

Article

Not peer-reviewed version

Study on Object-Based High-Resolution Remote Sensing Image Classification of Crop Planting Structure in the Loess Plateau of Eastern Gansu Province

[Rui Yang](#) , [Yuan Qi](#) ^{*} , Hui Zhang , [Hongwei Wang](#) , Jinlong Zhang , Xiaofang Ma , Juan Zhang , Chao Ma

Posted Date: 5 June 2024

doi: 10.20944/preprints202405.2156.v1

Keywords: Gaofen images; Object-based image classification; Random forest; Convolutional neural network; Crop classification



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

aArticle

Study on Object-Based High-Resolution Remote Sensing Image Classification of Crop Planting Structure in the Loess Plateau of Eastern Gansu Province

Rui Yang, Yuan Qi *, Hui Zhang, Hongwei Wang, Jinlong Zhang, Xiaofang Ma, Juan Zhang and Chao Ma

Key Laboratory of Remote Sensing of Gansu Province, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; yangrui@lzb.ac.cn (R.Y.); 1027042853@qq.com (H.Z.); wanghw@lzb.ac.cn (H.W.); zjinling@lzb.ac.cn (J.Z.); 1561156172@qq.com (X.M.); zhangjuan@mails.ucas.ac.cn (J.Z.); machao23@mails.ucas.ac.cn (C.M.)

* Correspondence: qiyuan@lzb.ac.cn

Abstract: Timely and accurate acquisition of information on distribution of the crop planting structure in Loess Plateau of eastern Gansu Province, one of the most important agricultural areas in western China, is crucial for promoting fine management of agriculture and ensuring food security. In this study, the Object-Based Image Classification (OBIC), the Random Forest (RF) and Convolutional Neural Network (CNN) models were employed to classify the crop planting structure of four representative test areas in Qingyang City in a precise manner. Firstly, different optimal segmentation scales for various crops were selected using the Estimation of Scale Parameter 2 (ESP2) tool and the Ratio of Mean Difference to Neighbors (ABS) to Standard Deviation (RMAS) model. The images were then segmented through multiresolution segmentation combined with the Canny Edge Detection algorithm. Secondly, the L1 regularized logistic regression model was utilized to select and optimize 39 spatial feature factors including spectral, textural, geometric, and index features, in conjunction with phenological factors. Lastly, under the multi-level classification framework, the Random Forest (RF) classifier and Convolutional Neural Network (CNN) model, combined with object-based multiresolution segmentation, was used to classify the crop planting structure. The findings show that: Thanks to the Canny Edge Detection algorithm, we can obtain a more complete boundary of the segmented objects and improve the separability. The optimal segmentation scales for corn, vegetables, and buckwheat were found to be 55, 70, and 35, respectively, while wheat and apple had optimal segmentation scales of 65. In addition to phenological characteristics, the number of selected spatial features for corn, vegetables, buckwheat, wheat, and apple were 9, 7, 16, 12, and 10, respectively. The CNN model demonstrated high consistency with the RF model in the classification results, but the accuracy of RF model is higher than that of CNN model on the whole. The overall accuracy of classification using RF model in four test areas registered 91.93%, 94.92%, 89.37% and 90.68%, respectively. This paper introduced crop phenological factors, effectively improving the extraction precision of shattered agricultural planting structure in the Loess Plateau of eastern Gansu Province. Its findings have important application value in crop monitoring, management, food security and other related fields.

Keywords: gaofen images; object-based image classification; random forest; convolutional neural network; crop classification

1. Introduction

Located in the northwest of China, Loess Plateau is a typical traditional agricultural area subject to less precipitation, where agricultural development is dominated by dry farming[1]. Loess Plateau of eastern Gansu Province, as a typical rain-fed agricultural region, is located in a fragile zone of ecological environment in China, where the precipitation is unevenly distributed in time and space, with a large annual difference and unbalanced annual distribution. Corresponding to a gradual

increase in precipitation from northwest to southeast, the crop distribution in this region is subject to an obvious gradient difference from southeast to northwest.[2]To name a few, apples that need plenty of sunshine and moisture are mainly distributed in the southeast; buckwheat that is cold and drought-tolerant and has a short growing period is mainly distributed in the northwest; corn and wheat are the main crop types in the Loess Plateau of eastern Gansu Province, accounting for more than 50% of the total area. They are distributed across the plateau region, more in the southeast and less in the northwest. Fine classification of crop planting structure is of great significance for agricultural decision-making, sustainable agricultural development and food security[3].

Boasting short detection period, low cost, wide coverage and many other advantages, remote sensing technologies have become one of the most popular methods for mapping crop spatial distribution at regional scale[4–6].At present, the scholars at home and abroad have conducted in-depth and extensive studies on crop planting structure recognition mainly around key issues such as the number of remote sensing images, image object feature[7–10], and classifier selection [11,12].

Among the classification methods based on the number of remote sensing images, crop recognition methods are roughly divided into two categories based on single-temporal images and time-series images, respectively. The former is designed to identify the key growing season for crops[13], and then obtain the spatial distribution information of crops[14–16]. At present, medium spatial resolution images represented by Landsat and Sentinel are among the most widely used remote sensing data[17]. With suitable spatial resolution and spectral band, they contribute to an increase in the accuracy of regional crop identification. However, as the spectrum of crop is affected by many factors such as crop type, soil background, farming activities, as well as cloudy and rainy weather and the information obtained from single-temporal images is limited, misclassification or absent classification in crop identification results frequently emerges, which can not meet the needs of high-precision crop mapping. Compared with single-temporal image classification, multitemporal remote sensing images can map crop classifications more accurately based on a full use of the phenological characteristics of crops. China has implemented a major project for a high-resolution earth observation system, and successfully launched a line of domestic high-resolution satellites (GF-1, GF-2and GF-6)[18,19]. Collaboration of domestic high-resolution satellites can effectively improve the ability to monitor crops, and provide great potential for fine identification of crops subject to fragmented distribution in complex areas of the Loess Plateau. For example, Zhou et al. [20] found that it's able to effectively extract the information of crops in rainy areas from multi-temporal GF-1 images; Xia et al. [21]used multi-temporal GF-6 images for intensive observation of crops in cloudy and rainy areas, and worked out the rule of seasonal changes of typical crops, thus achieving high-precision mapping of crop planting structure.

Crop classification methods based on object features mainly include pixel-level classification and Object-Based Image Classification. Pixel-level classification focuses on locality and ignores the correlation between objects. Salt-and-pepper noise due to the spectrum difference of in-class pixels as well as "mixed pixels" as a result of interclass pixel proximity effect have compromised the classification accuracy of crops to a big extent, making it difficult to meet the need for practical applications[22,23]. With the segmented object as its basic unit, Object-Based Image Classification is more targeted and has the advantages of reducing the amount of data involved in the operation, smoothening the image noise, and further collecting the crop information to classify by introducing the features of object shape, texture, etc. For example, Su et al. [24]made full use of 58 spatial feature factors of crops, and completed high-precision classification of crops in the region using the Object-Based Image Classification method. Karimi et al. [25]enabled high-precision crop classification of multi-temporal Landsat remote sensing images with the object-oriented classification method. Common crop remote sensing classifiers include but are not limited to Maximum Likelihood Method, Support Vector Machine (SVM), and Random Forest [26–28]. In recent years, deep learning has gradually become the mainstream algorithm in the image pattern recognition area by virtue of its hierarchical representation of features, efficient operation and end-to-end automatic learning[29,30]. Convolutional Neural Network (CNN), as one of the fastest-developing algorithms in deep learning, has been widely used in crop classification tasks [31–33]. Both models demonstrate good applicability

in interpreting crop planting structures, and the classification results largely correspond to the actual distribution. The classification results exhibit good separability, with fewer instances of misclassification and omission, allowing for better extraction of crop information in areas with higher heterogeneity.

As one of the prerequisites of object-based image classification, the optimal scale segmentation of crops has a direct impact on the size of generated objects and the precision of crop extraction[34]. ESP2(Estimation of scale parameter 2, ESP 2) is the most commonly used optimal segmentation scale tool that is able to screen out the optimal segmentation scale of different crops[35]. However, in light of the small spectral difference between different types of crops and the great influence of noise in the segmentation results, it is difficult to separate crops correctly only using the Multi Resolution Segmentation technology. Cannyedge detection is able to obtain more accurate edge information and filter noise with less time.[36,37]. In addition to combining multiple feature learning,, Object-Based Image Classification also increases the dimension of feature space and reduces the efficiency of data processing, of which secondary features may result in noise and even lower the classification accuracy. For this reason, the efficient construction of feature space has become a key factor of Object-Based Image Classification. For example, F. Low et al. [38] used Random Forest (RF) to obtain the best features of crop classification, improving the computational efficiency and classification accuracy. Chen Zhulin et al. [39]used three feature dimensionality reduction methods for dimensionality reduction of crop features, i.e., Random Forest (RF), Muti-Information (MI) and L1 Regularization. It's found that L1 Regularization feature dimensionality reduction method performed best in crop classification.

In conclusion, object-based classification and multi-temporal feature classification of high-resolution images are the mainstream direction of studies on crop remote sensing recognition at present. How to integrate the two is the current key scientific problem and technical difficulty. Based on the time series remote sensing data sources of GF-1, GF-2 and GF-6, this paper conducted a case study on four representative test areas from southeast to northwest of the Loess Plateau of eastern Gansu Province, constructed the optimal segmentation scale set of different crops using the ESP2 tool, RMAS model and Canny Edge Detection, performed feature selection with L1 Regularization Logistic Regression Model and sophisticated classification of crops in the test areas using the object-based and random forest methods, and used Convolutional Neural Network to cross-verify the results, with a view to providing new ideas and methods for regional classification of rain-fed crops in the Loess Plateau of eastern Gansu Province.

2. Materials and Methods

2.1. Overview of the Study Area

Qingyang City is located in the Shaan-Gan-Jin plateau zone of the Loess Plateau in China, which is named as the Loess Plateau of eastern Gansu Province. Located in the east of Gansu Province, it is the loess plateau gully area in the middle reaches of the Yellow River, with a geographical location of 106°20' -108°45' E and 35°15' n-37°10' N, and a total area of 27119 km²[40]. As a typical dry farmland subject to rainfed agriculture in northwest China, the southern region of Qingyang City has a mild climate that features good coordination between light, heat and water resources, making it the best climate area for developing agricultural production in eastern Gansu Province, known as the "Granary of eastern Gansu". Timely and accurate acquisition of crop type, area and spatial distribution information is of great significance for improving the crop planting structure, and implementing precise management of agricultural production.

The spatial distribution of precipitation in the Loess Plateau of eastern Gansu Province is characterized by more in the southeast and less in the northwest. As a result, the spatial distribution of crop planting structure also follows certain rules, with buckwheat and other typical crops tolerant to cold and drought distributed in the northwest, and apples and winter wheat in the southeast[41]. Phenology refers to different growth periods of crops in the growing season, during which the growth state and physiological characteristics of crops are translated to different remote sensing

features. These features include reflectivity, vegetation index, etc., which can be used for the recognition of different types of crops[42]. Therefore, in crop classification and recognition, the accuracy and reliability of crop classification can be improved by using phenological changes. Under the double effect of climate and crop cycle, different crops vary greatly in phenological periods. To name a few, winter wheat is usually sown in mid-September and ripened in late June of the following year. Corn is sown in mid-April and ripened in late September. Buckwheat is sown in mid-July and ripened in early September. Apples blossom in April-May and ripen in September-October (Figure 1).

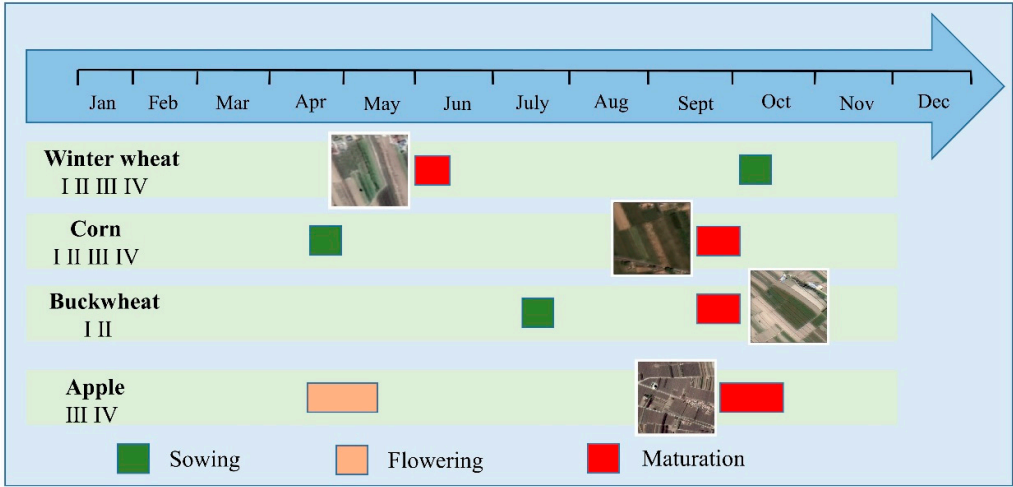


Figure 1. Phenological Periods of Different crops (I, II, III, IV represent the test areas where staple crops are located).

