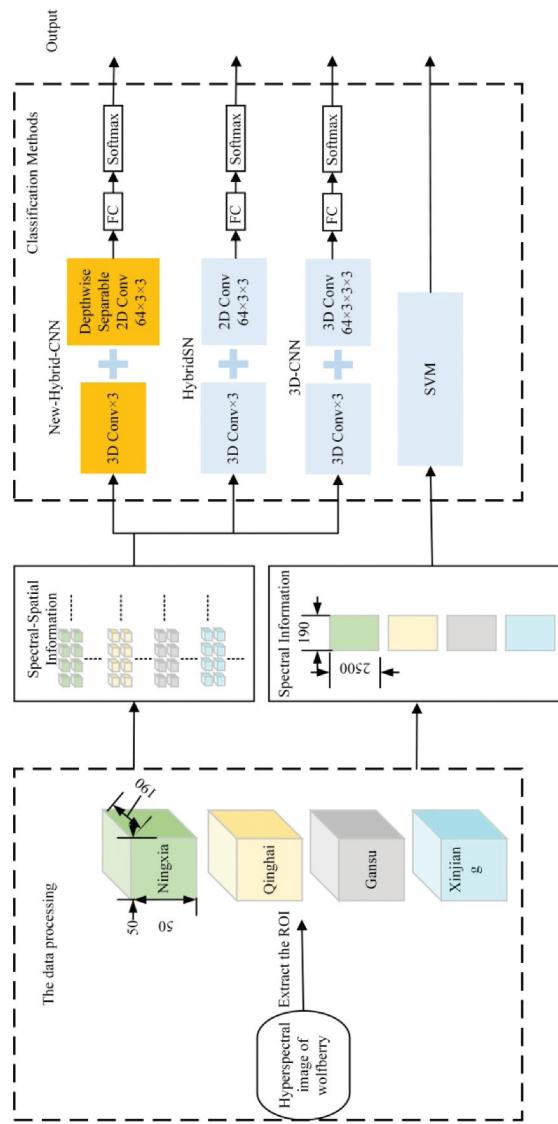


Figure 3. Experimental flow chart.

New-Hybrid-CNN architecture

In this paper, a New-Hybrid-CNN classification framework is established to realize the classification of the origin of Chinese wolfberry. First, the wolfberry HSI data cube is divided into small overlapping 3D patches, where the label of the central pixel determines the truth label. The size of the overlapping 3D patches is $S \times S \times B$. The small 3D patches cover the $S \times S$ spatial range and all B spectrum bands with the target pixel as the center. The patch-based classification method is performed to fully use the pixel information, dividing the HSI cubes into 2500 small 3D patches of $11 \times 11 \times 190$. So, there are 2500 samples of wolfberry from each origin. The samples of wolfberry from different origins are randomly divided into a training set (70%) and a test set (30%). Then 20% of the training set is randomly selected as the verification set to verify the generalization degree of the algorithm.

These small 3D patches are input into the classification algorithm, which first uses the homogeneous 3D convolution architecture, and the size of the 3D convolution filter is $3 \times 3 \times 3$. Then the output feature maps from the 3D convolution are reshaped and fed to a depth separable 2D convolution to learn spatial information. The output of the deep separable 2D convolutional layer is flattened and passed to the fully connected layer. After the fully connected layer, a softmax classifier for multi-classification tasks is added to output the classification results. The architecture is shown in Figure 4.

3DCNN Block: Usually, the purpose of a one-dimensional convolution operation is to extract spectral features. The purpose of 2D convolution operation is to extract the spatial features, and 3D convolution operation can simultaneously extract spectral and spatial features from high-dimensional information.^[30]

2D-CNN has made outstanding achievements in computer vision and image classification. The traditional 2D-CNN can be expressed as:

$$v_{ij}^{xy} = f\left(\sum_m \sum_{h=0}^{H_i-1} \sum_{w=0}^{W_i-1} k_{ijm}^{hw} V_{(i-1)m}^{(x+h)(y+w)} + b_{ij}\right) \quad (2)$$

where i represents the i -th layer in the network, and j is the j -th feature map of the layer. v_{ij}^{xy} represents the value at (x, y) in the j -th feature map in the i -th layer, b_{ij} is the offset, and m represents the index of all feature maps connected to the current layer in the $(i-1)$ layer. k_{ijm}^{hw} represents the value of the convolution kernel at (h, w) . H_i and W_i represent the height and width of the corresponding convolution kernel of the layer.

If the original hyperspectral data is directly used for 2D-CNN, each input band needs to be two-dimensionally convolved. Therefore, each input band needs a set of convolution kernels is required for each input band, leading to massive network parameters. In addition, this can lead to increased calculation and severe over-fitting problems.^[31]

Hyperspectral images are 3D hypercubes with spectral and spatial continuity. Combining spectral and spatial information is becoming more and more popular because it can fully use the original data information. Using 3D convolution to extract features is a simple method for hyperspectral image classification, which can simultaneously process 3D regions with joint spatial-spectral information.^[32]

The 3D convolution operation convolves the input data in spatial and spectral dimensions and outputs 3D data. This 3D data saves the spectral information input by the hyperspectral data. The formula for the 3D convolution operation is:

$$v_{ji}^{xyz} = f\left(\sum_{h=0}^{H_i-1} \sum_{w=0}^{W_i-1} \sum_{s=0}^{S_i-1} k_{ijm}^{hws} V_{(i-1)j}^{(x+h)(y+w)(z+s)} + b_{ij}\right) \quad (3)$$

where S represents the size of the 3D convolution kernel in the spectral dimension, i represents the number of feature blocks on the network, and j represents the number of convolution kernels in this layer. Correspondingly, the output of the layer (i -th layer) includes $i \times j$ 3D feature values.

