An Object-Based Approach for Mapping Crop Coverage Using Multiscale Weighted and Machine Learning Methods

Zengwei Tang, Hong Wang, Xiaobing Li, Xiaohui Li, Wenjie Cai, and Chongyuan Han

Abstract—**Accurate mapping of crop distribution on Earth's surface aids in predicting grain production. Pattern classification along with remote sensing imagery can facilitate traditional manual field measurement techniques using machine learning. With the rapid increase in satellite sensor resolution, the object-based classification paradigm has increasingly been applied. However, scale parameter selection is always a difficult part of the object-based classification. Based on ensemble learning, this study proposes a classification method using the multiscale object-based weighted method which includes manual digitizing of crop distribution in the southern region of Jishan County, Shanxi Province, China, applying Gaofen-2 (GF-2) images. This method initially uses estimations of the scale parameter (ESP) tool to select "good" scales, defined here as "preferred" scales, after which feature subsets are screened by each preferred scale as the input of multiple classifiers and classifies. Finally, all classification results are then fused. Our research results indicate that: 1) Feature importance values are sorted differently at different preferred scales; 2) accuracy differences become clear when different preferred scales are combined with different classifiers, and determining the "best" single appropriate scale is generally difficult; 3) accuracy of the multiscale weighted classification method is higher compared to the single preferred scale approach. Furthermore, ensemble learning can be achieved using this method on multiple scales and on multiple classifiers. With this method, procedures that necessitate the selection of segmentation scales and the selection and optimization of classifiers can be skipped altogether.**

Index Terms—**Ensemble learning, feature selection, GF-2, multilayer perceptron (MLP), object-based image analysis (OBIA), preferred scales, support vector machine (SVM).**

I. Introduction

CROP coverage information on the Earth's surface is very important for grain security and crop monitoring [1]–[5]. The accurate and timely mapping of surface crop distribution as well as the application of sound management policie, are helpful for the production and prediction of grain crops [6], [7].

Manuscript received May 19, 2019; revised December 21, 2019, January 20, 2020, and February 22, 2020; accepted March 17, 2020. Date of publication April 16, 2020; date of current version May 5, 2020. This work was supported in part by the National Key Basic Research Program of China under Grant 2016YFC0500502, in part by the National Innovation Team Project under Grant 41321001, and in part by the Ministry of Education Innovation Team Project under Grant IRT_15R06. *(Corresponding author: Hong Wang.)*

Zengwei Tang, Hong Wang, Xiaobing Li, Xiaohui Li, and Wenjie Cai are with the Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China (e-mail: tangzengwei@mail.bnu.edu.cn; wanghong@bnu.edu.cn; 80263051@qq.com; lixiaohui1820@163.com; 741075286@qq.com).

Chongyuan Han is with the Shanxi Investigation Team, National Bureau of Statistics, Taiyuan 030001, China (e-mail: 810784926@qq.com).

Digital Object Identifier 10.1109/JSTARS.2020.2983439

Many methods have been developed to map the distribution of crops on the Earth's surface [8]–[12], wherein methods that classify crops on the Earth's surface using remote sensing and machine learning theory have become popular. These methods greatly reduce the workload of field measurement. There are primarily two reasons for this: First, the development of "big data" and high-performance graphics processing units (GPUs), wherein machine learning has been widely applied to fields such as expert system, cognitive simulation, data mining, natural-language understanding, network information service (NIS), remote sensing image classification [13], [14], etc.; second, remote sensing techniques have progressively developed due to the many high-resolution remote sensing satellites that have launched in recent years [5]. Such developments have greatly increased access to basic agricultural data [15], [16] while reducing the cost of data acquisition and subsequently popularizing agricultural remote sensing mapping.

In terms of the unit size of the analysis, methods that classify earth surface coverage categories comprise two categories: Pixel-based image analysis (PBIA) and object-based image analysis (OBIA) [17]. The most significant difference between the two approaches is the different units they use during image processing: PBIA uses pixels as units while OBIA uses objects as units, the latter being homogeneous and consistent in using the accumulation of pixels as a unit. For a long time, PBIA has primarily been used for remote sensing image classification, but OBIA has progressively become more popular in the past ten years [18], [19]. Hay and Castilla [20] hypothesized that pixels are not isolated. Objects that comprise of a selection of pixels subdivide data in maps into homogeneous units representing actual surface features [21]. Compared to PBIA, OBIA has the advantage of being able to acquire spatial (location, size, shape, etc.) information of more objects [5], [18], [20], [22]–[25] and can effectively reduce the problem of spectrum heterogeneity [26], [27] and the "salt and pepper effect" following image classification [28]. Many studies have reported that OBIA is better at achieving higher classification accuracies compared to PBIA [29]–[33]. Platt and Rapoza [34] classified surface coverage of an urban-suburb-agriculture mixed area located in Gettysburg, Pennsylvania, based on IKONOS satellite images. The results indicated that the object-based nearest neighbor

(NN) classifier achieved the best classification result with an accuracy as high as 78%, but the highest value achieved for pixel-based classification accuracy was only 64% after the machine learning classifier (MLC) was used. Duro *et al.* [17] used SPOT satellite data to compare OBIA and PBIA results with different classifiers. The results indicated that the accuracy of OBIA classification was much higher compared to the accuracy of PBIA classification for the decision tree, random forest (RF) and support-vector machine (SVM) classifiers. Belgiu and Csillik [35] used Sentinel-2 satellite data to respectively draw maps of crop coverage in three different study areas using both PIBA and OIBA time dynamic weighting methods. The three study areas were respectively located in Romania, Italy and the USA. The results indicated that OBIA results were on average 2% higher compared to those of PBIA. With the development of high-resolution imagery, advantages of OBIA technology will become progressively more clear in earth surface coverage monitoring.

Image segmentation is not only the core technology of OBIA [17], it is also the first step of OBIA. Segmentation divides the image into homogeneous units that exhibit similar spectra and adjacent spaces [36]. These units show real surface features which to a certain extent have real meaning. The segmentation scale parameter determines the size of the object after segmentation. It is often determined according to the trial-and-error method as well as by subjective perception [25]. Thus, excessive or insufficient segmentation often results. The features of objects segmented out cannot describe the attributes of true objects on the Earth's surface [37], [38]. Additionally, such features affect classification accuracy [39], [40]. Thus, many studies have offered different methods on how to choose the best segmentation scale. These methods are both supervised [41], [42] and unsupervised [43], [46]. However, these methods still have certain defects with respect to the selection of the segmentation scale and final image classification: 1) Most of these methods are too focused in determining a single optimal scale to segment the whole image while assuming that the segmentation effect of the image along with the optimal scale is the best means forward. Earth's surface coverage categories are often very complex. Segmentation scales that correspond to different earth surface coverage categories may often differ [47]. When using a single optimal scale, information on a certain land type may prominent, but information on other land types will be unavoidably damaged; thus, it is very difficult to define a single or an optimal segmentation scale parameter [45]. Perhaps a better strategy is to name the "optimal" scale as the "preferred" scale. Different preferred scales correspond to different semantically significant regions. These preferred scales can be used to segment images "well" if not "best". For image classification, however, it is not necessary to pay too much attention to scale selection, which results from immediate process runs, if the final classification accuracy is high enough or even allows for multiple preferred scales to appear simultaneously during the classification process. 2) There is too little information regarding the importance in the performance of classification features of objects at different segmentation scales. 3) Finally, there is also too little information regarding the performance of all categories of machine learning algorithms at different scales.

Based on the above information, the purpose and contextual structure of this study are as follows: 1) We discuss the importance of classification features at different preferred scales; 2) we also discuss all categories of machine learning algorithms at different preferred scales; 3) we propose the use of a multiscale object-based weighted classification method which initially uses segmentation results at multiple preferred scales as the inputs of classifiers, then obtains weighted values according to classification results of each scale and lastly fuses classification results at different preferred scales according to the weighted value on a pixel level to obtain the final classification results; 4) we verify that the accuracy of the multiscale weighted classification method is higher than single-scale classification accuracy, which does not specifically mean that the final segmentation and classification results will be the "best" when single-scale accuracy is evaluated as being the "best" as we have previously affirmed; and 5) we map the main crops in the southern region of Jishan County, Shanxi Province, China, using Gaofen-2 (GF-2) image data and the multiscale weighted classification method to provide decision support for local agricultural management and sustainable development strategies as well as a reference for the classification of earth surface coverage categories of agricultural areas in other regions.

II. Materials and Methods

A. Study Area

The study area (35°22′–35°48′ N, 110°48′–111°54′ E; as shown in Fig. 1) is in Taiyang Township in the southern region of Jishan County, Shanxi Province, China. It is part of the tributary basin of the Yellow River, with an elevation of 500 m a.s.l. The study area is under the influence of a temperate continental monsoon climate, with an annual average temperature of 13 °C, namely, −4 °C in January and 27 °C in July. Annual rainfall is 483 mm. The frost period lasts from mid-October to mid-April of the following year. There are 220 days without frost in total. The terrain is flat, land is fertile and the average annual sunshine hoursare up to 2040 h within the study area. The study area abounds with grain crops and cash crops such as corn, wheat, walnut, etc., and is renowned as a model county for leisure agriculture and rural tourism in Shanxi Province. Our research group set up a test base in this location and established a large number of quadrats to monitor local agricultural planting distribution.