According to the characteristics of crop distribution, four representative test areas in Qingyang City were selected, which were located in Huanxian County (I), Zhenyuan County (II), Xifeng District (III) and Ningxian County (IV) respectively, shown by the gradient distribution from northwest to southeast. Among them, crops in test areas I and II were mainly corn, wheat and buckwheat, while those in test areas III and IV were mainly corn, wheat and apple (Figure 2).

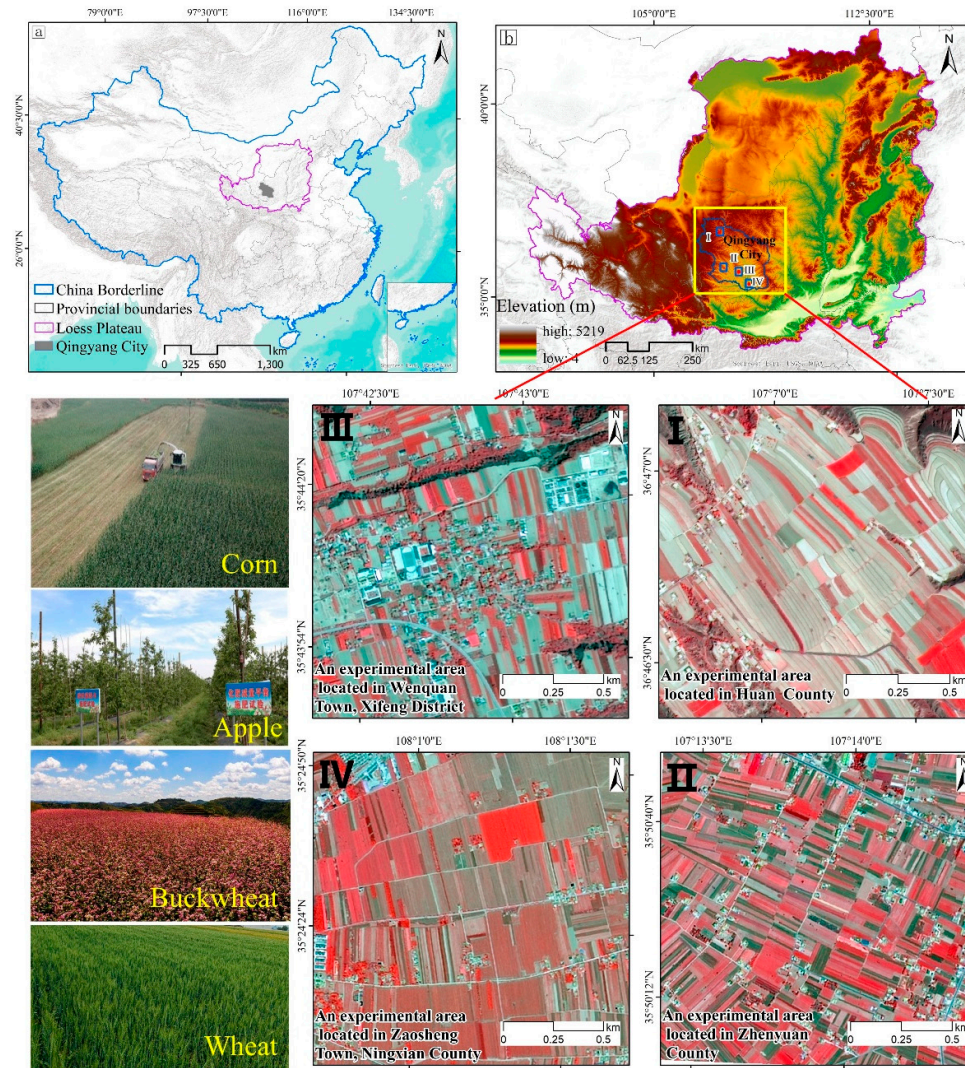


Figure 2. Study Area (a) Geographical distribution of Loess Plateau and Gansu Province in China; (b) Digital Elevation Model of Loess Plateau of eastern Gansu Province and distribution of representative test areas; I, II, III and IV represent locations in Huanxian County, Zhenyuan County, Xifeng Hot Spring Town and Zaosheng Town, Ningxian County, respectively.

2.2. Data

High-resolution remote sensing images have attracted wide attention by virtue of their ability to provide more detailed information on objects. Gaofen-1, 2 and 6 are the high-resolution remote sensing satellites launched and operated by China in recent years. These satellites capable of acquiring high precision remote sensing data [43,44] can be used to study the classification and identification of planting structure of crops in Qingyang City. This experiment conducted a study on crop classification with high-resolution images. The Gaofen-1 Satellite was launched in 2013, with a panchromatic image resolution of 2m and a single satellite imaging swath width of more than 60 km. Gaofen-2 Satellite was launched in 2014, with a panchromatic image resolution of 1m and a single satellite imaging swath width of more than 45 km. Gaofen-6 Satellite was launched in 2018, with a panchromatic image resolution better than 2m and a width larger than 90 km (Table 1). The parameters of China's high-resolution remote sensing satellites include spatial resolution, width, etc. Boasting good technical indicators and data quality, it is able to meet the needs of remote sensing classification and recognition of crop planting structure in Qingyang City.

Table 1. Image Data of High-resolution Satellites (I-IV represent 4 experimental areas respectively).

Name of satellite	Sensor (PMS)	Spatial resolution (m)	Image quantity			
			I	II	III	IV
GF-1	PMS1/PMS2	2	/	2	1	/
GF-2	PMS1/PMS2	1	4	1	1	4
GF-6	PMS1/PMS2	<2	/	1	2	/

2.3. Methodology

This experiment classified four typical crops of corn, wheat, buckwheat and apples using data sources from high-resolution images. In addition to crop types, four test areas also accommodate some natural vegetation and artificial construction land, without lake water systems. In light of different phenological periods of the crops, in order to classify crops based on full use of the phenological characteristics of crops, this study proposes a multi-level crop classification method that conducts multi-level segmentation of images with different segmentation scale parameters and then forms an inherited hierarchical structure between object layers. By combining phenological characteristics information of each crop in different periods, the classification of corresponding crops at different levels is enabled.

The main steps for multi-level crop classification are as follows: First, create Level 1 based on scale-1 segmentation, and distinguish artificial construction land, bare land and vegetation areas according to the $NDVI < 0.15$ threshold. In vegetation areas, classify the crops with the images of crops at corresponding phenological periods as the base layer; in areas other than those where the crops have been classified, re-distinguish artificial construction land, bare land and vegetation areas according to the threshold conditions, successively create Level 2 and Level 3 in the vegetation areas, and circulate in order until the classification of different crops is completed (Figure 3)

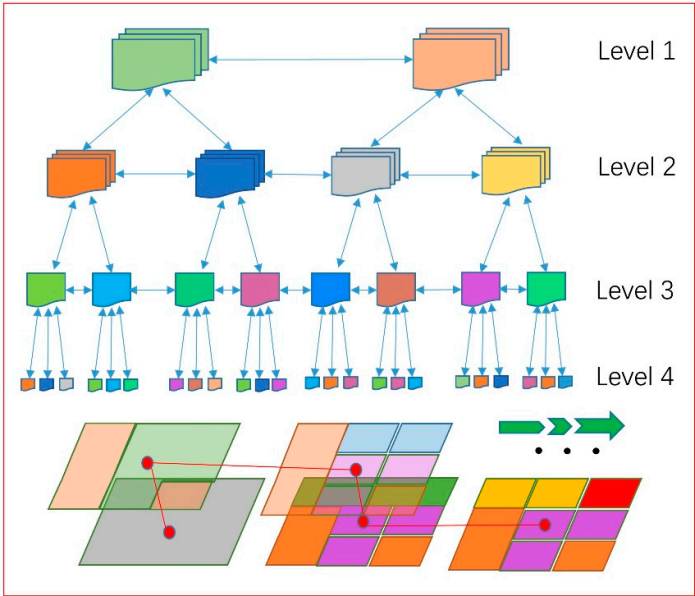


Figure 3. Hierarchical network diagram of image objects.

On the basis of multi-level classification, the classification steps at each layer mainly include to determine the optimal segmentation scale parameters of different crops using the ESP2 optimal scale evaluation tool and RMAS model based on eCognition software and data sources from high-resolution images, and to segment different crops combining the Canny Edge Detection results of

different phenology periods. This paper selected 40 feature factors and took the changes in phenological periods of different crops (NDVI ratio in phenological periods) as the phenological feature factor in view of the great difference in phenological periods between crops, selected the optimal spatial feature for 39 spatial feature factors other than phenological feature factors using the L1 regularized logistic regression model, and at last used the phenological feature factors and the selected feature factors to facilitate subsequent classification of crops.

This study selected 150 samples of corn, wheat, buckwheat and apple, and 40 samples of natural vegetation. In order to ensure sufficient samples for verification, the samples were divided into the training and test sets in the ration of 1:1. It employed the RF algorithm to interpret and classify the planting structure of crops in four test areas of Qingyang City, in conjunction with CNN and object-based optimal segmentation scale. Cross-validation was made for the results of two methods, and the confusion matrix evaluation method as well as the overall accuracy and Kappa coefficient accuracy evaluation parameters were used to evaluate the accuracy of two model algorithms. Figure 4 provides the technical flow chart of this test.

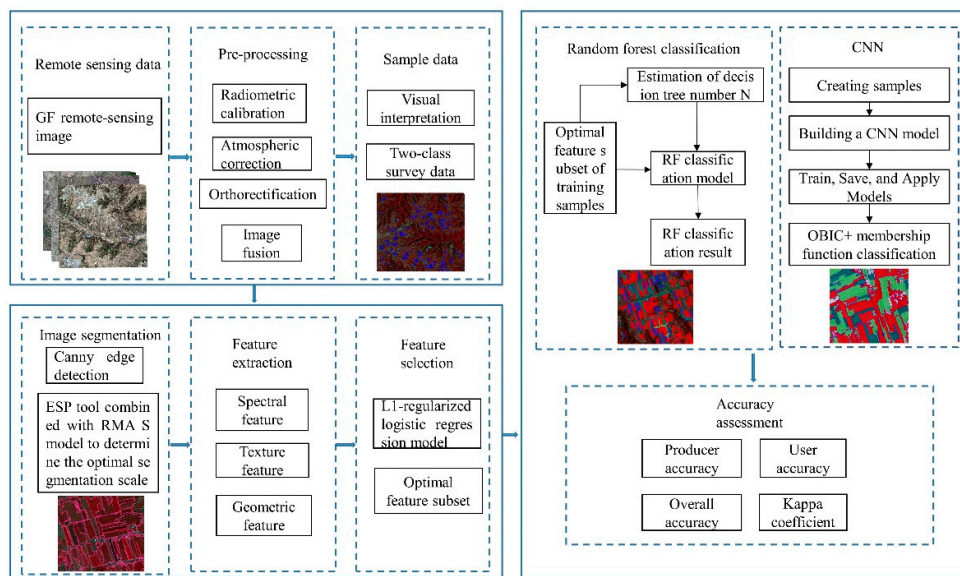


Figure 4. Flowchart.

2.3.1. Canny Edge Detection

The Canny Edge Detection algorithm seeks to ensure the detected edge is close to the actual edge, and to maintain the continuity of the detected edge. It creates as few breakpoints as possible and effectively controls the probability of miss and false detection. The steps of the algorithm are as follows:

There is usually noise in the images, which, along with the gradient value of nearby pixels, are relatively large, and easy to be mistakenly identified as an edge. To solve this, the Gaussian smoothing filter is employed to lower the noise[45]. The core of Gaussian smoothing filter is Gaussian function, whose formula is expressed as follows:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (1)$$

Wherein, (x,y) denotes the coordinates of a convolution with a Gaussian kernel and σ denotes the standard deviation.