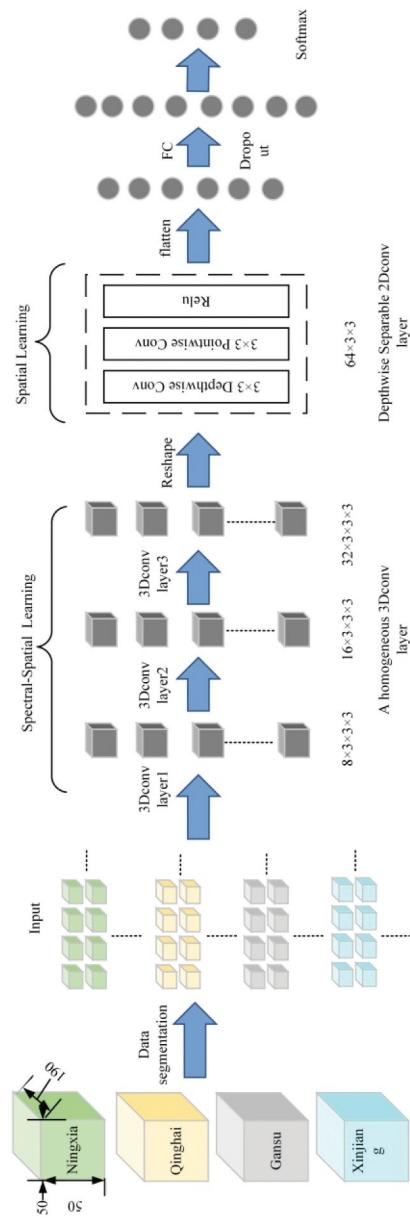


Figure 4. The New-Hybrid-CNN network architecture diagram.

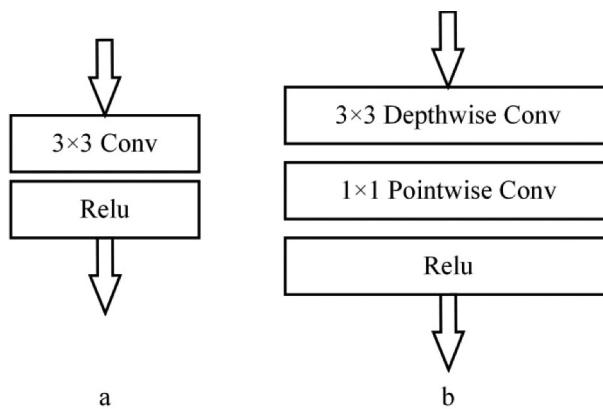


Figure 5. Comparison of standard convolution (a) and depth separable convolution (b).

Depth wise Separable Convolution Block: Figure 5(b) shows that the DSC separates the convolution channel by channel, mainly divides into two steps: depthwise convolution and pointwise convolution. Depthwise convolution uses a filter on each input channel, and Pointwise convolution uses a 1×1 convolution kernel for channel fusion. Pointwise convolution helps preserve spatial information, thereby optimizing the performance of convolutional networks. DSC is a transformation form of ordinary 2D convolution.^[33] Compared with the ordinary convolution operation, the DSC is lighter (fewer trainable weight parameters), and the calculation cost is less expensive.

Parameter setting: Compared with the image-level classification algorithm, the input data space size used in the pixel-level classification algorithm is relatively small. Using a convolution kernel with a smaller space can avoid excessive loss of input information.^[31] Through the study of 2D-CNN, it is found that a convolution kernel with a size of 3×3 usually produces better results. Some researchers used $3 \times 3 \times 3$ size 3D convolution kernels for spatiotemporal feature learning^[34] and common hyperspectral data sets.^[21] In the New-Hybrid-CNN, all 3D convolution filters are $3 \times 3 \times 3$ with a step size of one in the spatial and spectral dimensions.

As for the network structure of 3D-CNN, HybridSN, and New-Hybrid-CNN in this study, the first three layers of 3D convolution in these three classification algorithms all used the homogeneous 3D convolution architecture: all the 3D convolution filters were $3 \times 3 \times 3$; the step size was $1 \times 1 \times 1$; the number of filters was 8, 16, and 32. The fourth layer in the 3D-CNN was 3D convolution kernels of $64 \times 3 \times 3 \times 3$. The fourth layer of the HybridSN was ordinary 2D convolution kernels of $64 \times 3 \times 3$. The fourth layer of the proposed New-Hybrid-CNN was a deep separable convolution. It was divided into two steps: The first step was the deep convolution operation, and each convolution kernel only convolved one channel of the input layer (the size of each filter is 3×3). The second step used a 1×1 size convolution kernel to perform pixel-by-pixel convolution, and the number of filters was 64. The fully connected layer and classifier of these three classification algorithms remained the same.

Studies have shown that appropriately increasing the size of the input data can help improve classification performance. However, a large data window area will generate additional noise.^[35] In addition, too much input will significantly increase the amount of calculation in the training and prediction process of the network and increase training time. In the experiment, the input data window size was $W \times W$, and the step size was 1. The experimental performance of 9×9 , 11×11 , 13×13 , and 15×15 was evaluated, and the result showed that the optimal window size was 11×11 .

Table 1 shows the type of each layer of the New-Hybrid-CNN structure, the dimensionality of the output mapping, and the number of parameters. Among them, `separable_Conv2d` represents DSC, including deep convolution and pixel-by-pixel convolution.

**Table 1.** The New-Hybrid-CNN algorithm structure.

Layer(type)	Output Shape (height, width, depth, numbers of feature map) for 3D data	Parameters Number
input_1(InputLayer)	(11, 11, 190, 1)	0
conv3d_1(Conv3D)	(9, 9, 188, 8)	224
conv3d_2(Conv3D)	(7, 7, 186, 16)	3472
conv3d_3(Conv3D)	(5, 5, 184, 32)	13856
reshape	(5, 5, 5888)	0
separable_Conv2d_1(Depthwise separable Conv2D)	(3, 3, 64)	429888
Flatten	(576)	0
dense_1(Dense)	(256)	147712
Dropout(Dropout)	(256)	0
dense_2(Dense)	(4)	1028

For the established wolfberry data set, the total number of training parameters for New-Hybrid-CNN is 596180. All weights were randomly initialized and trained using the backpropagation algorithm with the Adam optimizer by using the softmax loss. Minbatch was 128. Since the training set was relatively small, the batch size was set to 100. Rectified linear unit (Relu) was used as the activation function. In order to prevent the model from over-fitting, the dropout method was used to temporarily discard the neural network unit from the network according to a certain probability. The dropout was defined as 0.3. According to the changes in accuracy and loss during training, the learning rate was chosen to be 0.001.