B. Images and Samples

Cloudless 4th issue GF-2 images from 2017 (June 20th, July 9th, December 4th , and January 13th) were used as remote sensing base maps (as shown in Fig. 2). The image selection takes into account the phenological period of the main crops. The jointing period of corn is in the middle of June, and the tasseling period is in the middle of July. The overwintering period of wheat is in the early December. The spring shoots of walnuts

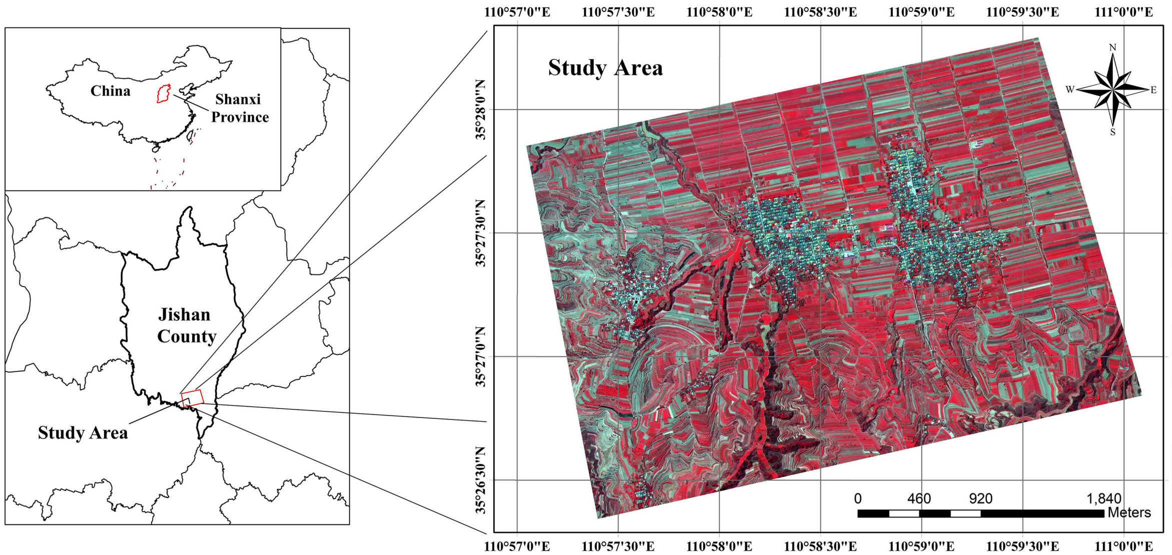


Fig. 1. Location of the study area.

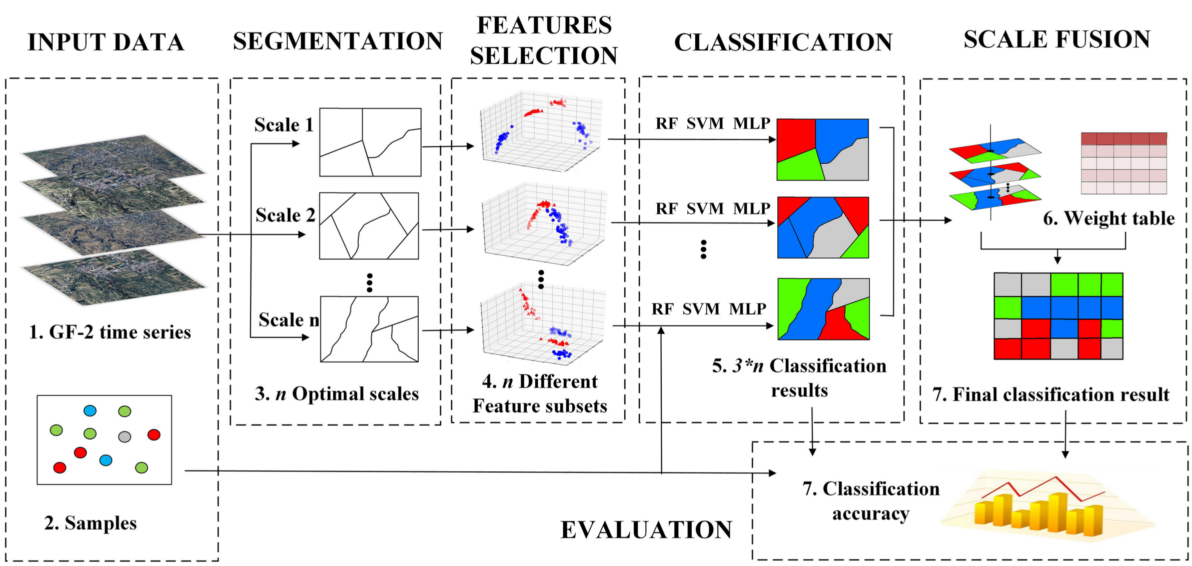


Fig. 2. Research flow chart.

slowing down period are between mid-June and mid-July. GF-2 is the first independently researched civil optical remote sensing satellite issued by China, with a spatial resolution of 0.8 m. Its observation width is 45 km, which is the highest among international sub-meter resolution satellites. GF-2 can be widely applied, namely, for land resource monitoring, exploitation of mineral resources, urban fine management, reconstruction of disaster areas, etc. The multispectral sensor carried by GF-2 can provide images within the four wave bands of red (0.45–0.52 μm), green (0.52–0.59 μm), blue (0.63–0.69 μm) and near-infrared (0.77–0.89 μm). In this study, we first preprocessed the 16 images (4 periods $\times$ 4 bands), including radiation correction, orthorectification, registration, fusion with the panchromatic band and clipping, and then stacked them into a time sequence layer used as the input of image segmentation processing.

Samples from the study area were from data obtained from a 2017 field investigation. We randomly established quadrats in the study area each year; thus, there are a great quantity of sample data for use in our study. These quadrats are in the form of polygons (not points), and their size corresponds to actual farmland borders. We also added some samples by visual interpretation to ensure both adequate and uniform distribution of samples. In total, there were 893 experimental quadrats. For cross validation, we randomly and uniformly divided the samples into sample set A and sample set B.

C. Classification System

The study area mainly comprises of corn, wheat, walnut and "other crops." We also added the other land types within the study area into our classification system. There are eight categories in total: corn, wheat, walnut, other crops, building, thinned forest, grassland, and bare land.

D. Image Segmentation

Preferred scale parameters were selected prior to image segmentation, and we applied the estimation of scale parameter (ESP) tool, which indicates the object segmentation effect by calculating the rates of changes (ROC) of local variance (LV) of image object homogeneity at different segmentation scale parameters [44], [48]. When the peak value of ROC of LV (ROC-LV) appears, the segmentation scale corresponding to the point is considered a "preferred" segmentation scale. The tool calculates results by setting the initial segmentation scale and increasing the step length. Given that we wanted to obtain the segmentation results of multiple scales, we selected n peak value points for this experiment and defined the scale parameters corresponding to these peak value points as the "preferred" scales (not "optimal" scales). This is different from the methods used by previous studies that only selected one peak value point [35], [49], [50].

We used multi-resolution algorithm [51] to segment images in eCognition Developer 9.1 [52] after obtaining *n* preferred scale parameters. The algorithm is a top-bottom method, achieving image segmentation on the basis of the region of merging technology under the precondition of ensuring minimum average heterogeneity among objects and maximum homogeneity among the internal pixels of the objects [52], [53]. The algorithm is widely believed to be one of the best methods for solving the OBIA problem [41]. Besides controlling the relative size of the generated object with a scale parameter, the algorithm also uses a spectrum parameter and a shape parameter to represent the homogeneity of the segmented object. The weighted sum of the two is 1.0. The shape is represented by smoothness and compactness, whereby the weighted sum is also 1.0.

We respectively set the spectrum parameter and shape parameter to 0.7 and 0.3 because spectrum features play a key role in differentiating different objects in the study process [54], and the study area mainly comprises of regular shaped farmland. Given that any detection of differences largely failed when the compactness parameter was adjusted, we respectively set the same weight (0.5) to both the compactness and smoothness parameters. With the ESP tool, scales less than 50 resulted in obvious excessive segmentation, and scales greater than 140 resulted in obvious insufficient segmentation. Thus, we did not use scales less than 50 or greater than 140. Accordingly, we set the preferred scale range from 50 to 140 (initial scale: 50; step length: 1)

E. Feature Set Construction and Feature Extraction

The object feature can be selected based on user experience and user recognition [17], [55] or through relevant algorithms [29], [56] during OBIA. We first selected a large number of features based on user knowledge to construct an initial feature set, and then this feature set was screened using an algorithm to obtain the final feature subsets for classification. Generally, there was a great amount of redundancy in the initial feature sets. Such redundancy greatly impacted the running efficiency of the program and led to its overfitting and low generalization. Therefore, it was necessary to select important feature from the initial feature sets.

TABLE I
VEGETATION INDEX FEATURES

Vegetation indices (VI)	VI adapted to GF-2	Layers
VIgreen [57]	(band3-band2)/(band3+band2)	4Dates
RVI [58]	band4/band3	4Dates
DVI [59]	band4-band3	4Dates
NDVI [60]	(band4-band3)/(band4+band3])	4Dates

TABLE II
TEXTURE FEATURES

Texture Features	Description	Layers
Homog (Homogeneity)	Measurement of partial gray homogeneity of image	4Dates*4bands
Con (Contrast)	Measurement of image definition and texture groove depth	4Dates*4bands
Correl (Correlation)	Measure the similarity of image gray level in row or column	4Dates*4bands
ASM (Angular second moment)	Reflect image gray distribution homogeneity and texture thickness	4Dates*4bands
Ent (Entropy)	Measure the randomness of information quantity contained in image	4Dates*4bands

Features related to the spectrum, vegetation index, texture and shape were used to construct the initial feature set of the image object, wherein the spectrum feature contained the 16 wave bands of the four periods of the input image. Given that the GF-2 sensor comprises of the blue, green, red and near-infrared wave bands, and on the basis of results from previous studies, we selected the four vegetation index features, namely, the visible atmospherically resistant indices green (VIgreen), the ratio vegetation index (RVI), the difference vegetation index (DVI) and the normalized difference vegetation index (NDVI) (Table I) as well as the most common texture and geometric features (shown in Tables II and III).