Calculate the first-order derivative of an image in X and Y directions using a 2x2 convolution kernel to determine the gradient amplitude of the pixel and obtain the gradient of the image in the X and Y directions. E_X E_Y [46] Then the gradient magnitude M and gradient direction θ of each pixel can be calculated with the formula as follows:

$$M = \sqrt{E_x^2 + E_y^2} \quad (2)$$

$$\theta = \arctan \frac{E_y}{E_x} \quad (3)$$

In order to ensure the accuracy of edge detection, it's necessary to conduct Non-Maximum Suppression. [47] This process seeks to segment the gradient values, and look for the maximum value in the entire gradient magnitude map. Near these maxima, there may be areas called "Ridge Strip", and these non-maxima need to be set to zero, with only the local gradient maximum retained. We enable non-maximum suppression by calculating the local maximum of the pixel and finally identify the pixel with the largest local gradient for retaining. Edge detection that observes the above steps paints the true edges of an image with relative precision, but there are still edge pixels affected by many factors. In order to avoid false edges, the Hysteresis Threshold Detection Strategy is adopted. This strategy uses dual threshold detection to filter out some unnecessary edge information for pixels[48,49].

2.3.2. Image Segmentation

Compared with single-scale segmentation, Multi Resolution Segmentation (MRS) is a kind of image segmentation algorithm that forms multiple geographic object layers by fully utilizing multiple information such as object structure, texture, spatial characteristics and the relationship between adjacent objects, in a bid to better reflect the actual physical form of land cover corresponding to remote sensing images[50,51]. This study employs the MRS algorithm of eCognition Developer 10.3 software, of which the core parameters include layer weights, segmentation scale and homogeneity criterion combination parameters, with the latter two having the most significant influence on segmentation results. In the execution process, the MRS algorithm starts with merging a single pixel from the bottom up, and fully optimizes the image segmentation object according to homogeneity and heterogeneity, so as to realize pixel aggregation at different scales in the same image.

2.3.3. Optimal segmentation scale estimation

Developed by Drgu et al., Estimating the scale parameter 2 (ESP2) is a tool used to estimate the maximum heterogeneity in image segmentation results with different scales.[52] This tool makes an evaluation with the change rate of the segmentation scales based on the mean of global and local variances of the image segmentation results, and the peak value of the change rate curve is the maximum heterogeneity of the whole image object. The diversity of images often results in multiple values of the optimal segmentation scale calculated by ESP2[53]. To this end, quantitative evaluation is required for segmentation results using the ROC calculation method, so that the reliability and accuracy of segmentation results can be better understood. When the ROC curve reaches a maximum value, this scale is the optimal segmentation scale. The formula for ROC is as follows:

$$ROC = \left[\frac{L - (L - 1)}{L - 1} \right] \times 100 \quad (4)$$

Wherein, L denotes the local variance of the target layer; L-1denotes the local variance of the layer next to the Ltarget layer as the base.

For the spatial and spectral feature information of medium and high-resolution remote sensing images, measure the internal homogeneity and heterogeneity of objects based on standard deviation and absolute mean difference, and quantify the methods that affect the quality of image segmentation results using the Ratio of Mean Difference to Neighbors (ABS) to Standard Deviation (RMAS) between segmentation objects and neighborhoods[54]. With RMASas the objective function, this method selects the optimal scale from the segmentation results of different sizes. When the internal segmented object is the identical object that is similarly spectral and textural, its internal standard deviation value is small, largely distinguishable from the adjacent objects. At this time, RMAS is the largest, and this scale is regarded as the optimal segmentation scale.

$$RMAS = \sum_{i=1}^f w_L \frac{\Delta C_L}{S_L} \quad (5)$$

$$\Delta C_L = \frac{1}{k} \sum_{i=1}^m \sum_{j=1}^m k_{ij} |\bar{C}_L - C_{Li}| \quad (6)$$

$$S_L = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (C_{Li} - \bar{C}_L)^2} \quad (7)$$

Wherein: L denotes the number of bands of remote sensing images, w_L denotes band weights corresponding to different bands, f denotes the number of bands, ΔC_L denotes the absolute difference between the image object of the L th band and the mean value of the neighborhood, S_L denotes the standard deviation of the image object of the L th band, C_L denotes the gray value of the i th pixel of the L th band, (C_L) denotes the band mean of the L th band, N denotes the number of pixels in the segmented object, m denotes the number of objects adjacent to the segmented object, k denotes the boundary length of the segmented object, k_{ij} denotes the length of the common side of the i th image object and the j th adjacent object.

2.3.4. Feature factor

Due to the diversity of crops, they often exhibit the same spectral characteristics at certain stages. By using the Normalized Difference Vegetation Index (NDVI) as phenological characteristics, it can not only screen the baseline date images for specific crops, but also be used as a feature factor to effectively distinguish different crops at specific stages[55]. In this paper, the phenological characteristics of different crops were introduced as key factors for classification. Moreover, the study established a complete object feature set by synthesizing spectral, texture and geometric features of images[56,57], and preliminarily selected a total of 39 spatial feature factors covering four kinds of features, i.e., spectrum, texture, geometry and index. The details are as follows:

- (1) Spectral feature (SPEC): Including the mean of four bands of visible spectrum, namely, the mean of red band (Mean_R), the mean of green band (Mean_G), the mean of blue band (Mean_B), the mean of near infrared band (Mean_NIR), and the maximum difference value (Max_diff), the brightness value (Brightness) and the Standard Deviation (Std) of different bands.
- (2) Texture features (GLCM, GLDV): Texture feature refers to the spatial relationship between gray levels of adjacent pixels, which reflects a regional feature rather than that of a single pixel. It is determined by the distribution of a given pixel and its adjacent pixels. The most common methods for texture features include Gray level co-occurrence matrix (GLCM) and Gray level difference vector (GLDV). This paper selects All dir GLCM Mean, GLCM Ent, GLCM Homo, GLCM Std, GLCM Dissim, GLCM Contrast, GLCM Ang, 2nd Moment and GLCM Corr; all dir. GLDV and GLDV Mean, GLDV Ent, GLDV Contrast, GLDV Ang, 2nd Moment.
- (3) Geometric features (GEOM): A total of 13 shape and scope features of objects, including Area, length/Width, length, Width, Border length, Shape Index, Density, Asymmetry, Roundness, Boundary Index, Compactness, Ellipse Fitting and Rectangle Fitting.
- (4) Index features (INDE): Including Enhanced Vegetation Index (EVI), Normalized Difference Vegetation Index (NDVI), Red/Green ratio (Red/Green, R/G), and Ratio Vegetation Index (RVI).

The details are shown in Table 2:

Table 2. Spatial feature factor information.

Feature category	Feature variable	Total/number
Spectral feature	Mean_R, Mean_G, Mean_B, Mean_NIR, Max_diff、Briahtness and Standard Deviation (four bands)	10
Texture features	GLCM Mean, GLCM Ent, GLCM Homo, GLCM Std, GLCM Dissim, GLCM Contrast and GLCM Ang. 2nd Moment, GLCM Corr, GLDV Mean, GLDV Ent, GLDV Contrast and GLDV Ang. 2nd Moment	12
Geometric features	Area, length/Width, length, Width, Border Length, Shape Index, Density, Asymmetry, Roundness, Boundary Index, Compactness, Ellipse Fitting, Rectangle Fitting	13
Index features	EVI, NDVI, R/G and RVI	4

This study selects texture, shape, geometry, index and other feature factors, refers to the actual crop types of different crop segmentation results, and usesL1 regularized logistic regression model to optimize the features of different crops. By constraining the weight of the model, it also generates the sparse weight matrix and lowers the order of the polynomial to a reasonable range. In regularization models, variables with coefficients other than zero are considered as important ones. This regularization method helps effectively reduce the complexity and instability caused by high-dimensional input, mitigate the risk of overfitting, and screen out important variables of practical significance. For a given dataset $D=\{(x_1,y_1)\} ,\{(x_2,y_2)\} ,\{(x_m,y_m)\}$, an embedded feature selection model based on L1 regularization can be expressed as:

$$\min_{\omega} \sum_{i=1}^m (y_i - \omega^T x_i)^2 + \lambda \|\omega\|_1 \tag{8}$$

Wherein, ω is the coefficient vector for a linear fit, λ is the regularization parameter, and $\|\omega\|_1$ is the first-order norm of the coefficient vector.

2.3.5. Random Forest Classification

The random forest classifier is a parameterless classifier based on random sampling technique and multiple independent decisions. The classifier predicts the target category by building several independent decision trees. Each decision tree is built based on a randomly-selected subset of samples and a subset of features. Each decision tree is built based on the sub-training set extracted through bootstap sampling of all training samples. The high error within classification of the decision tree is calculated using the samples that fall outside the subtraining set. In the classification process, the RF classifier synthesizes the independent prediction results of all decision trees, and determines the most likely category of the classification target by means of voting. The following two parameters are required for building a RF classifier: (1) The number of features n selected when building each decision tree; (2) The total number of decision trees k . In this experiment, k is set to 100, and n is equal to a single randomly splitting variable; the goal is to minimize generalization errors and correlations between trees, and to prevent overfitting in the classification process [58].

2.3.6. Convolutional Neural Network

Convolutional Neural Network (CNN) is a mathematical processing model that is generated by simulating the biological neural network structure and function. With the shared convolutional kernel, CNN has good high-dimensional data processing capacity, and reduces complex operations such as data preprocessing and additional feature extraction. It boasts efficient automatic feature extraction, strong self-learning ability and high prediction precision and accuracy[59].

The CNN network is mainly comprised of the convolutional layer, pooling layer and fully connected layer. As the core layer of the CNN network, the convolutional layer performs convolution

operations of the input matrix through several different convolutional kernels, covering most of the computation in the network, which obtains feature information of corresponding data via activation function [60]. Then the ReLU (Rectified Linear Unit) activation function for the n th element of the i th layer is expressed as:

$$X_n^i = f\left(\sum_{t \in M_n} X_t^{i-1} K_{tn}^i + b_n^i\right) \quad (8)$$

Wherein: M_n denotes the n th convolution region of the i -1st layer, X_t^{i-1} denotes an element in this region; K_{tn}^i denotes the weight matrix of convolution kernel; b_n^i denotes the bias term for the network.

The pooling layer, i.e., subsampling layer, is used to reduce the model weight parameters, enable network lightweighting on the premise of retaining features, speed up the operation, avoid overfitting, and output new feature information with less data [61]. Features extracted from sampling at the pooling layer are expressed as follows:

$$X_n^i = f(w_n^i \text{down}(X_t^{i-1}) + b_n^i) \quad (9)$$

Wherein: w_n^i denotes the weight matrix of the network, and $\text{down}(X_t^{i-1})$ denotes the subsampling function. After extracting features of data information through multiple convolutional layers and pooling layers, the fully connected layer assembles and classifies all local features by weight matrix. Hidden layers contained in the fully connected layer can effectively improve the model's generalization ability.

50% samples of the test area are randomly selected as training samples according to the size of the test area, and 64x64 pixels are set as the size of a single sample. In order to maintain the high-resolution features of images, it only rotated the images to increase the number of samples [58]. The rotation angle was rotated every 30° in a range between 0-330°. It's found after many tests that the optimal effect can be achieved by setting the number of hidden layers to 3 and the convolution kernel size to 5. Upon extraction of crops using CNN, we obtain the heatmap of classification results, set the Membership Function range to (0-1) and then get the final results of crop classification by using the CNN heatmap through membership degree classification in conjunction with object-based multiresolution segmentation.

2.3.7. Accuracy evaluation

In order to evaluate the results of two models, this paper adopts both qualitative and quantitative dimensions. Qualitative evaluation compares the two results and cross-verifies the quality of the two models, while quantitative evaluation evaluates the accuracy of classification results with no less than 50% of the field validated samples in the segmentation results of the test area.

Due to the uneven quantity of samples under different categories, the confusion matrix based on classification results was used to evaluate the accuracy, and the Overall Accuracy (OA) and Kappa coefficient indexes were calculated. The Kappa coefficient and OA reflect the overall classification effect, and the values of these indicators can be calculated with the confusion matrix. The greater the indicator value, the higher the classification accuracy of the category [62]. These indicators are calculated using the following formula:

$$OA = \frac{\sum_{i=1}^K N_{ii}}{N_{total}} \times 100\% \quad (10)$$

$$Kappa = \frac{N_{total} \sum_{i=1}^K N_{ii} - \sum_{i=1}^K N_{i+} N_{+i}}{N_{total}^2 - \sum_{i=1}^K N_{i+} N_{+i}} \quad (11)$$

Wherein, N_{ii} denotes the number of samples correctly classified, $N_{+(i)}$ denotes the number of real samples of Category i , $N_{(i)+}$ denotes the predicted number of samples of Category i , N_{total} denotes the total number of samples and K denotes the total number of categories.

