



Article

Investigating the Potential of Sentinel-2 MSI in Early Crop Identification in Northeast China

Mengfan Wei ^{1,2}, Hongyan Wang ¹, Yuan Zhang ¹, Qiangzi Li ^{1,*}, Xin Du ¹, Guanwei Shi ^{1,2} and Yiting Ren ^{1,2}

¹ Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; weimengfan19@mails.ucas.ac.cn (M.W.); wanghy@aircas.ac.cn (H.W.); zhangyuan@aircas.ac.cn (Y.Z.); duxin@aircas.ac.cn (X.D.); shiguanwei20@mails.ucas.ac.cn (G.S.); renyiting20@mails.ucas.ac.cn (Y.R.)

² University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: liqz@radi.ac.cn

Abstract: Early crop identification can provide timely and valuable information for agricultural planting management departments to make reasonable and correct decisions. At present, there is still a lack of systematic summary and analysis on how to obtain real-time samples in the early stage, what the optimal feature sets are, and what level of crop identification accuracy can be achieved at different stages. First, this study generated training samples with the help of historical crop maps in 2019 and remote sensing images in 2020. Then, a feature optimization method was used to obtain the optimal features in different stages. Finally, the differences of the four classifiers in identifying crops and the variation characteristics of crop identification accuracy at different stages were analyzed. These experiments were conducted at three sites in Heilongjiang Province to evaluate the reliability of the results. The results showed that the earliest identification time of corn can be obtained in early July (the seven leaves period) with an identification accuracy up to 86%. In the early stages, its accuracy was 40~79%, which was low, and could not reach the satisfied accuracy requirements. In the middle stages, a satisfactory recognition accuracy could be achieved, and its recognition accuracy was 79~100%. The late stage had a higher recognition accuracy, which was 90~100%. The accuracy of soybeans at each stage was similar to that of corn, and the earliest identification time of soybeans could also be obtained in early July (the blooming period) with an identification accuracy up to 87%. Its accuracy in the early growth stage was 35~71%; in the middle stage, it was 69~100%; and in the late stage, it was 92~100%. Unlike corn and soybeans, the earliest identification time of rice could be obtained at the end of April (the flooding period) with an identification accuracy up to 86%. In the early stage, its accuracy was 58~100%; in the middle stage, its accuracy was 93~100%; and in the late stage, its accuracy was 96~100%. In terms of crop identification accuracy in the whole growth stage, GBDT and RF performed better than other classifiers in our three study areas. This study systematically investigated the potential of early crop recognition in Northeast China, and the results are helpful for relevant applications and decision making of crop recognition in different crop growth stages.

Keywords: automatic sample generation; historical crop maps; feature optimization; classifier; early crop identification



Citation: Wei, M.; Wang, H.; Zhang, Y.; Li, Q.; Du, X.; Shi, G.; Ren, Y. Investigating the Potential of Sentinel-2 MSI in Early Crop Identification in Northeast China. *Remote Sens.* **2022**, *14*, 1928. <https://doi.org/10.3390/rs14081928>

Academic Editors: Thomas Alexandridis, Christos Chalkias, Konstantinos X. Soulis, Dionissios Kalivas and Emmanouil Psomiadis

Received: 12 March 2022

Accepted: 13 April 2022

Published: 15 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Crop identification and preliminary estimation of crop planting areas in the early stage can provide valuable and timely information for individual farmers regarding farmland management, agricultural insurance, and agricultural subsidy policies. It is also of great significance for agricultural management departments to make reasonable decisions [1–3]. Remote sensing technology can map the latest and most detailed crop-type mapping in a timely and accurate manner, which has been proven to be one of the most effective means for obtaining precise crop information [4–7].

At present, using remote sensing technology to identify crops is mainly focused on the middle or late stage of crop growth [8–10], and early crop recognition refers to identifying crops as early as possible from crop emergence to preharvest [9–12]. The crop information is too weak at this time, compared with the middle or late stage [10–12]. Moreover, the ground sample work has not been collected. Therefore, the sample problem is one of the challenges faced by early identification. All of these factors have resulted in a relatively few studies on early identification. Although many challenges make early-stage crop recognition difficult, previous studies have attempted to solve these problems with different methods. It can be divided into the following two methods: (1) Using the key features in the key phenological periods. For example, during the transplanting period of rice, there is often 2–15 cm water in the paddy field, and the surface is a mixture of rice and water. Based on this a priori knowledge, Wei et al. [13] used the Land Surface Water Index (LSWI) threshold method to realize rice identification in the transplanting period. Wang et al. [14] used multi-temporal images of winter wheat, including the periods of sowing, seedling emergence, tillering, and wintering, and then took the multi-scale segmentation object as the basic classification unit. Finally, the early recognition of winter wheat was realized by constructing a hierarchical decision threshold. However, the above-mentioned studies identified crops by constructing decision thresholds according to the different spectral characteristics of ground objects in the key periods, and the determination of threshold value was both time consuming and laborious. Thus, the efficiency of automatic crop extraction was greatly reduced. Furthermore, the determination of the same crop's threshold was difficult to apply to other regions due to the influence of soil, vegetation type, and image brightness [15–19]. (2) Using the image data or ground survey data in historical years. This type of method learned knowledge from historical year data and applied to the target year with the help of transfer learning technology. You et al. [20] trained the classifier by using the early season image time series and field samples collected in 2017, and then the classifier was migrated to the corresponding time series image data of the target year in 2018 to realize the early recognition of crops. In this paper, a single binary classifier (target ground object and non-target ground object) was trained for the recognition of each crop, but when there were many target objects to extract, the method was time-consuming. Vorobiova et al. [21] used historical data and the cubic spline function to generate Normalized Difference Vegetation Index (NDVI) time series to realize early crop recognition in the current year. However, they only considered a single NDVI in their study, and the inversion information of a single index was limited, so the identification accuracy of crops had some constraints [22–26].

All the above studies demonstrated that remote sensing technology has the potential to identify crops in the early stage; however, they placed too much focus on a single crop type, and the features used in their studies were relatively few. All of them lacked a systematic summary and research on how to obtain training samples quickly in the early stage, how early different crops can achieve a satisfactory level of monitoring accuracy, and what are the key identification features used to achieve crop identification, as well as what accuracy level each stage can achieve. So, it is difficult to obtain a comprehensive answer to the above questions.

Therefore, our objectives are: (i) Explore the feasibility of the sample automatic generation method based on the historical crop maps and image remote sensing images; (ii) Summarize and analyze the optimal features in different stages; (iii) Explore how early crops can be identified, what the accuracy variation characteristics of crops at different stages are, and the differences among different classifiers. Specifically, first, the automatic sample generation method was used to obtain the training samples based on the historical crop maps from 2019 and remote sensing images from 2020. Second, the feature optimization method was used to calculate feature importance. Therefore, the optimal feature sets in different stages can be achieved. Third, four classifiers were used to identify crops in different periods. Through the above three parts, the potential of early-stage crop recognition in Northeast China are systematically explored.

2. Study Area and Data

2.1. Study Area

Heilongjiang is an important commodity grain base in China. Three typical study areas with different characteristics were selected in Heilongjiang (Figure 1) as our study areas. They basically included the main crop types combinations of bulk crops in Heilongjiang, which provided good conditions for the extrapolation of research methods and the objectivity of research conclusions.

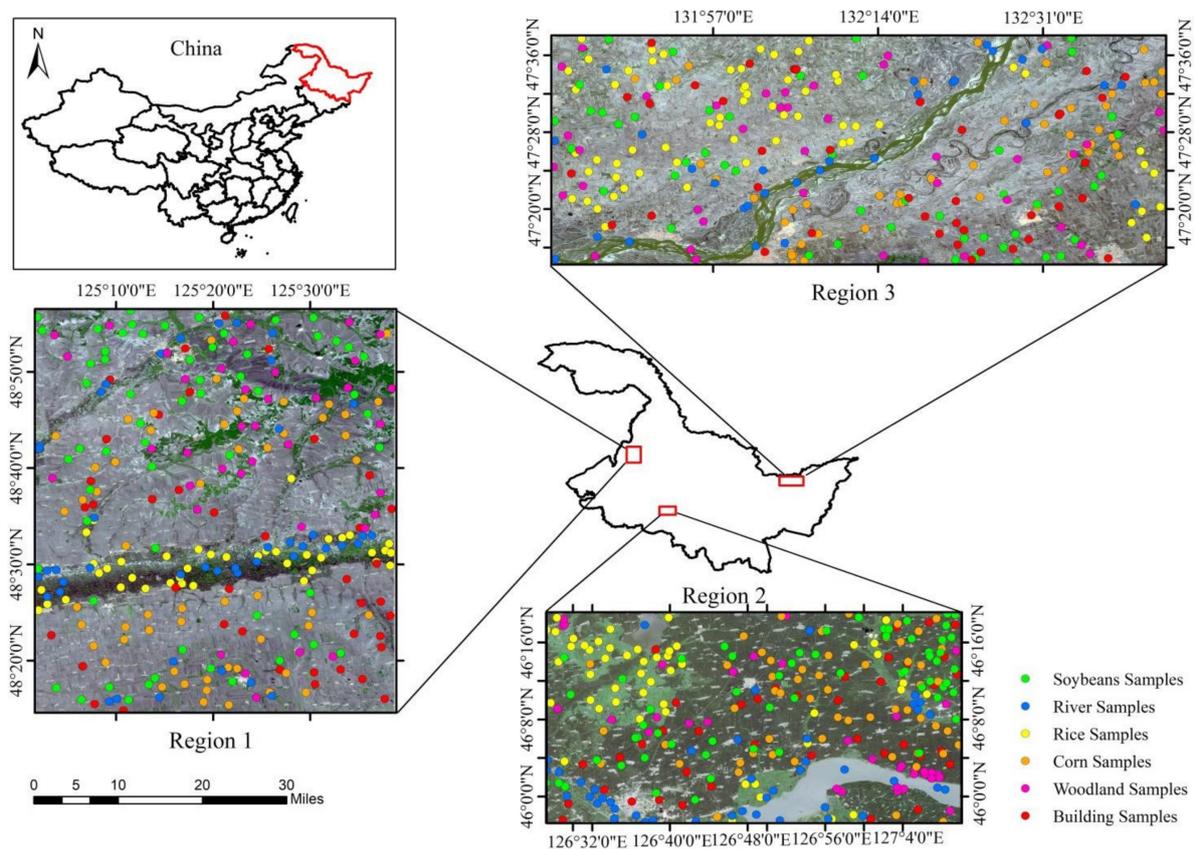


Figure 1. Location of the three regions.

Region 1 ($125^{\circ}02' E$ – $127^{\circ}64' E$, $48^{\circ}24' N$ – $48^{\circ}94' N$) is located on the Songnen Plain. The crop planting structure in Region 1 is relatively complex and fragmented, with intercropped corn and soybeans being the major crops in this region. Rice is mainly distributed along the river.

Region 2 ($126^{\circ}45' E$ – $127^{\circ}16' E$, $45^{\circ}95' N$ – $46^{\circ}31' N$) is also located on the Songnen Plain, the crops in this area are mainly corn, followed by rice, and soybeans are planted sparsely.

Region 3 ($131^{\circ}67' E$ – $132^{\circ}72' E$, $47^{\circ}23' N$ – $47^{\circ}63' N$) is located on the Sanjiang Plain. The region's major crop is rice, followed by corn and soybeans, and both of them are covered with large areas. In the western part of the region, some soybeans and corn are around rice.

The three crops are harvested only once per year due to low sunshine hours and accumulated heat. Their growing seasons are concentrated from April to early October, and all of them are sown in April. At this time, the water information on rice is more prominent in the image, whereas the corn and soybeans are in the state of bare land. In the early June, the rice is in the reviving period, and it is a mixture of water and rice in the image. Corn and soybeans are in the seedling period and show low vegetation. From July to August, the crops gradually enter vigorous growth, and the spectral characteristics of each crop are more obvious with a certain degree of separability in the image. In the September to early October, all crops gradually enter the mature and harvest period, soybeans are harvested first, followed by corn, and rice is harvested in early October. Based on the performance of

different crops in all the available images, we divided the entire growth stage into three stages: the early stage (April–June), the middle stage (July–August), and the late stage (September–early October). These crops show different characteristics in different growth stages, and the crop types can be identified as early as possible according to these key characteristics [20].

2.2. Data

2.2.1. Sentinel-2 Data

Sentinel-2 is equipped with the state-of-the-art Multi-Spectral Instrument (MSI), whose revisit period is 2–5 days depending on latitude, and the width is 290 km. It can provide optical images with a minimum resolution of 10 m, and the spectral bands cover 13 bands from visible light, near-infrared, red-edge to shortwave infrared (Table 1). Sentinel-2 has been proven to be an effective means of monitoring different crop types [27–31]. The Google Earth Engine (GEE) cloud platform provides an efficient environment to process Sentinel-2 data and can effectively implement various satellite-based remote sensing applications and research activities. The Sentinel-2 dataset in this platform contains completed data preprocessing work, such as radiometric calibration, terrain correction, and geometric correction. The pre-processing work we performed included selecting the images with cloud cover less than 10% from April to early October in 2020, and using Q60A band to remove thick clouds in some images. In addition, image cutting work was also necessary.