Support vector machine

SVM is a traditional supervised learning algorithm that can perform classification and regression analysis, widely used in pattern recognition algorithms for spectral data analysis.^[36] As for linearly separable problems, SVM seeks a hyperplane to maximize the distance between different categories of samples. As for the linearly inseparable issue, SVM maps the original data to a higher dimension space to make the samples linearly separable. SVM explores a set of hyperplanes in a new space to maximize the sample spacing.^[13] It introduces kernel functions to effectively process nonlinear data or indivisible linear data based on structural risk minimization. Appropriate kernel function can improve the performance of the model. Radial basis function (RBF) has been widely used in the field of spectral analysis.

In the experiment, 3D ($50 \times 50 \times 190$) cubes of four wolfberry origins were first mapped to a 2D (2500×190) matrix, and the 2D matrix was used as the input data. The row (representing pixel) in the matrix was defined as a sample, so there were 2500 samples of wolfberry from each place. Based on the Python scikit-learn library, the grid search method is used to determine the optimal parameters of the SVM (penalty coefficient “c” and kernel width “g”).

Evaluation metrics

The overall accuracy (OA), kappa coefficient (Kappa) evaluation measures were used to judge the classification performance of HSI. OA measures the number of correctly classified samples in the total test sample; Kappa is a measurement metric calculated by weighted measurement accuracy.

$$OA = \frac{\sum_{i=1}^n a_{ii}}{N} \times 100\% \quad (4)$$

where N is the number of all samples, n is the number of categories, and a_{ii} is the diagonal element of the corresponding confusion matrix. Kappa is a measure of the accuracy of classification. The expression is

$$Kappa = \frac{N \sum_{i=1}^n a_{ii} - \sum_{i=1}^n (a_{i+} a_{+i})}{N^2 - \sum_{i=1}^n (a_{i+} a_{+i})} \quad (5)$$

where a_{i+} is the sum of the i -th row, and a_{+i} is the row of the i column. In a word, the larger the value of OA and Kappa, the better the performance of the classification algorithm.

RESULTS AND DISCUSSION

Data analysis of wolfberry samples

Figure 6 shows the original spectra of wolfberries from four origins in the ROI. It can be seen that wolfberries from different geographical origins have similar spectral curves. There are absorption peaks at approximately 1000 nm, 1200 nm, and 1465 nm. The absorption peak near 1000 nm is due to the second vibration of the N-H bond in the protein or amino acid.^[37] The reason for the absorption peak near 1200 nm is the secondary stretching vibration of the C-H bond in protein, starch, or lipid,^[38,39] and the absorption peak near 1465 nm is the sensitive area of water absorption.^[40]

Through further analysis of the average spectral curves of wolfberry from different origins, as shown in Figure 7, the trends of the average spectral curves of wolfberry from four different origins are generally similar. Still, the reflectance values of the four curves are different. These differences are caused by internal factors such as different regions and climates, indicating that the origin of wolfberry can be classified by using reflectance spectra. Among them, the average spectral curves of Qinghai wolfberry and Xinjiang wolfberry overlap. Therefore, machine learning methods are used to identify the origin of wolfberry further.

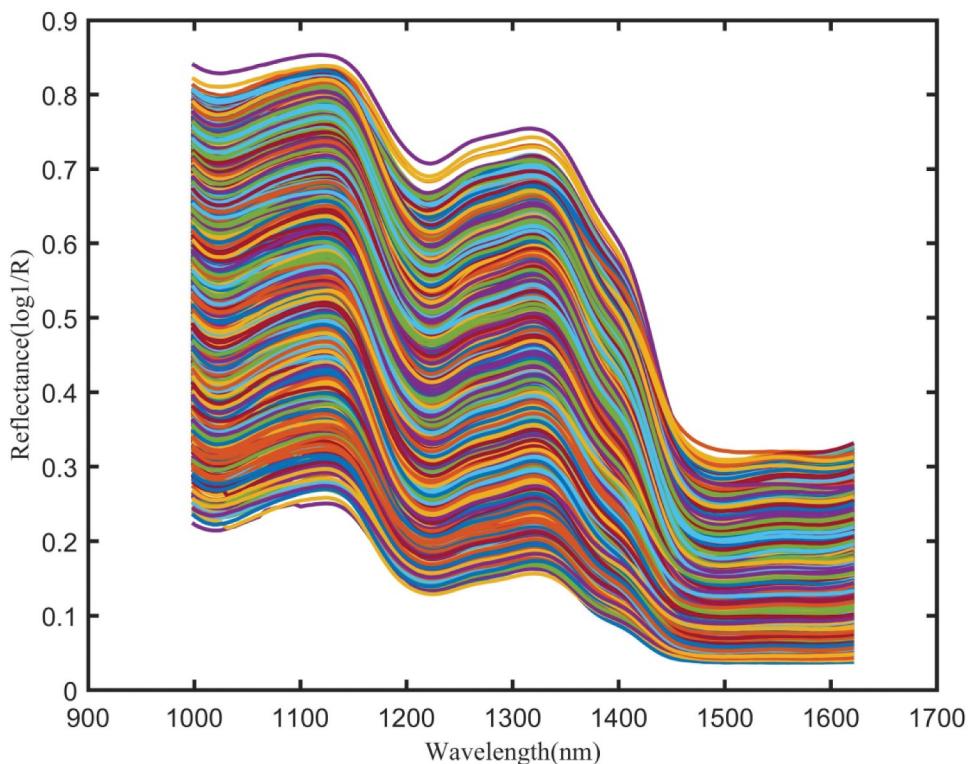


Figure 6. The spectra of wolfberries from different geographical origins in the range of 998–1622 nm.

