Computers and Electronics in Agriculture 176 (2020) 105618

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yi'an County, Heilongjiang province, China

Hongyan Zhang[a,*], Jinzhong Kang[a], Xiong Xu[b,*], Liangpei Zhang[a]

[a] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei 430079, China
[b] College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China

ARTICLE INFO

Keywords:
Sentinel-2
Machine learning
Support vector machine
Random forest
Short-wave infrared
Crop type mapping
Feature selection

ABSTRACT

Crop type mapping visualizes the spatial distribution patterns and proportions of the cultivated areas with different crop types, and is the basis for subsequent agricultural applications. Understanding the effectiveness of different temporal and spectral features in detailed crop classification can help users optimize temporal window selection and spectral feature space construction in crop type mapping applications. Therefore, in this study, we used time-series Sentinel-2 image data from Yi'an County, Heilongjiang province, China, to analyze the effectiveness of the temporal and spectral features used in three common machine learning classification methods: classification and regression tree (CART) decision tree, Support Vector Machine (SVM), and random forest (RF). For CART and SVM classifiers, the relative importance of the features was reflected by the order and frequency of attributes selected as the node and the square of the model weight. In RF, the change in prediction error as calculated by out of bag data is taken as the measure of feature importance. The standard deviation of the average value of all labeled pixels was used to evaluate the correctness of the unanimous conclusions drawn by these three methodologies. The quantitative evaluation results given by the confusion matrix show that random forest achieved the best overall accuracy, while support vector machine ranked second, and the decision tree algorithm yielded the least accurate classification results. From the perspective of feature importance, making full use of the discriminative information between different crops, and constructing a rational feature space, can help to improve classification accuracy significantly. In detail, the discriminative information between the different crop types is as follows: 1) images at the peak of the crop growth period are crucial in the classification of different crops; 2) the short-wave infrared bands are particularly suitable for fine crop classification; and 3) the red edge bands can effectively assist classification. Finally, our study achieved crop type mapping in the study area with an overall accuracy of 97.85% and a Kappa coefficient of 0.95.

1. Introduction

Agriculture is the foundation for social and economic development (Awokuse, 2009), food security (Lu et al., 2016; Gilbertson et al., 2017) and land resource management (Huang et al., 2016; Lebourgeois et al., 2017). Crop type mapping can help us to obtain the spatial distribution patterns and proportions of the cultivated areas of different crop types, and is the basis for yield estimation (Bolton and Friedl, 2013; Song et al., 2017; van der Velde et al., 2019), water resources management (Vogels et al., 2019), and disaster assessment (Zhang, 2004). However, the traditional methods of obtaining and updating the crop type and planting area information are mainly based on sampling surveys and statistical reports (Hu et al., 2017), which have problems such as strong subjectivity, time-consuming, labor-intensive, delayed updating, and the lack of spatial distribution information (Zhong et al., 2016; Zhang et al., 2020). In contrast, remote sensing technology has delivers wide coverage, timely data acquisition, fast and dynamic updating, and thus has become an increasingly powerful tool for crop type mapping (Wei et al., 2018; Zhai et al., 2019a).

However, crop type mapping using remote sensing data is challenging for several reasons. One reason is the mutual constraint between remote sensing imaging technology and the data cost (Peña-Barragán et al., 2011; Zhang et al., 2019). An appropriate combination of spatial, spectral, and temporal resolutions is required (Peña-Barragán et al., 2011), then low-cost but of good quality data is always preferred. For example, Moderate-Resolution Imaging Spectroradiometer (MODIS)

* Corresponding authors.
E-mail addresses: zhanghongyan@whu.edu.cn (H. Zhang), xvxiong@tongji.edu.cn (X. Xu).

https://doi.org/10.1016/j.compag.2020.105618
Received 6 February 2020; Received in revised form 25 May 2020; Accepted 2 July 2020
Available online 17 July 2020

images have been extensively used for these reasons over the last decade (Gallego et al., 2014; Hu et al., 2017; King et al., 2017; Massey et al., 2017; Song et al., 2017). The corresponding products are also appropriate for studying land-use changes caused by rapid agricultural development (Teluguntla et al., 2018). However, the spatial resolution of these products is too coarse to characterize individual farmland plots. Such products are only appropriate for parcels larger than 32 ha (Wardlow and Egbert, 2008), which limits their usefulness in assessing small cultivated plots. Although the Landsat satellite has alleviated this problem, the spatial resolution of its scale still cannot capture the actual spatial distribution patterns of fragmented farmland (with field sizes less than 0.5 ha), particularly in China, where the average field size is less than 1 ha (Samberg et al., 2016). In addition, the 16-day revisit cycle of Landsat is insufficient to capture the different phenological information of crops. High and very high resolution (less than 1 m) images contain rich crop texture and structure information, and thus have the ability to be used for crop type mapping under complex terrain and planting conditions, but their high cost, large data storage and computational requirements limit their application (Wang et al., 2019). Fortunately, with the operation of the Sentinel-2 series of satellites, the spatial resolution of medium-resolution images can reach 10 m, which can more accurately represent the spatial distribution patterns of different crop types. Meanwhile, the revisit cycle of Sentinel-2 is only five days, which is helpful for capturing the phenological information of crops during the growing season. This type of bi-directionally enhanced data provides us with a new opportunity for precise crop type mapping (Kang et al., 2018, Lupia and Antoniou, 2018).

Another reason that crop type mapping using remote sensing data is a challenging task can be attributed to agronomic factors. Different crop types may have similar phenological calendars and spectral responses (interclass similarity). At the same time, crop sowing times and growth states can be varied for the same crop type (intraclass variability) on account of the complex environmental and climatic conditions. Both these issues make it even more challenging to identify specific crop types, compared with land cover mapping.

The spectral feature is the most basic information of remote sensing data and needs to be considered thoroughly in such complex circumstances (Zhong et al., 2014; Zhai et al., 2019b). The spectral reflectance of crops is influenced by the vegetation status, pigmentation, leaf water content, residue cover, etc. (Bolton and Friedl, 2013). Different crop types therefore have different responses in different spectral regions. Thus, spectral indices were designed to amplify these differences. The normalized difference vegetation index (NDVI) utilizes the different increase rate of reflectivity between the red and near-infrared bands to characterize the vegetation status and canopy structure, and has been demonstrated to be a very effective way to distinguish different crops (Wardlow and Egbert, 2008). Despite this, some crops exhibit unique phenological stages that are distinct from those of other crops, which means that the extensively used NDVI may not be the optimal spectral feature for identifying these crops (Hu et al., 2017). For instance, for rice, irrigation information is a more effective discriminative attribute than canopy structure when differentiating rice from other crop types (Wang et al., 2015; Dong et al., 2016). Therefore, an index that is sensitive to soil moisture and leaf water content, such as the land surface water index (LSWI), may be more effective in extracting rice in periods of transplanting and flooding. Furthermore, a single spectral index such as NDVI or LSWI may be unable to capture some of the vital phenological stages when multiple crop types are mapped simultaneously, and they cannot characterize complex crop development and environmental conditions. Hence, it is essential to understand the relative effectiveness of the different spectral features for crop type mapping, and to evaluate whether the red edge bands in Sentinel-2 are important for accurate crop classification.