3. Results and Analysis

3.1. Segmentation Results Combined with Canny Edge Detection

When the object information is complex, mutiresolution segmentation often has a poor effect. For example, the segmentation results of crop contain more noise and a fuzzy outline. To solve this, the edge information of objects detected by Canny operator is used in our study as a feature factor to participate in mutiresolution segmentation, and the segmentation results are compared with the single mutiresolution segmentation results without Canny detection.

This method preliminarily determines the optimal segmentation scale range of crops using the ESP2 tool prior to crop segmentation, and calculates the optimal segmentation scale of each crop with the RMAS model. The ESP2 tool mainly uses the object homogeneity Local Variance (LV) and its Rates Of Change (ROC) as the quantitative evaluation indicator for screening, while considering the shape factor and compactness factor of 0.1 and 0.5, respectively. In order to evaluate the relative contribution of each band and edge detection of images, the weights of R, G, B, NIR bands and Canny Edge Detection are set to be 1. When the ROC of local variance reaches the maximum, more reasonable segmentation results can be obtained. Therefore, this paper selected the segmentation scale corresponding to the peak rate of ROC value as the optimal segmentation scale for the images, while giving priority to the LV indicator (Figure 5). The optimal segmentation scales for preliminary screening were 35, 45, 70, 95, 105, 125, 135, 170 and 220. Within the range of the segmentation scale, it increases successively with an increment of 5. It's determined after repeated tests that the range of crop segmentation scale value is between 35-95. Within this range, we successively calculate the RMAS value of different crops, and take the maximum RMAS value as the optimal segmentation scale, as shown in Figure 6. Upon final calculation, the optimal segmentation scale of buckwheat is 35, that of wheat and apple is 65, and that of corn is 55.

Mutiresolution segmentation results before and after edge information detection by Canny are integrated based on the optimal segmentation scale (Figure 7). It's concluded from the comparison that under the same segmentation scale, the image objects obtained only by mutiresolution segmentation cannot effectively distinguish adjacent features. In the results with object edge participating in segmentation, the object is more complete, and has a clear outline, stronger separability and better segmentation quality, in addition to reducing salt-and-pepper noise.

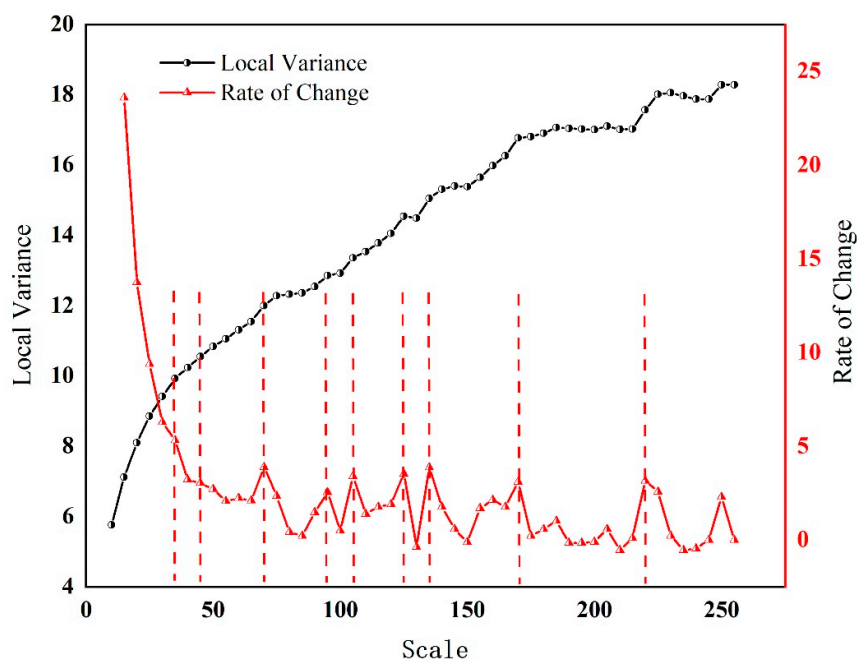


Figure 5. Evaluation of optimal segmentation scale using ESP2 tool.

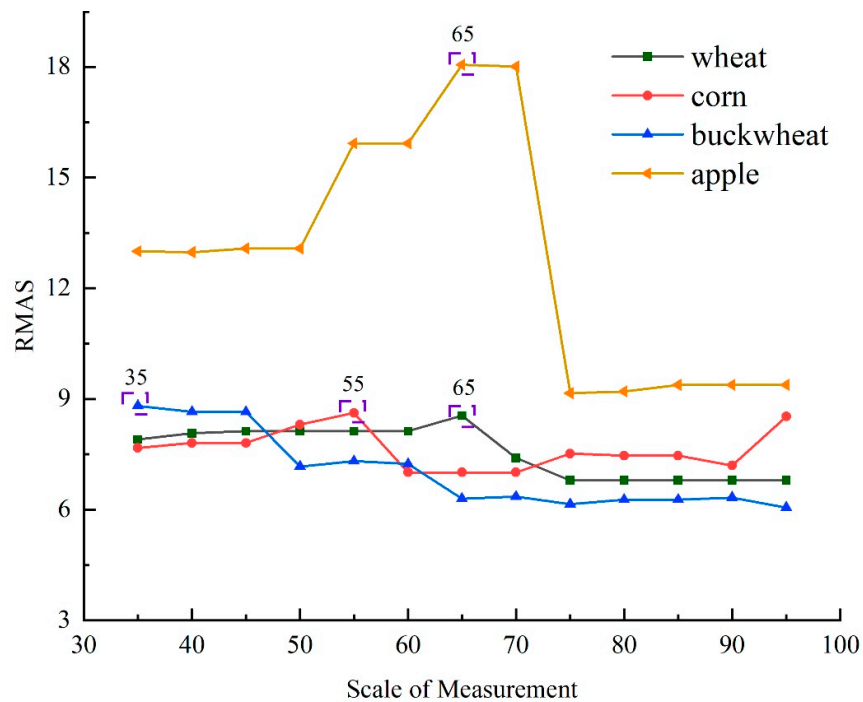


Figure 6. RMAS values of different crops.

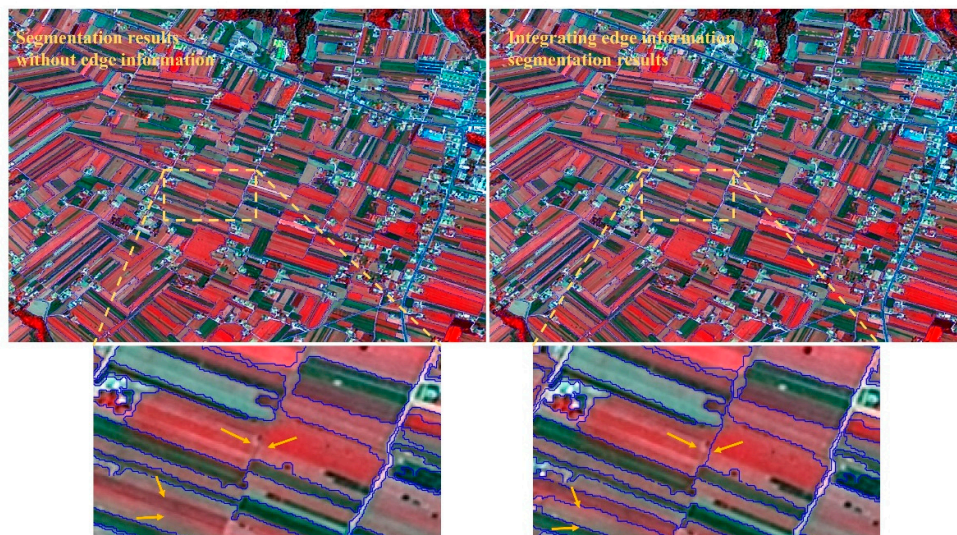


Figure 7. Segmentation map before and after integrating edge information.

3.2. Feature Factor Optimization

This study gathered statistics of the NDVI values of different crops and the NDVI ratios of adjacent months to learn about the growth status and change of crops, and took the changing trend of NDVI and the month with the largest change rate as the reference image for crop classification. As indicated by statistical calculation (Figure 8), the reference images of wheat, corn, apple and buckwheat were selected in May, July, August and September, respectively. On this basis, NDVI ratios in different phenological periods was used as the feature factor in order to facilitate the classification and interpretation of crops and improve the classification accuracy.

As the features of a higher dimension are selected, it leads to redundancy among the features. In order to reduce the redundant computation of data, this paper uses the L1 regularized logistic regression model[63] to conduct quantitative analysis of 39 features. It's found after many tests, better

results can be obtained by setting regularization parameter C to 0.9 and iteration times to 1000. In view of the fact that the feature factors vary between crops in the classification process, we optimize the feature factors of each crop, and make statistics of the contribution rate of feature factors of different categories of crops. Figure 9 shows the feature factors of different crops screened using the L1 regularized logistic regression model.

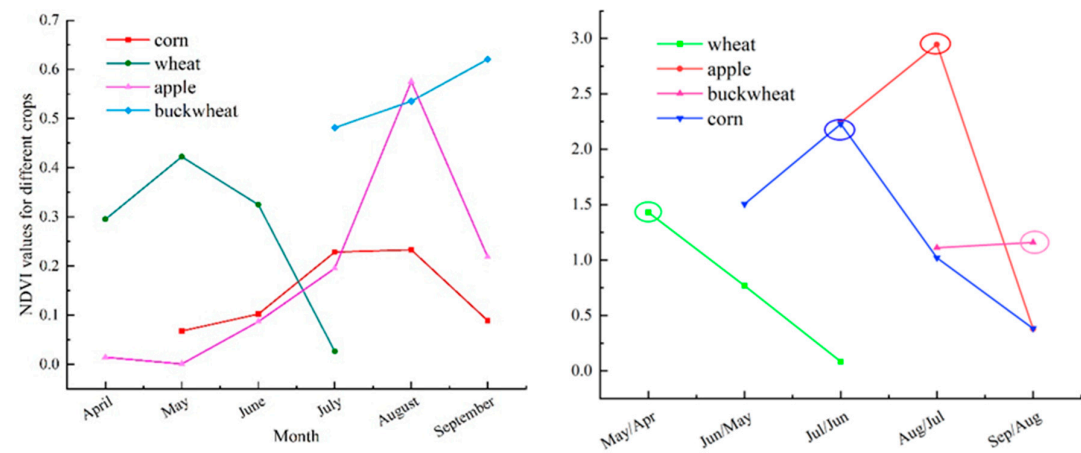


Figure 8. NDVI variation trend of different crops and its variation rate in adjacent months.