During feature screening, we used the SelectKBest method, one of the single variable feature selection methods in the Scikit-learn base to remove 50% of features in the initial feature set. The principle is as follows: The mutual information score of each variable is separately calculated by the chi-square test, and the higher the score is, the more important the feature will be; and then all features of which the scores are listed after the top *K* are subsequently removed.

F. Classifier Model

We used three typical machine learning algorithms, namely, RF, SVM and multilayer perceptron (MLP) for image object classification. In the whole flow (Fig. 2), segmentation objects at each preferred scale were classified using the three algorithms, and 3∗n times of classification were carried out. All processes

TABLE III
GEOMETRIC FEATURES

Geometric Features	Description [53]	Layers
DS (Density)	The closer the object shape is to a square shape, the larger the index will be	1Scale
LW (Length/Width)	Length/width of the outer-wrapped rectangle of object	1Scale
BI (Border Index)	Describes the saw teeth of the image object; the more irregular they are, the larger the index will be	1Scale
CP (Compactness)	The more compact the image object is, the smaller the index will be	1Scale
RD (Roundness)	Describes the similarity between an image object and the ellipse	1Scale
RF (Rectangular Fit)	Describes the matching degree between the image object and the rectangle of similar size and proportion	1Scale

were performed based on python programming. We also used the same classifier parameters at each preferred scale to verify that the accuracy of the final classification result using multiple preferred scales was better than that of the single optimal scale.

The RF classifier randomly establishes multiple decision trees which subsequently form a decision tree forest, and then the decision is made through the voting of multiple trees [61]. The classifier of the category must have good robustness [62], [63]. Its basic procedures are as follows: A plurality of sample sets is generated by resampling the existing samples and then used to simulate the randomness in data, and the influence of the randomness is considered in the final result. This method not only samples training samples but also samples features. It sufficiently ensures the independence of all constructed trees and ensures nonbiased voting results. To train the RF classifier, the quantity of decision trees (Ntree) and the optimal rate (Mtry) of the training set and test set are required, which are formed by dividing the feature variables in forest growth and the quality function (criterion) for segmentation measurement. We used default values (Ntree = 100, Mtry = sqrt (n_features), criterion = gini) in this study.

The SVM was proposed by Vapnik [64]. Its basic procedures are as follows: Convert the nonlinear problem in the former space into the linear problem in the new space by introducing feature transformation. The SVM must choose different parameters according to different core functions. In this study we used the default RBF core and optimized its core function width. The SVM is applicable for the binary classification problem; thus, we must also set the classification strategy of models in multiple classifications, namely, use the one-vs.-rest or one-vs.-one strategies. We used the former in this study, which is simple and quick [65].

The MLP is a feed-forward neural network in which information only flows unidirectionally. Forward information moves from the input layer and then passes through the hidden layer and output layer [66]. After model testing and optimization, the three hidden layers were finally set, and each layer comprised of 60 nodes. We used the backpropagation algorithm (BP algorithm) for training. The objective function of the algorithm is the mean square error of the prediction output and expectation output of the neural network on all training samples. Minimization of the objective function was achieved using the stochastic gradient descent (SGD) method through the adjustment of weights in all layers.

G. Multiscale Weighted Model

This study proposes a multiscale weighted model that fuses the abovementioned 3∗n classification results at the pixel layer on the basis of the scoring mechanism (Fig. 2). The benefit of this model is that it skips the need of having to select optimal scales and classifiers. All processes were performed based on python programming. Its steps are as follows:

1) Select the classifier results of which the overall accuracy is the highest at n different scales for voting and scoring respectively; there are n results.
2) Construct the scoring table (Fig. 2); calculate the classification accuracy of the eight categories respectively at n different scales; and for each scale, the category of which the accuracy is highest is scored as n while the second highest is scored as $n-1$, and so on, until reaching the lowest accuracy, which is scored as 1.
3) For the first pixel of the image, there is one judgment score for each of the eight attribution categories, and the judgment score is initialized as 0 (Fig. 3). Assume the pixel is judged as category C_1 at the first scale, determine the score of the C_1 category at the first scale according to the scoring table, and accumulate the score into the judgment score table of the C_1 category. Traverse the category attribution of the pixel at n scale classification results using the same method and determine the final judgment score of each category according to the scoring table.
4) Take the category of which the final judgment score of the pixel is highest as the final attribution.
5) Repeat steps 3 and 4 according to the same method and judge the final attribution category of each pixel.

H. Experimental Procedures

The experimental flow of the study mainly comprised of the following seven procedures (Fig. 2):

1) Create input data: Include the time series which are stacked with images from 16 wave bands of four periods and sample data.
2) Image segmentation: Choose n preferred scales as the scale parameter of image segmentation. Segment the image for n times and obtain n segmentation results.
3) Features selection: Determine the initial set of classification features of the image object. Filter the initial feature set and determine the classification feature subset of n segmentation results.
4) Object classification: Use different classifiers, take sample set A for training and sample set B for testing to classify segmentation result objects at multiple scales.

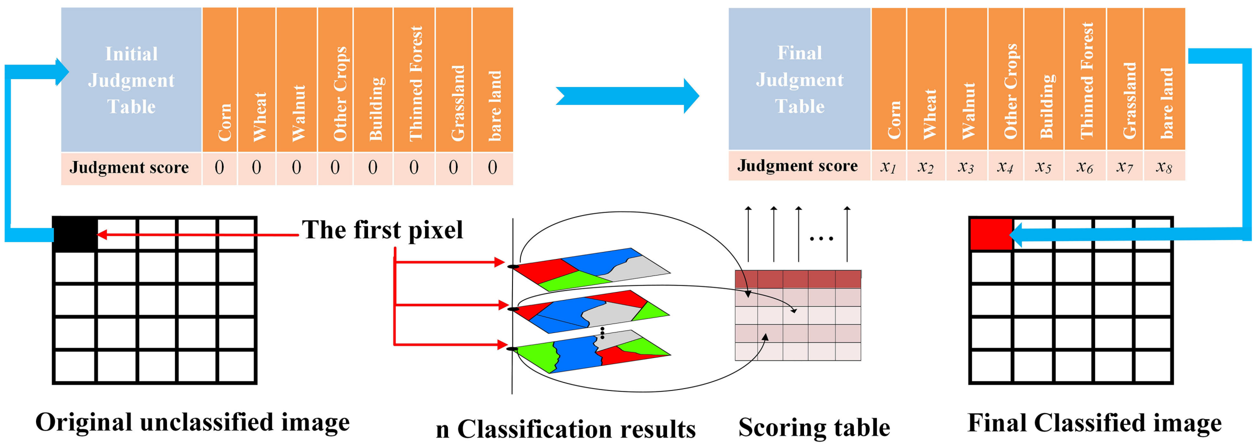


Fig. 3. Each pixel has a judging table with an initial value of 0. We take the category of which the final judgment score of the pixel is highest as the final attribution (the category corresponding to the maximum value in x1 to x8).

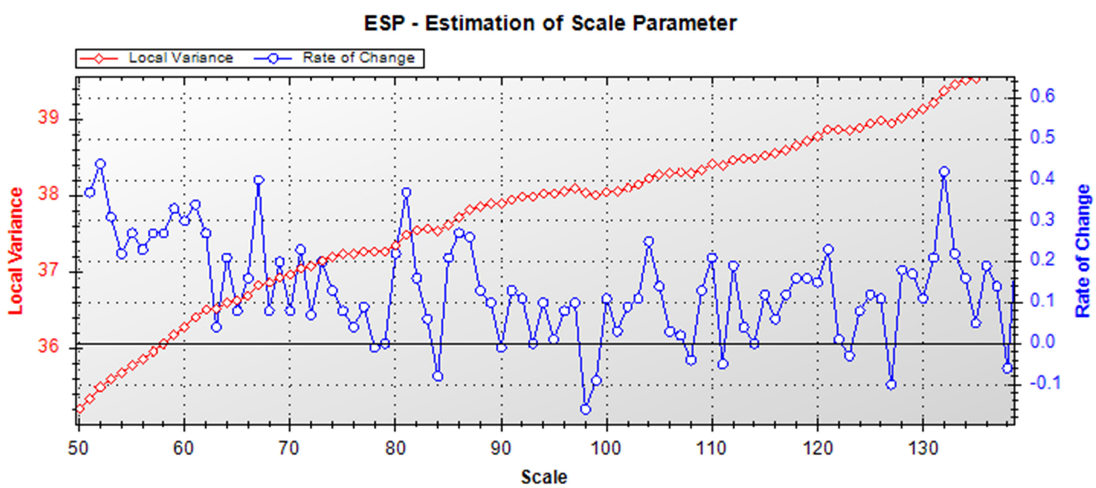


Fig. 4. Preferred scale results, wherein the red curve is the local variance (LV) and the blue curve is the rates of changes (ROC) of LV (i.e., ROC-LV). The scales corresponding to the wave peak of the ROC-LV represent potential preferred scales. We selected eight preferred scales, including 52, 61, 67, 81, 86, 104, 121, and 132, according to the peak positions of waves.

5) Scale weighting: Calculate the accuracy of classification results at different scales with different classifiers. For each scale, select the optimal classification result of the scale, and construct a better classification result set. Create the scoring table and use the weight values in the scoring table to fuse the result set to act as the final classification result.
6) Assess the accuracy of the final classification result of step 5.
7) Exchange sample set A and sample set B; namely, use sample set B for training and sample set A for testing. Repeat steps 5 and 6.