Temporal attributes are another primary factor relevant to crop type mapping (Foerster et al., 2012), and can help distinguish expansive landforms, vegetation types, soil associations, and other natural resources. Some crop types such as corn and soybean however, may have similar phenological stages (e.g. sowing, emergence, three leaves, and senescence) in some periods, but they also have their own unique phenological properties (Löw et al., 2013). Phenological differences thus can be effectively exploited to distinguish one crop type from natural vegetation and other crops. A widely used approach in crop classification relies on a single-date image that corresponds to the time most suited for identification based on specific phenological events. For example, a SPOT 5 image from May 2006 was used by Yang et al. (2011) to identify different crop types in South Texas, United States, since most of the fields had reached or were approaching their peak canopy growth period. The “single-date image” method is both time-efficient and user-friendly (Hu et al., 2017). However, it is impossible to guarantee that an image at the appropriate identification time is available due to cloud cover, which can often prevent or delay the image acquisition (Carrão et al., 2008; Zhai et al., 2019). Moreover, a single image may fail when identifying several crop types with similar spectral responses at that date, thus leading to an inaccurate classification result (Hu et al., 2017). To overcome this problem, multi-temporal information has been extensively used to identify different crop types in agricultural applications (Wardlow and Egbert, 2008). Time-series images are able to cover the different phenological stages of crops and can describe the spectral distribution and variance of a pixel over a year or season. Although many studies have illustrated the high performance of using time-series images for crop type mapping (Peña-Barragán et al., 2011; Tatsumi et al., 2015; Hu et al., 2017), some limitations still exist. For example, not every temporal image provides useful information, and the utilization of all the time-series images does not necessarily lead to the highest accuracy (Löw et al., 2013). Other studies have indicated that the discriminative information between different land covers lies in a low-dimensional feature space, and that additional images generally add little information, but increase the complexity of the computation (Hu et al., 2017). In addition, the classification accuracy may even drop when the number of images used increases, which is known as the “Hughes effect” (Löw et al., 2013). The distinct disadvantages of crop type mapping based on single-date images and time-series data bring the need for better utilization of the temporal information to the fore. Therefore, a comprehensive understanding of the optimal temporal window for crop type mapping is also urgently needed.

In this paper, within this context, we analyze the relative usefulness of the temporal and spectral features employed in three common methods of machine learning classification. Yi’an County in Heilongjiang province, China, was taken as a case study, and time-series Sentinel-2 data from 2018 were used. The primary goal of our study was to provide further insights into the feature space construction of crop type mapping from multi-temporal Sentinel-2 data using the classical classifiers of support vector machine (SVM), random forest, and decision tree. In detail, the following research questions are specifically addressed in this paper:

- Which temporal window can capture the crucial phenological stages most effectively?
- What spectral feature is most discriminative?
- Whether the red edge bands of Sentinel-2 are important for accurate crop identification?

The rest of this paper is structured as follows. Section 2 provides detailed descriptions of the study area, the multi-temporal Sentinel-2 images, the preprocessing methods, the reference ground-truth data, and the experimental setup. Section 3 presents the experimental results, including a qualitative presentation and a quantitative evaluation. In Section 4, we answer our research questions based on the analysis of the relative usefulness of the temporal and spectral features for each of the three classification methods, and we verify these conclusions based on a statistical analysis. Section 5 summarizes and concludes the work.