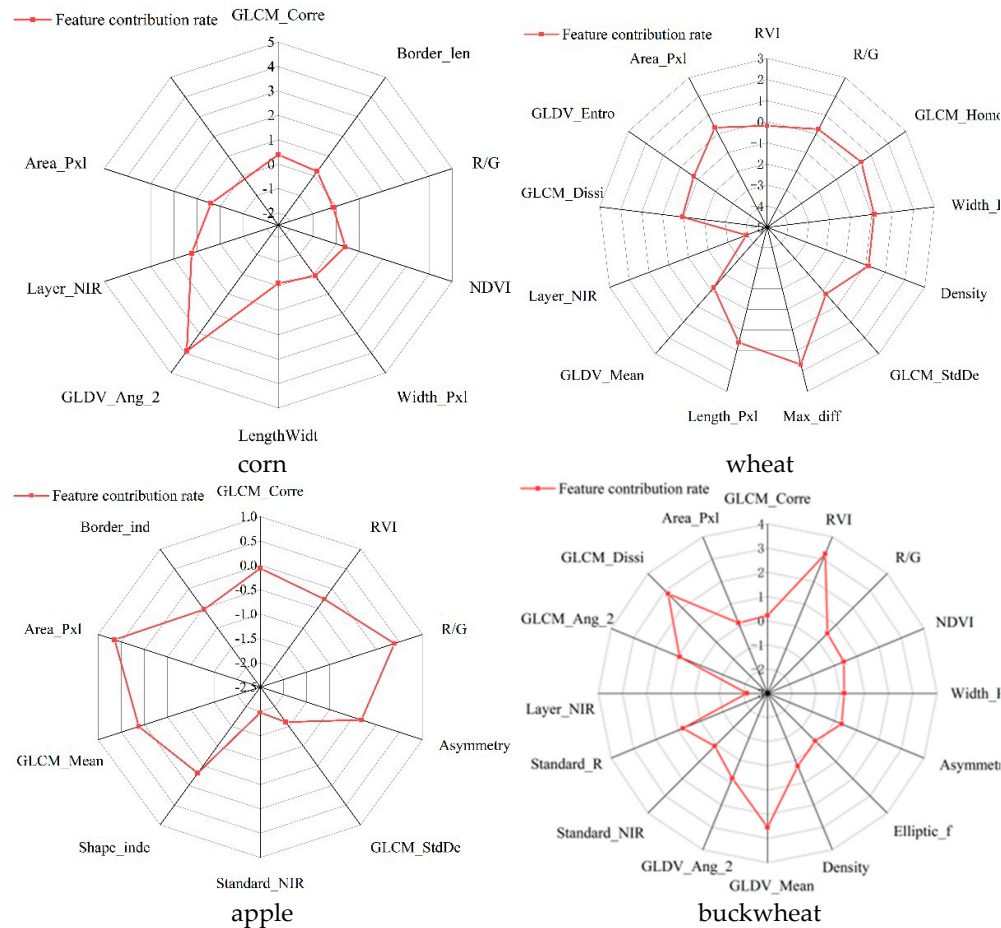
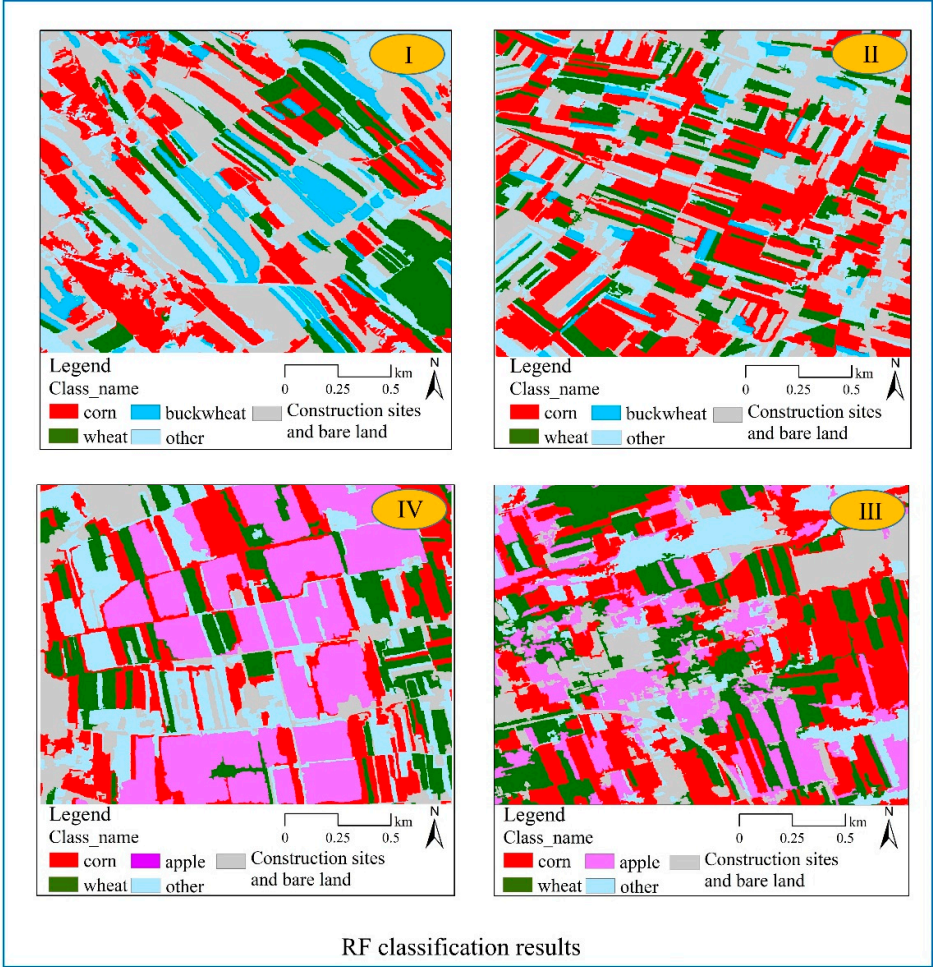


Figure 9. Preferred feature factors of different crops.

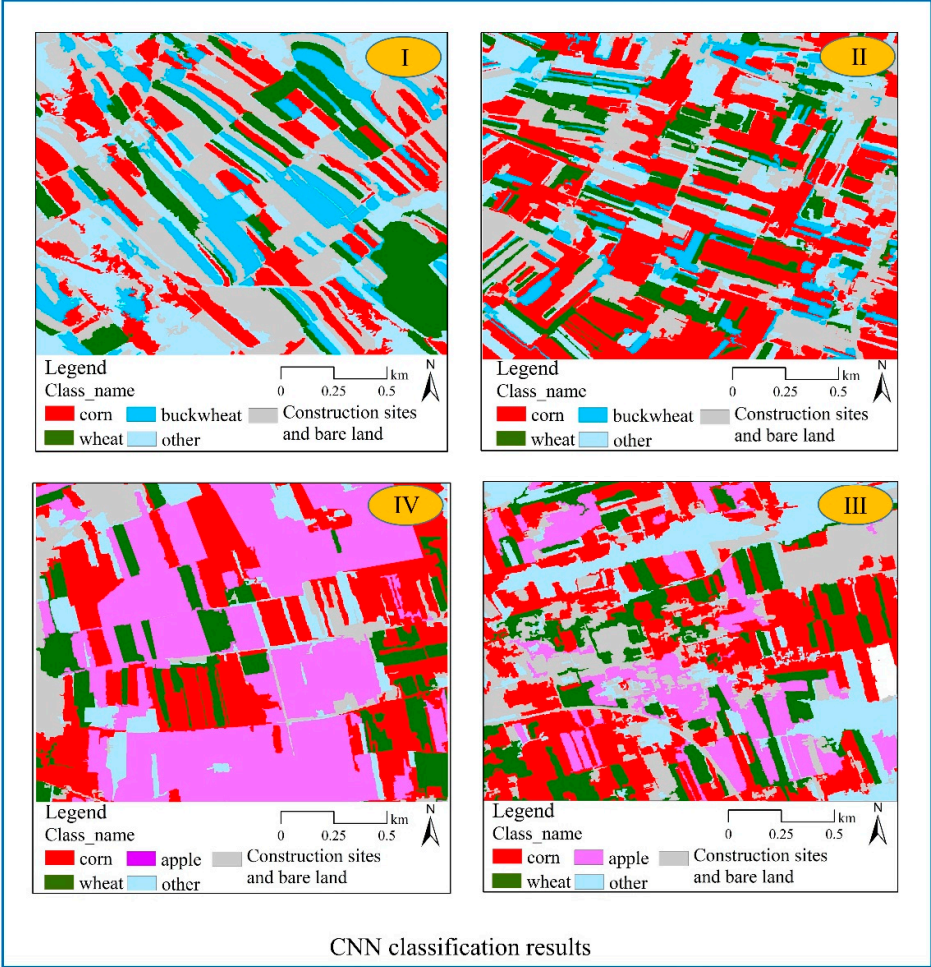
It can be seen from statistics that the number of feature factors after optimization of the four crops is greatly reduced, with texture and spectral features accounting for a large proportion, followed by geometric and exponential features. Among them, the number of feature factors selected for corn, wheat, apple and buckwheat are: 9, 12, 10 and 16, respectively. As can be seen from the Figure, corn has a total of 9 auxiliary feature factors, dominated by texture and spectral feature factors, namely GLDV_Ang_2 and Mean_NIR, with weights of 3.89 and 1.24 respectively. Since corn ripens mainly in July and August, during this period, corn has a deep red false-color image and obvious texture features. GLCM Ang. 2nd Moment reflects the uniformity of texture in the image and can be used as a cofactor to identify corn crops together with the reflection characteristics of near-infrared band. There are a total of 13 auxiliary classification feature factors for wheat, dominated by spectral and texture factors. Its main spectral features include Mean_NIR and Max_diff, with weights of -3.95 and 1.69 respectively. During the growing season of wheat, which mainly ranges between March and May, wheat is mainly manifested as light red in color and has even texture, making it easier to distinguish than other crops. In addition to using high near-infrared band characteristics, we can effectively assist wheat classification with the maximum band difference value, and the texture features are mainly GLDV_Mean_, with a weight of -1.18. The features of apple are dominated by GLCM_StdDe and Standard_NIR, with the weights of -1.61 and -1.97, respectively. During the growing season of apple, which mainly ranges between August and September, the texture of apple is more significant than other crop features, and the information change in the band standard deviation constitute a good reflection of the information on apple. The dominant factors of buckwheat are RVI, GLDV_Mean_, Layer_NIR and GLCM_Dissi, whose weights are 3.23, 2.54, -2.14 and 2.82, respectively. This is because buckwheat mainly ripens in September and October, which are also the ripe season of apple and corn. Buckwheat, on the other hand, is more delicate in texture than other crops, looking pink in false-color images, while other crops are darker in color. RVI that reflects the growth state of crops has the highest weight. Besides, if participating in classification, Mean_NIR among spectral features has a better effect, because the reflectance of crops is usually higher in near-infrared bands.

3.3. Object-Based Crop Classification Results

The classification results from the four typical test areas (Figure 10) indicate that both the RF and CNN models categorize crops into regular blocks, which aligns with the orderly distribution of crop planting structures in the test plots. To further assess the classification accuracy of different models, field sample data was used for accuracy validation. The confusion matrix evaluation method was employed, utilizing overall accuracy and Kappa coefficient as evaluation parameters to assess the accuracy of the Random Forest (RF) algorithm and Convolutional Neural Network (CNN). The accuracy validation results for the four test areas are shown in Table 3.



(a) RF classification results



(b) CNN classification results

Figure 10. Classification results of RF and CNN models

Table 3. Accuracy validation of the classification results through random forest and deep learning.

Test area	Type of Crops	Kappa coefficient of each crop		Kappa coefficient of overall classification results		Overall Accuracy	
		RF Model	CNN Model	RF Model	CNN Model	RF Model	CNN Model
I	Wheat	0.92	0.90				
	Corn	0.85	0.81	0.89	0.87	0.92	0.91
	Buckwheat	0.96	0.93				
II	Wheat	0.93	0.89				
	Corn	0.91	0.87	0.91	0.88	0.95	0.93
	Buckwheat	0.86	0.88				
III	Wheat	0.87	0.89				
	Corn	0.84	0.81	0.85	0.84	0.89	0.89
	Apple	0.85	0.80				
IV	Wheat	0.86	0.79				
	Corn	0.78	0.86	0.86	0.85	0.91	0.90
	Apple	0.93	0.89				

According to Table 4, both RF and CNN models feature a high accuracy in the results of classification of crops in the test area I. Overall, regardless of the Kappa coefficient or overall classification accuracy, in the four test areas, the classification results obtained by RF model are higher than those obtained by CNN model. In the test area I, it can be seen through calculating the Kappa coefficient of each category of crops that the one calculated by RF model is higher than that by the CNN model. Overall, the Kappa coefficients calculated by the RF model and CNN model are 0.89 and 0.87, respectively. The overall accuracy of the calculation by two models both exceeds 90%. Test areas II and I have the same category of crops, but the former has a smaller planting area of buckwheat than the latter. The Kappa coefficient of two models for buckwheat classification is smaller than that of the other two crops, because the buckwheat is planted in a small area where there are not enough samples for training. Due to too few samples, there emerges the phenomenon of model overfitting, resulting in low accuracy[64]. However, the RF and CNN models have a high classification accuracy overall, which reaches 94.92% and 93.43%, respectively. The test areas III and IV dominated by wheat, corn and apple feature a high classification accuracy of the two test areas overall. In the test area III, the overall classification accuracy of RF and CNN models reached 89.37% and 88.94%, respectively, while that of the two models in the test area IV was 90.68% and 90.18%, respectively.