III. Results

A. Image Segmentation Results

The preferred results of segmentation scales are shown in Fig. 4. The ROC-LV curve shows a plurality of wave peaks that correspond to significant semantic regions [50] and are therefore potential preferred scales. Assuming an ideal condition whereby all ground classes correspond to a segmentation scale, we finally selected eight preferred scales (52, 61, 67, 81, 86, 104, 121, and 132) according to the wave peak position in the ROC-LV curve.

These abovementioned eight scale parameters were used for multiresolution segmentation from which we obtained the final segmentation results. Details of two different sites in the study area were selected and are shown in Fig. 5. As indicated by this figure, the object number of the image gradually decreased with an increase in scale. The first segmentation detail in Fig. 5(a) shows that segmentation results of scales 67 and 81 were good. Obvious excessive segmentation (such as seen in the black box of scale 61 in Fig. 5(a)) will occur when segmentation is lower than the two scales. On the other hand, obvious insufficient segmentation (such as seen in the black box of scale 132 in Fig. 5(a)) will occur when segmentation is higher than the two scales. Although obvious excessive segmentation can be seen

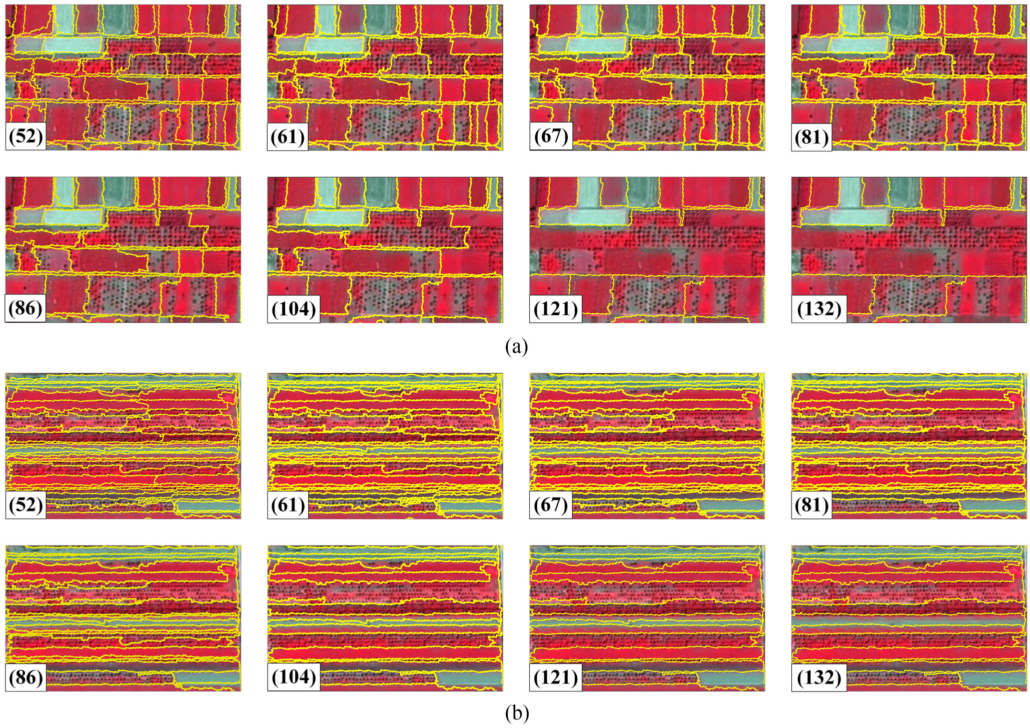


Fig. 5. Details of image segmentation wherein (a) and (b), respectively represent segmentation details of different parts of the same image. The number in the lower- left corner of the image represents the scale. The black box of scale 61 in (a) shows obvious excessive segmentation. The black box of scale 132 in (a) shows obvious insufficient segmentation.

in scales 67 and 81 in Fig. 5(b), the segmentation effects of scales 121 and 132, which we at first believed were excessive segmentation scales in Fig. 5(a), were good. Note that these two segmentation detail sets (i.e., a and b in Fig. 5) are segments (bands) taken from different parts of the same image.

B. Feature Extraction Results

In the process of constructing the object feature set, we have established a total of 118 features, including 16 spectrum features (4 periods $\times$ 4 bands), 6 geometric features, 16 vegetation index features (4 periods $\times$ 4 categories) and 80 texture features (4 periods $\times$ 4 bands $\times$ 5 categories). We selected the features from which the top 60 scores were ranked according to the single variable feature selection method at each scale and obtained eight feature subsets as classifier inputs. Here, we only show the feature selection results of scales 52 and 132 (Fig. 6).

C. Classification Results at Different Scales

This study used a sample set A for training and sample set B for testing. Additionally, the classification procedure was repeated 24 times wherein 24 classification results were obtained with RF, SVM and MLP classifiers, respectively, at eight scales. To verify the feasibility of the follow-up model, we exchanged sample sets A and B. Therefore, sample set A was used for testing and sample set B was used for training using the same classification procedure (24 repetitions). Classification results are shown in Figs. 7 and 8, while Table IV shows classification accuracy (Bolded numerals represent classification accuracy of the optimal classifier at the corresponding scale). Here we should pay attention to our accuracy assessment method. Our preferred scale weighting method was carried out by using the pixel as a unit; thus, our accuracy assessment also applies to the pixel layer. All classifiers at different scales performed differently. We found that RF, SVM and MLP classifiers all appeared after counting

TABLE IV
CLASSIFICATION ACCURACY AT EACH PREFERRED SCALE

	based on sample set A used for training and sample set B used for testing(%)			based on sample set B used for training and sample set A used for testing(%)		
	RF	SVM	MLP	RF	SVM	MLP
Scale52	83.19	**85.19**	84.60	82.71	80.82	**83.35**
Scale61	82.80	83.57	**84.62**	**83.06**	80.02	82.90
Scale67	82.82	**83.23**	82.87	**83.04**	79.85	82.57
Scale81	82.36	82.74	**84.30**	79.80	81.82	**83.23**
Scale86	83.57	83.30	**84.15**	79.16	81.14	**82.37**
Scale104	80.86	79.40	**83.79**	79.11	**79.54**	78.64
Scale121	**79.22**	75.27	78.27	78.79	78.67	**80.93**
Scale132	76.78	75.71	**79.16**	76.95	77.82	**79.51**
Average	81.45	81.05	82.72	80.32	79.96	81.68

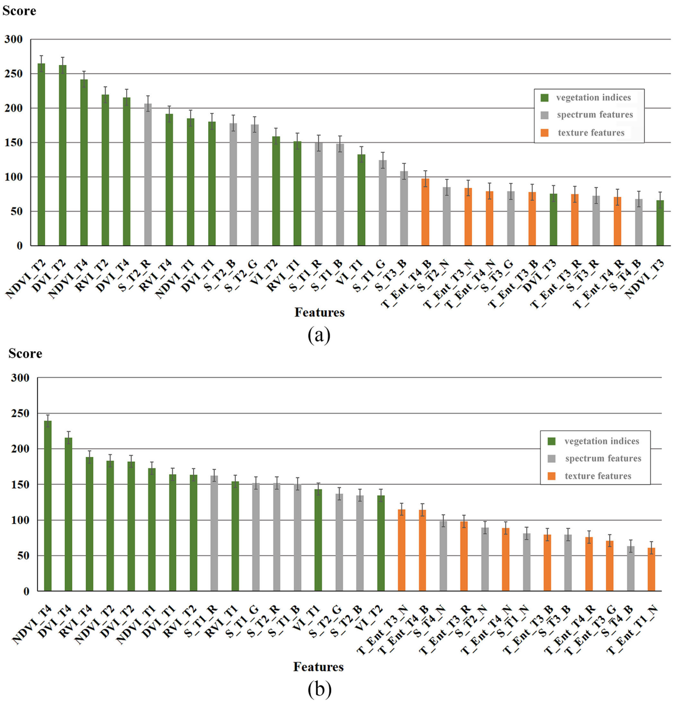


Fig. 6. Feature selection results (a) and (b) respectively show feature selection results of the segmentation results at scales 52 and 132. Limited by the figure size, we only show the features from which the top 30 scores were ranked. "S" represents spectrum features; "T" and its abbreviation represents texture features; corresponding abbreviations represent vegetation indices. T1, T2, T3 and T4 respectively represent images from June, July, December and January. B, G, R, and N respectively represent blue, green, red and near-infrared wave bands.

TABLE V

Scoring Table (Based on Sample Set A Used for Training and Sample Set B Used for Testing)

Scoring Table	Corn	Wheat	Walnut	Other Crops	Building	Thinned Forest	Grassland	Bare land
Scale52	4	3	6	1	8	5	7	2
Scale61	4	3	7	2	8	5	6	1
Scale67	4	2	5	3	8	7	6	1
Scale81	7	3	5	2	8	1	6	4
Scale86	5	2	6	4	8	1	7	3
Scale104	6	1	7	2	8	3	6	4
Scale121	7	4	3	2	6	1	5	8
Scale132	4	2	5	7	6	1	3	8

TABLE VI

Scoring Table (Based on Sample Set B Used for Training and Sample Set A Used for Testing)

Scoring Table	Corn	Wheat	Walnut	Other Crops	Building	Thinned Forest	Grassland	Bare land
Scale52	6	3	8	5	1	4	7	2
Scale61	8	4	7	2	6	1	5	3
Scale67	8	3	6	2	5	1	4	7
Scale81	6	2	5	8	7	4	1	3
Scale86	5	1	7	8	6	3	2	4
Scale104	7	3	6	2	8	1	5	4
Scale121	6	3	7	2	8	1	4	5
Scale132	7	4	3	2	8	1	5	6

the classifiers of which accuracy was highest at each preferred scale; however, the MLP classifier appeared most often, and its average accuracy was higher than those of the other two (i.e., RF and SVM). From a scale aspect, we found that classification accuracy at 104, 121 and 132 was not ideal for all sample sets used for training. When sample set A was used for training, the SVM classifier had the highest accuracy (85.19%) at scale 52, and the SVM classifier had the lowest accuracy (75.27%) at scale 121. When sample set B was used for training, the MLP classifier had the highest accuracy (83.35%) at scale 52, and the RF classifier had the lowest accuracy (76.95%) at scale 132.