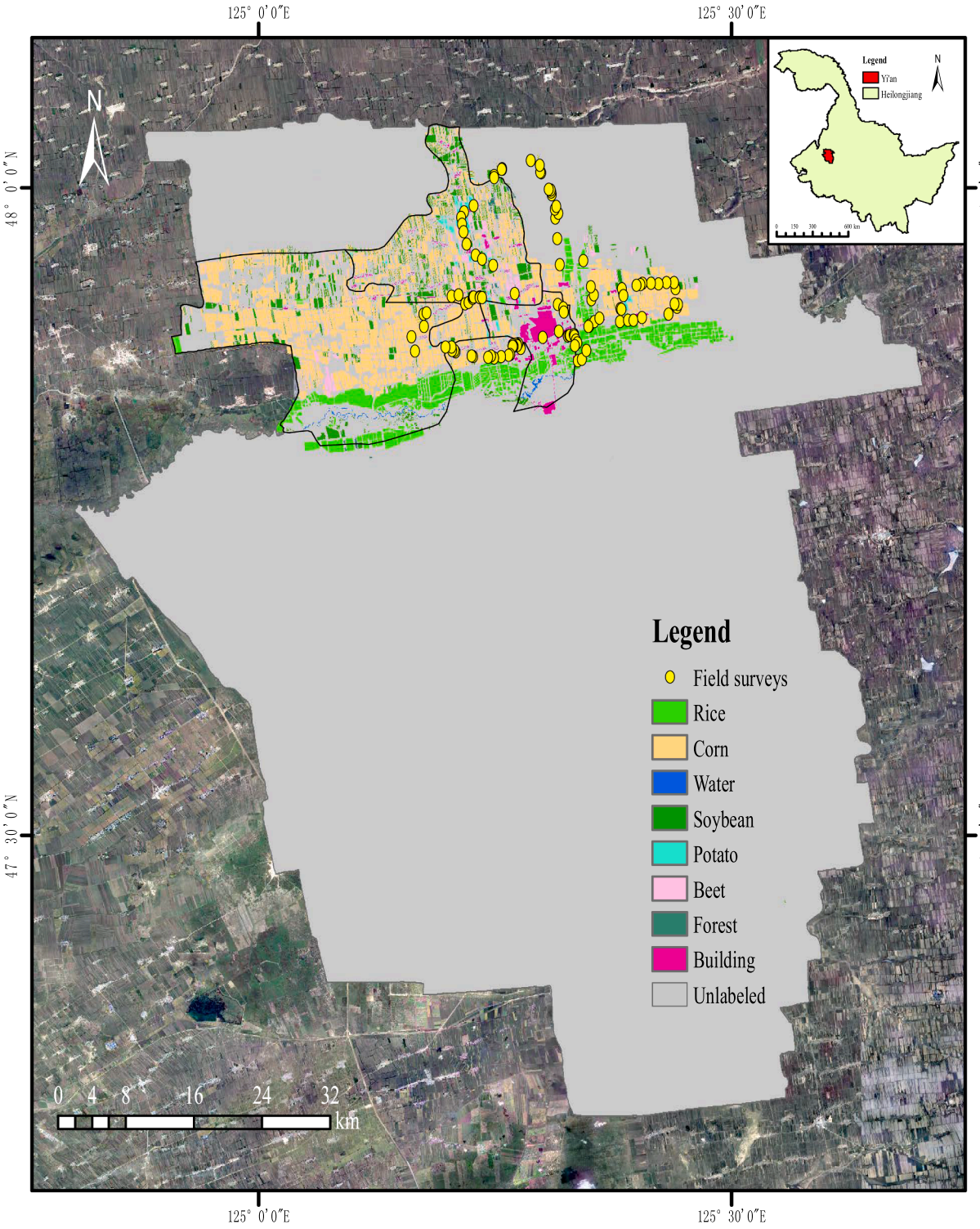

Fig. 1. Location of Yi'an County, Heilongjiang Province, China and the sample data. The study area is represented in gray and the sample data was manually labeled by visually interpreting three towns (Yi'an, Hongxing and Xintun) around the 115 field sampling points, which were selected in the field from July 16th to 20th, 2018. In total, 2,766,108 pixels were labeled as eight categories.

2. Materials and methodology

2.1. Study area

The study area of Yi'an County is a county-level administrative region in Heilongjiang province, China. It is located between 124°50′–125°42′E and 47°16′–48°02′N, with 70 km in length and 55 km in width, covering an area of 3,685 km^2, as shown in Fig. 1.

The area shown in Fig, 1 is the transition zone between the Kebai Plain and the Songnen Plain. The terrain gradually slopes from northeast to southwest, with an average elevation of 220 m. The area is thus suitable for the mechanized cultivation of crops. The climate is a temperate continental monsoon climate, which is windy in spring, warm and rainy in summer, cool in autumn, and cold in winter. As a result, the crops in this region are mainly cultivated in spring and harvested in autumn, due to the cold winter. The average temperature during April

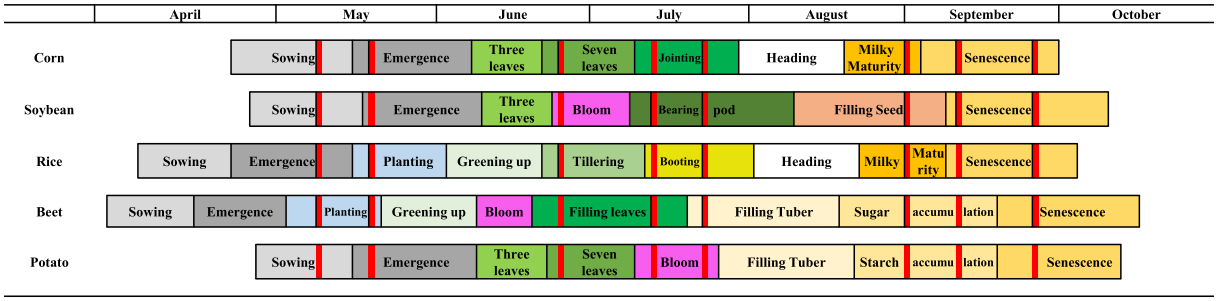

Fig. 2. Phenological calendars for the main crop types, and the dates of the selected images.

to October is 15.4 °C and the average precipitation is 402.9 mm, while the annual "effective accumulated temperature" is 2513.3 °C. Most of the precipitation is concentrated in summer, accounting for 69.4% of the annual precipitation. July has the most precipitation, with an average of 130 mm, accounting for 29.2% of the annual precipitation.

Heilongjiang province is known for its rich land and the high levels of organic matter in the soil, which is very suitable for the growth of crops. Thus, Yi'an is one of the major commodity grain counties of China. The main crop types in this region are rice, corn, soybeans, potatoes, and beets. Their specific phenological calendars are shown in Fig. 2.

The red vertical lines in Fig. 2 represent the dates of the image used in our study. The growth periods of these crops are all from April to October, growing the fastest during the end of July to the beginning of September (Hu et al., 2017). It can be observed that the phenological calendars of corn, soybeans and potatoes are very similar. They have many phenological stages during the same period, however, their phenological differences are concentrated from the end of June to the beginning of September. Beets have the longest growth cycle with unique phenological planting and flowering attributes. The flowering periods of soybeans, beets and potatoes are different, so we can use their phenological differences for classification.

2.2. Sentinel-2 data and preprocessing

The bi-directionally enhanced Sentinel-2 image data was used in this study. The Copernicus Sentinel-2 mission comprises a constellation of two polar-orbiting satellites placed in the same sun-synchronous orbit, phased at 180° to each other. The Sentinel-2 mission is aimed at monitoring the variability in land surface conditions. Its wide swath width (290 km) and short revisit time (10 days at the equator with one satellite, and five days with two satellites under cloud-free conditions, which results in revisit cycle 2–3 days at the mid-latitudes thus supporting the monitoring of Earth surface changes. The details of all the bands of the Sentinel-2 satellites are provided in Table 1.

We can see in Table 1 the spatial resolution of the red, green and blue bands reaches a scale of 10 m, more accurately representing the spatial distribution patterns of different crop types. Meanwhile, there are three addition vegetation red edge bands, which is far different from Landsat and MODIS. One of our research questions was to evaluate whether the red edge bands of Sentinel-2 are appropriate accurate crop identification.

Table 1
Spectral bands of the Sentinel-2 sensors.

Band name	Central wavelength (μm)	Resolution (m)
Band 1 – Coastal aerosol	0.443	60
Band 2 – Blue	0.490	10
Band 3 – Green	0.560	10
Band 4 – Red	0.665	10
Band 5 – Vegetation red edge	0.705	20
Band 6 – Vegetation red edge	0.740	20
Band 7 – Vegetation red edge	0.783	20
Band 8 – NIR	0.842	10
Band 8A – Narrow NIR	0.865	20
Band 9 – Water vapor	0.945	60
Band 10 – SWIR – Cirrus	1.375	60
Band 11 – SWIR	1.610	20
Band 12 – SWIR	2.190	20

The available images for the study area in 2018 were downloaded from the USGS Earth Explorer data portal (https://earthexplorer.usgs.gov/). To reduce the impact of summer rainfall on the data, all the available data was filtered with cloud content of no more than 20%, and eight temporal phases in the growing season were finally obtained. The corresponding dates are May 9th, May 19th, June 23rd, July 13th, July 23rd, September 1st, September 11th, and September 26th. The L1C-level atmospheric apparent reflectance data was converted to L2A-level surface reflectance data using the Sen2Cor-2.5.0 plug-in. Since the research area covers two tiles, each phase was obtained by cutting and seamlessly mosaicking the two tiles with an administrative vector. The images for each phase were resampled to a 10-m spatial resolution in the Sentinel Application Platform (SNAP) 6.0. All the images were unified into the WGS 1984 coordinate system with UTM 51 N projection. At this point, each time phase was made up of 12 bands, with the exception of the 10th band (cirrus band), which formed part of the input to the classifier as the spectral features.