This paper conducts a comparative analysis of the accuracy of the results using the two methods and makes statistics on the area and proportion of four test areas (Figures 11–14). In the test area I, among wheat, corn and buckwheat, corn has the largest planting area in the classification results, among which the classification area of RF model is 0.56 km², accounting for 22.74%, and that of CNN model is 0.46 km², accounting for 18.54%. The test area II has the identical crop planting structure as the test area I. Similarly, corn has the largest area in the classification results, among which the corn area of classification using RF model is 0.86 km², accounting for 34.84%, and that of CNN model is 0.92 km², accounting for 37.19%. Buckwheat has the smallest area, and the area classified by RF and CNN models is 0.08 km² and 0.20 km², accounting for 3.12% and 7.94%, respectively. In the test area III, corn has the largest area compared to any other crop. The area classified by RF and CNN is 0.73 km² and 0.77 km², respectively, and the area of apple classified using these two models is larger, which is 0.43 km² and 0.33 km², respectively. In the results of two models, the area of corn classified by RF model is larger than that of CNN model, while the area of apple classified by RF model is smaller than that of CNN model. According to field investigation statistics, there are more planting structures of corn and apple in this area. The main reason is that both corn and apples mature in August and September, both of which are similar in false-color remote sensing images during this period. Compared with RF model, the CNN model can better adapt to complex data patterns. In addition to the crops studied, other vegetation also exists in the test area. Due to the interference of diverse and complex vegetation on the classification process, the RF model is not effective in fully classifying crops. In the test area IV, apple has the largest area. The area classified by RF and CNN is 0.80 km² and 1.02 km², respectively. As can be seen from the classification result graph, the area of independent apple blocks classified by CNN model is larger. Seen from the field investigation and interpretation of remote sensing image satellites, the planting structures in this test area were regularly distributed in blocks. As the CNN model divided adjacent crops into a unified entirety, this may be the main reason why CNN model has a larger classification area for apples than RF model. Overall, the classification results of four test areas vary as two models use different algorithms. The difference is particularly significant in the test areas where corn and apple are planted. However, in terms of spatial distribution and quantity classification of crops, the classification results of CNN model and RF model are highly consistent and integral in spatial distribution.

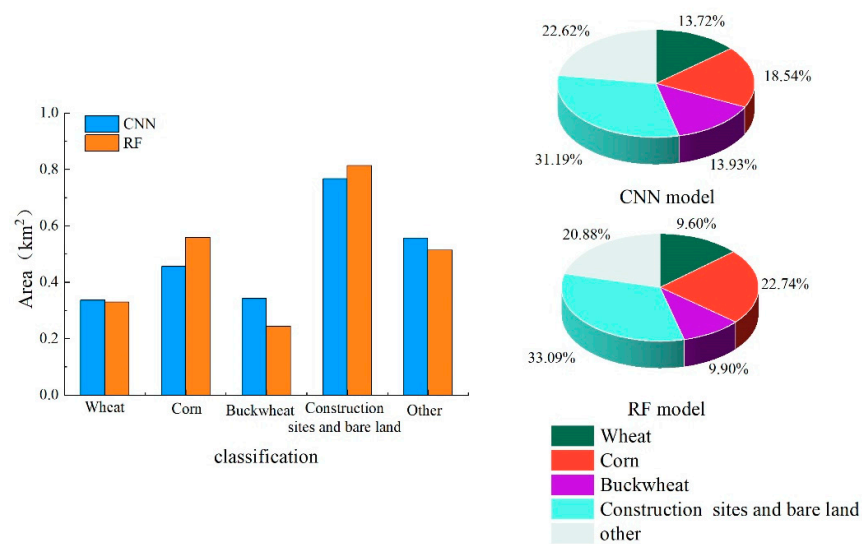


Figure 11. Area and proportion of classification results by two models in the test area I (Huanxian County).

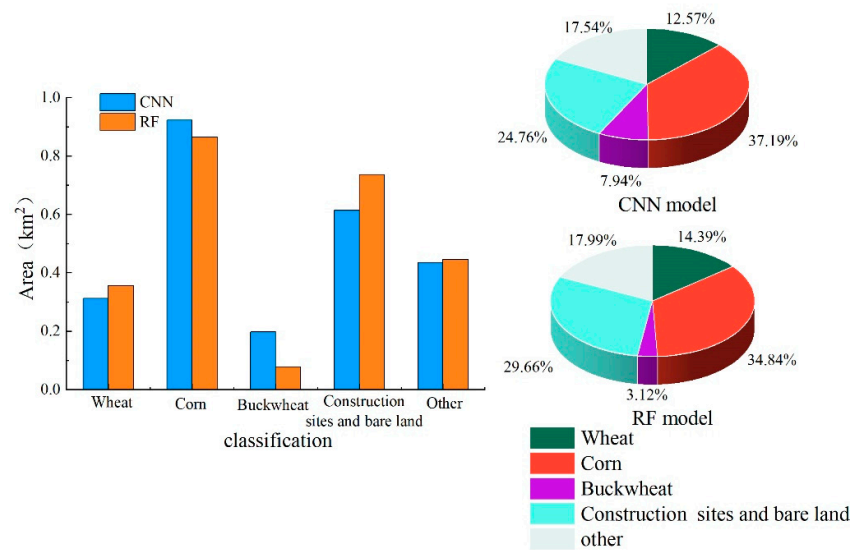


Figure 12. Area and proportion of classification results by two models in the test area II (Zhenyuan County).

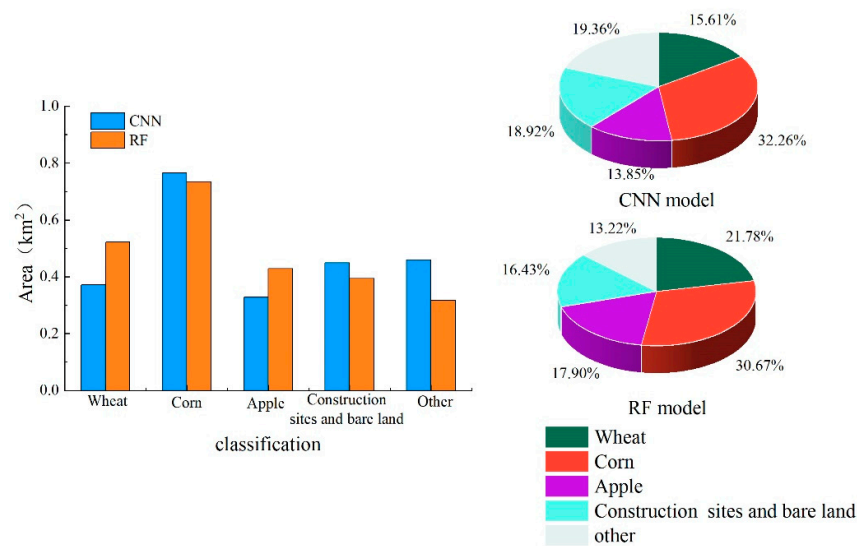


Figure 13. Area and proportion of classification results by two models in the test area III (Xifeng County).

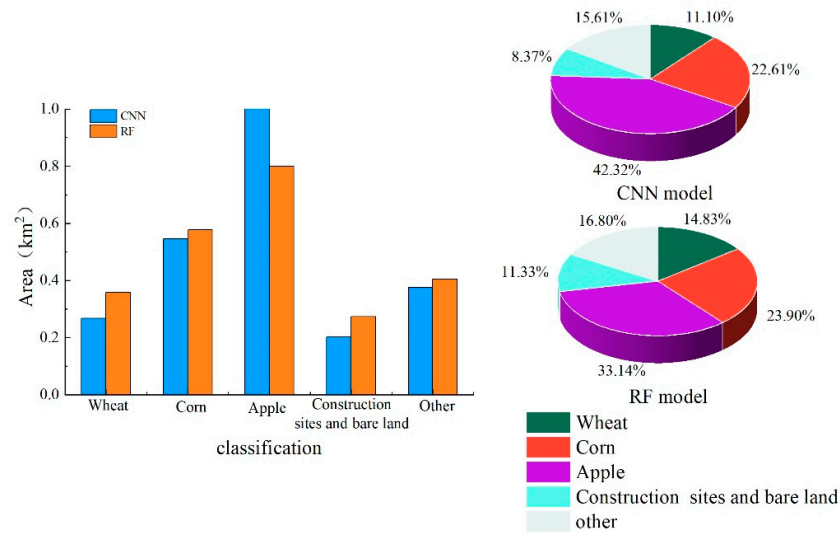


Figure 14. Area and proportion of classification results by two models in the test area IV (Ningxian County).

4. Discussion

4.1. Evaluation of RF and CNN Model Results

By virtue of excellent robustness, the RF model helps effectively process noise data. Albeit accurate crop classification results obtained by this model [65,66], crops in the study area are usually simple in planting structure or identical in the growing periods, so the classification targets are usually easy to distinguish. This study has, based on scale segmentation and selection of feature factors, newly introduced the phenological feature factor to effectively distinguish crop planting structures in the study area. Similarly, CNN model with strong learning capacity not only enables autonomous learning of crop features, but also captures complex nonlinear relationships for effective classification of crops with complex features [67,68]. Nonetheless, the CNN model still requires

improvement in accurately delineating crop boundaries at different scales. By combining the CNN model with multi-scale segmentation, we achieved fine classification of crop planting structures.

Both models exhibit good applicability in interpreting crop planting structures, and the classification results largely correspond to the actual distribution. The classification results demonstrate good separability, with fewer instances of misclassification and omission. This allows for better extraction of crop information in areas with higher heterogeneity. In comparison, the accuracy of RF model is slightly higher than that of CNN model on the whole, and the classification results based on RF are closer to actual distribution of crops in the study area. The results obtained through these two approaches have a good overall effect of effectively distinguishing the categories of corn and apple with little spectral difference and overcoming the "salt-and-pepper noise", so that the boundary information of different crops is clearer and more accurate results are obtained. At the same time, the CNN model has a higher demand for computing power and sample size, and is more suitable for vegetation classification issues with a larger scale and more sample data. Limited by the small size of the test area and insufficient data sets of samples, the accuracy of CNN remains low. Moreover, Lin Yi et al[69] also found that in the fine classification of vegetation in small-scale areas and small-size samples, the traditional machine learning classification method has a better effect than the deep learning method. Although the RF model has classification results better than the CNN model in four test areas, this study still has some limitations such as small sample size and lack of transferable remote sensing image datasets. As a result, the fine classification of crops in small areas based on the CNN model fails to exert its advantages. However, deep learning approaches have great potential of application in image classification[70]. On the one hand, traditional machine learning approaches require manual selection of samples for feature extraction, while deep learning allows to automatically extract high-level semantic features and performs better on large-scale data sets [71]. On the other hand, traditional machine learning approaches require manual selection of samples for training during each classification and have poor generalization ability, while the features learned by deep learning approaches have better generalization ability and stronger robustness. Therefore, we can expand the scope of crop classification data sets in the future, and conduct research based on better deep convolutional networks in order to leverage the value of deep learning approaches for crop classification in more regions.

4.2. Evaluation of the Effect of Combining Scale Segmentation with Feature Optimization

Compared with traditional pixel-based crop classification studies, Object-Based Image Classification takes into full account the spectral, texture and geometric features of different crops, greatly improving the classification accuracy [72,73]. However, in order to accurately improve the features of crops, it's important that we better segment and effectively distinguish different crops in the image. In spite of good results in image segmentation through segmentation algorithms at multiple scales, there are still some phenomena such as misclassification, absent classification and loss of detail caused by low image resolution [74,75], and only single-cropping crops in the study area are classified. This study proposes to integrate the new crop edge profile information obtained by the Canny Edge Detection Algorithm into the segmentation process. Crop edge profile information as a cofactor can be used to classify crop planting structure more accurately, but different crops have different segmentation scales. However, crop segmentation results under the same segmentation scale often appear over-segmentation or under-segmentation. The ESP2 optimal scale evaluation tool combined with RMAS model can effectively segment the contour information of different crops. Figure 15 shows the false-color original images of different crops. Figure 16 shows the results under the optimal segmentation scale of different crops combined with Canny edge detection. It can be seen from the figure that the contour edges of different crops are relatively clear, and the results are basically consistent with the real ground objects. The classification results obtained based on the characteristics of multiresolution segmentation are further improved under the multi-level classification framework.

