**A Band Influence Algorithm for Hyperspectral Band Selection to Classify Moldy Peanuts**

**ZIWEI LIU**, **JINBAO JIANG**, **YING DU**, AND **ZHIFANG XU**

College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing 100083, China

Corresponding author: Jinbao Jiang (jjb@cumtb.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 41871341.

**ABSTRACT** Moldy peanuts are often found in harvested and stored peanuts. Aflatoxins in moldy peanuts pose a potential risk to food safety. Hyperspectral imaging techniques is often used for rapid nondestructive testing of food. However, the information redundancy of hyperspectral data has a negative effect on the processing speed and classification accuracy. In this study, a novel band selection method, namely the band influence algorithm (BIA), was proposed to extract key features and remove redundancy for the classification of moldy peanuts. Firstly, hyperspectral images of moldy, healthy and damaged peanuts were collected with 128 bands ranging from 400 nm to 1000 nm. Secondly, the BIA method was used to extract feature bands according to the influence of each band subset on the accuracy of the classification model. The effectiveness of BIA method was compared with five representative band selection methods on four classification models: decision tree (DT), k-nearest neighbor (KNN), support vector machine (SVM), and ShuffleNet V2. The experimental results show that the BIA method performs superior effect and stability than other methods on all classification models. The integration of BIA and ShuffleNet V2 achieved the best classification effect. Especially when using 10 feature bands on ShuffleNet V2, the average accuracy, F1 score, and kappa coefficient of BIA reached 97.66%, 0.977, and 0.963 respectively.

**INDEX TERMS** Hyperspectral image, band selection, band influence algorithm, classification, moldy peanuts.

**I. INTRODUCTION**

Although hyperspectral data contains abundant information, it also contains a lot of data redundancy. The large amount of data of hyperspectral images will not only occupy more computing resources, but also affect the timeliness of the application. Therefore, it is necessary to reduce the amount of data on the basis of ensuring accuracy. Both feature selection and feature extraction can reduce the amount of data. Because the method based on feature extraction needs to transform the data, it takes more time than the method of feature selection and the key discriminant information may be damaged. Therefore, the focus of this paper is feature selection.

Feature selection, or band selection, is to select the feature band with the highest correlation with the target in the data by some method. According to learning strategies, it can be divided into three categories: supervised [1], [2], semi-supervised [3]–[5] and unsupervised [6]–[9]. In recent research, band selection methods mainly use the following theories.

The first category is based on search or goal optimization [7], [10]–[14]. These methods are usually optimized on the basis of existing search algorithms such as artificial bee colony, cuckoo search method and gravitational search algorithm, or a new search algorithm is proposed to find the effective feature bands. However, the process of search or optimization is often complex.

Band selection methods based on rough set theory [15]–[17] and information theory [18]–[20] have made new developments in recent years. The essence of rough set theory based method is to apply the idea of mathematical set to hyperspectral band selection. In the method based on information theory, researchers usually use mutual information, entropy and other information metrics combined with spectral and spatial information to select feature bands.

Compared with the common methods, the method based on deep learning has gained more research enthusiasm. The band selection methods using deep learning mainly include convolutional neural network-based [3], [21]–[24], attention...
mechanism [25]–[28] and autoencoder-based [27], [29]. For example, Feng et al. [21] proposed a method of using band wise independent convolution and hard threshold. The convolution weight of the unselected band was reset to zero through the threshold, and a network model was constructed for hyperspectral classification. Zhan et al. [23] proposed a band selection method based on convolutional neural network (CNN) and distance density. In this method, the band combination was randomly constructed according to the distance density, and the best band was selected according to the classification accuracy. In addition, Lorenzo et al. [28] performed hyperspectral image classification by combining attention-based CNN and anomaly detection. Singh & Karthikeyan [29] combined auto-encoders and genetic algorithm to select the most relevant bands from hyperspectral images.

Overall, the method based on deep learning generally constructs the band selection method from the aspects of convolution structure, feature extraction block or network structure. These methods are usually model-driven and need to build a new band selection model, but the model design is generally complex.

Furthermore, researchers have applied band selection technology to food detection. For example, Shuaibu et al. [30] proposed an unsupervised feature selection method based on orthogonal subspace projection (OSP) for Apple Marssonina blotch detection. Deng et al. [31] proposed a band selection method based on entropy distance and sequence backward selection for the detection of citrus Huanglongbing. Yuan et al. [32] fused genetic algorithm and successive projection algorithm for band selection to identify moldy peanuts. These methods usually construct specific band selection methods for specific research objects. However, traditional classification models such as k-nearest neighbor (KNN) [33] and support vector machine (SVM) [34] were usually used to verify the effectiveness of the band selection method. There is a lack of verification research for band selection method on deep learning classification model.