2.3. Sample data

The sample data used in this study was manually labeled by visually interpreting the local area around the field sampling points in the study area. These 115 field sample points were selected in the field from July 16th to 20th, 2018. These data points were geo-referenced to remote sensing images by latitude and longitude coordinates. In total, 2,766,108 pixels in three towns, including Yi'an, Hongxing, and Xintun were labeled into eight categories, i.e. rice, corn, soybean, potato, beet, water, forest, and building, as shown in Fig. 1. The specific sample proportions are shown in Fig. 3.

As shown, corn, rice, soybean, potato and beet samples account for 54.22%, 21.72%, 13.13%, 1.91% and 3.88% of the total. Buildings, water and forest samples accounted for 5.13% of the labeled samples in

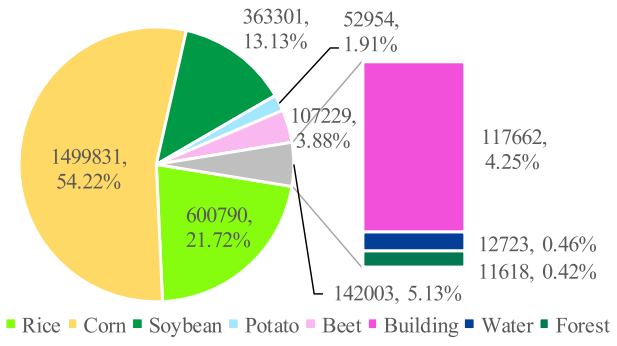

Fig. 3. Sample data category settings and proportions.

Table 2
Distribution of the training and test data (Unit: pixel).

Category	Rice	Corn	Soybean	Water	Potato	Building	Beet	Forest
Training	578	1533	359	12	63	106	103	12
Test	600,212	1,498,298	362,942	12,711	52,891	117,556	107,126	11,606

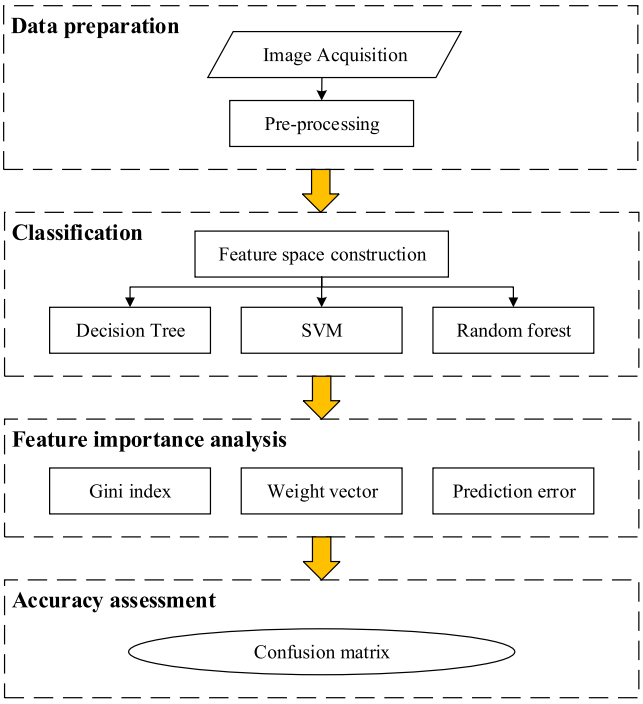

Fig. 4. The whole flow chart of our research.

total, and were 4.25%, 0.46% and 0.42%, respectively. The unbalanced distribution of the samples is caused by the uneven distribution of land cover and with the operation of labeling samples by administrative areas, one-thousandth of the pixels were randomly selected for classifier model training. The others were used to verify the classification accuracy. Their specific distribution is shown in Table 2.

2.4. Methodology

Our research consisted of four parts: data preparation, classification, feature importance analysis and accuracy assessment, illustrated by the whole flow chart in Fig. 4.

Firstly, remote sensing image preprocessing was implemented on the time series Sentinel-2 data that we downloaded. A feature space of five spectral indexes and 12 spectral bands for each temporal phase were constructed, and superimposed in chronological order. These features were fed into three common machine learning classifiers, i.e. decision tree, SVM and random forest. Three methodologies were used to calculate feature importance for these three classification algorithms. A confusion matrix of the classification results were used to evaluate the classification precision. Details of some parts will be given point by point in the following sections.

2.4.1. Feature space construction

In addition to the 12 original bands, five spectral indices were also calculated for each phase. We divided these spectral indices into three sets, according to the spectral bands they use, i.e., 1) visible-NIR region indices; 2) visible-SWIR region indices; and 3) SWIR-SWIR region indices. The formula for each index derived from Sentinel-2 wavebands are listed in Table 3.

The NDVI was selected first as it has been widely used in classification applications, and an effective discriminator for distinguishing between crops (Wardlow and Egbert, 2008). Other indices have also been used in agricultural studies, and have shown the potential to detect certain field crop properties that indicate crop differentiation (Peña-Barragán et al., 2011; Hu et al., 2017). For example, Kang et al. (2018) indicated that the normalized difference residue index (NDRI), the normalized differential senescent vegetation index (NDSVI), and the normalized differential tillage index (NDTI) make full use of the shortwave infrared bands to characterize the water content and vegetation residue coverage, so they play a key role in discrimination between corn and soybeans. In the meantime, the perpendicular moisture index (PMI), newly proposed by Zhang et al. (2018), has exhibited a good performance in rice extraction when combined with NDVI. Therefore, these five indices were selected, considering the crop structure in the study area. The two parameters in the formula of PMI were the same as the original paper, i.e. $G = 0.5$ and $M = 1.13$. Together with the 12 spectral bands for each phase that are mentioned before, there were 96 spectral bands and 40 index images for eight temporal phases. All 136 features were superimposed in chronological order and were used as the input feature spaces for the machine learning classification methods.