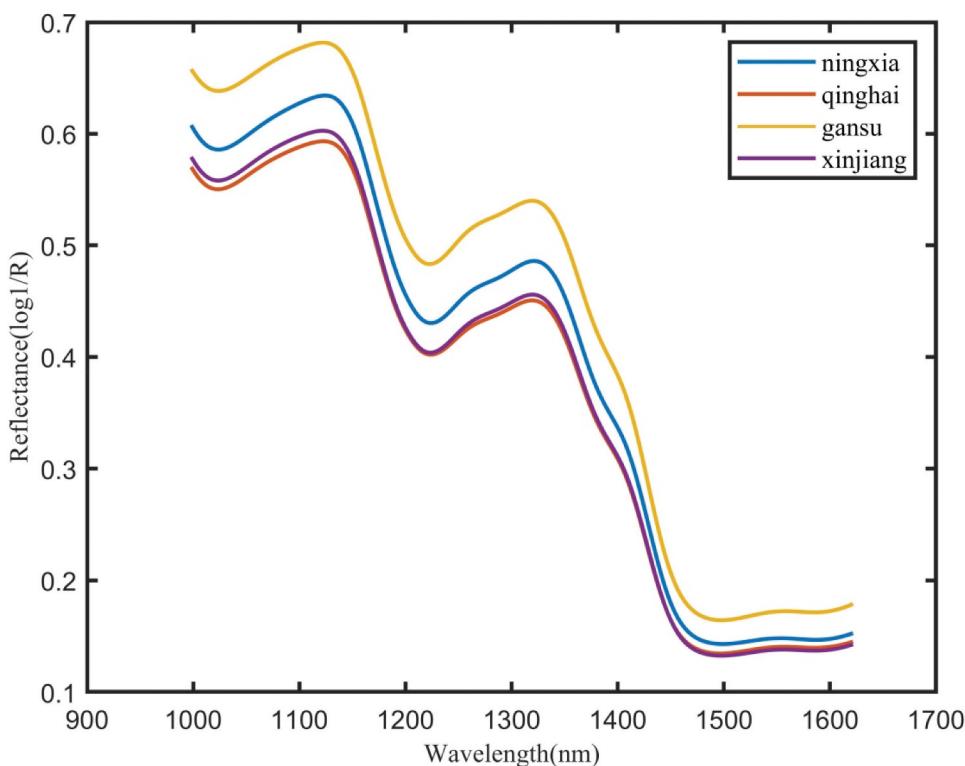


Figure 7. Mean reflectance spectra of wolfberries from different geographical origins in the range of 998–1622 nm.

Classification results

In the experiment, the New-Hybrid-CNN, HybridSN, 3D-CNN, and SVM parts were programmed in Python 3.8 and implemented based on Tensorflow and Keras open-source framework. The operating platform was on a PC with Intel (R) Core (TM) i5-5200 U CPU with 2.20 GHz and 4 GB RAM. All the classification algorithms were established using the full spectrum (998–1622 nm) and set the same parameters (window size, training sample, testing sample) for a fair comparison. Since the hyperspectral data of wolfberry from each origin was divided into 2500 samples during the experiment, the total number of wolfberry samples from the four origins was 10000. The number of samples of the training set, validation set, and test set input in different classification methods is shown in [Table 2](#).

First, for the SVM using full spectrum (998–1622 nm), the optimal parameters (c, g) were (1000, 0.01). Based on the SVM discriminant model of optimal parameters, the OA of the four origins is 95.36%, and the Kappa coefficient is 93.82%. A confusion matrix visualization method is used to compare the classification performance of the four classification algorithms on the test set, as shown in [Figure 8](#). It shows that the SVM has the worst performance capability. For the classification accuracy of wolfberry from different origins, the SVM classification algorithm has the highest correct recognition rate for Ningxia wolfberry, and the recognition rate is 99.7%. The other three classification algorithms

Table 2. The number of training, validation, and testing samples.

classification Method	Training samples	Validation samples	Testing samples
New-Hybrid-CNN	1400×4	350×4	750×4
HybridSN	1400×4	350×4	750×4
3D-CNN	1400×4	350×4	750×4
SVM	1400×4	350×4	750×4

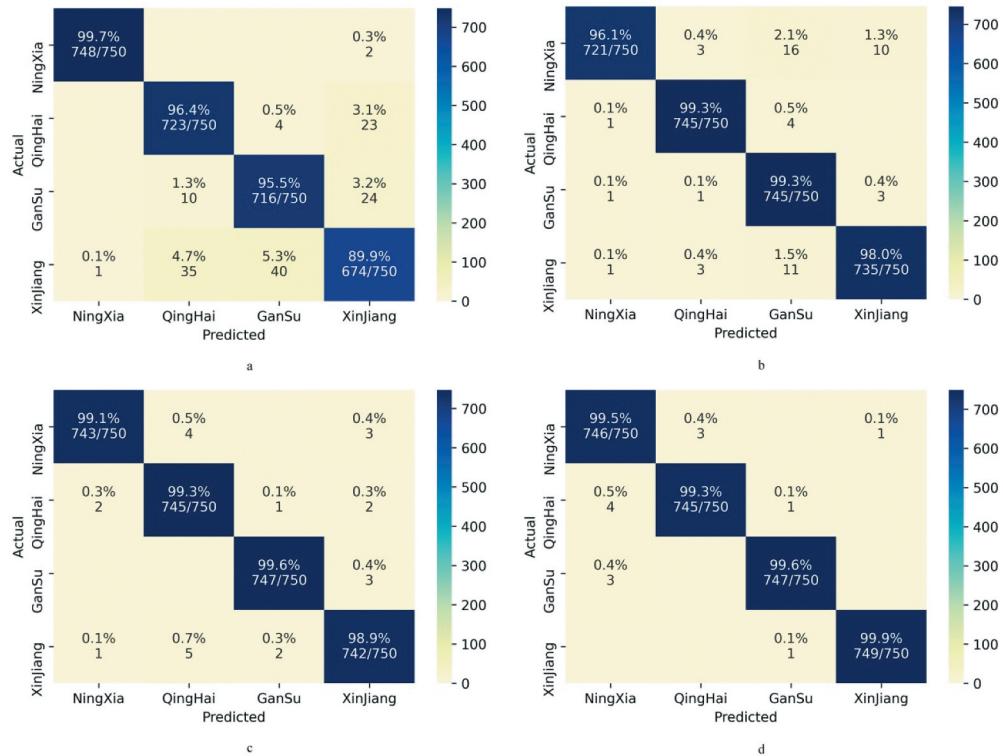


Figure 8. The confusion matrix of SVM (a), 3D-CNN (b), HybridSN (c) and New-Hybrid-CNN (d).

using spectral-spatial information have the highest recognition rate of Qinghai wolfberry. The HybridSN and New-Hybrid-CNN classification algorithms have the highest correct recognition rate of 99.6% for Gansu wolfberry. For Xinjiang wolfberry, the recognition rate of New-Hybrid-CNN is up to 99.9%.

Compared with the three classification algorithms based on spectral-spatial joint information, the classification accuracy of SVM is the worst. It shows that the classification method combining spectral-spatial joint information has improved classification performance than the classification method using spectral information.