Table 1. Sentinel-2 Band Information.

Band Number	S2A		S2B		Spatial Resolution (m)
	Center Wavelength (nm)	Band Width (nm)	Center Wavelength (nm)	Band Width (nm)	
B1	443.9	27	442.3	45	60
B2	496.6	98	492.1	98	10
B3	560	45	559	46	10
B4	664.5	38	665	39	10
B5	703.9	19	703.8	20	20
B6	740.2	18	739.1	18	20
B7	782.5	28	779.7	28	20
B8	835.1	145	833	133	10
B8A	864.8	33	864	32	20
B9	945	26	943.2	27	60
B10	1373.5	75	1376.9	76	60
B11	1613.7	143	1610.4	141	20
B12	2202.4	242	2185.7	238	20
QA60					60

The availability of useful data is limited due to clouds. Therefore, 34 good quality images in 2020 that could cover our study areas (Figure 2) were selected. Among them, 11, 11, and 12 images were used in Region 1, 2, and 3, respectively. The cloud coverage of these images was less than 10%, which could meet the monitoring requirements. The images collected in each study area can generally cover the period from crop sowing to harvest, which facilitates the objective results of the identification features and identification capabilities of each stage. Previous studies have shown that the use of single-date remote sensing images have limited crop identification accuracy in the early stage, and combining multi-date images can effectively improve the crop identification accuracy [3,20]. Therefore, some incremental designs were made for the available image used in each period. That is, when performing a supervised classification, every time that the image was used in a current period, all the previously available images also needed to be included. Finally, there

were 11, 11, and 12 periods for image-conducted increment designs in Region 1, Region 2, and Region 3, respectively. From the accuracy obtained at different periods in each study area, we explored how early can the crops be monitored, and what level of accuracy can crops achieve at different stages.

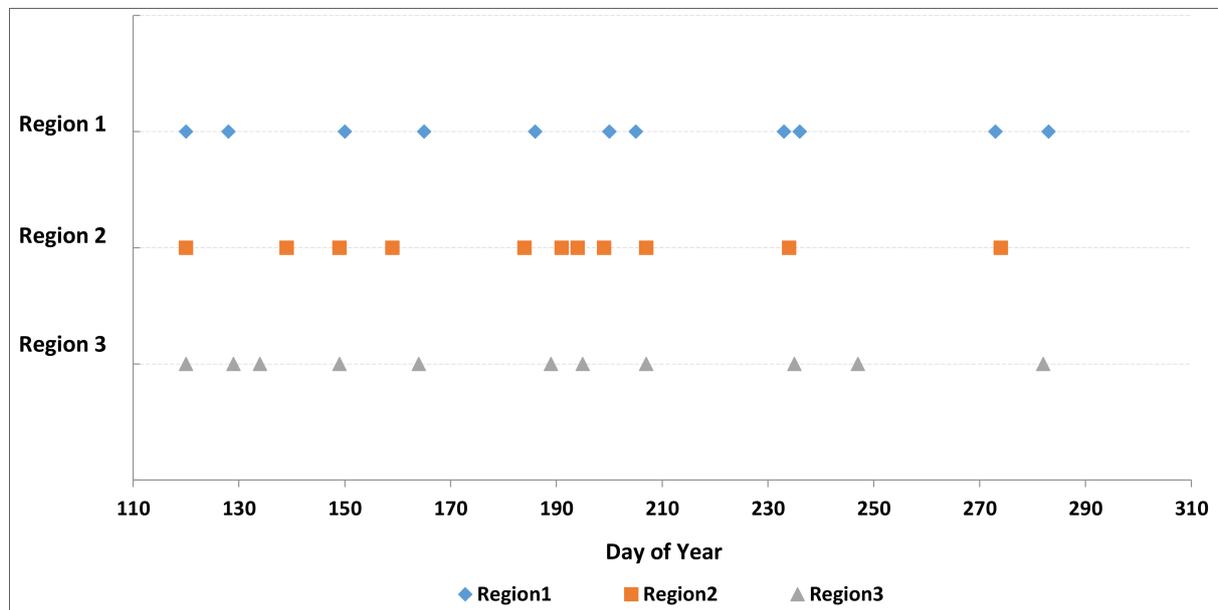


Figure 2. The image coverage of Sentinel-2 used in the three study areas.

2.2.2. Ground Survey Data

The ground samples are difficult to obtain in the early stage of crops. To reasonably and effectively evaluate the reliability of the training samples generated based on historical crop maps from 2019 and remote sensing images from 2020. The ground survey data were only used as validation samples to evaluate the classification results in our study. The Songnen Plain and Sanjiang Plain are our major ground survey areas. Rice, corn, and soybeans are major crops in these areas, so the latitude and longitude of these crops we collected were labeled, and the samples of woodlands, buildings, and rivers were labeled by combining Google Earth and Sentinel-2 images with visual interpretation. The usage of sample points in each study area is shown in Table 2.

Table 2. Validation samples in each study area.

	Corn	Soybeans	Rice	Building	Woodland	River
Region 1	56	60	40	40	38	43
Region 2	56	60	50	40	38	43
Region 3	50	50	63	40	40	30

2.2.3. Supplementary Data

The historical crop maps from 2019 of Heilongjiang Province were also collected, which could help effectively construct training sample generation work by combining remote sensing images from 2020, to simplify the workload of manual labeled samples and ensure the progress of the following research contents. The data came from the research results of Zhao et al. [32]. A binary indicator featuring high separability for single crop (corn, rice, and soybeans) based on Sentinel-2 images was proposed by them, and then they used the image adaptive threshold segmentation method to obtain crop-type mapping in Heilongjiang Province based on the binary indicative features of the three crops. With this method, the crop-type mapping was identified with an overall accuracy of 93.12% and

kappa coefficient of 0.90. A detailed review of this paper and accuracy information can be found in the article [32].

3. Method

At present, there are three problems that exist in early crop recognition: (i) The problem that the ground samples are difficult to obtain in real time. (ii) Which features are most important at different stages. (iii) What precision level can crops achieve at different stages, and the differences in early identification of crops among different classifiers, still lack a systematic summary. To solve the above problems, three parts were designed to study the early identification of crops in Northeast China. First, acquiring the training samples automatically, historical crop maps from 2019 and remote sensing images from 2020 were used to generate samples, which solved the problem that samples were difficult to obtain in real time. Second, selecting the optimal features, the Mean Decrease Accuracy (MDA) feature selection method was used to calculate feature importance. Therefore, the optimal feature set at different stages could be obtained. Third, identifying the crops with different classifiers and using the field samples for accuracy verification. Finally, the characteristic of crop accuracy levels in the different stages, the earliest identifiable time of each crop, as well as the difference among different classifiers were analyzed. The overall technical illustration is shown in Figure 3.

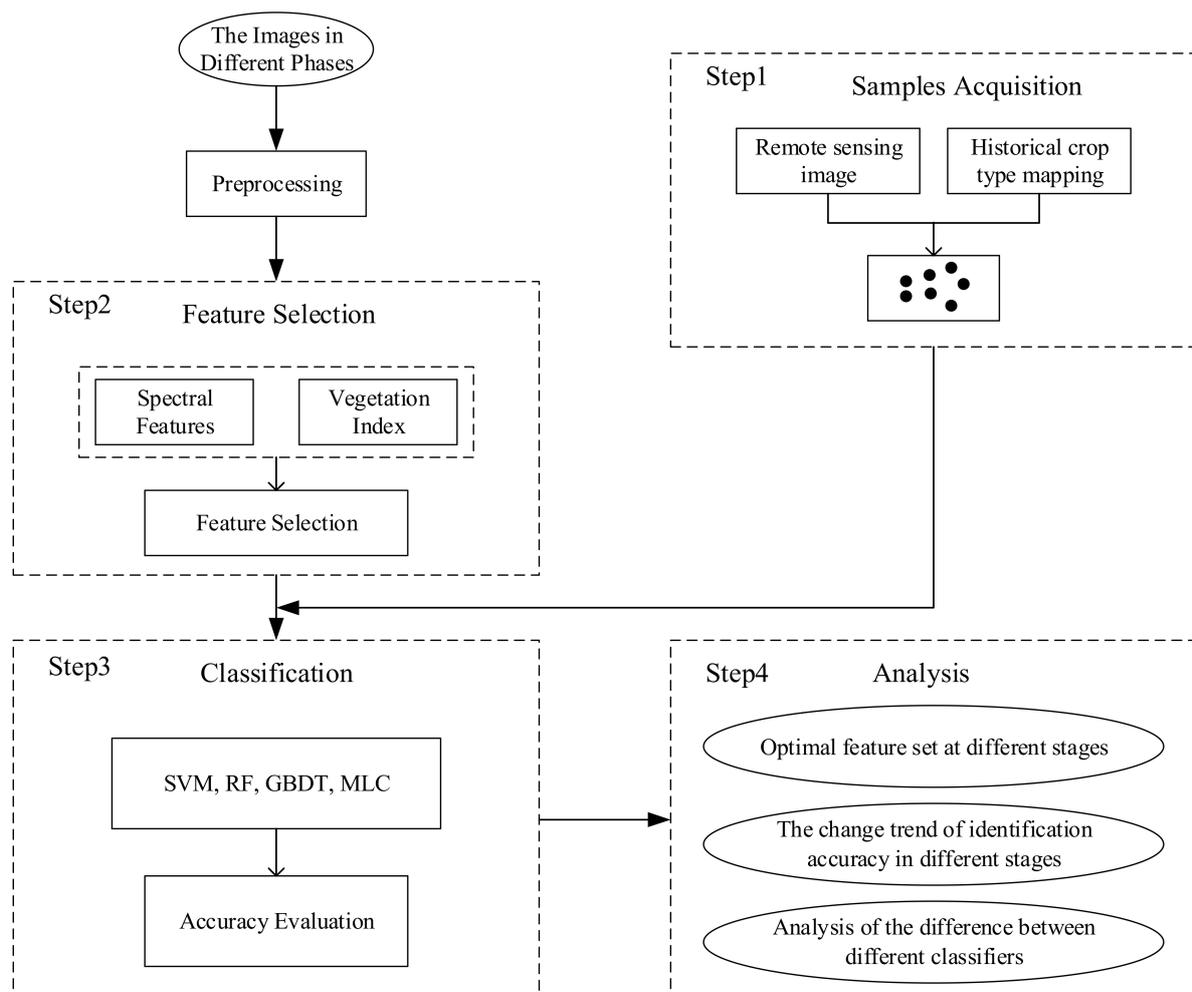


Figure 3. Illustration of early crop identification.

3.1. Automatic Sample Construction Based on Historical Crop Maps

Studies have shown that major crops vary little in neighboring years [7,20,30]. Based on the foundation, a method to automatically obtain current year samples based on historical crop type mapping was designed. First, the crop type mapping from 2019 was used to segment the remote sensing images from 2020, so that the candidate area image of each category could be obtained. Then, K-means clustering was used to obtain the clustering results in each candidate area image, and an appropriate threshold was set to obtain the pure sample area of each candidate area. Finally, training samples were randomly generated in the pure sample area (Figure 4). All the samples acquired based on the above method were used as training samples. The whole processing chain for automatic sample construction is summarized in the following sections.

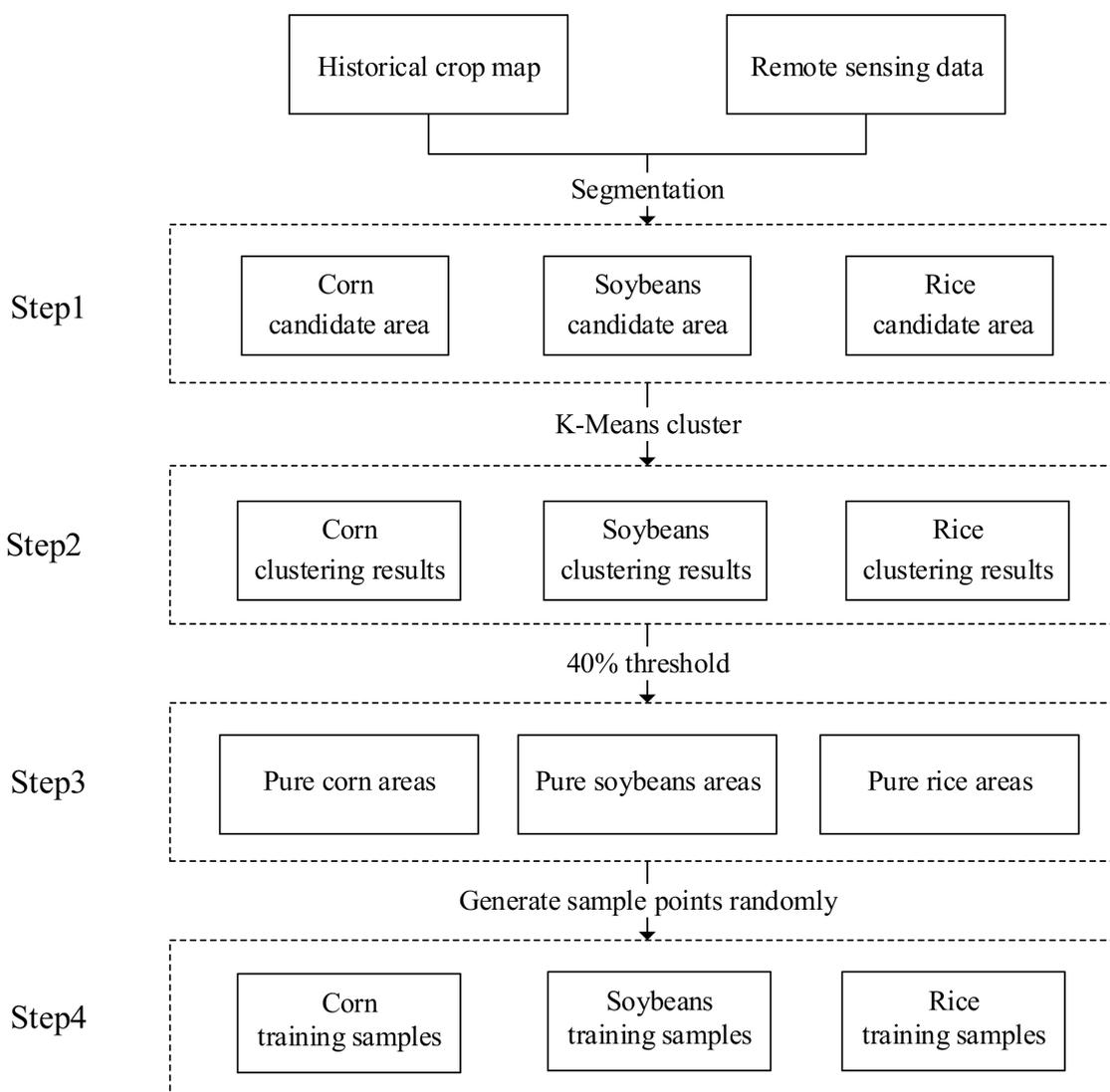


Figure 4. Technical process of automatic acquisition of training samples.