In this paper, different combinations of band selection method and classification model were used to test the recognition performance of peanuts. The research objectives of this paper were to: (1) explore a simple and adaptable band selection method for peanut feature extraction; (2) combine hyperspectral band selection and classification model to identify moldy peanuts; (3) compare the effect of the proposed method with other methods on different classification models including deep learning model.

II. MATERIALS AND METHODS
A. EXPERIMENTAL SAMPLE PREPARATION
Two batch of Haihua No.1 peanuts were purchased from the market. One batch was healthy peanuts, and another batch was manually screened residues. As shown in the lower part of Fig. 1, peanuts were divided into 3 classes with 6 subclasses. Among them, healthy peanuts are healthy peanuts with intact kernel. Because damaged peanuts often appear in actual production, the category of damaged peanuts was added to the experiment. Damaged peanuts include damaged seed coat and partial kernel, as shown in damaged-a and damaged-b, respectively. Moldy peanuts include intact moldy, damaged moldy and moldy partial kernel, as shown in moldy-a, moldy-b and moldy-c, respectively.

Aflatoxin B1 rapid test card (Shenzhen Finder Biotech Co., Ltd) was used to detect whether peanuts were moldy. Nine samples were taken from each kinds of peanuts and divided into three groups. Each group of peanuts was determined after grinding. The test paper showed that the content of aflatoxin in moldy peanut groups was more than 20 µg kg\(^{-1}\) (ppb), but not in healthy and damaged peanut groups. According to the national food safety standard (GB 2761-2017) of China, the content of aflatoxin B1 in peanuts and its products should not exceed 20 µg kg\(^{-1}\) (ppb). We think that moldy peanuts are contaminated by fungi. In addition, peanut samples were dried in a constant temperature oven at 60°C for 2 hours to remove the potential effect of peanut skin moisture on the spectrum.

B. DATA ACQUISITION AND PREPROCESSING
As shown in Fig. 1, the SOC710E portable hyperspectral imaging system was used to collect hyperspectral images. The wavelength range of the instrument is 400-1000 nm with the spectral resolution of 2.34 nm. The hyperspectral images were collected outdoors in sunny weather from 10 am. to 1 pm. In image acquisition, the black rubber belt was used as the background, and the gray panel was placed in the upper left corner of the field of view. No peanut kernels were scanned repeatedly. Finally, sixteen hyperspectral images with the size of 1040 × 1392 × 128 were obtained.
In image preprocessing, the pixels of gray panel was used to perform spectral correction and convert the digital value of the original image into reflectance. In the label making part, the mask of each image was made. The appropriate threshold was selected to segment peanut kernels. Next, the inaccurate peanuts were modified by manual labeling. The mask images were obtained after denoising. Then, the mask images was used to extract peanut kernels based on watershed segmentation algorithm. For each kernel, the background was set to zero, and the images were padded to $104 \times 104$. Finally, 2172 kernels of three kinds of peanuts were obtained (Table 1).

### TABLE 1. Experimental data.

| Data    | Subclass   | Image | Kernel |
|---------|------------|-------|--------|
| Healthy | Healthy    | 3     | 398    |
| Damaged | Damaged-a  | 2     | 227    |
|         | Damaged-b  | 2     | 302    |
| Moldy   | Moldy-a    | 3     | 454    |
|         | Moldy-b    | 2     | 265    |
|         | Moldy-c    | 2     | 260    |
| Mixed   | Mixed      | 2     | 266    |
| Total   |            | 16    | 2172   |

where $nd_i (i = 2, 3, \ldots, b - 1)$ indicates the subset that three consecutive bands of $(i - 1), i, (i + 1)$ in full-band data are zeroed, whereas the values of other bands remain unchanged. $ND_{s3}$ can be input into the trained model and generate the classification accuracy set $ACC$. The accuracy of $nd_i$ is regarded as the accuracy of the $i$ band, and the bands without accuracy is set to one.