D. *Final Classification Results*

Classification results of the optimal classifiers at the preferred scales of the eight category sets were counted before the accuracy of each category in each classifier was calculated. Scoring results are provided in Tables V and VI according to the weight setting method described in section 2.7. These scoring tables clearly show the preferred scale applicable to each category and the accuracy sequencing of each category at each scale. The final weighting was conducted according to the accuracy values of each category at different scales.

The eight classification result sets were weighted at a pixel scale by the scoring tables, and the results are shown in Figs. 7(i) and 8(i). Tables VII and VIII were generated after the accuracy assessment of the final fused results. Results showed that the final accuracy achieved after integrating classification results at multiple preferred scales in the experiment that used a sample set A for training and sample set B for testing was as high as 87.43%, which was higher than classification results from any single preferred scale. Compared to Table IV, our method achieved an accuracy 2.24% higher than the SVM classifier at a single scale of 52 as well as 12.16% higher than the SVM classifier at a single scale of 121. Moreover, the final accuracy achieved for the experiment that used sample set B for training and sample set A for testing also reached 86.20%. Compared to Table IV, our method achieved an accuracy of 2.85% higher than the MLP classifier at a single scale of 52 as well as 9.25% higher than the RF classifier at a single scale of 132.

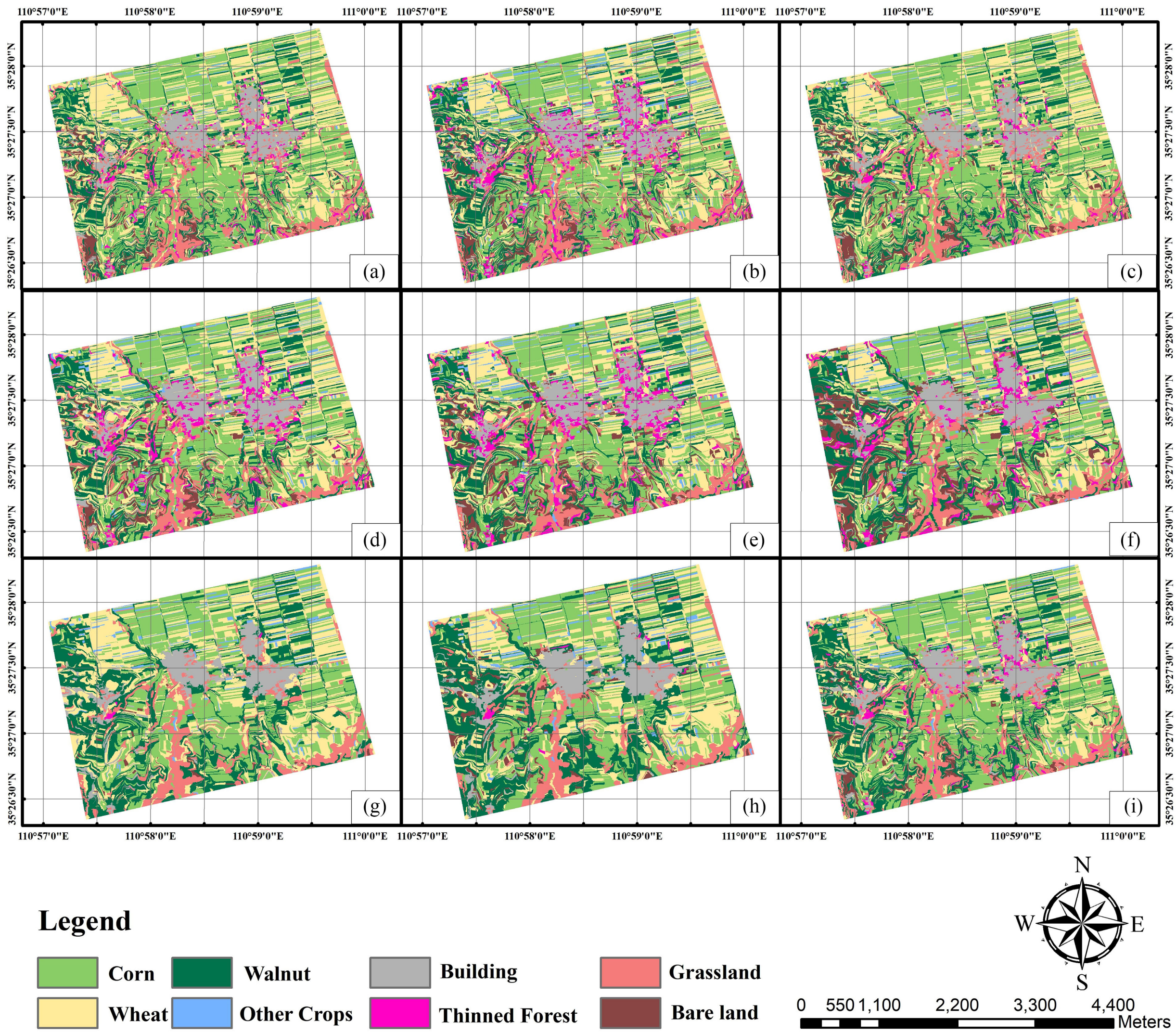


Fig. 7. Classification results when sample set A was used for training and sample set B was used for testing. (a) to (h) are respectively the classification results of the classifiers with the highest accuracy at scales 52, 61, 67, 81, 86, 104, 121, and 132. (i) is the result after (a) to (h) were weighted and fused, i.e., the final mapping result for different crop and land type distribution within the study area.

TABLE VII

MATRIX OF THE ACCURACY ASSESSMENT OF THE FINAL RESULTS (SAMPLE SET A WAS USED FOR TRAINING AND SAMPLE SET B WAS USED FOR TESTING)

	Corn	Wheat	Walnut	Other Crops	Building	Thinned Forest	Grassland	Bare land
Corn	133746	6040	10652	2231	1571	298	857	2545
Wheat	3990	189399	2912	1951	32	39	43	110
Walnut	3257	1607	190707	3	495	296	2112	73
Other Crops	1679	10912	516	20240	0	0	0	0
Building	145	661	311	780	58312	44	800	0
Thinned Forest	471	1055	8276	17	980	16503	7771	323
Grassland	6065	2993	7373	16	1227	466	102453	96
Bare land	2058	6773	1308	0	940	0	562	24597
Overall accuracy					87.438%			

IV. DISCUSSION

A. Preferred Scale and Image Segmentation Effect

We found that the ROC-LV curve had, in fact, multiple obvious wave peaks during the preferred scale process, which just proves that one image can be segmented through the usage of multiple scales. It is unsuitable to selectively select a wave peak as the optimal scale and assume that the image segmentation effect at this scale is the best. This is because coverage types on the Earth's surface are very diverse. Although one single optimal