2.4.2. Crop type mapping classifier selection

Three classical machine learning classification methods were selected in this study: decision tree, SVM, and random forest. The decision tree algorithm is an algorithm based on tree structure, which is used to make a decision and obtain the final decision result (Gallego et al., 2014). A recursive generation process of decision tree can be divided into ID3, C4.5, and classification and regression tree (CART) decision trees, according to the different methods of selecting the optimal partitioning attributes. Since the first two tested methods require different numbers of attribute values (Quinlan, 1993), the CART decision tree was used for the classification in this study. In addition, a post-pruning process was carried out on the generated decision tree. The decision tree algorithm is simpler to construct than the other machine learning methods. More significantly, a decision tree algorithm can reflect the importance of the feature variables by calculating the optimal partition attribute index, which is conducive to the analysis of the classification process (Gallego et al., 2014).

SVM is a supervised learning method based on statistical learning theory and structural risk minimization principle (Piiroinen et al., 2015). It has advantages when solving nonlinear problems, handling small sample sets, and processing high-dimensional data (Asgarian et al., 2016). SVM usually uses kernel functions to map the samples in the original feature space to a high-dimensional one, then seeks the partition hyperplane with the “maximum interval” (Zheng et al., 2015). The widely used kernels are linear, Gaussian (also known as the radial basis), polynomial, Sigmoid, and Laplace kernels. However, the classification result is influenced by the selection of the kernel function to some extent, when we do not know the form of the feature mapping. In this study, a Gaussian kernel was used, and its parameter γ determines how the data distribute in a new feature space. Furthermore, a penalty factor C is introduced to punish the samples, which are misclassified, in order to prevent overfitting, so that the optimization is more likely to achieve the desired goal. A larger value of C always achieves a better classification result on the training set, but the generated model has the risk of overfitting, which leads to a decline in the generalization

Table 3
The spectral indices used in this study.

Spectral index	Formula adapted to Sentinel-2	Related vegetation biochemical and structural properties	References
Visible-NIR			
Normalized difference vegetation index (NDVI)	(B8 - B4)/(B8 + B4)	Vegetation status, canopy structure	Rouse et al. (1974)
Perpendicular moisture index (PMI)	$G - \frac{1}{\sqrt{M^2}}(B8 + M * B4)$	Vegetation status, water content	Zhang et al. (2018)
Visible-SWIR			
Normalized differential senescent vegetation index (NDSVI)	(B11 - B4)/(B11 + B4)	Vegetation status, water content, residue cover	Qi et al. (2002)
Normalized difference residue index (NDRI)	(B4 - B12)/(B4 + B12)	Vegetation status, water content, residue cover	Gelder et al. (2009)
SWIR-SWIR			
Normalized differential tillage index (NDTI)	(B11 - B12)/(B11 + B12)	Non-photosynthetic components, residue cover	Van Deventer et al. (1997)

performance. Hence, the appropriate selection of C and γ is vital to achieve a good classification performance. The grid search method together with 10-fold cross-validation are widely used to determine the optimal combination values of C and γ. In this study, the obtained optimal values of C and γ were 64 and 0.5, used to train the model.

Random forest is a combinatorial classification algorithm of ensemble learning, which has the advantages of high accuracy, good robustness, and ease of use (Breiman, 2001). The core idea of random forest is to combine a set of base classifiers to obtain a classifier with a significantly superior classification performance. Each of the base classifiers can be an independent decision tree. The randomness of random forest is reflected in two key steps (Breiman, 2001). One is the random selection of samples, where an independent sampling method is adopted to carry out the random sampling with replacement. The other is the random selection of features. Some features are randomly selected from all the features as alternative features at each classification node, and then the optimal partitioning attribute is selected to construct each decision tree. The introduction of these two random processes ensures that the random forest avoids overfitting and has strong anti-noise capabilities (Breiman, 2001). Therefore, two parameters need to be set when using random forest. One is the number of base classifiers (trees), and the other is the number of features that need to be randomly selected at each node. It has been shown that the classification error converges with the increase of the trees, and 100 trees can usually obtain an accurate result (Tatsumi et al., 2015). Therefore, in this study, the number of trees was set to 100, and the number of random features was set as the square root of the number of total features.

2.4.3. Feature importance calculation

Three different methodologies were used to calculate the feature importance for these three classification algorithms. For the CART decision tree, the Gini index (Breiman et al., 1984) was adopted as the measurement for the selection of the optimal classification features in the decision tree. Specifically, the Gini value, which represents the purity of data set D, is calculated as follows:

$$Gini(D) = \sum_{k=1}^{|y|} \sum_{k' \neq k} p_k p_{k'} = 1 - \sum_{k=1}^{|y|} p_k^2 \quad (1)$$

In this formula, p_k represents the probability of an arbitrarily selected sample belonging to the kth class. Gini(D) expresses the probability of label inconsistency when we randomly select two samples from data set D. Therefore, the smaller the value of Gini(D), the higher the purity of data set D. Furthermore, if we consider that the number of samples at different branch nodes is different, we multiply the Gini value with a weight of $|D^V|/|D|$. The Gini index of attribute a can then be obtained:

$$Gini\_index(D, a) = \sum_{v=1}^{V} \frac{|D^V|}{|D|} Gini(D^v) \quad (2)$$

Thus, we select the attribute which has the smallest Gini index as the optimal partition attribute, i.e., $a_* = \underset{a \in A}{\arg\min}\ Gini\_index(D, a)$. The relative importance of the features can then be reflected by the order and frequency of the attribute being selected as the node (Gallego et al., 2014). However, a decision tree always contains complex hierarchical structures and numerous nodes. A detailed analysis of each branch node of the decision tree will limit the results to a specific study area, which is not conducive to understanding the discriminative information for different crops. Therefore, the frequency of the occurrence for the nodes is analyzed from both the spectral and temporal feature perspectives, in order to reflect the discriminative information between the different crops.

For SVM, the mathematical expression of the classification model is $g(x) = < w * x > + b$, where $w = \sum_{i=1}^{m} y^i \alpha^i x^i$ is the weight vector, m is the number of classes, x is the samples, α^i corresponds to the Lagrangian multiplier of the corresponding sample, and x^i is the support vector, while α^i is not zero. In this way, we can obtain the weight $w_j, j = 1, 2, ...,n$ of each feature from the SVM classification model where n is the number of features. We can take the weight as the importance measure of the j-th feature of the sample. However, the traditional SVM can only solve binary classification tasks, whereas the LIBSVM toolbox (Chang and Lin, 2011) we used uses a one-against-one approach to solve the classification of multiple classes. Therefore, LIBSVM carries out multiple binary classifications for a multi-class problem. The classifier then generates multiple groups of w. We express this as $w^k = [w_1^k, w_2^k, ...,w_n^k]$, $k = 1, 2, ...,l(l + 1)/2$, where l represents the number of classes and n represents the number of features. The number of one-to-one classifiers is then $l(l + 1)/2$. Finally, the weight w of all the features can be expressed as the following formula, in which $w_j, j = 1, 2, ...,n$is the sum of absolute values of w_j^k, $k = 1, 2, ...,l(l + 1)/2$, and the importance of the j-th feature is $c_j = w_j^2$, which is used as the criterion of feature importance.

$$w = [\sum_{k=1}^{l(l+1)/2} |w_1^k|, \sum_{k=1}^{l(l+1)/2} |w_2^k|, ..., \sum_{k=1}^{l(l+1)/2} |w_n^k|] = [w_1, w_2, ...,w_n] \quad (3)$$

Furthermore, in order to present the effectiveness of the different temporal and spectral features more intuitively, an adding operation was carried out on the importance of each feature. The importance of each feature were added up according to their spectral and temporal attributes, i.e. date or feature type, respectively. For example, the importance of the twelve spectral bands and five index images for each phase were added up as the importance of that image date. In the same way, the importance of eight phases for each spectral band and index were added up as a measure of importance of that feature type.