In this way, the lower the accuracy of the band, the higher the influence of the band on the classification. Inspired by partition strategy, the spectral range is divided into $P$ parts, and the feature bands are extracted from each part. The interval value is initialized as:

$$\text{Interval}_n = \begin{cases} 
\frac{b}{P} \times n, & n = 0, 1, \ldots, P - 1 \\
\frac{b}{P}, & n = P. 
\end{cases}$$

(2)

The minimum accuracy cannot be too large, otherwise the accuracy set cannot well reflect the band influence. Therefore, if the minimum accuracy is greater than the threshold $T$, the ND is redefine as:

$$ND_{s5} = nd_3, nd_4, \ldots, nd_j, \ldots, nd_{b-2}$$

(3)

In the task of classification or recognition, it is usually necessary to train the classification model. For full-band hyperspectral data $x_1, x_2, \ldots, x_b$, $b$ being the total number of bands, a new dataset ND is define as:

$$ND_{s3} = nd_2, nd_3, \ldots, nd_i, \ldots, nd_{b-1}$$

(1)
Algorithm 1 Proposed BIA Method

Input: Full-band hyperspectral data. The number of feature bands required RB. The number of partitions P. The threshold of minimum accuracy T.

1. Train the classification model with full-band data.
2. Generate ACC by inputting ND_{13} to the trained model.
3. if min(ACC) > T then
   4. Regenerate ACC by inputting ND_{15} to the trained model;
   5. end if
6. Smooth ACC; Initialize and fine-tune Interval_{n}.
7. Calculate SD by (4).
8. for p = 1 to P do
9.   if length (LM_{p}) \geq RB_{p} then
10.      sort(LM_{p}, ascending);
11.      select RB_{p} bands;
12.   else
13.      select RB_{p} bands by (6);
14.   end if
15. end for
16. Merge the selected feature bands in each part.
17. Output: RB selected bands.

where nd\_j(j = 3, 4, . . . , b – 2) indicates the subset that five consecutive bands of (j – 2), (j – 1), j, (j + 1), (j + 2) in the original data are zeroed, whereas the values of other bands remain unchanged. Then the new accuracy set is generated by ND_{15}.

The local minimum bands of each part in ACC_{p} are regarded as feature bands, namely LM_{p}, and sorted by ascending. There may be cases where adjacent values are equal in ACC, so data smoothing is used to smooth these values. In order to prevent the feature band from being on the divided interval, each interval is fine-tuned to the local maximum band in the accuracy set closest to the original interval value.

Next, it is need to calculate the number of bands extracted in each part. It is known that the difference between classes is the key information for classification. By calculating the average spectrum of each class, the absolute value of spectral difference SD between classes is obtained and summed up.

\[ SD = \sum_{i}^{c} |S_{ci} - S_{cj}| \]  \hspace{1cm} (4)

where S_{ci} and S_{cj} are the average spectra of two classes of peanuts.

The number of feature bands required is defined as RB. For each part, the number of feature bands to be extracted can be expressed as:

\[ RB_{p} = \frac{Interval_{p+1} SD}{\int_{p}^{b} SD} \times RB \]  \hspace{1cm} (5)

where p is the p-th part. For each part, if the number of LM_{p} is not less than RB_{p}, RB_{p} bands were selected as feature bands FB_{p}. Otherwise, the missing feature bands MFB is extracted in ACC_{p} by ascending with the step size of two and appended to LM_{p}. After removing the same band, RB_{p} bands are selected as the feature bands.

\[ FB_{p} = (LM_{p} \cup MFB) - (LM_{p} \cap MFB) \]  \hspace{1cm} (6)

The feature bands selected by each part FB_{p} are merged as the final selected feature bands. The whole process of the proposed BIA is illustrated in Algorithm 1.

2) METHODS FOR COMPARISON

In order to verify the effectiveness of different types of band selection methods in identifying moldy peanuts, five representative methods were selected for comparison. BS-Net [22] is an unsupervised band selection method based on CNN. Each band is given a weight, and the feature band is selected according to the weight. The version of fully connected networks (BS-Net-FC) was selected for comparison. EGCSR [35] combines the maximum ellipsoidal volume and orthogonal projection for band selection. The ranking-based strategy (EGCSR-R) was adopted for comparison. GSM [24] is a supervised learning model based on 1D-CNN. The feature band is extracted from the class saliency map generated by the middle layer of the trained model. MVPICA [36] is a classical band selection method. It constructs the load factor matrix for band selection by decomposing the eigenvalues and Eigenvectors obtained by the spectral matrix. SpaBS [37] is a band selection method based on sparse representation. In this method, k-means singular value decomposition algorithm [38] is used to decompose the image and the generated coefficient matrix histogram are used to extract the feature band.