First, the acquisition of candidate region image. Three remote sensing images with good visual separability from 2020 were used in our three study areas (August 22, August 21 and August 23), and they were segmented with historical crop maps from 2019, so that the 2020 corn candidate area, rice candidate area, soybeans candidate area, and other candidate areas of each study area could be obtained.

Second, the unsupervised clustering of images. K-means is a popular unsupervised clustering algorithm, which has the advantages of simple principle, fast processing, and

good effect [6]. Euclidean Distance (ED) is commonly used as an indicator to measure the similarity between data objects when using the K-means algorithm to classify remote sensing images, and the similarity is inversely proportional to the distance between data objects; that is, the smaller the distance between objects, the greater the similarity [33]. The main principle of obtaining the classification results of candidate region by K-means clustering method is: (1) In order to make the unsupervised clustering results more detailed, 12 initial points were selected as centroids for each candidate region image, and the pixels were classified into the most similar class by calculating the similarity between each pixel and the centroid; (2) The centroid of each class was recalculated, and this process was repeated until the centroid did not change, and which class each pixel belonged to was determined. Finally, the 12 sub-cluster results were obtained in each candidate areas.

Third, the determination of pure crop planting area. Its determination needed to be based on the classification results of the previous step. Specifically, the 12 sub-cluster results of each candidate region were ranked by area from the largest to smallest, and clusters ranked in the top 40% were selected as the candidate region category. The threshold was set at 40%, mainly considering that the principle of the sample could be missed, but that the error could not be allowed. If the threshold value is set too high, the target sample range increases, but other ground objects will be introduced, leading to an impure sample; if the threshold value is set too low, the target sample will not be representative. Therefore, 40% is a good trade-off, avoiding the introduction of other crop samples while maintaining the significant representativeness of samples.

Finally, the construction of sample set. The pure sample area was converted to vector data, and the vectorization results were composed of many sub-vector areas, then the area of these sub-vector regions was calculated and training sample points of the ground object in the top N sub-vector regions with the largest area were randomly generated (N represents the number of training samples of the ground object). Some ground features, such as buildings, woodlands, and rivers, were unified as an “other” class in the historical crop maps, which was not suitable for the method we proposed. Therefore, the acquisition of these samples needed to be combined with relevant interpretation knowledge to manually label the samples of these ground objects. All the training samples in the three regions are shown in Table 3.

Table 3. Training samples in the three study areas.

	Corn	Soybeans	Rice	Building	Woodland	River
Region 1	115	121	100	80	75	86
Region 2	88	82	88	84	78	80
Region 3	100	100	126	80	78	60

All generated training samples in the three study areas were used for training our classification model, and its effectiveness was verified from two parts:

(1) Jeffries–Matusita (J-M) distance was used to evaluate the sample separability between each two ground objects [34]; The J value ranges from 0 to 2. In general, the larger the value, the better the separability. When the value is less than 1, it is considered that the two types of samples are not separable. When the value is less than 1.8, it indicates that the sample features is moderately separable. However, when the value exceeds 1.9, it indicates that the sample features have good separability. Its formula is as follows:

$$J = 2\left(1 - e^{-B}\right) \quad (1)$$

J is the value of J-M distance, and B is the Bavarian distance between the two sample categories. Its calculation formula is

$$B = \frac{1}{8}(m_1 - m_2)^2 \frac{2}{\delta_1^2 + \delta_2^2} + \frac{1}{2} \ln\left(\frac{\delta_1^2 + \delta_2^2}{2\delta_1\delta_2}\right) \quad (2)$$

m_1 and m_2 represent the mean values of the spectral reflectance of classes 1 and 2, respectively. δ_1 and δ_2 represent the standard deviation of the spectral reflectance of classes 1 and 2, respectively.

(2) Random Forest (RF) has the advantage of fast training speed and relatively simple implementation. Therefore, we used these samples tried to perform a preliminary classification of RF, and then evaluated the quality of the samples based on the classification results.

3.2. Feature Preparation

To improve the identification ability of crops, 10 original bands, including B2, B3, B4, B5, B6, B7, B8, B8A, B11, and B12, and 16 most commonly used vegetation indices related to crop identification, were used (Table 4) in our study. The bands of B1, B9, and B10 were eliminated due to their coarser spatial resolution (60 m).

Table 4. Vegetation Index used in this study.

Vegetation Index	Formula	Reference
Normalized Difference Vegetation Index (NDVI)	$\frac{B_8 - B_4}{B_8 + B_4}$	Rouse et al. [35]
Land Surface Water Index (LSWI)	$\frac{B_8 - B_{11}}{B_8 + B_{11}}$	Xiao et al. [36]
Enhanced Vegetation Index (EVI)	$2.5 \frac{B_8 - B_4}{B_8 + 6B_4 + 1 - 7.5B_2}$	Huete et al. [37]
Modified Chlorophyll Absorption Ratio Index (MCARI)	$(B_8 - B_4 - 0.2(B_8 - B_3)) \frac{B_3}{B_4}$	Daughtry et al. [38]
Ratio Vegetation Index (RVI)	$\frac{B_8}{B_4}$	Deering et al. [39]
Difference Vegetation Index (DVI)	$B_8 - B_4$	Richardson et al. [40]
Triangular Vegetation Index (TVI)	$0.5(120(B_8 - B_3) - 200(B_8 - B_4))$	Rouse et al. [35]
Optimization Soil Adjusted Vegetation Index (OSAVI)	$(1 + 0.16) \frac{B_8 - B_4}{B_8 + B_4 + 0.16}$	Rondeaux et al. [41]
Green Chlorophyll Vegetation Index (GCVI)	$\frac{B_8}{B_3} - 1$	Gitelson et al. [42]
Red Edge Normalized Vegetation Index (RENDVI)	$\frac{B_8 - B_6}{B_8 + B_6}$	Gitelson et al. [43]
Normalized Difference Tillage Index (NDTI)	$\frac{B_{11} - B_{12}}{B_{11} + B_{12}}$	Deventer et al. [44]
Normalized Difference Senescent Vegetation Index (NDSVI)	$\frac{B_{11} - B_4}{B_{11} + B_4}$	Qi et al. [45]
Green Vegetation Index (VIgreen)	$\frac{B_3 - B_4}{B_3 + B_4}$	Peña-Barragán et al. [46]
Wide Dynamic Range Vegetation Index (WDRVI)	$\frac{0.2B_8 - B_4}{0.2B_8 + B_4}$	Gitelson et al. [47]
Green Normalized Difference Vegetation Index (GNDVI)	$\frac{B_8 - B_3}{B_8 + B_3}$	Gitelson et al. [48]
Normalized Difference Water Index (NDWI)	$\frac{B_3 - B_8}{B_3 + B_8}$	Gao et al. [49]

3.3. Crop Classification Model

In our study, Support Vector Machine (SVM), RF, Gradient Boosting Decision Tree (GBDT) and Maximum Likelihood Classification (MLC) were used to compare their differences in early-stage crop recognition. These four methods were selected on the basis of their wide application and reliability in land cover classifications. The parameter settings of different classifiers are shown in Table 5.

Table 5. The parameter settings of different classifiers.

Model	Parameter Settings
SVM	kernelType: RBF; gamma: 0.8; cost: 50
GBDT	numberOfTree: 100; samplingRate: 0.1; maxDepth: 6; shrinkage: 0.1
RF	ntree: 100; mtry: 10
MLC	data scale factor: 1.0

SVM has significant advantages in small samples, nonlinear problems, and high-dimensional data processing [50–52]. Usually, the SVM algorithm solves binary classification (two-dimensional) problems through finding the separation hyperplane that can divide the training dataset correctly and the geometric interval maximally; however, for multi-classifications, it is more complex. Specifically, in the n -dimensional space, an optimal segmentation surface (kernel function) is used to map it to the high-dimensional space to maximize the separation of the two types of samples, thus making multi-classification linear solvable. Finally, multi-classification is realized through combining multiple binary classifiers [53–55]. Radial Basis Function (RBF) is usually chosen as the kernel function when using SVM to classify remote sensing images, because it provides a trade-off between time efficiency and accuracy. There are two parameters in the RBF kernel that need to be optimized: the penalty coefficient (cost) and gamma, cost controls the complexity and generality of the model, and gamma determines the extent and width of the input space. Generally, a larger cost value will not only lead to overfitting but also increase the calculation time, and inappropriate gamma values will lead to inadequate model accuracy or introduce errors. After many experiments, a better accuracy can be achieved when the value of cost was set to 50 and the value of gamma was set to 0.8.

RF has been widely used due to its high speed, high accuracy, and good stability [56–60]. Which is an ensemble machine learning technique that combines multiple trees. Each tree uses the bootstrap sampling strategy to create about 2/3 of the training samples from the original dataset and generates a decision tree for each training samples separately, and the remaining about 1/3 of the training samples are used as Out-of-Bag (OOB) data for internal cross-checking. Finally, the final classification results are determined by voting according to the classification results of all trees. There are two key parameters in RF that need to be optimized: the number of randomly selected features used to segment each node (mtry), and the number of trees (ntree). To balance the accuracy and calculation time, ntree was set to 100 in our study. In general, mtry is set to the square root of ntree, which was 10.

GBDT is an iterative decision tree algorithm, which is composed of multiple decision trees, and the conclusions of all trees are accumulated to obtain the final result. Unlike RF, this algorithm builds a weak learner at each step of the iteration to compensate for the shortcomings of the original model. Therefore, it has the characteristics of strong generalization ability and high classification accuracy [61,62]. Its key parameters are: the number of decision trees to be created (numberOfTree), the learning rate (shrinkage), and the max depth of tree (maxDepth), the sampling rate for stochastic tree boosting (samplingRate), all of these optimized parameters can help improve the performance of the model in terms of speed and accuracy.

MLC is one of the most well-known and widely used classification algorithms in remote sensing, which has the advantage of being a simple principle and easy to implement [63–67]. It assumes that each type of statistic in each band is normally distributed, and then the likelihood that a given pixel belongs to a certain training sample is calculated, and finally the pixel merges into the class with the highest likelihood. However, this classification method is not encapsulated in GEE; the method was conducted in Environment for Visualizing Images (ENVI) software. The setting of the data scale factor parameter adopted default value 1.0.

3.4. Feature Optimization

The MDA feature optimization method was used to evaluate the importance of features at each stage. The MDA feature ranking method uses OOB error to evaluate the variable importance (VI) [68,69]. Its principle to evaluate the importance of features is that if random noise is added to the feature, the accuracy of OOB data classification will be greatly reduced, which indicates that this feature has a great impact on sample prediction, which in turn indicates that the feature has higher importance. The calculation formula of VI is as follows:

$$VI(M_A) = \frac{1}{N} \sum_{t=1}^N (B_{n_t}^{M_A} - B_{O_t}^{M_A}) \quad (3)$$

where VI represents the importance of different features, M_A represents a feature, N is the number of decision trees, $B_{n_t}^{M_A}$ represents the OOB error of the t -th decision tree when M_A is not added, and $B_{O_t}^{M_A}$ is the OOB error of the t -th decision tree when M_A is added with noise interference.

A previous study proved that there was no increase in crop identification accuracy when the number of features used exceeded 15 [70]. In order to trade-off between accuracy and computational cost, the top 15 features with the largest values were selected as the optimal features, and used for image classification.