Comparison of 3D-CNN, HybridSN and New-Hybrid-CNN

Then, the performance accuracy of the proposed algorithm was compared with the 3D-CNN and the HybridSN. The classification results are shown in Table 3. Under the same conditions as other variables, the OA and Kappa of the New-Hybrid-CNN classification algorithm are higher than the other two classification algorithms. Figure 9 summarizes the parameters and training time involved in the training of these three classification methods. It can clearly show that the proposed New-Hybrid-CNN has the least number of parameters, which is about 16% of the HybridSN method and 2.2% of

Table 3. Comparison of classification accuracy and computation cost with different classification algorithms.

classification Method	$W \times W$	Parameters	Train Times(s)	OA(%)	Kappa(%)
New-Hybrid-CNN	11×11	596,180	596.38	99.73	99.64
HybridSN	11×11	3,557,844	643.60	99.23	98.97
3D-CNN	11×11	26,911,188	1749.49	98.2	97.6

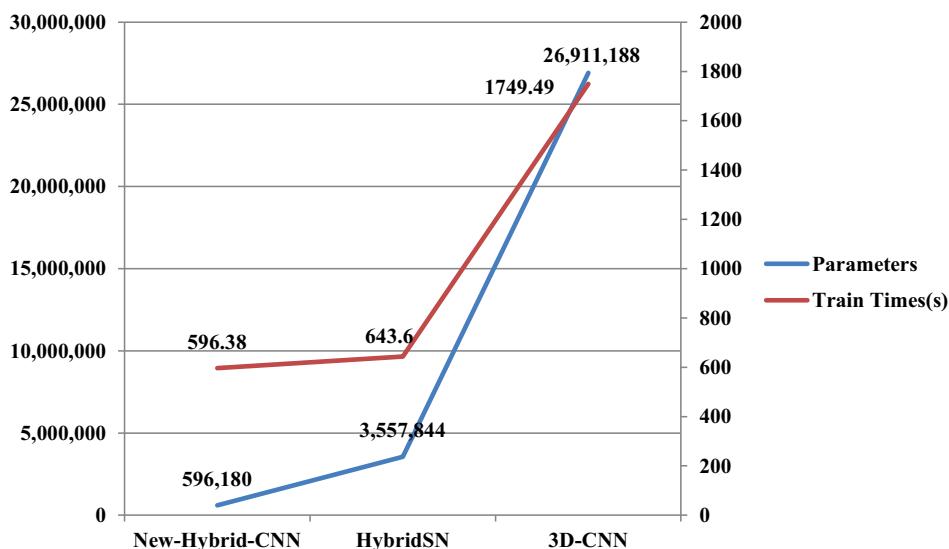


Figure 9. The parameters and train times of New-Hybrid-CNN, HybridSN, and 3D-CNN.

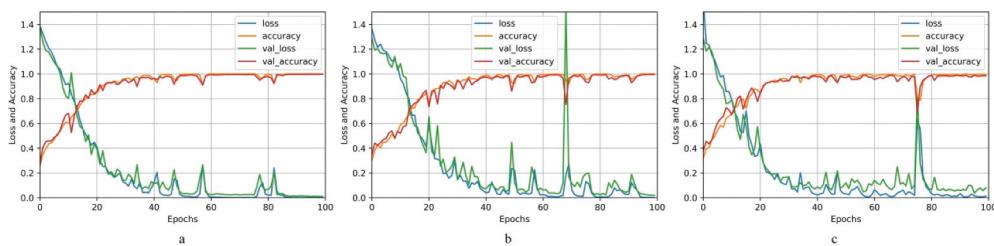


Figure 10. Loss and accuracy on the training and validation sets of New-Hybrid-CNN(a), HybridSN(b), and 3D-CNN(c).

3D-CNN. The reduction of parameters can effectively reduce the operating pressure of the computer. In addition, the training time is also the least, reduced by about 47 seconds compared to HybridSN, and significantly reduced by about 1153 seconds compared to 3D-CNN.

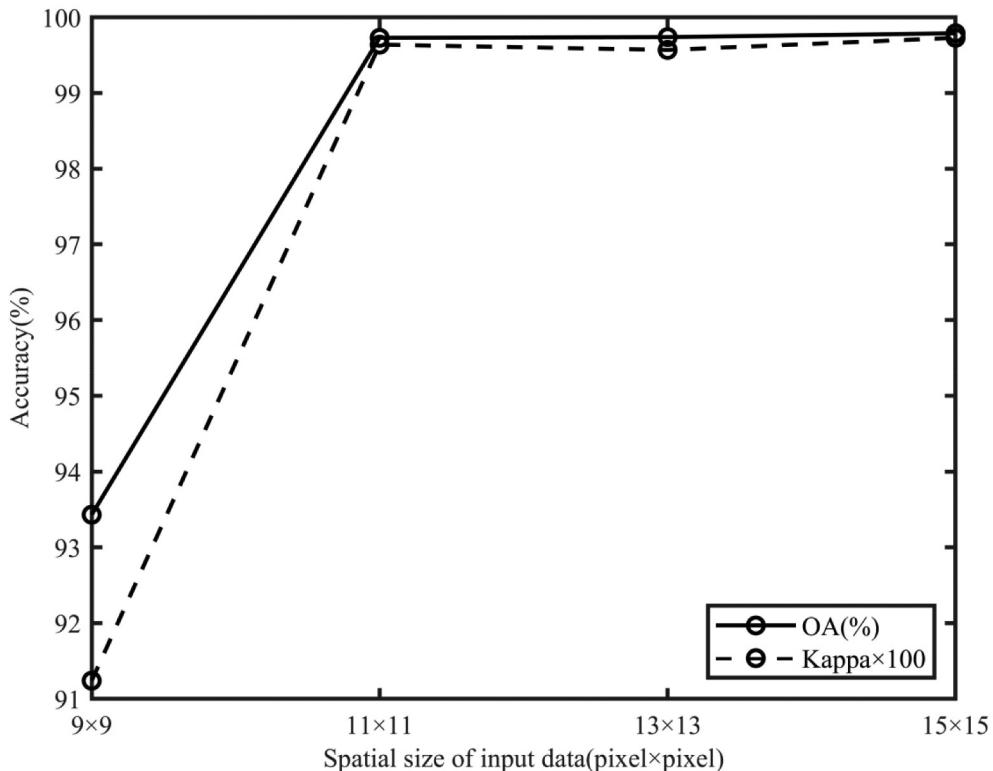
In order to further study of the stability of the three classification methods, the accuracy and loss rate curves of the training set and the validation set of the three different networks during the training process are compared. As shown in Figure 10, it can be seen from the figure that convergence begins when the epoch reaches about 40. After 100 iterations, the accuracy and loss of the model stabilize, and it is considered that the ideal classification accuracy has been achieved. Among them, the accuracy and loss rate curves of the training set and the verification set of Figure 10(a) are coincident with those of Figure 10(b) and Figure 10(c). There is no large-scale oscillation phenomenon. Therefore, the proposed New-Hybrid-CNN is stable. Moreover, it requires fewer parameters, a shorter time, and the highest classification accuracy, which has great practical application potential. Overall, the New-Hybrid-CNN classification algorithm shows the best classification performance.