D. EVALUATION METHOD

1) EVALUATION METRICS

In this paper, the average accuracy (AA), Kappa coefficient (Kappa) and F1 score (F1) were used as the evaluation metric. Their calculation formulas are as follows:

\[ Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \]  \hspace{1cm} (7)

\[ AA = \frac{1}{n} \sum_{i=1}^{n} Accuracy_{i}, i \in \{1, n\} \]  \hspace{1cm} (8)

\[ F1 = \frac{2TP}{2TP + FP + FN} \]  \hspace{1cm} (9)

where n represents the peanut category and n = 3. True positive (TP) indicates that positive samples are predicted as positive samples; false positive (FP) indicates that negative samples are predicted as positive samples; true negative (TN) indicates that positive samples are predicted as negative samples; false negative (FN) indicates that negative samples are predicted as negative samples.

\[ Kappa = \frac{p_0 - pe}{1 - pe} \]  \hspace{1cm} (10)

where p_0 and pe are the predicted and expected agreement, respectively, and −1 \leq Kappa \leq 1.
Z. Liu et al.: Band Influence Algorithm for Hyperspectral Band Selection

FIGURE 3. Average accuracy (AA) of different band selection methods on different validation models.

2) CLASSIFICATION MODELS
In order to verify the generality of band selection methods on different types of classification models, Decision tree (DT) [39], KNN, SVM and ShuffleNet V2 [40] were used to verify the effect of different band selection methods. Among them, DT, KNN and SVM are classic classifiers used to evaluate the effect of band selection [21], [25], [28], [41]. ShuffleNet V2 is a lightweight network that improves the computing speed on the basis of ensuring accuracy. The volume of peanuts is huge, which puts forward high requirements for the recognition speed of the model. Therefore, the lightweight model was selected to verify the effect of the band selection methods.

III. RESULTS AND DISCUSSION
A. EXPERIMENTAL CONFIGURATION
The experiment was implemented on Intel(R) Core(TM) i5-10400 CPU with 2.90GHz, 64G RAM and NVIDIA GeForce RTX 2060 Super GPU with 8G RAM. Sklearn package was used to call the DT, KNN and SVM algorithm. Pytorch 1.9.0 was adopted in ShuffleNet V2. One may need to refer to the demo code in Github (https://github.com/mepleleo/BIA-bandselection).

For the parameters of band selection methods, the P and T of BIA were set to 3 and 85% respectively. Because the BIA method requires the use of the classification model, the corresponding parameters are the same as the classification model. In EGCSR-R, the λ and k were set to 1e4 and 3 respectively. In SpaBS, the sparsity level was set to 0.5. In addition, the parameters of other compared band selection methods were set by the default settings or recommended settings.

For the parameters of classification models, the entropy was adopted as the function in DT to measure the quality of a split. In KNN, the number of neighbors was set to 5. In SVM, the radial basis function was set as the kernel function. The grid search was used to find the best hyper-parameters C and gamma. After many attempts, the ShuffleNet V2 model can be well fitted when the training reaches 50 epochs, so 50 epochs were set for training and the batch size was set to 32.

All peanut kernels in the mixed image were used for testing. In other images, the ratio of training set and test set was 1:3, and the data was randomly selected. Finally, 476 kernels were prepared for training and 1696 kernels were prepared for testing. The average spectrum of peanut kernels was calculated as a training unit. Moreover, some data enhancement strategies were adopted in ShuffleNet V2, including rotation and flip. In this way, the training data was expanded to 1904 images. In this classification task, these data were sufficient for model training. All processes were run independently 10 times to enhance the stability of the results, and the average value was taken as the experimental results.

B. CLASSIFICATION RESULTS
We verified the result that the number of feature bands ranges from 4 to 30 with an interval size of 2. At the same time, the results of all bands were also used for comparison. Fig. 3 shows the AA of different band selection methods. Firstly, in terms of the classification results of different band selection methods, BIA achieved the best overall effect on the four validation models. On DT and SVM models, BIA can
achieve the classification accuracy of all bands when using eight feature bands. BIA showed the most obvious superiority over other band selection methods on KNN model. Although BS-Net-FC performed competitive results on SVM, BIA shown better result when using less than 14 feature bands. When 10 feature bands were used on ShuffleNet V2, the AA of BIA reached 97.66%, which was 99.54% of the accuracy of all bands. In addition, MVPCA obtained a pool classification result on four validation models.

Secondly, in terms of the stability of the results, large fluctuations mean poor stability. BIA presented superior stability with less volatility in all classification models. BS-Net-FC and SpaBS reflected pool stability on DT and

FIGURE 4. F1 score (F1) of different band selection methods on different validation models.

FIGURE 5. Kappa coefficient (Kappa) of different band selection methods on different validation models.
SVM respectively. Although RGCSR-R and GSM had good stability, their band selection accuracy was low.