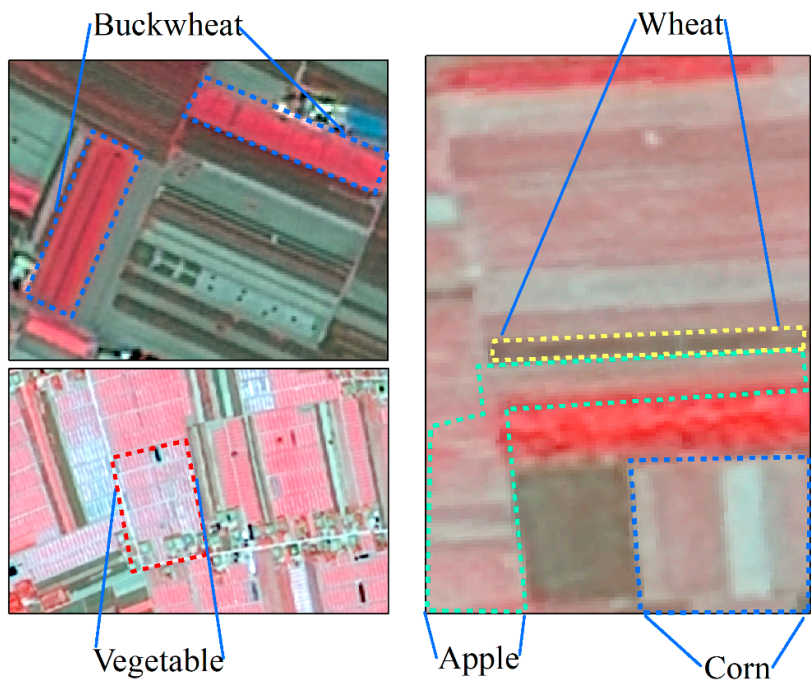


Figure 15. False-color images of different crops.

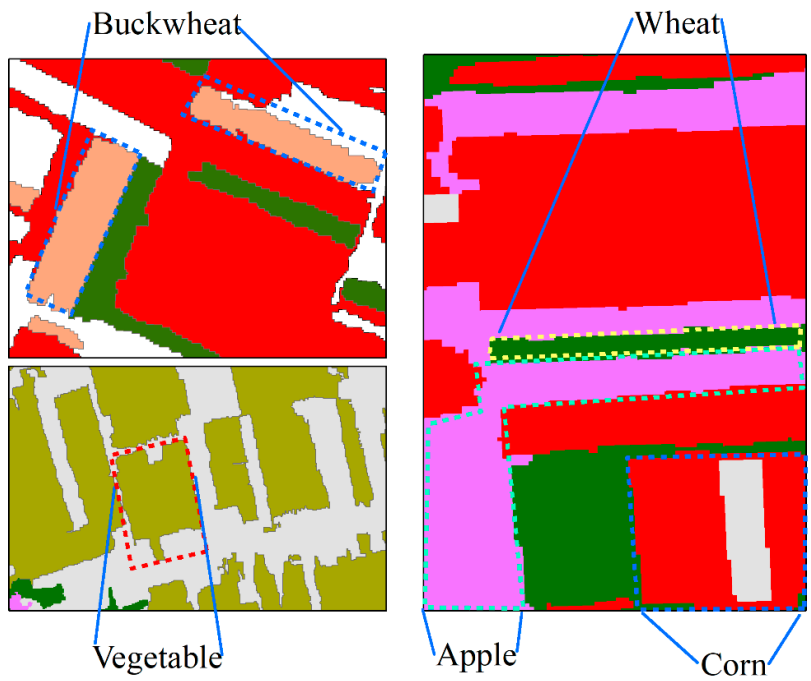


Figure 16. Segmentation results combined with Canny Edge Detection.

In light of similar false-color images of corn and apple, the texture features of both crops occupy a high proportion in the feature factors. According to the statistics on GLCM Ang.2nd Moment feature factor values of corn in the four test areas (Figure 17), the test area I has the smallest value, and the test area IV has the largest. This is because the Loess Plateau of eastern Gansu Province, as a rain-fed crop area, has more precipitation in the southeast region and better soil quality and fertility than the northwest region, creating better conditions for crops in the southeast region to grow and seeing a larger single crop area. Therefore, the area of a single corn block in test area I is smaller than

that in test area IV, the larger block contains more surface details and texture information, and the texture features are more rough [76,77]. The test area II demonstrates the consistency with test area III in terms of the size of blocks, and has an area of corn slightly smaller than that in test area II, resulting in a slightly larger feature factor value. The dominant texture feature factor of apple is GLCM_StdDe. Similarly, the test area IV is home to more pixels, with a big difference in texture features, and a relatively large standard deviation. Hence, the texture feature value of the test area IV is larger than that of the test area III.

Furthermore, crop features raise the interpretative nature of crops. Classification of crops based on the selection of better features can not only reduce the dimension of feature space to reduce the computational complexity, but also lift the classification accuracy [78]. However, most of the features after dimensionality reduction are jointly applied to the classification of various crops[73]. In this study, L1 regularized logistic regression model was used to select better feature factors for each crop, which, along with the phenological feature factor, have laid a foundation for using feature factors to assist crop classification.

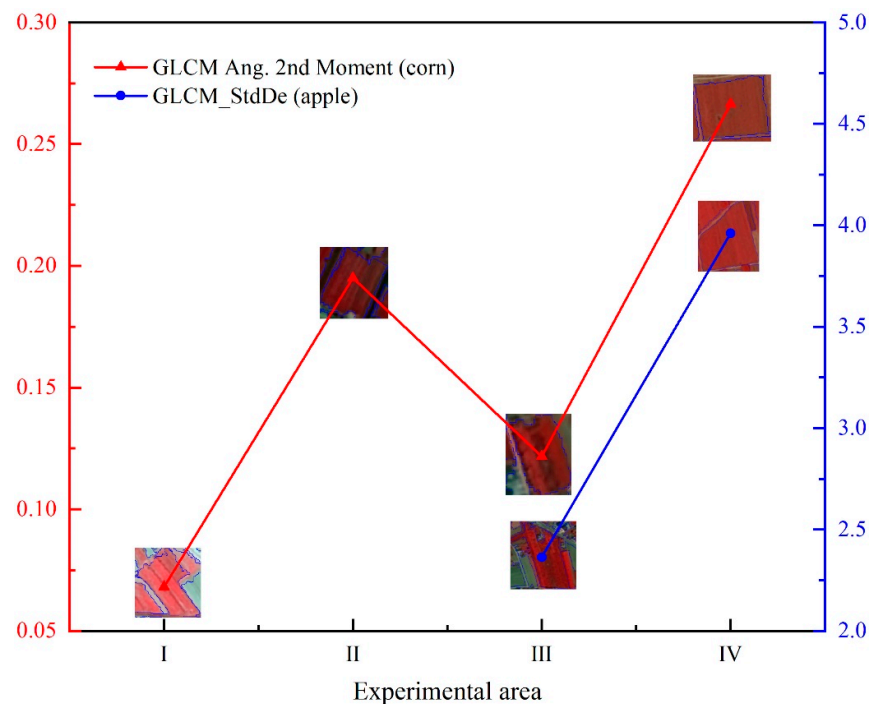


Figure 17. Statistics of texture features of corn and apple.

5. Conclusions

This paper selected four typical test areas in Qingyang City for research. It employed the optimal segmentation scale calculated by ESP2 tool and RMAS model, combined with Canny Edge Detection, to segment the crops in the study area, while selecting better spatial features with the L1 regularized logistic regression model. It also referred to the phenological feature information and analyzed the classification results by RF and CNN models under the OBIC framework on the basis of multi-level classification. The key conclusions are as follows:

- (1) Multiresolution segmentation that integrates the Canny Edge Detection algorithm helps improve the boundary integrity and separability of segmented objects. In addition, the best segmentation results of corn, buckwheat, wheat and apple are obtained at the segmentation scales of 55, 35, 65 and 65, respectively.
- (2) The redundancy of feature factors of different crops after optimization has been greatly reduced. The best classification results are available by combining the phenological feature factors and the reference images of different crops.

- (3) Two classification models under the multi-level classification framework ensure high accuracy, of which the RF model is overall superior to CNN model. In future studies, the focus can be placed on further refining the models and methods to improve the accuracy and applicability of crop classification.

Although this study has mitigated salt-and-pepper noise to some extent through multi-scale segmentation combined with Canny Edge Detection, the objects generated still diverge from crop "plots", and the classification is less automated. In the face of more complex crop areas with a larger area and subject to frequent cloud and rain conditions, it's important to continuously optimize the algorithms, improve the automation degree as well as the accuracy of remote sensing mapping of crops in complex areas, and expand the potential application value of high-resolution data in China.

Author Contributions: Conceptualization, Y.Q.; Methodology, R.Y. and H.Z.; Software, H.W.; Validation, J.Z.; Formal analysis, X.M.; Investigation, J.Z.; Data curation, C.M.; All authors have read and agreed to the published version of the manuscript.

Funding: Please add: This research was funded by the China High-resolution Earth Observation System (CHEOS) Major Project (Grant No. 92-Y50G34-9001-22/23).

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhu Y, Jia X, Qiao J, et al. What is the mass of loess in the Loess Plateau of China?[J]. *Science Bulletin*, 2019, 64(8): 534–539.
2. Zhang Q, Wei W, Chen L, et al. The Joint Effects of Precipitation Gradient and Afforestation on Soil Moisture across the Loess Plateau of China[J]. *Forests*, Multidisciplinary Digital Publishing Institute, 2019, 10(3): 285.
3. Li Chaoyang. Research of the Development of the Western Agriatral Industrization — — Taking Qingyang City of Gansu as An Example.[J]. *Journal of Northwest A&F University (Social Science Edition)*, 2010, 10(4): 37–41. Hu Q.
4. Sulla-Menashe D, Xu B, et al. A phenology-based spectral and temporal feature selection method for crop mapping from satellite time series[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2019, 80: 218–229.
5. Ozdarici-Ok A, Ok A O, Schindler K. Mapping of Agricultural Crops from Single High-Resolution Multispectral Images—Data-Driven Smoothing vs. Parcel-Based Smoothing[J]. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, 2015, 7(5): 5611–5638.
6. Blickensdörfer L, Schwieder M, Pflugmacher D, et al. Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany[J]. *Remote Sensing of Environment*, 2022, 269: 112831.
7. Ji S, Zhang C, Xu A, et al. 3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images[J]. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, 2018, 10(1): 75.
8. Yang S, Gu L, Li X, et al. Crop Classification Method Based on Optimal Feature Selection and Hybrid CNN-RF Networks for Multi-Temporal Remote Sensing Imagery[J]. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, 2020, 12(19): 3119.
9. Wang L, Ma H, Li J, et al. An automated extraction of small- and middle-sized rice fields under complex terrain based on SAR time series: A case study of Chongqing[J]. *Computers and Electronics in Agriculture*, 2022, 200: 107232.
10. Orynbaikyzy A, Gessner U, Mack B, et al. Crop Type Classification Using Fusion of Sentinel-1 and Sentinel-2 Data: Assessing the Impact of Feature Selection, Optical Data Availability, and Parcel Sizes on the Accuracies[J]. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, 2020, 12(17): 2779.
11. Johnson B A. High-resolution urban land-cover classification using a competitive multi-scale object-based approach[J]. *Remote Sensing Letters*, 2013, 4(2): 131–140.
12. Heumann B W. An Object-Based Classification of Mangroves Using a Hybrid Decision Tree—Support Vector Machine Approach[J]. *Remote Sensing*, Molecular Diversity Preservation International, 2011, 3(11): 2440–2460.