3.5. Accuracy Assessment

In this study, overall accuracy (OA), kappa, producer accuracy (PA), user accuracy (UA), and F_{score} were calculated according to the confusion matrix to evaluate the accuracy of the classification results.

The OA and kappa values represent the accuracy assessment of the global performance of our classification results [70]. Specifically, the OA represents the proportion of all samples that were correctly classified. kappa is a measure range 0–1 and quantifies the model prediction results and the actual classification results, while the PA, UA, and F_{score} values represent the accuracy assessment of individual crop type [71]. PA represents the proportion of the correct classification results of a certain class to its validation samples, and it reflects the omission. UA represents the proportion of the correct classification results of a certain class to the user's own classification results, and it reflects the misclassification. In general, PA and UA have contradictory relations, and they cannot achieve high classification accuracy at the same time. F_{score} is a comprehensive indicator composed of PA and UA, which can comprehensively evaluate the performance of the classifier in single crop [72]. Studies have shown that the 85% crop recognition accuracy can basically meet the agricultural applications of crop type mapping [73]. Therefore, the earliest identification time of each crop was defined as the first time when the F_{score} value was greater than 85% in this paper. The formula of F_{score} is:

$$F_{score} = 2 \times \frac{PA * UA}{PA + UA} * 100\% \quad (4)$$

4. Results and Discussion

4.1. Automatic Sample Construction and Analysis of Sample Separability

The distributions of training sample points, generated based on the historical crop maps and remote sensing images in our three study areas, are shown in Figure 5, and the classification results and accuracy of RF based on these samples are shown in Figures 6 and 7, respectively.

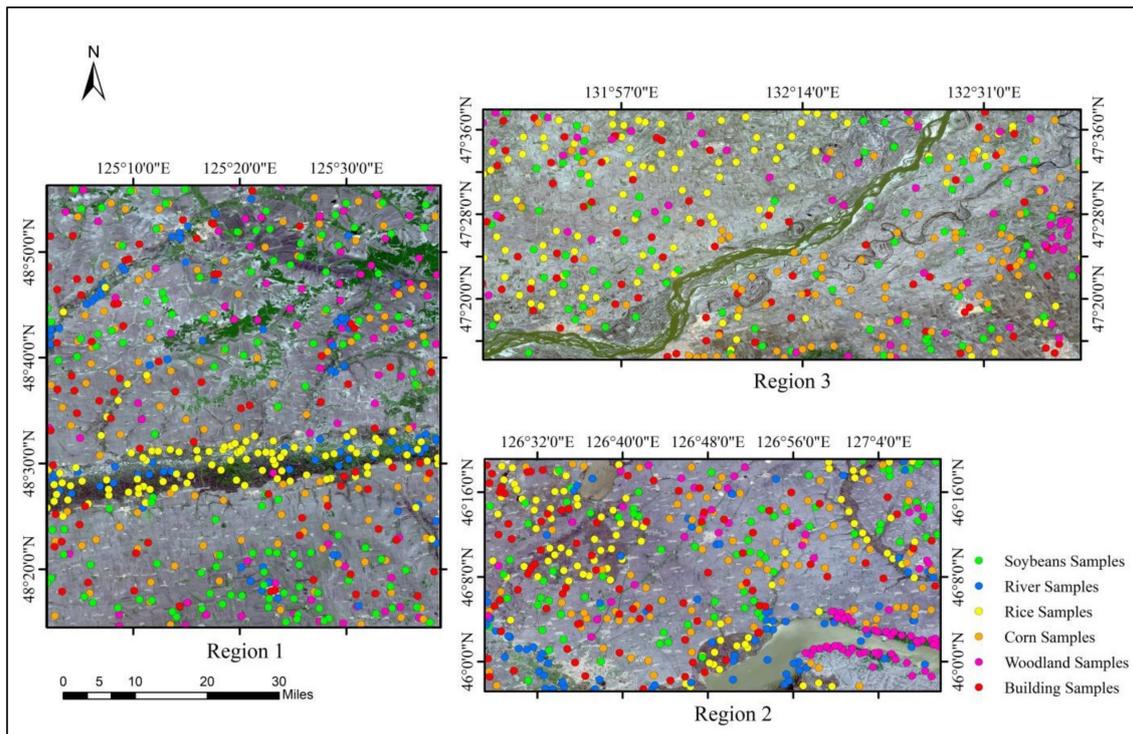


Figure 5. Distribution of training sample points in the three study areas.

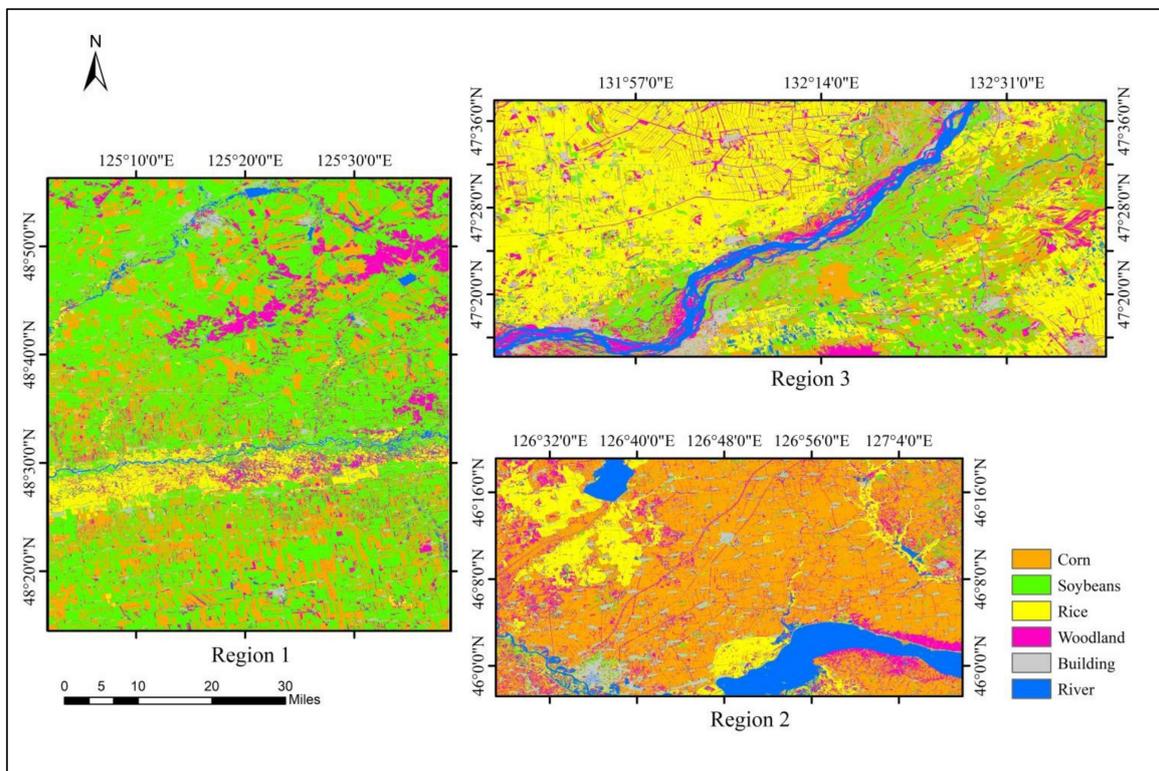


Figure 6. RF classification results in the three study areas.

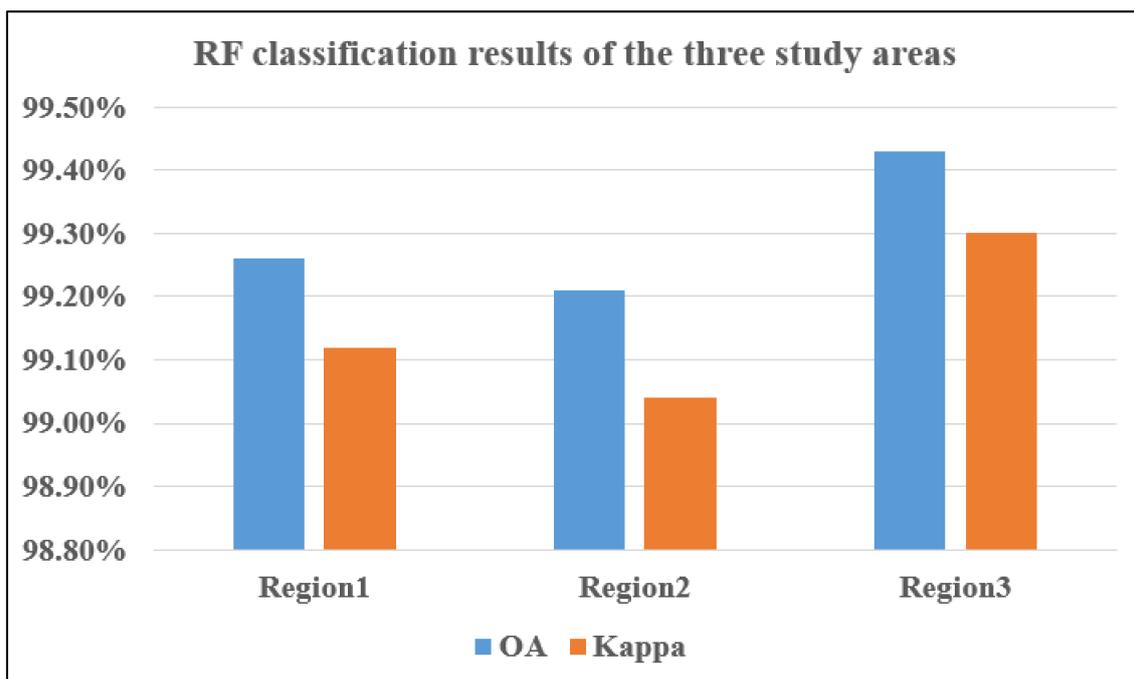


Figure 7. Overall accuracy and kappa of RF classification results in three study areas.

The J-M distance was used to measure the degree of separability between each pairwise crop, and all the J-M values in our study area were above 1.9, indicating these samples have good separability, and the OA and kappa coefficient of RF in the three study areas all reached about 99%. The above two results showed that the automatic generation method of sample points could be realized based on historical crop maps and remote sensing images, and the generated samples could be used for image classification work at different periods of the three study areas.

4.2. The Optimal Features and Changes in Different Stages

In this study, the MDA feature optimization method was used to obtain the optimal features of the different periods in our three study areas (Tables A1–A3). To analyze which features were important in the early stages (April–June), middle stages (July–August), and late stages (September–early October), the common top rankings feature of the different periods in each growth stage in our three study areas were summarized (Table 6).

Table 6. The key common features in different stages of the three study areas.

The Crop Growth Stages	The Common Features in the Three Study Areas
April–June	B12, NDTI, LSWI, NDSVI
July–August	B12, B11, B8, LSWI, NRED2, RENDVI
September–early October	NDTI, B11, NDSVI

At the early stage of crop growth (April–June), rice was basically covered with water; it was a mixture of water and rice in the image, while corn and soybeans were in a state of the sowing and emergence period. The height of the crops was low, and had low green vegetation and soil background in the image. Among the spectral characteristics, B12 was the most prominent, mainly because the short-wave infrared band was more sensitive to water information. The LSWI, which was composed of the short-wave infrared band, also contributed relatively high in this stage. In addition, traces of wheat stubble harvested from the previous season existed, and many farmlands were in a state of bare tillage. Therefore, the NDSVI and NDTI also played a key role in crop identification during this stage.

At the middle stage of crop growth (July–August), all crops in the study area gradually entered vigorous growth, and the vegetation coverage gradually reached the maximum. Due to the relatively high vegetation coverage and chlorophyll water content, the advantage of the red-edge vegetation index was also reflected. In general, the important features were B12, B11, B8, LSWI, NRED2, RENDVI, and so on, which, according to these features, could better identify crops at this stage.

At the late stage of crop growth (September–early October) into the late stage of the crop, the leaves gradually turned yellow and dried until harvest. In general, NDTI, B11, NDSVI, and other indicators made full use of the short-wave infrared band to represent the water and vegetation residual coverage. Therefore, they made a relatively large contribution to the identification of crops in this stage.

4.3. Variation Characteristics of Crop Identification Accuracy at Different Stages

In this study, the optimal feature sets for different periods (Tables A1–A3) were used as the input of four classifiers (SVM, GBDT, RF, and MLC) to identify crops. The crop recognition accuracy in different periods could be obtained (Figures 8–10). Figures 8 and 9 show that the accuracy trends of corn and soybeans in the three study areas were similar, which increased slowly in the initial stage and then entered a rapid increase period, after which, it gradually tended to be in a stable state after reaching the maximum value. However, the change in amplitude of rice identification accuracy was relatively smaller because of its higher initial value; after reaching the maximum value, it was in a state of fluctuating up and down (Figure 10). To analyze what accuracy each crop can achieve at different stages, the lowest and highest values for all periods at each stage in our three study areas were taken as the accuracy level of this stage.