Thirdly, in terms of classification accuracy of different classification models, the classification performance of the four models from high to low is ShuffleNet V2, SVM, KNN, and DT. Especially when using 10 feature bands on ShuffleNet V2, the AA of BIA exceeded the classification accuracy of all bands selection methods on DT, KNN, and SVM. Because ShuffleNet V2 can make full use of spectral and spatial information, it can achieve better classification results than other classifiers. DT and SVM can achieve the classification accuracy of all bands when using a small number of bands, which showed that these two models were more likely affected by the Hughes phenomenon. Through the combination of BIA band selection method and ShuffleNet V2 classification model, the best classification effect was achieved on peanuts dataset.

In Fig. 4 and Fig. 5, the curve of F1 and Kappa are similar to AA, which verifies the band selection effect of the proposed method again. When using 10 feature bands on DT, SVM, and ShuffleNet V2, the F1 and Kappa of proposed BIA reached local maximum. Moreover, the F1 and Kappa of proposed BIA on ShuffleNet V2 reached 0.977 and 0.963 respectively.

With the increase of the number of feature bands, the difference of classification results between different band selection methods tended to be smaller. However, the proposed BIA still maintained the optimal overall result. All in all, the proposed method achieved the best classification performance in the result of these four classification models.

C. BAND RESPONSE CHARACTERISTICS

The spectral information of peanuts is shown in Fig. 6 (a) and (b). According to the average spectra of three kinds of peanuts in Fig. 6 (a), it can be find that the key spectrum to distinguish damaged peanuts is around 450nm. The key spectrum to distinguish moldy peanuts is around 800nm. The key spectrum to distinguish healthy peanuts is around 980nm. In Fig. 6 (b), the spectral difference between classes shows the comprehensive weight of all classes. The weight of spectral bands tend to be locally maximum at 450nm and 760nm.

Fig. 6 (c) – (f) present the accuracy curve generated by accuracy set in BIA. Although the accuracy curves are different on different classification models, the key bands near 450 nm, 800 nm, and 980 nm are well reflected and included in the accuracy curves. The accuracy reduction of...
all classification models is in the range of 0-50%. Moreover, there are some unique characteristics on each model. When using DT as the classifier, the classification accuracy of 760nm to 830nm was reduced by about 50%, whereas the accuracy of other positions was reduced by less than 20%. When using KNN and SVM as classifiers, the classification accuracy was reduced in varying degrees. The accuracy curve fluctuates most when using ShuffleNet V2. This shows that the influence of different bands on the accuracy is different, which reflects the importance of the band. Meanwhile, it also illustrates that the method of zeroing the band subset is effective.

Fig. 6 (g) – (j) show the example of the feature bands used by DT, KNN, SVM, and ShuffleNet V2. Specifically, EGCSR-R, GSM and SpaBS extracted a large number of adjacent bands, whereas the bands selected by BIA were scattered and contained the key wavelengths mentioned above. It is known that the information of hyperspectral adjacent bands is similar. If the selected bands are too concentrated, a large amount of repeated information will be used, which is not conducive to classification. At the same time, the classification accuracy also illustrates this issue. Especially the classification results of MVPCA clearly reflect it.

D. VISUALIZATION OF CLASSIFICATION RESULTS

As illustrated in Fig. 7, the main type of misrecognition was that healthy peanuts were misidentified as damaged peanuts. Combined with Fig. 6 (a), it can be seen that the spectral reflectance differentiation of damaged peanuts and healthy peanuts is relatively low after 600nm. If the selected feature bands are mainly concentrated in this interval, it is more likely to lead to misclassification of healthy peanuts and damaged peanuts. In addition, the average spectrum of some slightly damaged peanuts may be closer to the healthy peanuts. This is also an important factor leading to misclassification. Combining all the image classification results, the proposed band selection method shows superior classification results on all classification models.

In actual production, peanut kernels can be located by peanut coordinates to realize rapid non-destructive testing. Furthermore, the shape of peanuts are still quite different after sieve screening. Watershed algorithm is a threshold-based segmentation algorithm, which leads to the need to adjust
the threshold to extract peanut kernels. This is disadvantageous to the image segmentation in complex situations. If the screening is carried out on the conveyor belt, a new or improved method is needed to optimize this process, so that the method combining hyperspectral band selection and deep learning classification technologies can be further applied. If the screening is carried out in a falling scene in a color sorter, peanut kernels are mostly independent. There is no need to worry about image segmentation.