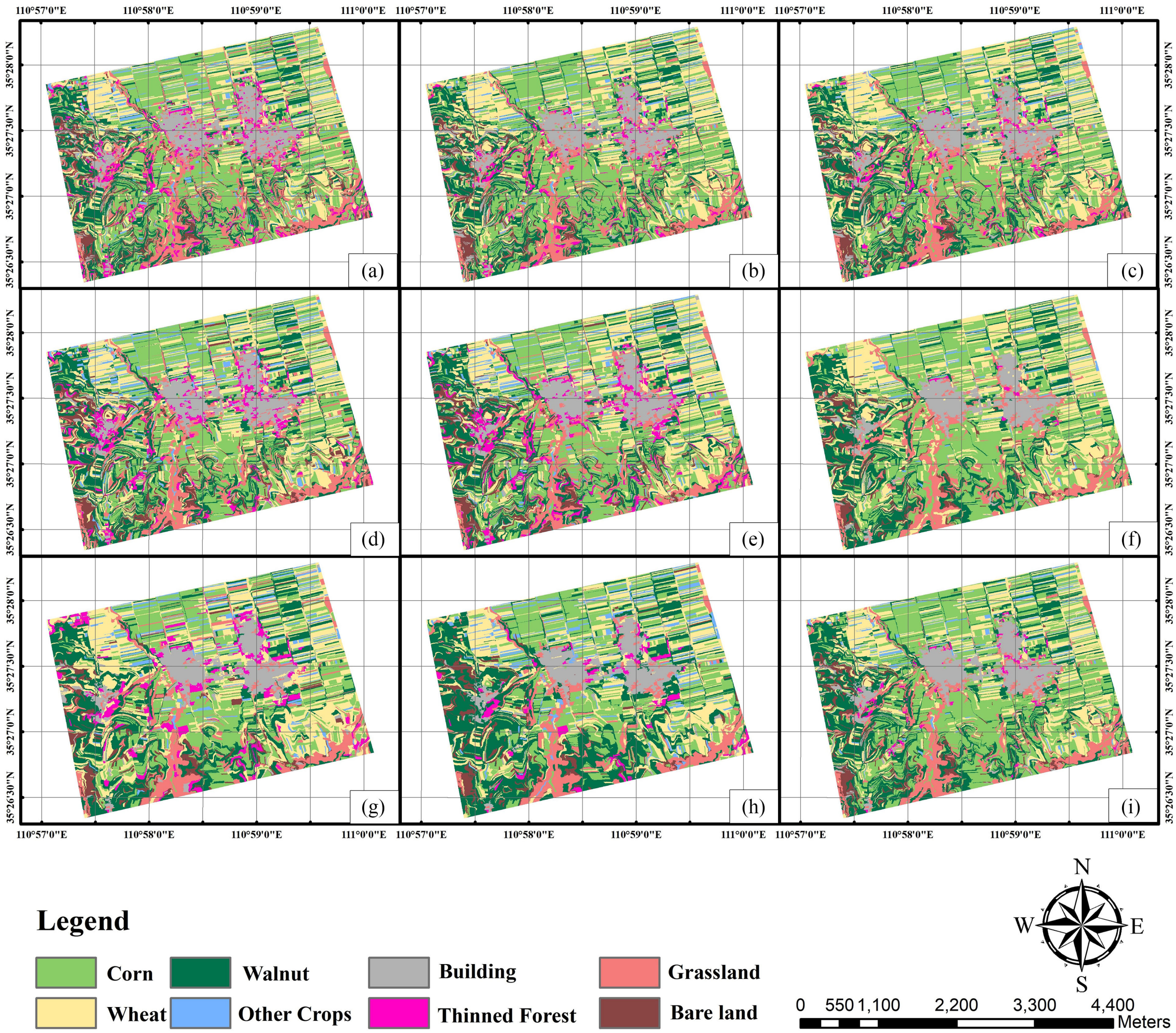


Fig. 8. Classification results when sample set B was used for training and sample set A was used for testing. (a) to (h) are respectively the classification results of the classifiers with the highest accuracy at scales 52, 61, 67, 81, 86, 104, 121, and 132. (i) is the result after (a) to (h) were weighted and fused.

TABLE VIII
MATRIX OF THE ACCURACY ASSESSMENT OF THE FINAL RESULTS (SAMPLE SET B WAS USED FOR TRAINING AND SAMPLE SET A WAS USED FOR TESTING)

	Corn	Wheat	Walnut	Other Crops	Building	Thinned Forest	Grassland	Bare land
Corn	177079	18244	4578	223	1277	0	467	2031
Wheat	6292	232831	3466	6427	497	0	1449	440
Walnut	2209	2594	189576	39	15	1774	1434	956
Other Crops	4470	7162	342	26382	0	0	0	0
Building	699	13	457	0	50910	25	912	0
Thinned Forest	1060	228	13998	67	2904	13856	7653	18
Grassland	5836	3806	3628	823	293	141	91029	633
Bare land	4847	9017	1009	13	3970	37	567	24447
Overall accuracy					86.201%			

scale can be used to highlight the information of a land type, it will also inevitably corrupt information of other land types. This point can be proven by Fig. 5(a) and (b): While the segmentation effect of scale 67 was good in the former, this same scale resulted in an excessive segmentation effect in the latter. Thus, we must emphatically emphasize that the wording of "optimal" scale is in itself improper. Generally, we can only say that the scales selected by us can segment the image well. Thus, we refer to the selected "good" scale as the "preferred" scale, and different preferred scales, in fact, correspond to different semantically significant region. Our study selected multiple preferred scales for the purpose of using the "useful" information of multiple scales. It is very similar to ensemble learning in the machine learning algorithm which uses a plurality of small classification models to obtain more robust classification results. Here we use the ESP tool in order to directly accelerate the process of scale selection. Some previous studies also selected multiple scales (e.g., [46], [67]–[69]), but most of these methods only selected scales equidistantly. Thus, there is no guarantee that each scale is "better". The information of some scales will be omitted if the distance is too long because segmentation results will significantly change even with a minor scale change [48].

B. Feature Importance Performance at Different Scales

In this study, we use two steps to obtain classification features. The first step is to construct an initial feature set based on the user knowledge. It is worth noting that this step only provides a range for important feature selection in the later stage and we do not use the entire initial feature set as a classified input. For each image object, the specific values of features were calculated by eCognition Developer 9.1. Nevertheless, there is a limitation in the construction of an initial feature set that we cannot guarantee that all the features are very proper for later classification. The reasons are as follows: 1) The construction of initial feature set is relatively subjective, and some useful features may be missed; 2) initial feature set determined the total number of features in classification processing; and 3) there was some redundancy in the initial feature sets, which will affect running efficiency of the program and lead to its over fitting. So, in order to solve the redundancy problem, the second step is to select important features from the initial feature set as the classified input. Feature selection results (Fig. 6) show that there were differences in our feature score sequences at different scales. The reason is that the scale parameter controls the final scale of the image object during image segmentation, and the value of the classification feature is calculated from all pixels in the image object. When the scale changes, the former object will either delete or add pixels to form a new object, and its feature values will change in succession, which then changes the feature importance score.

However, after excluding the differences, we found the vegetation index feature always ranked in the top no matter the scale used, followed by the spectrum feature, the texture feature and the geometric feature. At scale 52, we found that all features with scores in the top 5 were in fact vegetation index features, namely, the first texture feature T_Ent_T4_N ranked 19 while most spectrum features ranked from 10 to 20. Also, at scale 132, we found that all features with scores in the top 5 were also vegetation index features, while those whose scores ranked from 10 to 20 were mostly spectrum features, and most texture features ranked after 20. Only a few geometric features ranked near the top 60 in the eight feature selection result sets.

Rules that govern sequencing of the classification features at different scales showed that the vegetation index feature and the spectrum feature played crucial roles in identifying the eight land types in the study area. This was followed by the texture feature, while the geometric feature only played a minor role. The reasons are as follows: 1) The study area is mainly comprised of vegetation, and the vegetation index can well reflect vegetation growth; and 2) vegetation in the study area is mainly planted on customary farmland where geometric differences are not significant; thus, the significance of the geometric feature is generally low. It is important to emphasize that although during this stage we removed approximately 50% of features within the initial feature set using the SelectKBest method, we still cannot guarantee there is no redundancy in residual features. For parameter selection, the SelectKBest method is very subjective [70]. However, parameter selection is beyond the scope of this study. We were more focused on the sequencing differences in feature importance at different scales.

Differences are shown by specific features. For example, important scores of the feature of NDVI_T4 ranked differently at scale 52 and scale 132. However, differences are not obvious for certain types of features. For example, vegetation index features always rankedat the top no matter the scale.

C. Performances of All Classifiers at Different Scales

We respectively used RF, SVM and MLP classifiers to classify input features at eight preferred scales. These classifiers respectively represented three typical machine learning algorithms of ensemble learning, Vapnik-Chervonenkis (VC) dimension theory and neural network. Our initial feature set and classifier model parameter were the same for all preferred scales; thus, it sufficiently ensured that the change of classification accuracy is not caused by the change of model parameters.

Table IV shows that the same classifier performs differently at different scales. For example, when sample set A was used for training, the RF classifier reached the highest accuracy (83.57%) at scale 86 but less so (76.78%) at scale 132. The main reason for this is that different preferred scales results in different features which subsequently enter the classifier. We found that in general MLP was better than RF and SVM for the same scale but different classifiers whether sample set A or sample set B was used for training. When sample set A was used for training, we averaged classification accuracy of the three classifiers at eight preferred scales, and results showed that MLP achieved the highest accuracy (82.72%) followed by RF (81.45%) and SVM of (81.05%). Thus, different classifiers have different classification accuracy [42]. However, this does not mean that MLP performance was better than the other two because we cannot guarantee that our classifier parameters are unquestionably the best.

In summary, the combination of different scales and classifiers resulted in differences in classification accuracy (as seen in Table IV). Thus, as it relates to the problem of object-based classification, final classification results cannot be determined by one row or one column (Table V). This does not mean that the final segmentation and classification results are also optimal when a single scale reaches optimal (as identified by this study). For example, when sample set A was used for training, we achieved the highest classification accuracy (83.57%) at scale 86 only when the RF classifier was used and we assumed that this scale was best; however, this assumption is correct only if we neglect the fact that higher classification accuracy would be obtained at another scale if we used the MLP classifier. Thus, this study proposes the use of the multiscale weighted classification method to solve this problem.

D. Multiscale Weighted Classification Model and the Single Scale Classifier Model

The multiscale weighted classification model avoids the difficulty in both classifier selection and scale selection. It combines multiple preferred scales and multiple classifiers to achieve multiple classification results and then fuses these results to obtain the final result.

In the scoring shown in Table V, the corn category had the highest scores at scales 81 and 121, which means that the most suitable scales for the corn category were 81 and 121. The walnut category had the highest scores at scales 61 and 104, which means that the most suitable segmentation scales for the walnut category were 61 and 104. The wheat category always yielded low scores at all scales, which means no scale is suitable to segment wheat. The main reason for this is that the wheat category is scattered within the study area (see Figs. 7 and 8). The building category yielded high scores at all scales, which means that the building category is not sensitive to scale, and this also conforms to the actual conditions of the study area. The study area is mainly comprised of vegetation, and it is not difficult to differentiate buildings from all categories of crops. In other words, it only requires certain simple features to differentiate. Thus, the accuracy of the building category is always high no matter how it is segmented. In addition, we found that there were certain differences and undulations between Tables V and VI after comparison. For example, the most suitable scale to use for the other crop categories in Table V was 132, but it was 81 and 86 in Table VI. One reason for this phenomenon is the differences between sample set A and sample set B, which were divided from the samples as a whole, while other reasons were that we selected different classifiers for the two experiments and there were also accuracy differences among all categories.

In the final results, the accuracy reached 87.43% using the multiscale weighted classification method (see Table VII). But the accuracy using single preferred scale method was only 75.27% to 85.19% (see Table IV). Improvements range from 2.24% to 12.16%. The results we obtained were ideal and verify that the accuracy of the multiscale weighted classification method is higher than that of single scale classification. In fact, the multiscale weighted classification method is an ensemble learning method or can be understood as an image decision-level fusion process. This method combines multiple scales and classifiers to obtain a more robust classification result. The ensemble learning used in this study has two distinct meanings: 1) The first is an ensemble learning at different scales. A plurality of preferred scales is used to obtain multiple segmentation results, and then the useful information of segmentation results is transferred to the final classification result through the scoring table. 2) The second is learning on different classifiers. Different classifiers may perform differently for different study areas, for different images and for different samples. Both classifier and parameter selection are always difficult during classification. The multiscale weighted classification method reinforces the final classification result with a plurality of classifiers of simple parameters while avoiding the problems of model parameter selection and optimization.