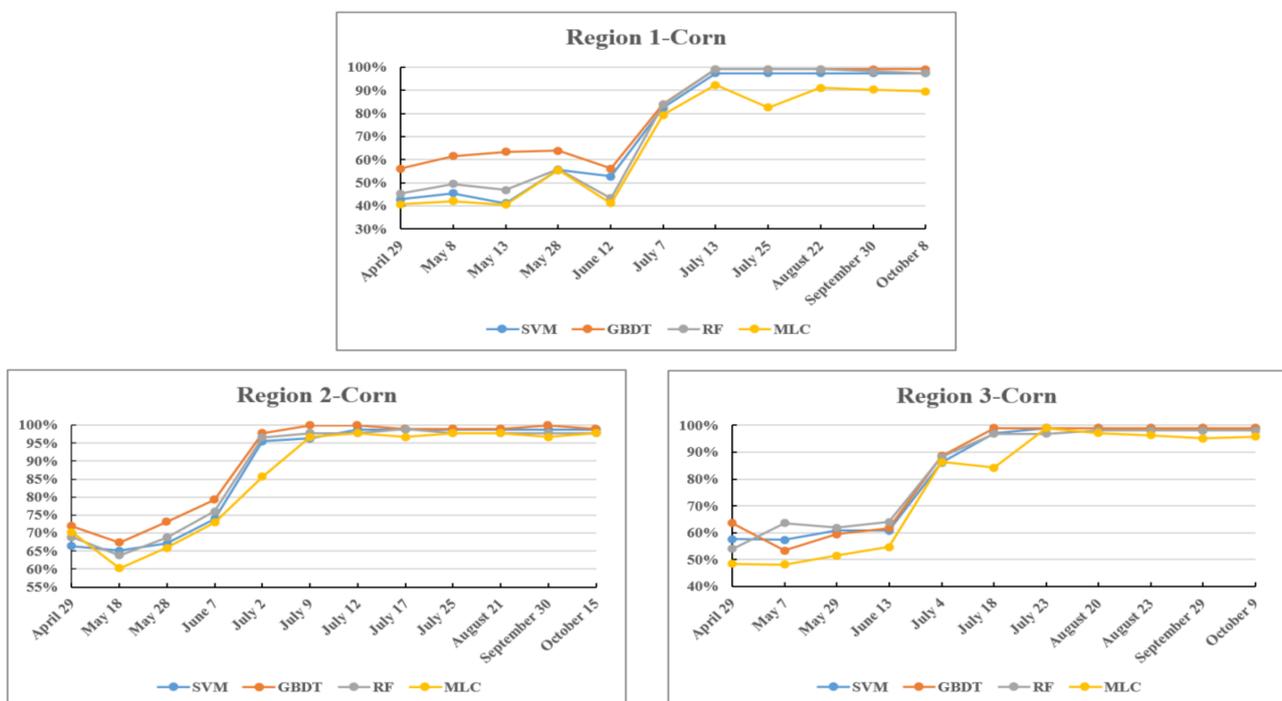


Figure 8. The recognition accuracy of corn in different periods in the three study areas.

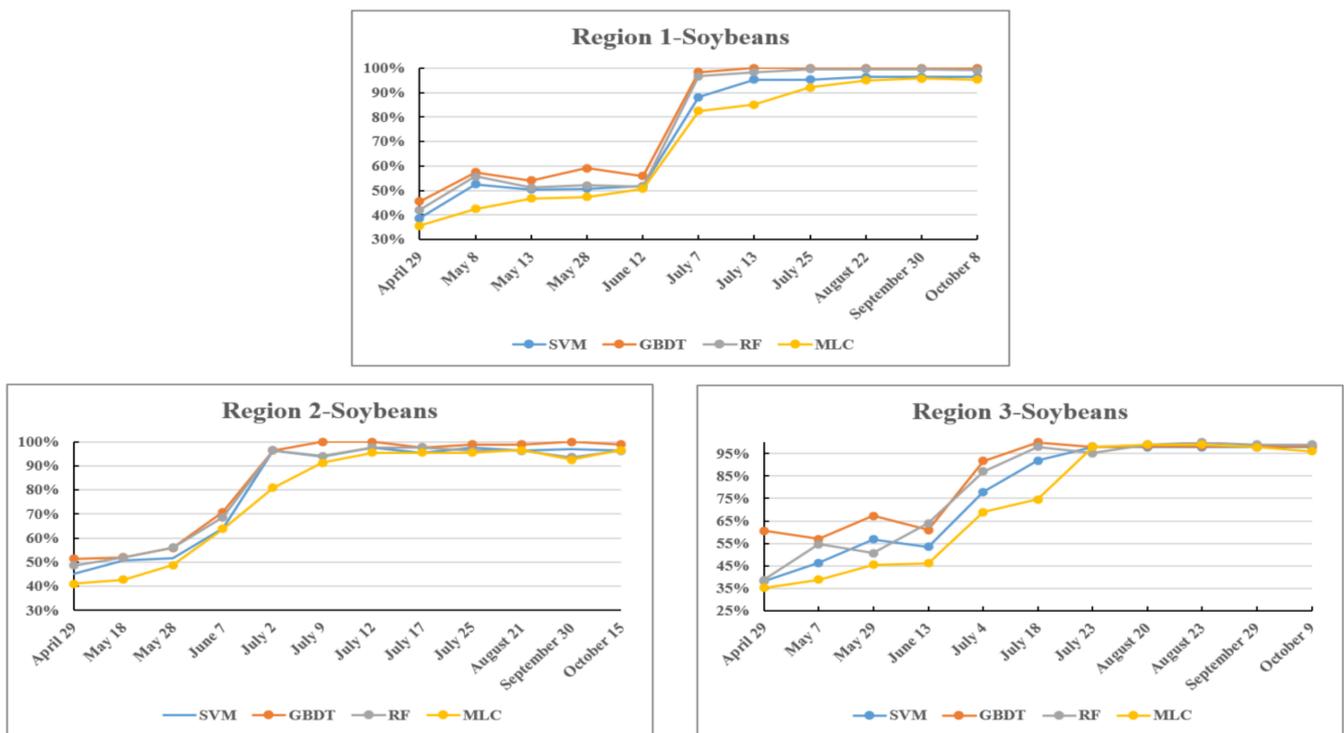


Figure 9. The recognition accuracy of soybeans in different periods in the three study areas.

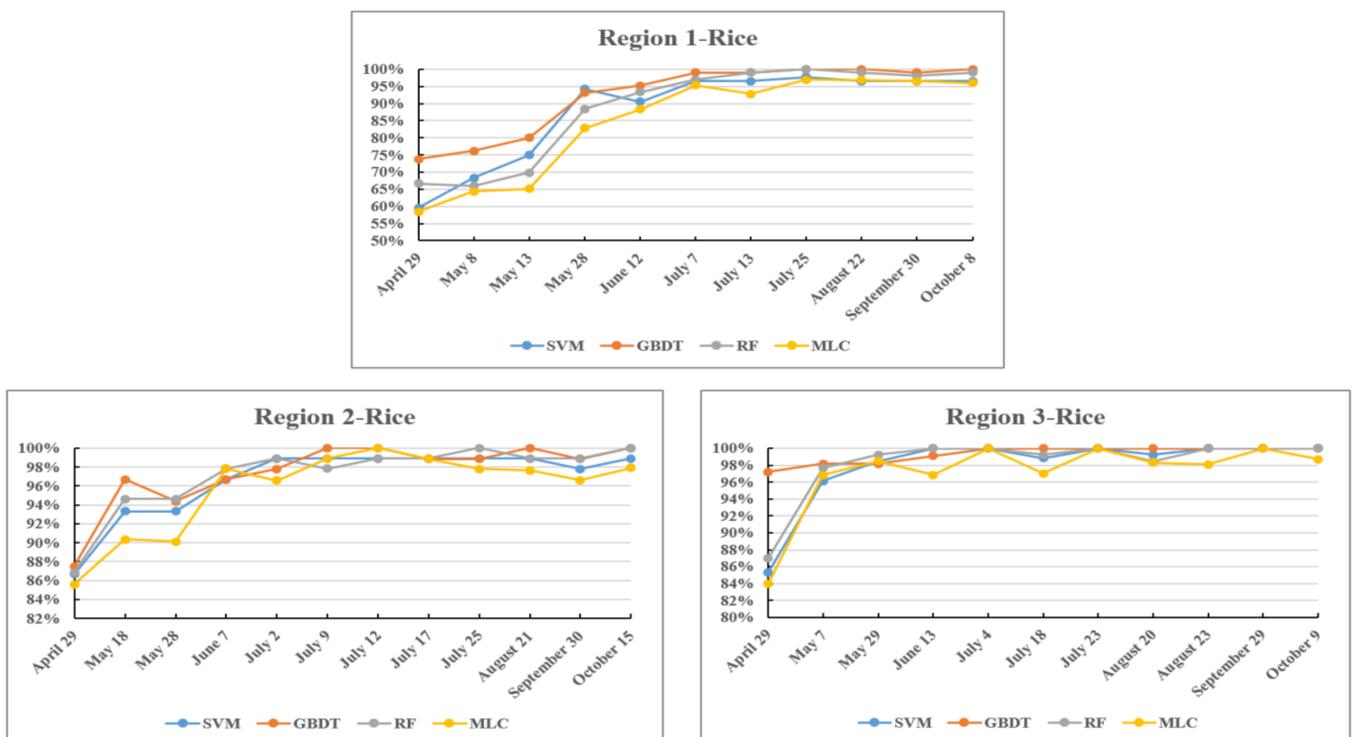


Figure 10. The recognition accuracy of rice in different periods in the three study areas.

4.3.1. The Recognition Accuracy Level of Corn at Different Stages

Figure 8 shows that the identification levels of corn at different phases, and its recognition accuracy level were analyzed as follows:

At the early stage of crop growth (April–June), the identification accuracy of soybeans was between 40% and 79%. Its accuracy was low and could not meet the basic requirements

of corn identification. With the addition of images, the recognition accuracy of GBDT was higher than that of other classifiers in Region 1 and Region 2, and it could achieve the highest accuracy of 79% with four images in Region 2 among the three study areas. In Region 3, RF performed better than that of other classifiers with 2–4 images, and SVM was slightly lower than RF. However, MLC performed poorly compared to other classifiers in the three study areas.

At the middle stage of crop growth (July–August), the identification accuracy of corn in the three study areas was between 79% and 100% during this stage. The accuracy in this stage basically reached the maximum value, which could meet the demand of crop mapping. In Region 2 and Region 3, the recognition accuracy of corn with four classifiers using five images was all greater than 86% on July 2. However, the earliest recognition time of corn in Region 1 was a little later than that of the other two study areas; its recognition accuracy of four classifiers with seven images was all greater than 93% as early as July 13. These results indicated that the high-precision identification of corn could be identified in early July at the earliest identification time, and the identification accuracy was above 86%. At this stage, GBDT performed better than other classifiers, SVM and RF performed after GBDT, and MLC performed worse in the three study areas.

At the late stage of crop growth (September–early October), corn was basically harvested. Its accuracy ranged between 90% and 100%, and basically remained in a stable state during this stage. The performance of the classifier in this stage was similar to that of the previous stage. The recognition accuracy of GBDT, RF, and SVM was comparable and higher than that of MLC.

4.3.2. The Recognition Accuracy Level of Soybeans at Different Stages

It can be seen from Figure 9, the identification levels of soybeans at different stages were analyzed as follows:

At the early stage of crop growth (April–June), the identification accuracy of soybeans was between 35% and 71%, which was lower and could not meet the basic requirements of soybeans mapping in this stage. Among all the classifiers, the accuracy of GBDT was higher than other classifiers in the three study areas, and it could achieve the highest recognition accuracy of 71% with four images in Region 2 among the three study areas. RF and SVM performed comparably in Region 1 and Region 2, and MLC performed worse than other classifiers in the three study areas.

At the middle stage of crop growth (July–August), the accuracy of soybeans in the three study areas ranged between 69% and 100%. With the addition of more images, the accuracy in this stage gradually reached saturation, and the identification accuracy of soybeans in many periods could achieve satisfactory accuracy during this stage. In Region 2, the recognition accuracy of soybeans with GBDT, RF, and SVM reached 96% with five images on July 2. In Region 3, the recognition accuracy of soybeans with GBDT and RF was 91% and 87% with five images on July 4, respectively. In Region 1, the recognition accuracy of soybeans with SVM, GBDT, and RF was all higher than 91% with six images on July 7. However, MLC needed to use six images to reach the earliest high-precision identification accuracy of 93% on July 9 (Region 2). These results showed that the early high-precision identification of soybeans could be obtained in early July, and the identification accuracy was as high as 87%.

At the late stage of crop growth (September–early October), soybeans were basically harvested. Its identification accuracy ranged between 92% and 100%. The recognition accuracy of the four classifiers was basically in a saturated state compared with the previous several stages, and the accuracy did not increase very much with more images added.

4.3.3. The Recognition Accuracy Level of Rice at Different Stages

It can be seen from Figure 10, the identification levels of rice at different stages were analyzed as follows:

At the early stage of crop growth (April–May), the recognition accuracy of rice ranged between 58% and 100%. In Region 2 and Region 3, the recognition accuracy of GBDT and RF could achieve 86% with one image as early as April 29. However, in Region 1, SVM, GBDT, and RF could use four images to realize early identification of rice as early as May 28 with an identification accuracy up to 88%. These results showed that the early high-precision identification of rice could be obtained as early as the end of April. Among all the classifiers, the initial recognition accuracy of GBDT was much higher than that of other classifiers, followed by RF and SVM, and MLC was lower than other classifiers in the three study areas.