IV. CONCLUSION

In this paper, a band influence algorithm for hyperspectral band selection was proposed to identify moldy peanuts. In BIA, the feature bands were extracted by the influence of each band on the classification accuracy. The effectiveness of this method was verified on four classification models of DT, KNN, SVM and ShuffleNet V2. The experimental results show that the proposed method achieves the best results than other advanced band selection methods. The best classification effect was achieved by the combination of BIA and ShuffleNet V2 classification model. Specially, when using 10 feature bands on ShuffleNet V2, the AA, F1, and BIA and ShuffleNet V2 classification model. Specially, when using 10 feature bands on ShuffleNet V2, the AA, F1, and classification accuracy. The effectiveness of this method was verified on four classification models of DT, KNN, SVM and ShuffleNet V2. The experimental results show that the proposed method achieves the best results than other advanced band selection methods. The best classification effect was achieved by the combination of BIA and ShuffleNet V2 classification model.

REFERENCES

[1] C. He, Y. Zhang, and D. Gong, “A pseudo-label guided artificial bee colony algorithm for hyperspectral band selection,” Remote Sens., vol. 12, no. 20, p. 3546, Oct. 2020.
[2] A. Paul and N. Chaki, “Supervised data-driven approach for hyperspectral band selection using quantization,” Geosci. Int., early access, pp. 1–11, Sep. 2020, doi: 10.1007/10106049.2018.122929.
[3] A. Sellami, M. Farah, I. R. Farah, and B. Solaiman, “Hyperspectral imagery classification based on semi-supervised 3D deep neural network and adaptive band selection,” Expert Syst. Appl., vol. 129, pp. 246–259, Sep. 2019.
[4] X. Cao, C. Wei, Y. Ge, J. Feng, J. Zhao, and L. Jiao, “Semi-supervised hyperspectral band selection based on dynamic classifier selection,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 12, no. 4, pp. 1289–1298, Apr. 2019.
[5] X. Cao, Y. Ji, T. Liang, Z. Li, X. Li, J. Han, and L. Jiao, “A semi-supervised spatially aware wrapper method for hyperspectral band selection,” Int. J. Rem. Sens., vol. 39, no. 12, pp. 4020–4039, Jun. 2018.
[6] R. Vaddi and P. Manoharan, “CNN based hyperspectral image classification using unsupervised band selection and structure-preserving spatial features,” Inf. Phys. Technol., vol. 110 Nov. 2020, Art. no. 103457.
[7] W. Chen, Z. Yang, J. Ren, J. Cao, N. Cai, H. Zhao, and P. Yuen, “MIMN-DPP: Maximum-information and minimum-noise determinant point processes for unsupervised hyperspectral band selection,” Pattern Recognit., vol. 102, Jun. 2020, Art. no. 107213.
[8] S. S. Sawant and P. Manoharan, “Unsupervised band selection based on weighted information entropy and 3D discrete cosine transform for hyperspectral image classification,” Int. J. Rem. Sens., vol. 41, no. 10, pp. 3948–3969, May 2020.
[9] W. Zhang, X. Li, Y. Dou, and L. Zhao, “A geometry-based band selection approach for hyperspectral image analysis,” IEEE Trans. Geosci. Remote Sens., vol. 56, no. 8, pp. 4318–4333, Aug. 2018.
[10] Q. Wang, F. Zhang, and X. Li, “Hyperspectral band selection via optimal neighborhood reconstruction,” IEEE Trans. Geosci. Remote Sens., vol. 58, no. 12, pp. 8465–8476, Dec. 2020.
[11] I. Tschannen, J. Ren, P. Yuan, G. Sun, H. Zhao, Z. Yang, Z. Wang, and S. Marshall, “MIMR-DGSA: Unsupervised hyperspectral band selection based on information theory and a modified discrete gravitational search algorithm,” Inf. Fusion, vol. 51, pp. 189–200, Nov. 2019.
[12] S. Sawant and P. Manoharan, “Hyperspectral band selection based on metaheuristic optimization approach,” Infr. Phys. Technol., vol. 107, Jun. 2020, Art. no. 103295.
[13] F. Xie, F. Li, C. Lei, J. Yang, and Y. Zhang, “Unsupervised band selection based on artificial bee colony algorithm for hyperspectral image classification,” Appl. Soft Comput., vol. 75, pp. 428–440, Feb. 2019.