Effect of input size of the sample

In order to find the best input size of training samples, we tested the proposed classification algorithm with input data of different sizes. The window sizes are set to 9×9 , 11×11 , 13×13 , 15×15 , and the test results are shown in Table 4. Figure 11 shows the effect of window size on OA and Kappa coefficients.

Table 4. The influence of window size on model classification performance.

Methord	$W \times W$	Parameters	Train Time(s)	OA(%)	Kappa(%)
New-Hybrid-CNN	9×9	456,108	278.8736	93.43	91.24
New-Hybrid-CNN	11×11	596,180	596.38	99.73	99.64
New-Hybrid-CNN	13×13	858,324	1039.1	99.74	99.57
New-Hybrid-CNN	15×15	1,251,540	1554.52	99.79	99.73

**Figure 11.** The influence of window size on OA and Kappa.

It can be seen from **Table 4** that as the data window becomes more extensive, the parameters and training time of the algorithm gradually increase. Moreover, **Figure 11** shows that an appropriate increase in the input data window size helps improve the classification performance. However, after the window size of 11×11 , the OA curve and the Kappa curve are gradually smoothed, which means that increasing the window size does not significantly improve the accuracy. In addition, the input window is too large to substantially increase the amount of calculation for network training, resulting in a longer training time. Based on the above experimental results, the window size of 11×11 can achieve better results.

CONCLUSION

In this study, hyperspectral imaging in the spectral region of 998–1622 nm was applied to identify the geographical origins of wolfberry by using four types of classifiers named New-Hybrid-CNN, HybridSN, 3D-CNN, and SVM. The SVM classifier was based on the spectral information of wolfberry samples. However, the New-Hybrid-CNN, HybridSN, and 3D-CNN were established on the spectral-spatial information of the samples. The experimental results revealed that the classifier



using spectral-spatial information was more effective than the SVM focuses on spectral features. In addition, the proposed New-hybrid-CNN classifier combined the advantages of homogeneous 3D convolution and DSC. Compared with the 3D-CNN and HybridSN, it had a lighter model, improved computational efficiency, and the highest classification accuracy. For future work, the New-Hybrid-CNN classification algorithm proposed in this paper is conducive to popularization and application to other samples. It helps to solve similar problems in food classification and agricultural applications.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Science and Technology Project of Hebei Education Department [ZD2018045]; Tianjin Enterprise Science and Technology Commissioner Project [18JCTPJC57500]; Tianjin Research Program of Application Foundation and Advanced Technology of China [15JCYBJC17100]; Tianjin Research Program of Application Foundation and Advanced Technology of China [15JCYBJC17100]; Tianjin Enterprise Science and Technology Commissioner Project [18JCTPJC57500];

References

- [1] Yang, Z.; Wang, Z.; Yuan, W.; Li, C.; Han, H. Classification of Wolfberry from Different Geographical Origins by Using Electronic Tongue and Deep Learning Algorithm. *IFAC-PapersOnLine*. **2019**, 52(30), 397–402. DOI: [10.1016/j.ifacol.2019.12.592](https://doi.org/10.1016/j.ifacol.2019.12.592).
- [2] Jin, M.; Huang, Q.; Zhao, K.; Biological Activities, S. P. Potential Health Benefit Effects of Polysaccharides Isolated from *Lycium Barbarum* L. *Int. J. Biol. Macromol.* **2013**, 54, 16–23. DOI: [10.1016/j.ijbiomac.2012.11.023](https://doi.org/10.1016/j.ijbiomac.2012.11.023).
- [3] Wang, Y. J.; Liang, X. J.; Guo, S. J.; Li, Y. K.; Zhang, B.; Yin, Y.; An, W.; Cao, Y. L.; Zhao, J. H. Evaluation of Nutrients and Related Environmental Factors for Wolfberry (*Lycium Barbarum*) Fruits Grown in the Different Areas of China. *Biochem. Syst. Ecol.* **2019**, 86, 103916. DOI: [10.1016/j.bse.2019.103916](https://doi.org/10.1016/j.bse.2019.103916).
- [4] Bondia-Pons, I.; Savolainen, O.; Toerroenen, R.; Martinez, J. A.; Poutanen, K.; Hanhineva, K. Metabolic Profiling of Goji Berry Extracts for Discrimination of Geographical Origin by Non-targeted Liquid Chromatography Coupled to Quadrupole Time-of-flight Mass Spectrometry. *Food Res. Int.* **2014**, 63(part B), 132–138. DOI: [10.1016/j.foodres.2014.01.067](https://doi.org/10.1016/j.foodres.2014.01.067).
- [5] Bleibaum, R. N.; Stone, H.; Tan, T.; Labreche, S.; Isz, S. Comparison of Sensory and Consumer Results with Electronic Nose and Tongue Sensors for Apple Juices. *Food Qual. Preference*. **2002**, 13(6), 409–422. DOI: [10.1016/S0950-3293\(02\)00017-4](https://doi.org/10.1016/S0950-3293(02)00017-4).
- [6] Meng, J.; Liu, Z.; Gou, C. L.; Rogers, K. M.; Yu, W. J.; Zhang, S. S.; Yuan, Y. W.; Zhang, L. Geographical Origin of Chinese Wolfberry (Goji) Determined by Carbon Isotope Analysis of Specific Volatile Compounds. *J chromatogr B*. **2019**, 1105, 104–112. DOI: [10.1016/j.jchromb.2018.12.011](https://doi.org/10.1016/j.jchromb.2018.12.011).
- [7] Mollazadeh, K.; Arefi, A. LightScatter: A Comprehensive Software Package for Non-destructive Monitoring of Horti-food Products by Monochromatic Imaging-based Spatially-resolved Light Scattering Technology. **2017**, 142(part B), 597–606. DOI: [10.1016/j.compag.2017.11.009](https://doi.org/10.1016/j.compag.2017.11.009).
- [8] Nie, P. C.; Zhang, J. N.; Feng, X. P.; Yu, C. L.; He, Y. Classification of Hybrid Seeds Using Near-infrared Hyperspectral Imaging Technology Combined with Deep Learning. *Sens. Actuators B Chem.* **2019**, 296, 126630.1–126630.12. DOI: [10.1016/j.snb.2019.126630](https://doi.org/10.1016/j.snb.2019.126630).
- [9] Goetz, A. F.; Vane, G.; Solomon, J. E.; Rock, B. N. Imaging Spectrometry for Earth Remote Sensing. *Science*. **1985**, 228(4704), 1147–1153. DOI: [10.1126/science.228.4704.1147](https://doi.org/10.1126/science.228.4704.1147).
- [10] Lack, M. B.; Aboonajmi, M. **2013**, Application of Spectrometry in Agricultural Practices. CSBE/SCGAB 2013 Annual Conference.
- [11] Qu, J. H.; Liu, D.; Cheng, J. H.; Sun, D. W.; Ma, J.; Pu, H. B.; Zeng, X. A. Applications of Near-infrared Spectroscopy in Food Safety Evaluation and Control: A Review of Recent Research Advances. *Crit. Rev. Food Sci. Nutr.* **2015**, 55(13), 1939–1954. DOI: [10.1080/10408398.2013.871693](https://doi.org/10.1080/10408398.2013.871693).
- [12] Mishra, P.; Nordon, A.; Tschanerl, J.; Lian, G.; Redfern, S.; Marshall, S. Near-infrared Hyperspectral Imaging for Non-destructive Classification of Commercial Tea Products. *J. Food Eng.* **2018**, 238, 70–77. DOI: [10.1016/j.jfooodeng.2018.06.015](https://doi.org/10.1016/j.jfooodeng.2018.06.015).