In the middle stage of crop growth (July–August), the recognition accuracy of rice ranged between 93% and 100% during this stage. In the late stage of crop growth (September–early October), the rice was basically harvested in early October. Its recognition accuracy ranged between 96% and 100% during this stage. In Region 2 and Region 3, the accuracy fluctuated upwards and downwards with more images added, GBDT and RF performed comparably in the three study areas, followed by SVM, and MLC performed lower than other classifiers in the three study areas.

In general, the identification accuracy of the three crops in the early stage of crop growth was less than that in the middle stage, and the accuracy in the late stage was the best and basically in a stable state. Corn and soybeans had similar phenological periods, and the accuracies of the two crops were similar. The early identification of the two crops could be realized in early July at the earliest identification time; they were in the seven leaves period and the blooming period, respectively. The early identification of rice could be realized at the end of April at the earliest identification time (the flooding period) due to its characteristics being different to those of the former two crops. In terms of the whole process of crop recognition, GBDT and RF performed better than SVM and MLC.

4.4. Potential Analysis of Early Crop Identification

Timely and accurately acquisition of early crop information has important scientific significance and practical value. It can not only provide basic information for agriculture-related decisions and applications, but also can be used to support national food security, market planning, and many other social economic activities. This study systematically evaluated the potential of early crop identification in Northeast China based on Sentinel-2 image data, and the recognition abilities of the three crops in different growth stages are presented below.

In the early stage of crop growth (April–June), the recognition accuracies of corn, soybeans and rice were 40~79%, 35~71%, and 58~99%, respectively. Corn and soybeans appeared as bare land, and their accuracies were low and could not meet the requirements of early recognition accuracy in this stage. Rice was basically covered by water in this stage, which was a mixture of water and rice in the image, and the short-wave infrared band was more sensitive to the water content information [74]. Therefore, B12, LSWI, and NDTI ranked relatively higher in the optimal feature set and played an important role in crop identification in this stage [75]. Combined with GBDT and RF classifiers, the earliest high-precision recognition date of rice occurred on April 29, and the accuracy was as high as 86%. Therefore, we made a conclusion that rice could be identified as early as the end of April (the flooding period). This conclusion was consistent with the early identification time of rice obtained in the literature [76,77]. In the middle stage of crop growth (July–August), the recognition accuracies of corn and soybeans were 79~100% and 69~100%, respectively. Corn and soybeans gradually entered a vigorous growth stage. During this stage, the cover density and the chlorophyll water content of soybeans were higher than those of corn [78]. Short-wave infrared bands, NRED2 and red-edge index of vegetation, played an important role in crop identification in this stage. Combining the four classifiers could identify corn and soybeans as early as July 2, and their recognition accuracy was as high as 86% and 87%, respectively. Therefore, we concluded that the earliest identifiable times of corn and soybeans occurred in early July, and the phenological stages were in the seven-leaf period

and flowering period, respectively. It has been mentioned in articles [72,79] that corn could be identified at the earliest period of seven leaves and soybeans at the earliest period of flowering. Moreover, some other findings were found when analyzing the experimental results, which are discussed below: from the recognition accuracy of corn in the whole period of the three regions, it can be found that the early stage was more susceptible to the influence of the crop planting structure, while the middle and later stages were less affected. Taking the recognition accuracy of corn in the early stage as an example, the identification accuracy of corn in Region 2 was higher than that in Region 3, and the Region 3 than that in Region 1 during this stage. There were also some differences in the earliest identified time of the three regions, and corn was the earliest identified crop in early July in Regions 2 and 3, while in Region 1, it was slightly later, which was in mid-July. Therefore, the early identification of crops was closely related to the planting structure of crops; that is, the planting structure was relatively simple, the farmlands were relatively regular, and the accuracy of crop identification was higher [80]. In contrast, the more complex the planting structure was, the lower the crop identification accuracy [81,82], thus delaying the time of early crop identification.

From crop emergence to harvest, with the addition of images, crop recognition accuracy gradually increases in the early stages, while in the middle and late stages, it gradually becomes stable. As mentioned above, in the early stage, the rice information is more prominent in this stage due to its own characteristics, while the information of corn and soybeans are weak, and the rice information can be enhanced by some key features, while the information of corn and soybeans is more difficult to capture. Therefore, the recognition accuracy of rice was higher than that of corn and soybeans in this stage. In the middle stage, the uniqueness of corn and soybeans gradually emerges, and these two crops' information can be well captured by some key separable features, and the addition of these features greatly improves the accuracy of crops. In the late stage, the crops enter a mature and harvest period, and the accuracy was in a saturated state. Due to the incremental design method used for the image data, some previous key features also played an important role on these periods in the late stage, so the accuracy in this stage was also high. Previous studies have shown that a high classification accuracy can be achieved by using the images in only a few optimal dates. However, the accuracy may change to a saturated or decreased state as more images are added [3,20,83], and our results were consistent with these conclusions.

The problem that early crop identification cannot be carried out due to a lack of samples. Considering that the historical crop map has a large amount of prior knowledge, if the knowledge was fully utilized, it will help to solve the sample problem. Therefore, we proposed an automatic sample generation method to generate training samples based on the historical crop map and remote sensing images, which solved the difficulty of obtaining samples in real time, and this also can be applied to other regions with little difference in crop planting structure from year to year. Then, the feature optimization method was used to calculate the feature importance in different periods; from which we summarized the key recognition feature in different stages to provide a reference for relevant crop recognition research in selecting effective recognition features and reducing the blindness in feature selection. Finally, four commonly used classifiers were used to identify crops in different periods. From which we summarized the accuracy variation characteristics of crops at different stages and the earliest identifiable date that each crop can reach. In addition, the difference in the performance of the classifiers was also analyzed, and it was found that GBDT and RF were better than SVM and MLC in terms of the recognition of crops at different stages. Although GBDT was not as popular as the other three classifiers, it showed better accuracy in our three study areas, and can provide a new reference for the selection of classifiers for crop identification research.

Limited by the resolution of Sentinel-2 images, the spectral and vegetation index features were only considered in our study, but texture features also contribute significantly to crop identification. Next, satellite data with higher spatial resolution will be considered

to comprehensively evaluate the performance of spectral, vegetation index, and texture features in different periods. Additionally, Deep Learning (DL) technology is more and more prominent in the field of remote sensing classification [84–86]. Therefore, research on the comparison between DL algorithms and traditional classifiers in terms of classification effect and model transfer ability is also the direction of future efforts.

5. Conclusions

In this study, three important contents of early crop identification at three sites in Heilongjiang Province were explored. First, the method of sample points generation based on historical crop maps solved the problem that the ground samples could not be obtained in real time in the early crop stage. Then, MDA was used to optimize the optimal feature sets at different phases. Finally, four classifiers were used to identify crops, and the variation characteristics of crop recognition accuracy at different phases and the differences among different classifiers were analyzed from this part. The main conclusions are below.

- (1) The identification accuracy of corn in the early growth stage was between 40% and 79%, while in the middle stage, it could reach 79~100%, and in the late stage, it was between 90% and 100%. The earliest identification time of corn could be obtained in early July (the seven leaves stage), and the identification accuracy was up to 86%. The identification accuracy of soybeans in the early growth stage was between 35% and 71%, while in the middle stage, it could reach 69~100%, and in the later stage, it was between 92% and 100%. The earliest identification time of soybeans could also be obtained in early July (the blooming stage), and the identification accuracy was up to 87%. The identification accuracy of rice in the early growth stage was between 58% and 100%, while in the middle stage, it could reach 93~100%, and in the late stage, it was between 96% and 100%. The earliest identification time of rice could be obtained at the end of April (the flooding period), and the identification accuracy was up to 86%.
- (2) GBDT and RF performed better in the whole growth phases and had higher recognition accuracy than other classifiers. Therefore, they are recommended crop early recognition research.
- (3) In the early stage, B12, NDTI, LSWI, and NDSVI played important roles in identifying the crops. In the middle stage, features such as B12, B11, B8, LSWI, NRED2, and RENDVI contributed greatly. In the late stage, NDTI, B11, and NDSVI were important in identifying the crops.
- (4) It was effective in acquiring training samples based on crop-type mapping and remote sensing data, which could effectively reduce the workload of manual sample selection, and it is of great significance for large area and real-time crop mapping.

Author Contributions: Conceptualization, M.W. and Q.L.; methodology, M.W., H.W. and Y.Z.; software, M.W.; validation, M.W. and H.W.; formal analysis, M.W.; investigation, M.W., H.W., G.S. and Y.R.; writing—original draft preparation, M.W. and H.W.; writing—review and editing, M.W.; visualization, M.W.; funding acquisition, Q.L. and X.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China (Grant No. 2021YFD1500103), the National Science Foundation of China (Grant No. 42071403), and the Key Program of High-resolution Earth Observation System (Grant No. 11-Y20A16-9001-17/18-4).

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

Appendix A

Table A1. Optimal feature sets at different periods in Region1.

Phase	April 29	May 8	May 13	May 28	June 12	July 7	July 13	July 25	August 22	September 30	October 8
Rank by feature importance	B12	NDTI	NDTI	LSWI_3	B12_4	B5_5	B12_6	B8_7	NRED2_8	B12_4	NDTI_8
	NDTI	LSWI	OSAVI_2	NDSVI_3	NRED2_3	B12_5	LSWI_4	B11_4	B12_4	B12_7	B5_7
	B7	RENDVI_1	LSWI	NDTI	B2_4	RENDVI_5	B5_6	B12_4	NDTI_7	B11_4	NDSVI_8
	NDSVI	NRED2_1	B2	NRED2_3	NRED1_3	B8_5	B11_6	B3_6	B5_8	NRED3_8	B12_6
	RENDVI	NDVI_1	NDSVI_1	B12_3	WDRVI_3	NRED3_5	B12_4	B2_6	B11_7	B5_7	B12_5
	LSWI	NRED2	LSWI_2	WDRVI_1	LSWI_3	B12_4	NDSVI_5	NDTI_7	B11_4	RENDVI_7	B8A_8
	NRED1	WDRVI_1	NDSVI_2	NRED1_3	NDVI_3	NDVI_5	B2_4	B12_6	B8A_8	B11_7	B11_7
	B11	NDSVI	NDVI_1	EVI_3	B4_2	B11_4	B11_4	NDTI_4	NRED2_8	B11_8	B12_3
	WDRVI	B11	B4_2	B11_2	TVI_2	RENDVI_5	NDSVI_6	B12_7	B5_8	B3_8	B11_9
	TVI	NRED1_1	B2_2	GNDVI_2	VIgreen_4	LSWI_5	B4_6	NDVI_7	LSWI_6	B3_6	B11_4
	EVI	TVI_1	WDRVI_2	B11_3	WDRVI_1	EVI_5	B11_3	B11_7	B11_6	EVI_3	B11_8
	NDVI	MCARI_1	NDVI_2	NDVI_3	B2_2	VIgreen_5	NDVI_6	B3_7	LSWI_4	NDSVI_4	MCARI_10
	MCARI	NDTI_1	TVI_1	MCARI_3	B4_4	NDSVI_5	B4_5	NDSVI_6	NDSVI_4	B12_5	B4_6
	B4	NRED3_1	B11	GCVI_3	NDSVI_4	B3_5	LSWI_6	NRED2_6	B12_6	B12_3	LSWI_9
	B5	GNDVI_1	RENDVI_1	NDTI_2	B11_4	OSAVI_2	B2_6	NDSVI_4	B3_6	OSAVI_7	RENDVI_8

Table A2. Optimal feature sets at different periods in Region 2.

Phase	April 29	May 18	May 28	June 7	July 2	July 9	July 12	July 17	July 25	August 21	September 30	October 15
Rank by feature importance	TVI	NDSVI_1	LSWI	LSWI_3	B11_4	B11_5	B12_6	B12_6	B8_6	B11_9	B12_6	NDTI_2
	LSWI	NDSVI	B11_1	B11_2	B11_4	NDSVI_4	B12_5	NDSVI_2	LSWI_8	B11_7	B11_7	B12_8
	NDTI	B2_1	B12_1	B12_2	EVI_4	NRED2_5	RENDVI_6	B11_7	B11_2	B11_3	NDSVI_2	B11_8
	RENDVI	LSWI_1	GNDVI	NDSVI_2	B2_4	B12_5	B2_5	B11_2	B5_7	NDSVI_6	B11_3	RENDVI_5
	NDSVI	TVI	B2_1	NDSVI_3	B12_4	B5_4	B11_5	B4_7	B12_5	B11_9	B12_7	NRED2_6
	B11	B6_1	OSAVI_1	B12_2	NRED1_4	B2_2	B5_5	NDSVI_5	NDSVI_7	B4_2	NRED3_8	B12_2
	VIgreen	GNDVI_1	GCVI_1	DVI_2	OSAVI_4	B12_5	B11_5	NRED2_7	B12_2	NRED3_6	NRED1_2	GCVI_8
	EVI	B12_1	RENDVI_1	NRED2	B12_4	B4_5	NDSVI_5	LSWI_2	B3_3	LSWI_2	B3_3	B12_9
	B12	NRED2_1	NRED1_1	TVI_1	NDSVI_3	LSWI_3	LSWI_3	NRED3_4	NDSVI_4	LSWI_3	LSWI_3	B11_2
	WDRVI	NDSVI_1	TVI	NRED1_2	NRED2_3	B11_2	B3_2	B12_2	B12_6	NDSVI_6	DVI_5	NDSVI_3
	NRED2	NRED3_1	NDSVI_2	NDTI_3	NDTI_4	NDSVI_2	WDRVI_7	B12_7	B11_2	B12_2	RVI_7	NRED3_7
	B4	NRED1_1	WDRVI	B3_3	RVI_4	NRED2_4	NDSVI_2	LSWI_4	NRED3_2	NDSVI_1	NDTI_2	NDSVI_3
	B2	B4_1	NDTI	WDRVI_3	LSWI_2	RVI_4	B5_6	VIgreen_5	EVI_7	B12_8	B12_4	B2_8
	B8A	B6_1	B12_2	B5_3	TVI_4	NDVI_2	B12_2	NDSVI_6	NDTI_2	NDTI_2	MCARI_8	B8A_2
	B3	NRED2	B2_2	B8A_3	B11_3	DVI_5	NRED3_6	B5_2	B11_6	NRED2_8	LSWI_7	DVI_8

Table A3. Optimal feature sets at different periods in Region 3.