[14] Z. Yong, H. Chun-Lin, S. Xian-Fang, and S. Xiao-Yan, “A multi-strategy integrated multi-objective artificial bee colony for unsupervised band selection of hyperspectral images,” Swarm Evol. Comput., vol. 60, Feb. 2021, Art. no. 100806.
[15] S. Patra and B. Barman, “A novel dependency definition exploiting boundary samples in rough set theory for hyperspectral band selection,” Appl. Soft Comput., vol. 99, Feb. 2021, Art. no. 106944.
[16] A. Ghorbanian, Y. Maghoush, and A. Mohammadzadeh, “Clustering-based band selection using structural similarity index and entropy for hyperspectral image classification,” Traitement du Signal, vol. 37, no. 5, pp. 785–791, Nov. 2020.
[17] R. M. Torres, P. W. T. Yuen, C. Yuan, J. Piper, C. McCullough, and P. Godfrey, “Spatial spectral band selection for enhanced hyperspectral remote sensing classification applications,” J. Imag., vol. 6, no. 9, p. 87, Aug. 2020.
[18] J. Feng, J. Chen, Q. Sun, R. Shang, X. Cao, X. Zhang, and L. Jiao, “Convolutional neural network based on bandwise-independent convolution and hard thresholding for hyperspectral band selection,” IEEE Trans. Cybern., vol. 51, no. 9, pp. 4414–4428, Sep. 2021.
[19] Y. Cai, X. Liu, and Z. Cai, “BS-Net: An end-to-end framework for band selection of hyperspectral image,” IEEE Trans. Geosci. Remote Sens., vol. 58, no. 3, pp. 1969–1984, Mar. 2020.
[20] Y. Chen, D. Hu, H. Xing, and X. Yu, “Hyperspectral band selection based on deep convolutional neural network and distance density,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 12, pp. 2365–2369, Dec. 2017.
[21] R. M. Torres, P. W. T. Yuen, C. Yuan, J. Piper, C. McCullough, and P. Godfrey, “Spatial spectral band selection for enhanced hyperspectral remote sensing classification applications,” J. Imag., vol. 6, no. 9, p. 87, Aug. 2020.
[22] J. Feng, J. Chen, Q. Sun, R. Shang, X. Cao, X. Zhang, and L. Jiao, “Convolutional neural network based on bandwise-independent convolution and hard thresholding for hyperspectral band selection,” IEEE Trans. Cybern., vol. 51, no. 12, pp. 8465–8476, Nov. 2019.
[23] Y. Zhan, D. Hu, H. Xing, and X. Yu, “Hyperspectral band selection based on deep convolutional neural network and distance density,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 12, pp. 2365–2369, Dec. 2017.
[24] L. Zhao, Y. Zeng, P. Liu, and X. Su, “Band selection with the explanatory gradient saliency maps of convolutional neural networks,” IEEE Geosci. Remote Sens. Lett., vol. 17, no. 12, pp. 2105–2109, Aug. 2020.
[25] P. R. Lorenzo, L. Tucezewicz, M. Marcinkiewicz, and J. Nalepa, “Band selection from hyperspectral images using attention-based convolutional neural networks,” 2018, arXiv:1811.02667. [Online]. Available: http://arxiv.org/abs/1811.02667.
[26] H. Zhang, J. Lan, and G. Wang, “A dense spatial–spectral attention network for hyperspectral image band selection,” Remote Sens. Lett., vol. 12, no. 10, pp. 1025–1037, 2021.
[27] Z. Dou, K. Gao, X. Zhang, H. Wang, and L. Han, “Band selection of hyperspectral images using attention-based autoencoders,” IEEE Geosci. Remote Sens. Lett., vol. 18, no. 1, pp. 147–151, Jan. 2021.
[28] P. R. Lorenzo, L. Tucezewicz, M. Marcinkiewicz, and J. Nalepa, “Hyperspectral band selection using attention-based convolutional neural networks,” IEEE Access, vol. 8, pp. 42384–42403, 2020.
[29] P. S. Singh and S. Karthikeyan, “Enhanced classification of remotely sensed hyperspectral images through efficient band selection using autoencoders and genetic algorithm,” Neural Comput. Appl., vol. 4, pp. 1–12, May 2021.
[30] M. Shuiabu, W. S. Lee, J. Schueller, P. Gader, Y. K. Hong, and S. Kim, “Unsupervised hyperspectral band selection for apple Marssonina blotch detection,” Comput. Electron. Agriculture., vol. 148, pp. 45–53, May 2018.
[31] X. Deng, Z. Huang, Z. Zheng, Y. Lan, and F. Dai, “Field detection and classification of citrus Huanglongbing based on hyperspectral reflectance,” *Comput. Electron. Agricult.*, vol. 167, Dec. 2019, Art. no. 105006.