As for random forest, the change of prediction error, which is calculated by the out-of-bag data, is taken as the measure of feature importance. For a given data set D with m samples, we now resample it to generate a new data set D'. In detail, we first randomly pick a sample from D and copy it into D'. We then put it back into the initial data set D, so that the sample can still be picked next time. After this process is repeated m times, we obtain data set D' containing m samples. Clearly,

some samples in D will be picked multiple times, while others will not be picked. If the probability of a sample never being picked in m sampling processes is $(1-\frac{1}{m})^m$, then the percentage of the samples in the initial data set D that do not appear in the sampled data set D' is the limitation of $(1-\frac{1}{m})^m$, i.e., 36.8%. Thus, we can use m training samples to train the model, and we still have about one-third of the total samples that are not in the training set to evaluate the model. This part of the samples is called the out-of-bag data. In each decision tree of the random forest, we can randomly disturb the value of the feature for each sample in the out-of-bag data, and the change in the prediction error caused by this change is taken as the measure of feature importance. The overall mean of this metric for each tree is then calculated, which is then finally divided by the overall standard deviation to obtain the coefficient of variation. This can be used to compare the discreteness of groups with significantly different mean values, and can eliminate the effect of different units and/or averages on the degree of variability of two or more groups. Meanwhile, an adding operation was again executed on the spectral bands by dates.

However, observations of each methodology may differ to some extent, so in order to evaluate the correctness of the unanimous conclusions drawn by these three methodologies, all 2,766,108 labeled pixels were also analyzed based on just their pixel values. Specifically, the standard deviation of the average value of the eight land cover types was taken to reflect the degree of differentiation. The larger the indicator, the easier it is to distinguish between different types of cropland. Furthermore, the image dates reflected by the different features can vary, making it impossible to select visually the best identification date. Therefore, the normalization operation shown in formula (4) was used to change the absolute value into a relative value. After normalizing the values of each feature, the values corresponding to each date and each feature were also added up to present the effectiveness of the different temporal and spectral features more intuitively.

$$x_i' = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}},\ i = 1, 2, \ldots, 8 \tag{4}$$

3. Results

3.1. Comparison of the different classification methods

In order to verify the effectiveness of the different classification methods, the same training samples and test samples were used to implement the decision tree, SVM, and random forest classifiers. The parameters of each method were set and the classification results of each method are shown in Fig. 5.

It can be seen that random forest achieves the highest overall classification accuracy of 97.85%. SVM ranks second, lower than random forest by 0.63%. The least effective performer is the decision tree classifier, with an overall classification accuracy of 95.92%, which is lower than SVM by 1.3%. In terms of method efficiency, SVM takes the longest time as multiple grid search operations are needed to determine the optimal combination of parameters. Random forest has the second-highest efficiency, on the premise of setting 100 trees. The decision tree classifier takes the shortest time as it only consists of one tree. Some additional operations are also needed to evaluate the importance of the features for random forest and SVM. However, once the decision tree is completed, the importance of the features can be visualized. Therefore, if the classification mechanism is needed to be understood thoroughly and the requirement for accuracy is not very high, the decision tree classifier is a good choice (Peña-Barragán et al., 2011). Otherwise, random forest is recommended for classification under most cases (Teluguntla et al., 2018).

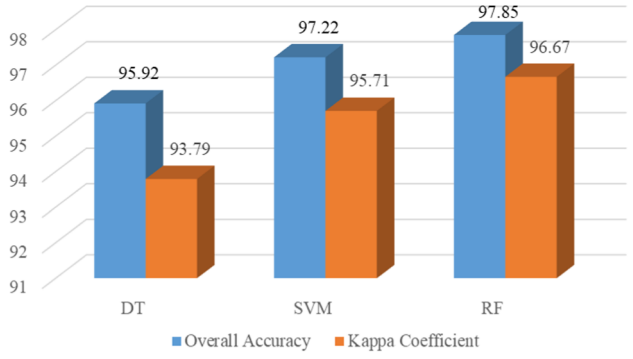

Fig. 5. Classification results of the different classification methods.

3.2. Classification results

The classification results presented in this section were obtained via the random forest classifier, i.e., the most accurate classifier. Fig. 6 intuitively presents the spatial distribution patterns and proportions of the cultivated areas of the different crops in the study area.

In detail, corn has the largest planting area in the study area, with a uniform and continuous distribution and a relatively large plot area. The planting area of rice ranks second, and its spatial distribution is relatively concentrated, along both sides of the river and around the water sources, reflecting the dependence of rice planting on water resources. The crop with the third-largest planting area is soybean, whose spatial distribution is discrete. Except for a few areas where the soybeans are planted in blocks, the other soybean plots all present a long and narrow rectangular shape along the north–south direction, and most are intercropped with corn. Beet and potato present a blocky distribution, and the planting area of beet is concentrated in the middle part of the area. The reason for this is that the soil in the middle part is close to the river, so it has high fertility. A strong water and fertilizer retention capacity and suitable irrigation conditions can result in beet having a higher sugar content and yield. The distribution of potatoes is concentrated in the middle of the northern part of the study area. In addition to the above five crops, there are three reservoirs in the area and a river running through the center from east to west, which together allow the irrigation of the crops. The distribution of the residential land presents a spot-like distribution, and most of the villages extend in an east–west direction. Since the terrain in the study area is flat and the soil is fertile, it is suitable for the cultivation of crops. As a result, the proportion of forest is very small, and is almost invisible in Fig. 6.

Furthermore, in order to show the intercropping phenomenon of corn and soybeans more clearly, we have enlarged the upper part of Fig. 6, indicated by the red rectangle, and the classification map is shown in Fig. 7. We can see that the classification result can effectively depict corn and soybean plots with only a few pixels' width.