13. Yang C, Everitt J H, Murden D. Evaluating high resolution SPOT 5 satellite imagery for crop identification[J]. *Computers and Electronics in Agriculture*, 2011, 75(2): 347–354.
14. Huawei Mou A, Huan Li B, Yuguang Zhou C, et al. Estimating winter wheat straw amount and spatial distribution in Qihe County, China, using GF-1 satellite images[J]. *Journal of Renewable and Sustainable Energy*, 2021, 13(1): 013102.
15. Meng S, Zhong Y, Luo C, et al. Optimal Temporal Window Selection for Winter Wheat and Rapeseed Mapping with Sentinel-2 Images: A Case Study of Zhongxiang in China[J]. *Remote Sensing, Multidisciplinary Digital Publishing Institute*, 2020, 12(2): 226.
16. Zhang P, Hu S, Li W, et al. Parcel-level mapping of crops in a smallholder agricultural area: A case of central China using single-temporal VHSR imagery[J]. *Computers and Electronics in Agriculture*, 2020, 175: 105581.
17. Zou J, Huang Y, Chen L, et al. Remote Sensing-Based Extraction and Analysis of Temporal and Spatial Variations of Winter Wheat Planting Areas in the Henan Province of China[J]. *Open Life Sciences, De Gruyter Open Access*, 2018, 13(1): 533–543.
18. Shan X, Zhang J. Does the Rational Function Model's Accuracy for GF1 and GF6 WFV Images Satisfy Practical Requirements?[J]. *Remote Sensing, Multidisciplinary Digital Publishing Institute*, 2023, 15(11): 2820.
19. Wu J, Li Y, Zhong B, et al. Integrated vegetation cover of typical steppe in China based on mixed decomposing derived from high resolution remote sensing data[J]. *Science of The Total Environment*, 2023, 904: 166738.
20. Zhou Q, Yu Q, Liu J, et al. Perspective of Chinese GF-1 high-resolution satellite data in agricultural remote sensing monitoring[J]. *Journal of Integrative Agriculture*, 2017, 16(2): 242–251.
21. Xia T, He Z, Cai Z, et al. Exploring the potential of Chinese GF-6 images for crop mapping in regions with complex agricultural landscapes[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 107: 102702.
22. Meng S, Zhong Y, Luo C, et al. Optimal Temporal Window Selection for Winter Wheat and Rapeseed Mapping with Sentinel-2 Images: A Case Study of Zhongxiang in China[J]. *Remote Sensing, Multidisciplinary Digital Publishing Institute*, 2020, 12(2): 226.
23. Zhang P, Hu S, Li W, et al. Parcel-level mapping of crops in a smallholder agricultural area: A case of central China using single-temporal VHSR imagery[J]. *Computers and Electronics in Agriculture*, 2020, 175: 105581.
24. Su T, Zhang S. Object-based crop classification in Hetao plain using random forest[J]. *Earth Science Informatics*, 2021, 14(1): 119–131.
25. Karimi N, Sheshangosht S, Eftekhari M. Crop type detection using an object-based classification method and multi-temporal Landsat satellite images[J]. *Paddy and Water Environment*, 2022, 20(3): 395–412.
26. Li J, Shen Y, Yang C. An Adversarial Generative Network for Crop Classification from Remote Sensing Timeseries Images[J]. *Remote Sensing, Multidisciplinary Digital Publishing Institute*, 2021, 13(1): 65.
27. Du X, Zare A. Multiple Instance Choquet Integral Classifier Fusion and Regression for Remote Sensing Applications[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(5): 2741–2753.
28. Li D, Yang F, Wang X. Study on Ensemble Crop Information Extraction of Remote Sensing Images Based on SVM and BPNN[J]. *Journal of the Indian Society of Remote Sensing*, 2017, 45(2): 229–237.
29. Hinton G E, Osindero S, Teh Y-W. A Fast Learning Algorithm for Deep Belief Nets[J]. *Neural Computation*, 2006, 18(7): 1527–1554.
30. Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84–90.
31. Penatti O A B, Nogueira K, Dos Santos J A. Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?[A]. 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)[C]. Boston, MA, USA: IEEE, 2015: 44–51.
32. LeCun Y, Boser B, Denker J S, et al. Backpropagation Applied to Handwritten Zip Code Recognition[J]. *Neural Computation*, 1989, 1(4): 541–551.
33. Lecun Y. Gradient-Based Learning Applied to Document Recognition[J]. *PROCEEDINGS OF THE IEEE*, 1998, 86(11).
34. Kanda F, Kubo M, Muramoto K. Watershed segmentation and classification of tree species using high resolution forest imagery[A]. IEEE International IEEE International IEEE International Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004[C]. Anchorage, AK, USA: IEEE, 2004, 6: 3822–3825.
35. Wei H, Hu Q, Cai Z, et al. An Object- and Topology-Based Analysis (OTBA) Method for Mapping Rice-Crayfish Fields in South China[J]. *Remote Sensing, Multidisciplinary Digital Publishing Institute*, 2021, 13(22): 4666.

36. Shao Peng. Study on Main Features Information Extraction Technology of High-Resolution Remotely Sensed Image Based on Multiresolution Segmentation.[D]. Jilin University, 2015.
37. Lu T, Zhang B, Hu Y, et al. Computed Tomography Imaging Based on Edge Detection Algorithm in Diagnosis and Rehabilitation Nursing of Stroke Patients with Motor Dysfunction[J]. G. Ramirez. Scientific Programming, 2021, 2021: 1–10.
38. Löw F, Michel U, Dech S, et al. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using Support Vector Machines[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2013, 85: 102–119.
39. Chen zhuling, Jia Kun, Li Qiangzi, et al. Hybrid feature selection for cropland identification using GF-5 satellite image.[J]. National Remote Sensing Bulletin, 2022, 26(7): 1383–1394.
40. Zhang Kexin, Zhang Moucao, Du Jun. Spatial and Temporal Variation Characteristics of Surface Humid Condition in Qingyang from 1981 to 2016[J]. Chinese Agriculture Science Bulletin, 2019, 35(13): 101–106.
41. Nolan S, Unkovich M, Yuying S, et al. Farming systems of the Loess Plateau, Gansu Province, China[J]. Agriculture, Ecosystems & Environment, 2008, 124(1): 13–23.
42. Zhao Y, Zhu W, Wei P, et al. Classification of Zambian grasslands using random forest feature importance selection during the optimal phenological period[J]. Ecological Indicators, 2022, 135: 108529.
43. Shan X, Zhang J. Does the Rational Function Model's Accuracy for GF1 and GF6 WFV Images Satisfy Practical Requirements?[J]. Remote Sensing, Multidisciplinary Digital Publishing Institute, 2023, 15(11): 2820.
44. Ghimire P, Lei D, Juan N. Effect of Image Fusion on Vegetation Index Quality—A Comparative Study from Gaofen-1, Gaofen-2, Gaofen-4, Landsat-8 OLI and MODIS Imagery[J]. Remote Sensing, Multidisciplinary Digital Publishing Institute, 2020, 12(10): 1550.
45. Saleh M A, Ameen Z S, Altrjman C, et al. Computer-Vision-Based Statue Detection with Gaussian Smoothing Filter and EfficientDet[J]. 2022.
46. Zhao J, Yu H, Gu X, et al. The Edge Detection of River model Based on Self- adaptive Canny Algorithm And Connected Domain Segmentation[J]. .
47. Zhao Wenqing, Yan Hai, Shao Xuqing. Object detection based on improved non-maximum suppression algorithm[J]. Journal of Image and Graphics, 2018, 23(11): 1676–1685.
48. Jiang F. Application of canny operator threshold adaptive segmentation algorithm combined with digital image processing in tunnel face crevice extraction[J]. .
49. Li P. Design of Threshold Segmentation Method for Quantum Image[J]. International Journal of Theoretical Physics, 2020.
50. Tab F A, Naghdy G, Mertins A. Scalable multiresolution color image segmentation[J]. Signal Processing, 2006, 86(7): 1670–1687.
51. Dian Y Y, Fang S H and Yao C H. . Change detection for high-resolution images using multilevel segment method. [J]. Journal of Remote Sensing, 2016, 20(1): 129–137.
52. Drăguț L, Csillik O, Eisank C, et al. Automated parameterisation for multi-scale image segmentation on multiple layers[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 88: 119–127.
53. Ma Haoran. Object-based Remote Sensing Image Classification of Forest Based on Multi-level Segmentation[D]. Beijing Forestry University, 2014.
54. Zhang Jun, Wang Yunjia, Li Yan, et al. An Object-Oriented Optimal Scale Choice Method for High Spatial Resolution Remote Sensing Image[J]. Science and Technology Review, 2009, 27(21): 91–94.
55. Ashourloo D, Shahrabi H S, Azadbakht M, et al. A novel method for automatic potato mapping using time series of Sentinel-2 images[J]. Computers and Electronics in Agriculture, 2020, 175: 105583.
56. Wang M, Fei X, Zhang Y, et al. Assessing Texture Features to Classify Coastal Wetland Vegetation from High Spatial Resolution Imagery Using Completed Local Binary Patterns (CLBP)[J]. Remote Sensing, Multidisciplinary Digital Publishing Institute, 2018, 10(5): 778.
57. Zhu Yongsen, Zeng Yongnian, Zhang Meng. Extract of land use/cover information based on HJ satellites data and object-oriented classification[J]. Transactions of the Chinese Society of Agricultural Engineering, 2017, 33(14): 258–265.
58. Fu T, Ma L, Li M, et al. Using convolutional neural network to identify irregular segmentation objects from very high-resolution remote sensing imagery[J]. Journal of Applied Remote Sensing, SPIE, 2018, 12(2): 025010.
59. Hubel D H, Wiesel T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. The Journal of Physiology, 1962, 160(1): 106–154.

60. Uchida K, Tanaka M, Okutomi M. Coupled convolution layer for convolutional neural network[J]. *Neural Networks*, 2018, 105: 197–205.
61. Fortuna-Cervantes J M, Ramírez-Torres M T, Mejía-Carlos M, et al. Texture and Materials Image Classification Based on Wavelet Pooling Layer in CNN[J]. 2022.
62. Więckowska B, Kubiak K B, Józwiak P, et al. Cohen's Kappa Coefficient as a Measure to Assess Classification Improvement following the Addition of a New Marker to a Regression Model[J]. *Int. J. Environ. Res. Public Health*, 2022.
63. An B, Zhang B. Logistic regression with image covariates via the combination of L1 and Sobolev regularizations[J]. *M. Ahn. PLOS ONE*, 2020, 15(6): e0234975.
64. Ma M, Chen J, Liu W, et al. Ship Classification and Detection Based on CNN Using GF-3 SAR Images[J]. *Remote Sensing, Multidisciplinary Digital Publishing Institute*, 2018, 10(12): 2043.
65. Tatsumi K, Yamashiki Y, Canales Torres M A, et al. Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data[J]. *Computers and Electronics in Agriculture*, 2015, 115: 171–179.
66. Wang Z, Zhao Z, Yin C. Fine Crop Classification Based on UAV Hyperspectral Images and Random Forest[J]. *ISPRS International Journal of Geo-Information, Multidisciplinary Digital Publishing Institute*, 2022, 11(4): 252.
67. Wang Y, Zhang Z, Feng L, et al. A new attention-based CNN approach for crop mapping using time series Sentinel-2 images[J]. *Computers and Electronics in Agriculture*, 2021, 184: 106090.
68. Zhao H, Chen Z, Jiang H, et al. Evaluation of Three Deep Learning Models for Early Crop Classification Using Sentinel-1A Imagery Time Series—A Case Study in Zhanjiang, China[J]. *Remote Sensing, Multidisciplinary Digital Publishing Institute*, 2019, 11(22): 2673.
69. Lin Yi, Zhang Wenhao, Yu Jie, et al. Fine Classification of Urban vegetation based UAV images[J]. *China Environmental Science*, 2022, 42(6): 2852–2861.
70. Ismail Fawaz H, Forestier G, Weber J, et al. Deep learning for time series classification: a review[J]. *Data Mining and Knowledge Discovery*, 2019, 33(4): 917–963.
71. Mishra C, Gupta D L. Deep Machine Learning and Neural Networks: An Overview[J]. *International Journal of Hybrid Information Technology*, 2016, 9(11): 401–414.
72. Zou B, Xu X, Zhang L. Object-Based Classification of PolSAR Images Based on Spatial and Semantic Features[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 609–619.
73. Kavzoglu T, Tonbul H, Yildiz Erdemir M, et al. Dimensionality Reduction and Classification of Hyperspectral Images Using Object-Based Image Analysis[J]. *Journal of the Indian Society of Remote Sensing*, 2018, 46(8): 1297–1306.
74. Ding Z, Wang T, Sun Q, et al. Adaptive fusion with multi-scale features for interactive image segmentation[J]. *Applied Intelligence*, 2021, 51(8): 5610–5621.
75. Jiang N, Shi H, Geng J. Multi-Scale Graph-Based Feature Fusion for Few-Shot Remote Sensing Image Scene Classification[J]. *Remote Sensing, Multidisciplinary Digital Publishing Institute*, 2022, 14(21): 5550.
76. Chen S, Useya J, Mugiyo H. Decision-level fusion of Sentinel-1 SAR and Landsat 8 OLI texture features for crop discrimination and classification: case of Masvingo, Zimbabwe[J]. *Heliyon*, 2020, 6(11): e05358.
77. Khojastehnazhand M, Roostaei M. Classification of seven Iranian wheat varieties using texture features[J]. *Expert Systems with Applications*, 2022, 199: 117014.
78. Zhang D, Ying C, Wu L, et al. Using Time Series Sentinel Images for Object-Oriented Crop Extraction of Planting Structure in the Google Earth Engine[J]. *Agronomy, Multidisciplinary Digital Publishing Institute*, 2023, 13(9): 2350.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.