[32] D. Yuan, J. Jiang, X. Qi, Z. Xie, and G. Zhang, “Selecting key wave-lengths of hyperspectral image for nondestructive classification of moldy peanuts using ensemble classifier,” *Inf. Phys. Technol.*, vol. 111, Dec. 2020, Art. no. 103518.

[33] J. Goldberger, G. E. Hinton, S. Roweis, and R. R. Salakhutdinov, “Neighbourhood components analysis,” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 17, 2004, pp. 1–8.

[34] F. Melgani and L. Bruzzone, “Classification of hyperspectral remote sensing images with support vector machines,” *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[35] Y. Cai, Z. Zhang, X. Liu, and Z. Cai, “Efficient graph convolutional self-representation for band selection of hyperspectral image,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4869–4880, 2020.

[36] C.-I. Chang, Q. Du, T.-L. Sun, and M. L. G. Althouse, “A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 6, pp. 2631–2641, Nov. 1999.

[37] S. Li and H. Qi, “Sparse representation based band selection for hyperspectral images,” in *Proc. 18th IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 2693–2696.

[38] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[39] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. New York, NY, USA: Springer, 2009, p. 33, doi: 10.1007/978-0-387-84858-7.

[40] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, “ShuffleNet V2: Practical guidelines for efficient CNN architecture design,” in *Computer Vision—ECCV 2018*. Cham, Switzerland: Springer, 2018, pp. 122–138, doi: 10.1007/978-3-030-01264-9_8.

[41] Q. Wang, F. Zhang, and X. Li, “Optimal clustering framework for hyperspectral band selection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5910–5922, Oct. 2018.

**Ziwei Liu** received the B.S. and M.S. degrees from Shandong Agricultural University, in 2015 and 2017, respectively. He is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. His research interests include deep learning, computer vision, and hyperspectral image processing.

**Jinbao Jiang** is currently a Professor and a Doctoral Supervisor with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing. His main research interests include resources and environment remote sensing, agricultural remote sensing, and deep learning.

**Ying Du** received the B.S. and M.S. degrees from Yangzhou University, in 2016 and 2019, respectively. She is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. Her research interests include deep learning, agricultural remote sensing, and hyperspectral image processing.

**Zhifang Xu** received the M.S. degree from Henan Polytechnic University, in 2017. He is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. His research interests include remote sensing application, image processing, and deep learning.

---

**Ziwei Liu** received the B.S. and M.S. degrees from Shandong Agricultural University, in 2015 and 2017, respectively. He is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. His research interests include deep learning, computer vision, and hyperspectral image processing.

**Jinbao Jiang** is currently a Professor and a Doctoral Supervisor with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing. His main research interests include resources and environment remote sensing, agricultural remote sensing, and deep learning.

**Ying Du** received the B.S. and M.S. degrees from Yangzhou University, in 2016 and 2019, respectively. She is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. Her research interests include deep learning, agricultural remote sensing, and hyperspectral image processing.

**Zhifang Xu** received the M.S. degree from Henan Polytechnic University, in 2017. He is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. His research interests include remote sensing application, image processing, and deep learning.

---

**Ziwei Liu** received the B.S. and M.S. degrees from Shandong Agricultural University, in 2015 and 2017, respectively. He is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. His research interests include deep learning, computer vision, and hyperspectral image processing.

**Jinbao Jiang** is currently a Professor and a Doctoral Supervisor with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing. His main research interests include resources and environment remote sensing, agricultural remote sensing, and deep learning.

**Ying Du** received the B.S. and M.S. degrees from Yangzhou University, in 2016 and 2019, respectively. She is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. Her research interests include deep learning, agricultural remote sensing, and hyperspectral image processing.

**Zhifang Xu** received the M.S. degree from Henan Polytechnic University, in 2017. He is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. His research interests include remote sensing application, image processing, and deep learning.

---

**Ziwei Liu** received the B.S. and M.S. degrees from Shandong Agricultural University, in 2015 and 2017, respectively. He is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. His research interests include deep learning, computer vision, and hyperspectral image processing.

**Jinbao Jiang** is currently a Professor and a Doctoral Supervisor with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing. His main research interests include resources and environment remote sensing, agricultural remote sensing, and deep learning.

**Ying Du** received the B.S. and M.S. degrees from Yangzhou University, in 2016 and 2019, respectively. She is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. Her research interests include deep learning, agricultural remote sensing, and hyperspectral image processing.

**Zhifang Xu** received the M.S. degree from Henan Polytechnic University, in 2017. He is currently pursuing the Ph.D. degree with the College of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China. His research interests include remote sensing application, image processing, and deep learning.