Phase	April 29	May 7	May 29	June 13	July 4	July 18	July 23	August 20	August 23	September 29	October 9
Rank by feature importance	NDTI	NDTI	B12_2	NDSVI_3	GNDVI_4	B3_5	NDSVI_6	LSWI_7	RENDVI_8	NDSVI_3	NDSVI_3
	TVI	NDTI_1	OSAVI_2	NRED2_3	B12_4	B11_5	NRED2_6	B5_7	B8_2	B6_7	NDTI_7
	LSWI	NDSVI_1	NDTI_2	B11_3	B11_4	B11_5	B12_6	NDTI_7	B11_7	B11_6	EVI_2
	B12	LSWI_1	NDTI_2	RVI_2	WDRVI_4	NRED2_4	B5_2	NDSVI_7	NDWI_8	NRED3_5	NDSVI_8
	NDSVI	RENDVI	LSWI_2	NDWI_2	NDSVI_2	B12_2	NDTI_6	NDSVI_3	NRED2_2	NDVI_7	NDTI_8
	RENDVI	TVI_1	NDSVI	DVI_2	NRED2_2	B11_5	B11_1	B11_7	B11_7	OSAVI_8	B6_8
	B5	WDRVI	TVI_2	NRED1_1	NRED1_2	B5_5	B11_6	B12_7	B12_8	B5_6	MCARI_8
	NRED1	B11_1	B11	B2	B3_4	B5_5	B11_4	NRED1_7	B12_7	NDSVI_2	WDRVI_8
	NRED2	B3_1	NDTI_1	NRED2_2	NDSVI_1	VIgreen_5	B3_5	B11_5	B3_7	NRED1_7	B3_1
	B11	NRED1	VIgreen_2	MCARI	VIgreen_2	NDSVI_5	B3_6	B12_7	B11_8	B11_3	OSAVI_8
	VIgreen	TVI	EVI_2	OSAVI_3	NDWI_4	NDSVI_5	NDSVI_4	EVI_6	B12_8	NRED2_5	NRED3_8
	NDVI	NDVI	NRED2	B12_2	NRED3_4	B12_5	B4_6	B5_7	B8_6	EVI_2	B12_8
	EVI	NDVI_1	NRED1_1	NRED1_2	RENDVI_3	EVI_5	B2_6	RVI_7	B12_6	B7_7	NDSVI_5
	NRED3	RVI	RVI_2	NRED1_1	EVI_4	RENDVI_5	B2_6	RENDVI_6	NDSVI_2	B12_7	B11_7
GNDVI	RENDVI	NDTI_1	B12	B11_2	B12_5	NRED1_6	RENDVI_6	TVI_3	NRED3_8	B8A_7	

Note: This study uses an incremental image data design. The features of the first period have no suffix, and the suffixes of the features of each subsequent period are incremented by 1. For example, _1 represents the features used in the second period of each region, _2 represents the features used in the third period in each region.

References

- Valero, S.; Arnaud, L.; Planells, M.; Ceschia, E. Synergy of Sentinel-1 and Sentinel-2 Imagery for Early Seasonal Agricultural Crop Mapping. *Remote Sens.* **2021**, *13*, 4891. [\[CrossRef\]](#)
- Demarez, V.; Helen, F.; Marais-Sicre, C.; Baup, F. In-Season Mapping of Irrigated Crops Using Landsat 8 and Sentinel-1 Time Series. *Remote Sens.* **2019**, *11*, 118. [\[CrossRef\]](#)
- Hao, P.Y.; Zhan, Y.L.; Wang, L.; Niu, Z.; Shakir, M. Feature Selection of Time Series MODIS Data for Early Crop Classification Using Random Forest: A Case Study in Kansas, USA. *Remote Sens.* **2015**, *7*, 5347–5369. [\[CrossRef\]](#)
- Chen, Z.; Ren, J.; Tang, H.; Shi, Y.; Liu, J. Progress and perspectives on agricultural remote sensing research and applications in China. *J. Remote Sens.* **2016**, *20*, 748–767.
- Jia, K.; Li, Q.Z. Research status and prospect of feature variable selection for crop remote sensing classification. *Resour. Sci.* **2013**, *35*, 2507–2516.
- Baldeck, C.A.; Asner, G.P. Estimating Vegetation Beta Diversity from Airborne Imaging Spectroscopy and Unsupervised Clustering. *Remote Sens.* **2013**, *5*, 2057–2071. [\[CrossRef\]](#)
- Fekri, E.; Latifi, H.; Amani, M.; Zobeidinezhad, A. A Training Sample Migration Method for Wetland Mapping and Monitoring Using Sentinel Data in Google Earth Engine. *Remote Sens.* **2021**, *13*, 4169. [\[CrossRef\]](#)
- Wei, M.F.; Qiao, B.J.; Zhao, J.H.; Zuo, X.Y. The area extraction of winter wheat in mixed planting area based on Sentinel-2 a remote sensing satellite images. *Int. J. Parallel Emergent Distrib. Syst.* **2020**, *35*, 297–308. [\[CrossRef\]](#)
- Skakun, S.; Roger, J.-C.; Vermote, E.; Franch, B.; Becker-Reshef, I.; Justice, C.O.; Masek, J.G. Combined Use of Landsat-8 and Sentinel-2 Data for Agricultural Monitoring. In Proceedings of the AGU Fall Meeting Abstracts, New Orleans, LA, USA, 11–17 December 2017; p. EP22C-07.
- Gallego, F.J.; Kussul, N.; Skakun, S.; Kravchenko, O.; Shelestov, A.; Kussul, O. Efficiency assessment of using satellite data for crop area estimation in Ukraine. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *29*, 22–30. [\[CrossRef\]](#)
- Ibrahim, E.S.; Rufin, P.; Nill, L.; Kamali, B.; Nendel, C.; Hostert, P. Mapping Crop Types and Cropping Systems in Nigeria with Sentinel-2 Imagery. *Remote Sens.* **2021**, *13*, 3523. [\[CrossRef\]](#)
- Rao, P.; Zhou, W.; Bhattarai, N.; Srivastava, A.K.; Singh, B.; Poonia, S.; Lobell, D.B.; Jain, M. Using Sentinel-1, Sentinel-2, and Planet Imagery to Map Crop Type of Smallholder Farms. *Remote Sens.* **2021**, *13*, 1870. [\[CrossRef\]](#)
- Wei, X.C. Extraction of Paddy Rice Planting Area Based on Environmental Satellite Images—Taking Jiangnan Plain Area as Example. Mather's Thesis, Hubei University, Wuhan, China, 2013.
- Wang, L.M.; Liu, J.; Yang, F.G.; Fu, C.H.; Teng, F.; Gao, J.M. Early recognition of winter wheat area based on GF-1 satellite. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 194–201.
- Tian, H.; Wang, Y.; Chen, T.; Zhang, L.; Qin, Y. Early-Season Mapping of Winter Crops Using Sentinel-2 Optical Imagery. *Remote Sens.* **2021**, *13*, 3822. [\[CrossRef\]](#)
- Pan, L.; Xia, H.; Zhao, X.; Guo, Y.; Qin, Y. Mapping Winter Crops Using a Phenology Algorithm, Time-Series Sentinel-2 and Landsat-7/8 Images, and Google Earth Engine. *Remote Sens.* **2021**, *13*, 2510. [\[CrossRef\]](#)
- Liu, J.K.; Zhong, S.Q.; Liang, W.H. Extraction on crops planting structure based on multi-temporal Landsat8 OLI Images. *Remote Sens. Technol. Appl.* **2015**, *30*, 775–783.
- Chen, L.; Lin, H.; Sun, H.; Yan, E.P.; Wang, J.J. Studies on information extraction of forest in Zhuzhou city based on decision tree classification. *J. Cent. South Univ. For. Technol.* **2013**, *33*, 46–51.
- Varela, S.; Dhodda, P.R.; Hsu, W.H.; Prasad, P.V.V.; Assefa, Y.; Peralta, N.R.; Griffin, T.; Sharda, A.; Ferguson, A.; Ciampitti, I.A. Early-Season Stand Count Determination in Corn via Integration of Imagery from Unmanned Aerial Systems (UAS) and Supervised Learning Techniques. *Remote Sens.* **2018**, *10*, 343. [\[CrossRef\]](#)
- You, N.S.; Dong, J.W. Examining earliest identifiable timing of crops using all available Sentinel 1/2 imagery and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **2020**, *161*, 109–123. [\[CrossRef\]](#)
- Vorobiova, N.; Chernov, A. Curve fitting of MODIS NDVI time series in the task of early crops identification by satellite images. *Procedia Eng.* **2017**, *201*, 184–195. [\[CrossRef\]](#)
- Guan, X.; Huang, C.; Liu, G.; Meng, X.; Liu, Q. Mapping Rice Cropping Systems in Vietnam Using an NDVI-Based Time-Series Similarity Measurement Based on DTW Distance. *Remote Sens.* **2016**, *8*, 19. [\[CrossRef\]](#)
- Li, Q.T.; Wang, C.Z.; Zhang, B.; Lu, L.L. Object-Based Crop Classification with Landsat-MODIS Enhanced Time-Series Data. *Remote Sens.* **2015**, *7*, 16091–16107. [\[CrossRef\]](#)
- Mondal, S.; Jeganathan, C. Mountain agriculture extraction from time-series MODIS NDVI using dynamic time warping technique. *Int. J. Remote Sens.* **2018**, *39*, 3679–3704. [\[CrossRef\]](#)
- Jia, K.; Liang, S.L.; Zhang, L.; Wei, X.Q.; Yao, Y.J.; Xie, X.H. Forest cover classification using Landsat ETM+ data and time series MODIS NDVI data. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *33*, 32–38. [\[CrossRef\]](#)
- Atzberger, C.; Rembold, F. Mapping the Spatial Distribution of Winter Crops at Sub-Pixel Level Using AVHRR NDVI Time Series and Neural Nets. *Remote Sens.* **2013**, *5*, 1335–1354. [\[CrossRef\]](#)
- Xiao, W.; Xu, S.; He, T. Mapping Paddy Rice with Sentinel-1/2 and Phenology-, Object-Based Algorithm—A Implementation in Hangjiahu Plain in China Using GEE Platform. *Remote Sens.* **2021**, *13*, 990. [\[CrossRef\]](#)
- Chakhar, A.; Hernández-López, D.; Ballesteros, R.; Moreno, M.A. Improving the Accuracy of Multiple Algorithms for Crop Classification by Integrating Sentinel-1 Observations with Sentinel-2 Data. *Remote Sens.* **2021**, *13*, 243. [\[CrossRef\]](#)

29. Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.R.; Murayama, Y.; Ranagalage, M. Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sens.* **2020**, *12*, 2291. [[CrossRef](#)]
30. Hu, Y.; Zeng, H.; Tian, F.; Zhang, M.; Wu, B.; Gilliams, S.; Li, S.; Li, Y.; Lu, Y.; Yang, H. An Interannual Transfer Learning Approach for Crop Classification in the Hetao Irrigation District, China. *Remote Sens.* **2022**, *14*, 1208. [[CrossRef](#)]
31. Praticò, S.; Solano, F.; Di Fazio, S.; Modica, G. Machine Learning Classification of Mediterranean Forest Habitats in Google Earth Engine Based on Seasonal Sentinel-2 Time-Series and Input Image Composition Optimisation. *Remote Sens.* **2021**, *13*, 586. [[CrossRef](#)]
32. Zhao, L.C. Research on Indicative Image Identification Feature of Major Crops and Its Application Method. Mather's Thesis, Aerospace Information Research Institute China Academy of Sciences, Beijing, China, 2020.
33. Yu, L.; Su, J.; Li, C.; Wang, L.; Luo, Z.; Yan, B. Improvement of Moderate Resolution Land Use and Land Cover Classification by Introducing Adjacent Region Features. *Remote Sens.* **2018**, *10*, 414. [[CrossRef](#)]
34. He, Z.X.; Zhang, M.; Wu, B.F.; Xing, Q. Extraction of summer crop in Jiangsu based on Google Earth Engine. *J. Geo-Inf. Sci.* **2019**, *21*, 752–766.
35. Rouse, J.W.; Haas, R.H.; Schell, J.; Deering, D. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Remote Sensing Center: College Station, TX, USA, 1973.
36. Xiao, X.; Boles, S.; Frolking, S.; Li, C.; Babu, J.Y.; Salas, W.; Moore, B., III. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sens. Environ.* **2006**, *100*, 95–113. [[CrossRef](#)]
37. Huete, A.; Justice, C.; Van Leeuwen, W. MODIS vegetation index (MOD13). *Algorithm Theor. Basis Doc.* **1999**, *3*, 295–309.
38. Daughtry, C.S.; Walthall, C.; Kim, M.; De Colstoun, E.B.; McMurtrey, J.E., III. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [[CrossRef](#)]
39. Deering, D.W. *Rangeland Reflectance Characteristics Measured by Aircraft and Spacecraftsensors*; Texas A&M University: College Station, TX, USA, 1978.
40. Richardson, A.J.; Everitt, J.H. Using spectral vegetation indices to estimate rangeland productivity. *Geocarto Int.* **1992**, *7*, 63–69. [[CrossRef](#)]
41. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [[CrossRef](#)]
42. Gitelson, A.A.; Viña, A.; Arkebauer, T.J.; Rundquist, D.C.; Keydan, G.; Leavitt, B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys. Res. Lett.* **2003**, *30*, 1248. [[CrossRef](#)]
43. Gitelson, A.; Merzlyak, M.N. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B Biol.* **1994**, *22*, 247–252. [[CrossRef](#)]
44. Deventer, V.A.P.; Ward, A.D.; Gowda, P.H.; Lyon, J.G. Using thematic mapper data to identify contrasting soil plains and tillage practices. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 87–93.
45. Qi, J.; Marsett, R.; Heilman, P.; Bieden-bender, S.; Moran, S.; Goodrich, D.; Weltz, M.J.E. RANGES improves satellite-based information and land cover assessments in southwest United States. *Eos Trans. Am. Geophys. Union* **2002**, *83*, 601–606. [[CrossRef](#)]
46. Peña-Barragán, J.M.; Ngugi, M.K.; Plant, R.E.; Six, J. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sens. Environ.* **2011**, *115*, 1301–1316. [[CrossRef](#)]
47. Gitelson, A.A. Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. *J. Plant Physiol.* **2004**, *161*, 165–173. [[CrossRef](#)] [[PubMed](#)]
48. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [[CrossRef](#)]
49. Gao, B.C. Normalized difference water index for remote sensing of vegetation liquid water from space. In Proceedings of the Imaging Spectrometry, Orlando, FL, USA, 17–21 April 1995; pp. 225–236.
50. Bhatt, P.; Maclean, A.; Dickinson, Y.; Kumar, C. Fine-Scale Mapping of Natural Ecological Communities Using Machine Learning Approaches. *Remote Sens.* **2022**, *14*, 563. [[CrossRef](#)]
51. Zhang, H.X.; Wang, Y.J.; Shang, J.L.; Liu, M.X.; Li, Q.Z. Investigating the impact of classification features and classifiers on crop mapping performance in heterogeneous agricultural landscapes. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102388. [[CrossRef](#)]
52. Rana, V.K.; Suryanarayana, T.M.V. Performance evaluation of MLE, RF and SVM classification algorithms for watershed scale land use/land cover mapping using sentinel 2 bands. *Remote Sens. Appl. Soc. Environ.* **2020**, *19*, 100351. [[CrossRef](#)]
53. Asgarian, A.; Soffianian, A.; Pourmanafi, S. Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: A case of central Iran using multi-temporal Landsat 8 imagery. *Comput. Electron. Agric.* **2016**, *127*, 531–540. [[CrossRef](#)]
54. Lary, D.J.; Alavi, A.H.; Gandomi, A.H.; Walker, A.L. Machine learning in geosciences and remote sensing. *Geosci. Front.* **2016**, *7*, 3–10. [[CrossRef](#)]
55. Aneece, I.; Thenkabail, P.S. Classifying Crop Types Using Two Generations of Hyperspectral Sensors (Hyperion and DESIS) with Machine Learning on the Cloud. *Remote Sens.* **2021**, *13*, 4704. [[CrossRef](#)]
56. Pelletier, C.; Valero, S.; Inglada, J.; Champion, N.; Dedieu, G. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. *Remote Sens. Environ.* **2016**, *187*, 156–168. [[CrossRef](#)]
57. Hao, P.; Di, L.; Zhang, C.; Guo, L. Transfer Learning for Crop classification with Cropland Data Layer data (CDL) as training samples. *Sci. Total Environ.* **2020**, *733*, 138869. [[CrossRef](#)] [[PubMed](#)]

58. Kang, Y.; Hu, X.; Meng, Q.; Zou, Y.; Zhang, L.; Liu, M.; Zhao, M. Land Cover and Crop Classification Based on Red Edge Indices Features of GF-6 WFV Time Series Data. *Remote Sens.* **2021**, *13*, 4522. [[CrossRef](#)]
59. Maponya, M.G.; Van Niekerk, A.; Mashimbye, Z.E. Pre-harvest classification of crop types using a Sentinel-2 time-series and machine learning. *Comput. Electron. Agric.* **2020**, *169*, 105164. [[CrossRef](#)]
60. Luo, H.X.; Dai, S.P.; Li, M.F.; Liu, E.P.; Zheng, Q.; Hu, Y.Y.; Yi, X.P. Comparison of machine learning algorithms for mapping mango plantations based on Gaofen-1 imagery. *J. Integr. Agric.* **2020**, *19*, 2815–2828. [[CrossRef](#)]
61. Yang, L.B.; Mansaray, L.R.; Huang, J.F.; Wang, L.M. Optimal segmentation scale parameter, feature subset and classification algorithm for geographic object-based crop recognition using multisource satellite imagery. *Remote Sens.* **2019**, *11*, 514. [[CrossRef](#)]
62. Mansaray, L.R.; Zhang, K.; Kanu, A.S. Dry biomass estimation of paddy rice with Sentinel-1A satellite data using machine learning regression algorithms. *Comput. Electron. Agric.* **2020**, *176*, 105674. [[CrossRef](#)]
63. Šiljeg, A.; Pađa, L.; Domazetović, F.; Marić, I.; Gašparović, M.; Borisov, M.; Milošević, R. Comparative Assessment of Pixel and Object-Based Approaches for Mapping of Olive Tree Crowns Based on UAV Multispectral Imagery. *Remote Sens.* **2022**, *14*, 757. [[CrossRef](#)]
64. Le Quilleuc, A.; Collin, A.; Jasinski, M.F.; Devillers, R. Very High-Resolution Satellite-Derived Bathymetry and Habitat Mapping Using Pleiades-1 and ICESat-2. *Remote Sens.* **2022**, *14*, 133. [[CrossRef](#)]
65. Sakamoto, M.; Ullah, S.M.A.; Tani, M. Land Cover Changes after the Massive Rohingya Refugee Influx in Bangladesh: Neo-Classic Unsupervised Approach. *Remote Sens.* **2021**, *13*, 5056. [[CrossRef](#)]
66. Barber, M.E.; Rava, D.S.; López-Martínez, C. L-Band SAR Co-Polarized Phase Difference Modeling for Corn Fields. *Remote Sens.* **2021**, *13*, 4593. [[CrossRef](#)]
67. Ha, N.T.; Manley-Harris, M.; Pham, T.D.; Hawes, I. A Comparative Assessment of Ensemble-Based Machine Learning and Maximum Likelihood Methods for Mapping Seagrass Using Sentinel-2 Imagery in Tauranga Harbor, New Zealand. *Remote Sens.* **2020**, *12*, 355. [[CrossRef](#)]
68. Ren, T.; Liu, Z.; Zhang, L.; Liu, D.; Xi, X.; Kang, Y.; Zhao, Y.; Zhang, C.; Li, S.; Zhang, X. Early Identification of Seed Maize and Common Maize Production Fields Using Sentinel-2 Images. *Remote Sens.* **2020**, *12*, 2140. [[CrossRef](#)]
69. Obata, S.; Cieszewski, C.J.; Lowe, R.C., III; Bettinger, P. Random Forest Regression Model for Estimation of the Growing Stock Volumes in Georgia, USA, Using Dense Landsat Time Series and FIA Dataset. *Remote Sens.* **2021**, *13*, 218. [[CrossRef](#)]
70. Fan, D.D.; Li, Q.Z.; Wang, H.Y.; Du, X. Improvement in recognition accuracy of minority crops by resampling of imbalanced training datasets of remote sensing. *J. Remote Sens.* **2019**, *23*, 730–742.
71. Tuvdendorj, B.; Zeng, H.; Wu, B.; Elnashar, A.; Zhang, M.; Tian, F.; Nabil, M.; Nanzad, L.; Bulkhbai, A.; Natsagdorj, N. Performance and the Optimal Integration of Sentinel-1/2 Time-Series Features for Crop Classification in Northern Mongolia. *Remote Sens.* **2022**, *14*, 1830. [[CrossRef](#)]
72. Luo, H.; Li, M.; Dai, S.; Li, H.; Li, Y.; Hu, Y.; Zheng, Q.; Yu, X.; Fang, J. Combinations of Feature Selection and Machine Learning Algorithms for Object-Oriented Betel Palms and Mango Plantations Classification Based on Gaofen-2 Imagery. *Remote Sens.* **2022**, *14*, 1757. [[CrossRef](#)]
73. Zhang, H.X.; Li, Q.Z.; Wen, N.; Du, X.; Tao, Q.S.; Tian, Y.C. Important factors affecting crop acreage estimation based on remote sensing image classification technique. *Remote Sens. Land Resour.* **2015**, *27*, 54–61.
74. Dong, J.; Xiao, X.; Menarguez, M.A.; Zhang, G.; Qin, Y.; Thau, D.; Biradar, C.; Moore, B., III. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sens. Environ.* **2016**, *185*, 142–154. [[CrossRef](#)]
75. Yin, L.K.; You, N.S.; Zhang, G.L.; Huang, J.C.; Dong, J.W. Optimizing feature selection of individual crop types for improved crop mapping. *Remote Sens.* **2020**, *12*, 162. [[CrossRef](#)]
76. Zhang, X.; Zhang, P.; Shen, K.; Pei, Z. Rice identification at the early stage of the rice growth season with single fine quad Radarsat-2 data. In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII*; International Society for Optics and Photonics: Bellingham, WA, USA, 2016; Volume 9998, p. 99981J.
77. Singha, M.; Dong, J.; Zhang, G.; Xiao, X. High resolution paddy rice maps in cloud-prone Bangladesh and Northeast India using Sentinel-1 data. *Sci. Data* **2019**, *6*, 1–10. [[CrossRef](#)]
78. Shen, Y.; Li, Q.; Du, X.; Wang, H.Y.; Zhang, Y. Indicative features for identifying corn and soybean using remote sensing imagery at the middle and later growth season. *J. Remote Sens.* **2021**. [[CrossRef](#)]
79. Veloso, A.; Mermoz, S.; Bouvet, A.; Le Toan, T.; Planells, M.; Dejoux, J.-F.; Ceschia, E. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sens. Environ.* **2017**, *199*, 415–426. [[CrossRef](#)]
80. Zhai, D.; Dong, J.; Cadisch, G.; Wang, M.; Kou, W.; Xu, J.; Xiao, X.; Abbas, S. Comparison of Pixel- and Object-Based Approaches in Phenology-Based Rubber Plantation Mapping in Fragmented Landscapes. *Remote Sens.* **2018**, *10*, 44. [[CrossRef](#)]
81. Feng, Z.; Huang, G.; Chi, D. Classification of the Complex Agricultural Planting Structure with a Semi-Supervised Extreme Learning Machine Framework. *Remote Sens.* **2020**, *12*, 3708. [[CrossRef](#)]
82. Zhu, S.; Zhang, J. Provincial agricultural stratification method for crop area estimation by remote sensing. *Trans. Chin. Soc. Agric. Eng.* **2013**, *29*, 184–191.
83. Yi, Z.; Jia, L.; Chen, Q. Crop Classification Using Multi-Temporal Sentinel-2 Data in the Shiyang River Basin of China. *Remote Sens.* **2020**, *12*, 4052. [[CrossRef](#)]

84. Ma, L.; Liu, Y.; Zhang, X.L.; Ye, Y.X.; Yin, G.F.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [[CrossRef](#)]
85. Reedha, R.; Dericquebourg, E.; Canals, R.; Hafiane, A. Transformer Neural Network for Weed and Crop Classification of High Resolution UAV Images. *Remote Sens.* **2022**, *14*, 592. [[CrossRef](#)]
86. Wang, D.; Cao, W.; Zhang, F.; Li, Z.; Xu, S.; Wu, X. A Review of Deep Learning in Multiscale Agricultural Sensing. *Remote Sens.* **2022**, *14*, 559. [[CrossRef](#)]