A quantitative evaluation was also carried out for the classification results and the confusion matrix of the model on the test data is shown in Table 4. The first column represents the classified categories, and the first row represents the true attribution of the pixels. Thus, the diagonal indicates the number of correct sample pixels while the others are misclassified pixels. User accuracy and mapping accuracy are located in the last column and last row. The overall accuracy appears in the lower right corner.

This table shows that the classification accuracies for rice, water, and buildings are all higher than 99%. This is due to the discriminant properties of these three types, including the water information contained in rice and water and because the reflectance properties of buildings differing from that of vegetation. This allows them to be clearly distinguished from the other types, A bit of confusion among the four summer crops of corn, soybean, potato, and beet, since these four crops are all dryland crops and their growth times are relatively close. In detail, the confusion between corn and soybean is the most serious, and can be attributed to two reasons. On the one hand, the

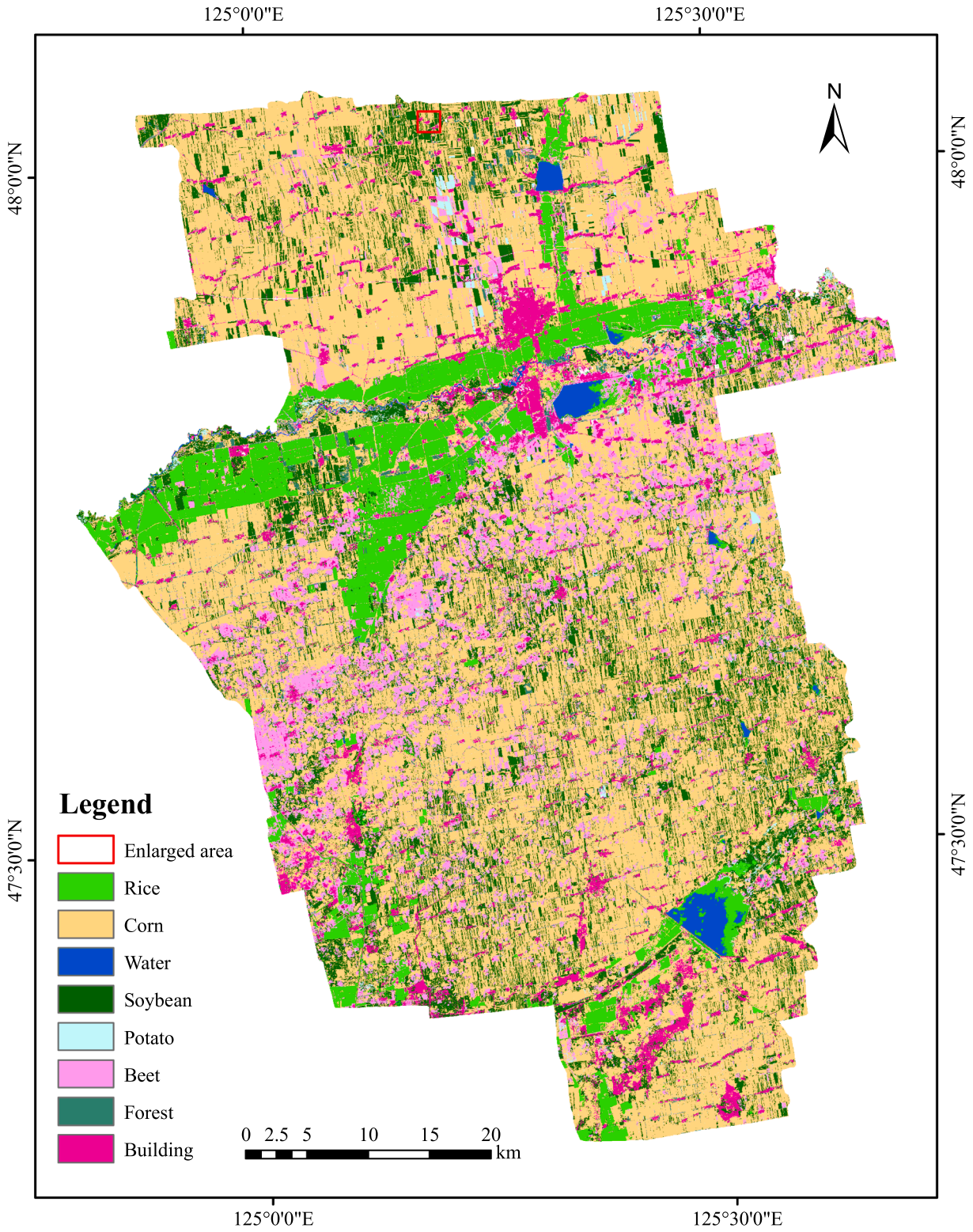

Fig. 6. Classification results of the random forest classifier.

intercropping phenomenon between corn and soybean is so common in our study area that they are nearly always confused. On the other hand, the phenological calendars of corn and soybean are very similar, so that their spectral characteristics during the growth period are also similar, making it difficult to distinguish them. The classification accuracy for potato and beet is about 5% lower than corn and soybean. When combined with the field survey analysis, we know that these two crops are minor crop types, compared to corn and soybean. Meanwhile, the proportion of the samples in the process of sample labeling is also consistent with the proportions of the crop field distribution, to some extent, leading to fewer training samples for these two types than the other major crops. A strategy of uniform sampling could solve this problem. However, the classification accuracies for the minor crops are higher than 90%, while the major crops have an even better classification accuracy, which can meet the requirements of most mapping tasks. Furthermore, the classification accuracy for forest is the worst, at

Fig. 7. Comparison of an area of zoomed classification results. (a) Labeled references. (b) Classification result.

only about 85.94%. This can be attributed to the real distribution of the land cover. The small coverage of forest in this large agricultural county results in an uneven distribution of samples. However, the obvious seasonal characteristics of forest can be used as a distinguishing feature. As we mainly focused on distinguishing different types of crops in this study, this was ignored. If necessary, the discrimination could be further improved by adding an image from a non-crop growing season. Furthermore, the mapping accuracy of these land covers in this study is consistent with the user's accuracy, so that the Kappa coefficient is as high as 0.9667.

4. Analysis of feature importance

In this section, we present the different results obtained by the three machine learning methods, with regard to the relative importance of all 136 features. Subsequently, we will answer our research questions based on the analysis from the spectral and temporal feature perspectives. Finally, the common conclusions drawn from the analysis are verified based on statistical methodology.

4.1. Importance of all the features

On the premise of using the minimum Gini index as the optimal partition attribute, the decision tree constructed in our study is shown in Fig. 8. Each node is represented by a rectangular box in the figure, where the upper row represents the image date while the lower row represents the feature. Meanwhile, the decision rules are represented by the connecting lines with corresponding values. Each category is represented by an ellipse of different color, and the subscript number indicates the number of subsets of that class. In addition, three border types are used to represent the different types of features, according to their different attributes.