- [13] Hong, Z. Q.; He, Y. **2020**, Rapid Nondestructive Discrimination of Geographical Origins of Longjing Tea Using Hyperspectral Imaging at Two Spectral Ranges Coupled with Machine Learning Methods. *Appl. Sci.* **10**(3), 1173. DOI: [10.3390/app10031173](https://doi.org/10.3390/app10031173).
- [14] Chu, B.; Yu, K.; Zhao, Y.; He, Y. Development of Noninvasive Classification Methods for Different Roasting Degrees of Coffee Beans Using Hyperspectral Imaging. *Sensors*. **2018**, *18*(4), 1259. DOI: [10.3390/s18041259](https://doi.org/10.3390/s18041259).
- [15] Wang, Y. G.; Gao, Y.; Yu, X. Z.; Wang, Y. Y.; Deng, S.; Gao, J. M. Rapid Determination of Lycium Barbarum Polysaccharide with Effective Wavelength Selection Using Near-Infrared Diffuse Reflectance Spectroscopy. *Food Anal. Methods*. **2016**, *9*(1), 131–138. DOI: [10.1007/s12161-015-0178-7](https://doi.org/10.1007/s12161-015-0178-7).
- [16] Gao, Z. M.; Shao, Y. Y.; Xuan, G. T.; Wang, Y. X.; Liu, Y.; Han, X. Real-time Hyperspectral Imaging for the In-field Estimation of Strawberry Ripeness with Deep Learning. *Artificial Intell. Agric.* **2020**, *4*, 31–38. DOI: [10.1016/j.aiia.2020.04.003](https://doi.org/10.1016/j.aiia.2020.04.003).
- [17] Yan, L.; Pang, L.; Wang, H.; Xiao, J. Recognition of Different Longjing Fresh Tea Varieties Using Hyperspectral Imaging Technology and Chemometrics. *J. Food Process Eng.* **2020**, *43*, e13378. DOI: [10.1111/jfpe.13378](https://doi.org/10.1111/jfpe.13378).
- [18] Oliveri, P.; Malegori, C.; Casale, M.; Tartacca, E.; Salvatori, G. An Innovative Multivariate Strategy for Hsi-nir Images to Automatically Detect Defects in Green Coffee. *Talanta*. **2019**, *199*, 270–276. DOI: [10.1016/j.talanta.2019.02.049](https://doi.org/10.1016/j.talanta.2019.02.049).
- [19] Mishra, P.; Nordon, A.; Asaari, M. M.; Lian, G.; Fusing Spectral, R. S. Textural Information in Near-infrared Hyperspectral Imaging to Improve Green Tea Classification Modelling. *J. Food Eng.* **2019**, *249*, 40–47. DOI: [10.1016/j.jfoodeng.2019.01.009](https://doi.org/10.1016/j.jfoodeng.2019.01.009).
- [20] Pu, H.; Sun, D. W.; Ma, J.; Cheng, J. H. Classification of Fresh and Frozen-thawed Pork Muscles Using Visible and near Infrared Hyperspectral Imaging and Textural Analysis. *Meat Sci.* **2015**, *99*, 81–88. DOI: [10.1016/j.meatsci.2014.09.001](https://doi.org/10.1016/j.meatsci.2014.09.001).
- [21] Al-Sarayreh, M.; Reis, M. M.; Yan, W. Q.; 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](https://doi.org/10.1016/j.foodcont.2020.107332).
- [22] Li, Y.; Zhang, H. K.; Shen, Q. **2017**, Spectral-spatial Classification of Hyperspectral Imagery with 3D Convolutional Neural Network. *Remote Sens.* **91**, 67. DOI: [10.3390/rs9010067](https://doi.org/10.3390/rs9010067).
- [23] Mirzaei, S.; Hamme, H. V.; Khosravani, S. Hyperspectral Image Classification Using Non-negative Tensor Factorization and 3D Convolutional Neural Networks. *Signal Proces. Image Commu.* **2019**, *76*, 178–185. DOI: [10.1016/j.image.2019.05.004](https://doi.org/10.1016/j.image.2019.05.004).
- [24] Hamida, A. B.; Benoit, A.; Lambert, P.; Amar, C. B. **2018**, 3-D Deep Learning Approach for Remote Sensing Image Classification. *IEEE Trans. Geosci. Remote Sens.* **56**, 4420–4434. DOI: [10.1109/TGRS.2018.2818945](https://doi.org/10.1109/TGRS.2018.2818945).
- [25] Roy, S. K.; Krishna, G Dubey, G.R.; Chaudhuri, B. B **2019**, HybridSN: Exploring 3D-2D CNN Feature Hierarchy for Hyperspectral Image Classification. *IEEE Geosci. Remote Sens. Lett.* **17**, 277–281. DOI: [10.1109/LGRS.2019.2918719](https://doi.org/10.1109/LGRS.2019.2918719).
- [26] Luo, Y.; Zou, J.; Yao, C.; Li, T.; Bai, G. HSI-CNN: A Novel Convolution Neural Network for Hyperspectral Image. *IEEE*. **2018**. DOI: [10.1109/ICALIP.2018.8455251](https://doi.org/10.1109/ICALIP.2018.8455251).
- [27] Kassani, S. H.; Kassani, P. H.; Wesolowski, M. J.; Schneider, K. A.; Deters, R. **2019**. Depthwise Separable Convolutional Neural Network for Skin Lesion Classification. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT). DOI: [10.1109/ISSPIT47144.2019.9001790](https://doi.org/10.1109/ISSPIT47144.2019.9001790).
- [28] Feng, Y. Z.; Sun, D. W. Application of Hyperspectral Imaging in Food Safety Inspection and Control: A Review. *Crit. Rev. Food Sci. Nutr.* **2012**, *52*(11), 1039–1058. DOI: [10.1080/10408398.2011.651542](https://doi.org/10.1080/10408398.2011.651542).
- [29] Cong, S.; Sun, J.; Mao, H.; Wu, X.; Wang, P.; Zhang, X. **2018**, Non-destructive Detection for Mold Colonies in Rice Based on Hyperspectra and GWO-SVR. *J. Sci. Food Agric.* **98**(4), 1453–1459. DOI: [10.1002/jsfa.8613](https://doi.org/10.1002/jsfa.8613).
- [30] Sellami, A.; Abbes, A. B.; Barra, V.; Farah, I. R. **2020**, Fused 3-D Spectral-spatial Deep Neural Networks and Spectral Clustering for Hyperspectral Image Classification. *Pattern Recogn. Lett.* **138**, 594–600. DOI: [10.1016/j.patrec.2020.08.020](https://doi.org/10.1016/j.patrec.2020.08.020).
- [31] Zhang, B.; Zhao, L.; Zhang, X. **2020**, Three-dimensional Convolutional Neural Network Model for Tree Species Classification Using Airborne Hyperspectral Images. *Remote Sens. Environ.* **247**, 111938. DOI: [10.1016/j.rse.2020.111938](https://doi.org/10.1016/j.rse.2020.111938).
- [32] Yu, C.; Han, R.; Song, M.; Liu, C.; Chang, C. I. **2020**, A Simplified 2D-3D CNN Architecture for Hyperspectral Image Classification Based on Spatial-spectral Fusion. *IEEE J. Selected Topics Appl. Earth Observa. Remote Sens.* **13**, 2485–2501. DOI: [10.1109/JSTARS.2020.2983224](https://doi.org/10.1109/JSTARS.2020.2983224).
- [33] Ma, S.; Liu, W.; Wei, C.; Shang, Z.; Liu, G. Lightweight Deep Residual CNN for Fault Diagnosis of Rotating Machinery Based on Depthwise Separable Convolutions. *IEEE Access*. **2019**, *7*, 57023–57036. DOI: [10.1109/ACCESS.2019.2912072](https://doi.org/10.1109/ACCESS.2019.2912072).
- [34] Tran, D.; Bourdev, L.; Fergus, R.; Torresani, L.; Paluri, M. Learning Spatiotemporal Features with 3D Convolutional Networks. *IEEE*. **2015**, 4489–4497. DOI: [10.1109/ICCV.2015.510](https://doi.org/10.1109/ICCV.2015.510).
- [35] Zhang, M.; Li, W.; Du, Q. **2018**, Diverse Region-based CNN for Hyperspectral Image Classification. *IEEE Trans. Image Proces.* **27**, 2623–2634. DOI: [10.1109/TIP.2018.2809606](https://doi.org/10.1109/TIP.2018.2809606).