V. Conclusions

This study mapped crop distribution in the southern region of Jishan County, Shanxi Province, China, using GF-2 high-resolution images and the multiscale object-based weighted classification method. We set multiple preferred scales to segment images during the study with the aim of using the useful information of each scale. Excessive computation loads and information redundancy will be caused if the information from all scales to be added into the final multiscale weighted classification model; thus, we primarily selected some scales which can segment images well using the ESP tool and defined them as "preferred" scales.

When selecting features, we found that feature importance scores at different scales sequence differently; thus, features that are finally inputted into the classifiers also differ. Differences are shown by specific features but are not clear for features of a certain category. In the study area we selected, the vegetation index feature was higher than the spectrum feature, the texture feature and the geometric feature.

We found significant differences in accuracy in classification results in the combinations of different preferred scales and different classifiers; thus, it is extremely difficult to determine a single suitable scale, namely, the "optimal" scale. Therefore, when a single scale was identified as optimal and then used by study, this did not mean that final segmentation and classification results were also optimal.

Accordingly, the intent of this study was to propose the multiscale weighted classification model which uses the concept of ensemble learning. We determined that the final classification accuracy of this model was higher than that at any single scale. It has two main advantages: First, it uses ensemble learning at different scales, and useful information at multiple scales can be applied; second, it is able to learn using multiple different classifiers, and the final classification result is reinforced by multiple classifiers of simple parameters.

The multiscale object-based weighted classification model provides a new solution for object-based classification. Procedures that include the selection of segmentation scales and the selection and optimization of classifiers can be skipped altogether. In this study, a number of scales were selected rather than just one for the process as a whole. Decisions were made for classification results at multiple scales to obtain more suitable results. However, certain components of the method are defective, which we intend to resolve in the future; for example, the highly subjectivity when establishing the methods and parameters for feature screening. Another uncertainty with this method that must be resolved is determining the number of suitable classifier types to use to set each scale. Although the number would be 1 in a limit state, such a number may omit certain information, while too many classifiers will result in an overburdened computation load which would reduce the benefits of using this method.

References

[1] S. Fritz *et al.*, "The need for improved maps of global cropland," *EOS, Trans. Amer. Geophys. Union*, vol. 94, no. 3, pp. 31–32, Jan. 2013.

[2] Y. Martin, H. Van Dyck, N. Dendoncker, and N. Titeux, "Testing instead of assuming the importance of land use change scenarios to model species distributions under climate change," *Global Ecol. Biogeography*, vol. 22, no. 11, pp. 1204–1216, Nov. 2013.

[3] M.-N. Tuanmu and W. Jetz, "A global 1-km consensus land-cover product for biodiversity and ecosystem modelling," *Global Ecol. Biogeography*, vol. 23, no. 9, pp. 1031–1045, Sep. 2014.

[4] C. Gardi, P. Panagos, M. Van Liedekerke, C. Bosco, and D. De Brogniez, "Land take and food security: Assessment of land take on the agricultural production in Europe," *J. Environ. Planning Manag.*, vol. 58, no. 5, pp. 898–912, May 2015.

[5] H. Costa, G. M. Foody, and D. S. Boyd, "Using mixed objects in the training of object-based image classifications," *Remote Sens. Environ.*, vol. 190, no. 3, pp. 188–197, Mar. 2017.

[6] X. Jiao *et al.*, "Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data," *ISPRS J. Photogramm. Remote Sens.*, vol. 96, no. 10, pp. 38–46, Oct. 2014.

[7] Y. Cai *et al.*, "A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach," *Remote Sens. Environ.*, vol. 210, no. 2, pp. 35–47, Feb. 2018.

[8] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens.*, vol. 28, no. 5, pp. 823–870, Mar. 2007.

[9] L. Yan and D. P. Roy, "Automated crop field extraction from multi-temporal Web Enabled Landsat Data," *Remote Sens. Environ.*, vol. 144, no. 3, pp. 42–64, Mar. 2014.

[10] R. Momeni, P. Aplin, and D. Boyd, "Mapping complex urban land cover from spaceborne imagery: The influence of spatial resolution, spectral band set and classification approach," *Remote Sens.*, vol. 8, no. 2, Jan. 2016, Art. no. 88.

[11] J. Xiong *et al.*, "Automated cropland mapping of continental Africa using Google Earth Engine cloud computing," *ISPRS J. Photogramm. Remote Sens.*, vol. 126, no. 4, pp. 225–244, Apr. 2017.

[12] W. Liu, J. Dong, K. Xiang, S. Wang, W. Han, and W. Yuan, "A sub-pixel method for estimating planting fraction of paddy rice in Northeast China," *Remote Sens. Environ.*, vol. 205, no. 9, pp. 305–314, Feb. 2018

[13] P. Ding, Y. Zhang, W. J. Deng, P. Jia, and A. Kuijper, "A light and faster regional convolutional neural network for object detection in optical remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 141, no. 5, pp. 208–218, May 2018.

[14] C. Zhang *et al.*, "Joint deep learning for land cover and land use classification," *Remote Sens. Environ.*, vol. 221, no. 2, pp. 173–187, Feb. 2019.

[15] S. Foerster, K. Kaden, M. Foerster, and S. Itzerott, "Crop type mapping using spectral–temporal profiles and phenological information," *Comput. Electron. Agriculture*, vol. 89, no. 11, pp. 30–40, Nov. 2012.

[16] A. Asgarian, A. Soffianian, and S. Pourmanafi, "Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: A case of central Iran using multi-temporal Landsat 8 imagery," *Comput. Electron. Agriculture*, vol. 127, no. 9, pp. 531–540, Sep. 2016.

[17] D. C. Duro, S. E. Franklin, and M. G. Dubé, "A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery," *Remote Sens. Environ.*, vol. 118, no. 3, pp. 259–272, Mar. 2012.

[18] M. A. Vieira, A. R. Formaggio, C. D. Rennó, C. Atzberger, D. A. Aguiar, and M. P. Mello, "Object based image analysis and data mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas," *Remote Sens. Environ.*, vol. 123, no. 8, pp. 553–562, Aug. 2012.

[19] T. Blaschke *et al.*, "Geographic object-based image analysis—Towards a New Paradigm," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, no. 1, pp. 180–191, Jan. 2014.

[20] G. J. Hay and G. Castilla, "Geographic object-based image analysis (GEOBIA): A new name for a new discipline," in *Object-Based Image Analysis*. Berlin, Germany: Springer, 2008, pp. 75–89.

[21] K. Johansen, L. A. Arroyo, S. Phinn, and C. Witte, "Comparison of geo-object based and pixel-based change detection of riparian environments using high spatial resolution multi-spectral imagery," *Photogramm. Eng. Remote Sens.*, vol. 76, no. 2, pp. 123–136, Feb. 2010.

[22] A. Darwish, K. Leukert, and W. Reinhardt, "Image segmentation for the purpose of object-based classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2004, vol. 3, pp. 2039–2041.

[23] U. C. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen, "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information," *ISPRS J. Photogramm. Remote Sens.*, vol. 58, nos. 3/4, pp. 239–258, Jan. 2004.

[24] H. M. A. van der Werff and F. D. van der Meer, "Shape-based classification of spectrally identical objects," *ISPRS J. Photogramm. Remote Sens.*, vol. 63, no. 2, pp. 251–258, Mar. 2008.

[25] A. S. Laliberte and A. Rango, "Texture and scale in object-based analysis of subdecimeter resolution unmanned aerial vehicle (UAV) imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 761–770, Mar. 2009.

[26] M. A. Vieira, A. R. Formaggio, C. D. Rennó, C. Atzberger, D. A. Aguiar, and M. P. Mello, "Object based image analysis and data mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas," *Remote Sens. Environ.*, vol. 123, no. 8, pp. 553–562, Aug. 2012.

[27] M. Hussain, D. Chen, A. Cheng, H. Wei, and D. Stanley, "Change detection from remotely sensed images: From pixel-based to object-based approaches," *ISPRS J. Photogramm. Remote Sens.*, vol. 80, no. 6, pp. 91–106, Jun. 2013.

[28] J. M. Peña-Barragán, M. K. Ngugi, R. E. Plant, and J. Six, "Object-based crop identification using multiple vegetation indices, textural features and crop phenology," *Remote Sens. Environ.*, vol. 115, no. 6, pp. 1301–1316, Jun. 2011.

[29] Q. Yu, P. Gong, N. Clinton, G. Biging, M. Kelly, and D. Schirokauer, "Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 7, pp. 799–811, Jul. 2006.

[30] I. L. Castillejo-González *et al.*, "Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery," *Comput. Electron. Agricultural*, vol. 68, no. 2, pp. 207–215, Oct. 2009.

[31] S. W. Myint, P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng, "Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery," *Remote Sens. Environ.*, vol. 115, no. 5, pp. 1145–1161, May 2011.

[32] H. Memarian, S. K. Balasundram, and R. Khosla, "Comparison between pixel- and object-based image classification of a tropical landscape using Système Pour l'Observation de la Terre-5 imagery," *J. Appl. Remote Sens.*, vol. 7, no. 1, Aug. 2013, Art. no. 073512.