- [36] Luts, J.; Ojeda, F.; Plas, R. V. D.; Moor, B. D.; Huffel, S. V.; Suykens, J. A. K. A Tutorial on Support Vector Machine-based Methods for Classification Problems in Chemometrics. *Anal. Chim. Acta.* **2010**, *665*(2), 129–145. DOI: [10.1016/j.aca.2010.03.030](https://doi.org/10.1016/j.aca.2010.03.030).
- [37] Kafle, G. K.; Khot, L. R.; Jarolmasjed, S.; Si, Y. S.; Lewis, K. Robustness of near Infrared Spectroscopy Based Spectral Features for Non-destructive Bitter Pit Detection in Honeycrisp Apples. *Postharvest. Biol. Technol.* **2016**, *120*, 188–192. DOI: [10.1016/j.postharvbio.2016.06.013](https://doi.org/10.1016/j.postharvbio.2016.06.013).
- [38] Nagao, R.; Ishii, K.; Matsui, D.; Awazu, K. Quantitative Evaluation of Lipid Volume Fraction in Atherosclerotic Plaque Phantoms by Near-infrared Multispectral Imaging at Wavelengths around 1200 nm. *Adv. Biomed. Eng.* **2015**, *4*, 158–163. DOI: [10.14326/abe.4.158](https://doi.org/10.14326/abe.4.158).
- [39] Ishii, K.; Kitayabu, A.; Nagao, R.; Awazu, K. Observation of Atherosclerotic Plaque Phantoms through Saline or Blood Layers by Near-infrared Hyperspectral Imaging. *Opt. Photonics J.* **2014**, *4*, 271–279. DOI: [10.4236/opj.2014.410026](https://doi.org/10.4236/opj.2014.410026).
- [40] Meijde, M.; Knox, N. M.; Cundill, S. L.; Noomen, M. F.; Werff, H. M. A.; Hecker, C. Detection of Hydrocarbons in Clay Soils: A Laboratory Experiment Using Spectroscopy in the Mid- and Thermal Infrared. *Int. J. Appl. Earth Observ. Geoinform.* **2013**, *23*, 384–388. DOI: [10.1016/j.jag.2012.11.001](https://doi.org/10.1016/j.jag.2012.11.001).