[33] D. G. Goodin, K. L. Anibas, and M. Bezymennyi, "Mapping land cover and land use from object-based classification: An example from a complex agricultural landscape," *Int. J. Remote Sens.*, vol. 36, no. 18, pp. 4702–4723, Sep. 2015.

[34] R. V. Platt and L. Rapoza, "An evaluation of an object-oriented paradigm for land use/land cover classification," *Professional Geographer*, vol. 60, no. 1, pp. 87–100, 2008.

[35] M. Belgiu and O. Csillik, "Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis," *Remote Sens. Environ.*, vol. 204, no. 10, pp. 509–523, Oct. 2018.

[36] S. Bontemps, P. Bogaert, N. Titeux, and P. Defourny, "An object-based change detection method accounting for temporal dependences in time series with medium to coarse spatial resolution," *Remote Sens. Environ.*, vol. 112, no. 6, pp. 3181–3191, Jun. 2008.

[37] M. Möller, L. Lymburner, and M. Volk, "The comparison index: A tool for assessing the accuracy of image segmentation," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 9, no. 3, pp. 311–321, Aug. 2007.

[38] D. Liu and F. Xia, "Assessing object-based classification: advantages and limitations," *Remote Sens. Lett.*, vol. 1, no. 4, pp. 187–194, Dec. 2010.

[39] E. A. Addink, S. M. de Jong, and E. J. Pebesma, "The importance of scale in object-based mapping of vegetation parameters with hyperspectral imagery," *Photogramm. Eng. Remote Sens.*, vol. 73, no. 8, pp. 905–912, Aug. 2007.

[40] M. Kim, M. Madden, and T. A. Warner, "Forest type mapping using object-specific texture measures from multispectral Ikonos imagery," *Photogramm. Eng. Remote Sens.*, vol. 75, no. 7, pp. 819–829, Jul. 2009.

[41] C. Witharana and D. L. Civco, "Optimizing multi-resolution segmentation scale using empirical methods: Exploring the sensitivity of the supervised discrepancy measure Euclidean distance 2 (ED2)," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, no. 1, pp. 108–121, Jan. 2014.

[42] L. Ma, L. Cheng, M. Li, Y. Liu, and X. Ma, "Training set size, scale, and features in geographic object-based image analysis of very high resolution unmanned aerial vehicle imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 102, no. 4, pp. 14–27, Apr. 2015.

[43] H. Zhang, J. E. Fritts, and S. A. Goldman, "Image segmentation evaluation: A survey of unsupervised methods," *Comput. Vision Image Understanding*, vol. 110, no. 2, pp. 260–280, May 2008.

[44] L. Drăguţ, D. Tiede, and S. R. Levick, "ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data," *Int. J. Geographical Inf. Sci.*, vol. 24, no. 6, pp. 859–871, Apr. 2010.

[45] A. Smith, "Image segmentation scale parameter optimization and land cover classification using the random forest algorithm," *J. Spatial Sci.*, vol. 55, no. 1, pp. 69–79, Jun. 2010.

[46] B. Johnson and Z. Xie, "Unsupervised image segmentation evaluation and refinement using a multi-scale approach," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 4, pp. 473–483, Jul. 2011.

[47] L. Yi, G. Zhang, and Z. Wu, "A scale-synthesis method for high spatial resolution remote sensing image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 4062–4070, Oct. 2012.

[48] L. Drăguţ, O. Csillik, C. Eisank, and D. Tiede, "Automated parameterisation for multi-scale image segmentation on multiple layers," *ISPRS J. Photogramm. Remote Sens.*, vol. 88, no. 2, pp. 119–127, Feb. 2014.

[49] T. Hellesen and L. Matikainen, "An object-based approach for mapping shrub and tree cover on grassland habitats by use of LiDAR and CIR orthoimages," *Remote Sens*., vol. 5, no. 2, pp. 558–583, Jan. 2013.

[50] E.-Q. Xu, H.-Q. Zhang, and M.-X. Li, "Object-based mapping of karst rocky desertification using a support vector machine," *Land Degradation Develop.*, vol. 26, no. 2, pp. 158–167, Feb. 2015.

[51] M. Baatz and A. Schaepe, "Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation," in *Proc. Angewandte Geographische Informationsverarbeitung XII*, 2000, pp. 12–23.

[52] Trimble, eCognition Developer 9.0 User Guide. Munich, Germany, 2014, pp. 320–357.

[53] Trimble, eCognition Developer 9.0 ReferenceBook. Munich, Germany, 2014, pp. 52–61.

[54] R. Pu and S. Landry, "A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species," *Remote Sens. Environ.*, vol. 124, pp. 516–533, Sep. 2012.

[55] A. S. Laliberte, E. L. Fredrickson, and A. Rango, "Combining decision trees with hierarchical object-oriented image analysis for mapping arid rangelands," *Photogramm. Eng. Remote Sens.*, vol. 73, no. 2, pp. 197–207, Feb. 2007.

[56] F. Van Coillie, L. Verbeke, and R. De Wulf, "Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium," *Remote Sens. Environ.*, vol. 110, no. 4, pp. 476–487, Oct. 2007.

[57] A. A. Gitelson, Y. J. Kaufman, R. Stark, and D. Rundquist, "Novel algorithms for remote estimation of vegetation fraction," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 76–87, Jan. 2002.

[58] G. S. Birth and G. R. McVey, "Measuring the color of growing turf with a reflectance spectrophotometer," *Agronomy J.*, vol. 60, no. 6, pp. 640–643, Jun. 1968.

[59] C. F. Jordan, "Derivation of leaf-area index from quality of light on the forest floor," *Ecology*, vol. 50, no. 4, pp. 663–666, Jul. 1969.

[60] A. Huete, K. Didan, T. Miura, E. Rodriguez, X. Gao, and L. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sens. Environ.*, vol. 83, nos. 1/2, pp. 195–213, Nov. 2002.

[61] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Nov. 2001.

[62] M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogramm. Remote Sens.*, vol. 114, no. 4, pp. 24–31, Apr. 2016.

[63] J. Niemeyer, F. Rottensteiner, and U. Soergel, "Contextual classification of lidar data and building object detection in urban areas," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, no. 1, pp. 152–165, Jan. 2014.

[64] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 988–999, May 1999.

[65] Y. Shao and R. S. Lunetta, "Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points," *ISPRS J. Photogramm. Remote Sens.*, vol. 70, no. 6, pp. 78–87, Jun. 2012.

[66] P. M. Atkinson and A. R. L. Tatnall, "Introduction neural networks in remote sensing," *Int. J. Remote Sens.*, vol. 18, no. 4, pp. 699–709, Mar. 1997.

[67] A. Stumpf and N. Kerle, "Object-oriented mapping of landslides using random forests," *Remote Sens. Environ.*, vol. 115, no. 10, pp. 2564–2577, Oct. 2011.

[68] B. Johnson and Z. Xie, "Classifying a high resolution image of an urban area using super-object information," *ISPRS J. Photogramm. Remote Sens.*, vol. 83, no. 9, pp. 40–49, Sep. 2013.

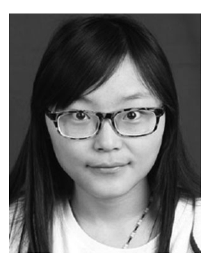
[69] X. Zhang and S. Du, "Learning selfhood scales for urban land cover mapping with very-high-resolution satellite images," *Remote Sens. Environ.*, vol. 178, no. 6, pp. 172–190, Jun. 2016.

[70] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.

Zengwei Tang received the B.S. in surveying and mapping engineering from the Department of Geoscience and Surveying Engineering, China University of Mining and Technology Beijing, Beijing, China in 2017. He is currently pursuing M.S. with the Faculty of Geographical Science, Beijing Normal University, Beijing, China. His research interests include vegetation remote sensing, machine learning, and UAV photogrammetry.

Hong Wang received the B.S. in cartography and geographic information systems from the Department of Urban and Resources, Northwest University, Xi'an, China in 2001, the M.S. in ecology from School of Natural Resources, Beijing Normal University, Beijing, China in 2005, and Ph.D. in land resource management from School of Natural Resources, Beijing Normal University, Beijing, China in 2008. He is currently a Fulltime Professor in the Faculty of Geographical Science, Beijing Normal University, Beijing, China. His research interests include photogrammetry and remote sensing, ecological remote sensing, and vegetation remote sensing.

Xiaobing Li received the B.S. in plant ecology from Inner Mongolia University, Hohhot, China in 1991, the M.S. in plant ecology from the Institute of Natural Resources, Inner Mongolia University, Hohhot, China in 1994, and Ph.D. in land resource management in Physical Geography from the School of Natural Resources, Beijing Normal University, Beijing, China in 1997. He is currently a Fulltime Professor in the Faculty of Geographical Science, Beijing Normal University, Beijing, China. His research interests include remote sensing of ecological environment, ecosystem services, plant physiology, and terrestrial ecosystem.



Xiaohui Li received the B.S. in geographic information system from the Department of Urban and Resources, Northwest University, Xi'an, China in 2016. She is currently pursuing M.S. in the Faculty of Geographical Science, Beijing Normal University, Beijing, China. Her research interests include ecological remote sensing, vegetation remote sensing and relevant parameter inversion based on hyperspectral data.

Wenjie Cai received the B.S. in geographical science (national training base) from the College of Geographical Sciences, Fujian Normal University, Fuzhou, Fujian, China in 2019. She is currently pursuing M.S. with the Faculty of Geographical Science, Beijing Normal University, Beijing, China. Her research interests include remote sensing application.

Chongyuan Han received the M.S. in natural resources from the Department of Resources Science and Technology, Beijing Normal University, Beijing, China in 2016. He is currently a Director in the Shanxi Investigation Team, National Bureau of Statistics, Taiyuan, China. His research interests include Agricultural statistics and Remote sensing application.