It can be clearly seen that the SWIR features take up almost half of the nodes, i.e. 16 of 38, which intuitively shows the importance of SWIR. Furthermore, the topmost branch is mainly used to distinguish water from forest, while the others are mainly used to distinguish the different crop types. It can also be seen that the subscript numbers for water, forest, building, and rice are all no greater than 3, while the subscript numbers for corn, soybean, beet, and potato are much greater. This hints that the distinction between the different crop types is more complicated than that between crops and non-crops.

Fig. 9 presents the SVM-based feature importance result. The horizontal and vertical axes represent the number of features and the importance of that feature. Numbers 1 to 12 are the 12 spectral bands on May 9th, and the number of the features is consistent with the order of the bands that we described previously. Similarly, numbers 13 to 24, 25 to 36, 37 to 48, 49 to 60, 61 to 72, 73 to 84, and 85 to 96 correspond to the 12 spectral bands on May 19th, June 23rd, July 13th, and July 23rd, September 1st, September 11th, and September 26th, respectively. Numbers 97 to 104, 105 to 112, 113 to 120, 121 to 128, and 129 to 136 correspond to NDRI, NDSVI, NDTI, NDVI, and PMI for the eight temporal phases.

Fig. 9 reveals that the use of the spectral indices obtains a significantly better effect than the original spectral bands, especially for

Table 4
Confusion matrix for the classification result.

	Rice	Corn	Water	Soybean	Potato	Beet	Forest	Building	UA/%
Rice	**595,898**	1908	0	1077	8	778	206	337	99.28
Corn	466	**1,485,305**	0	8555	90	2335	1452	95	99.13
Water	93	127	**12,126**	13	0	173	2	177	95.40
Soybean	220	16,119	34	**343,759**	256	2477	12	65	94.71
Potato	0	5206	0	3426	**41,743**	2511	0	5	78.92
Beet	239	4428	0	917	2906	**98,377**	1	258	91.83
Forest	58	1171	0	153	0	2	**10,222**	0	88.08
Building	45	399	49	361	2	115	0	**116,585**	99.17
MA/%	99.81	98.06	99.32	95.95	92.75	92.14	85.94	99.20	**97.85**

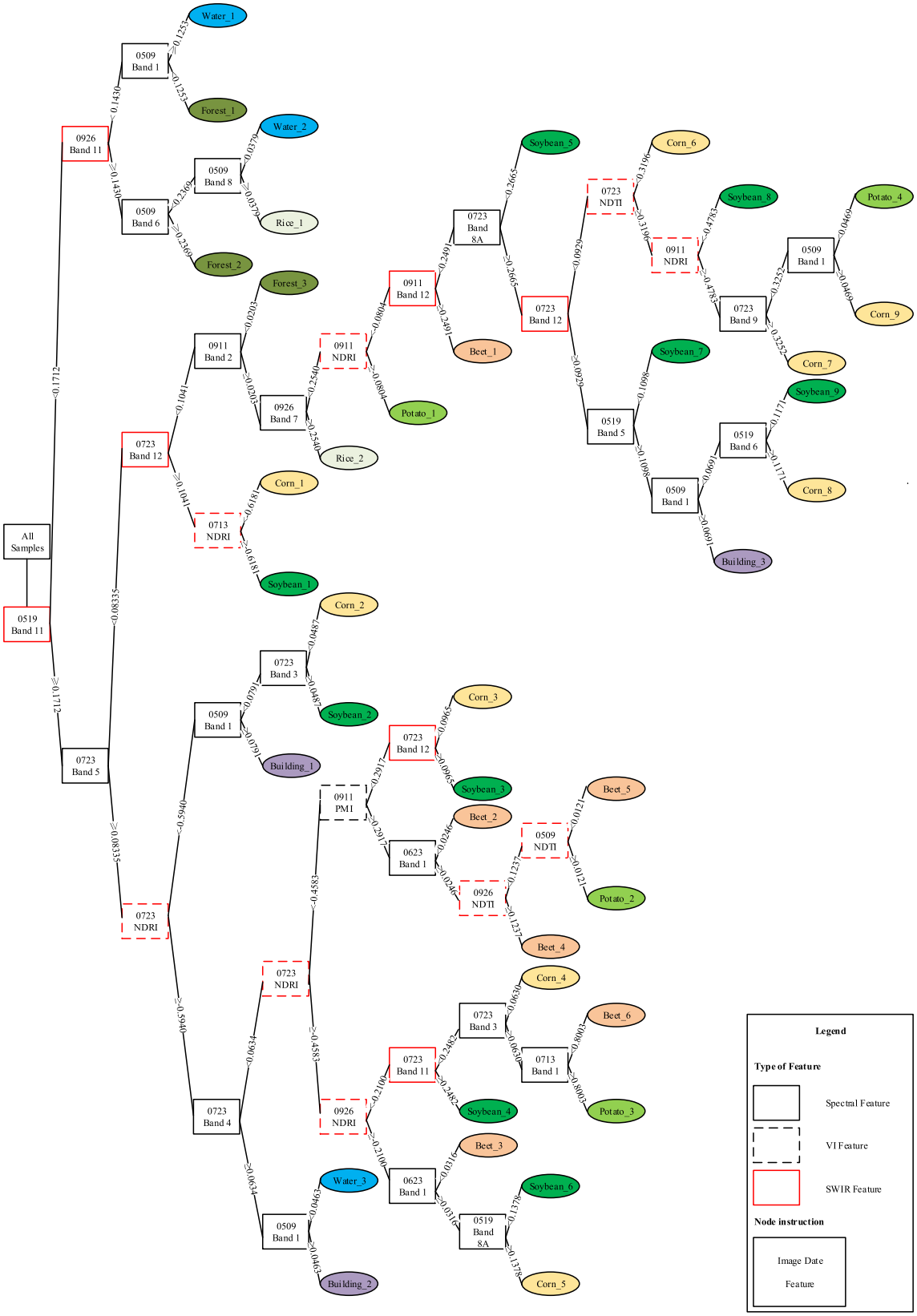

Fig. 8. The decision tree in our study.

NDRI, NDSVI, and NDVI. In addition, features 37 to 72 (July 13th to September 1st) exhibit more importance than those on the other dates.

While calculating the feature importance of random forest, the calculation was repeated 10 times in order to avoid the randomness of the result. Fig. 10 presents the average results. The number of the feature has the same meaning as in SVM, the vertical axis represents the percentage of the decrease in accuracy caused by changing that feature.

It can be seen that all the features have a positive effect on the

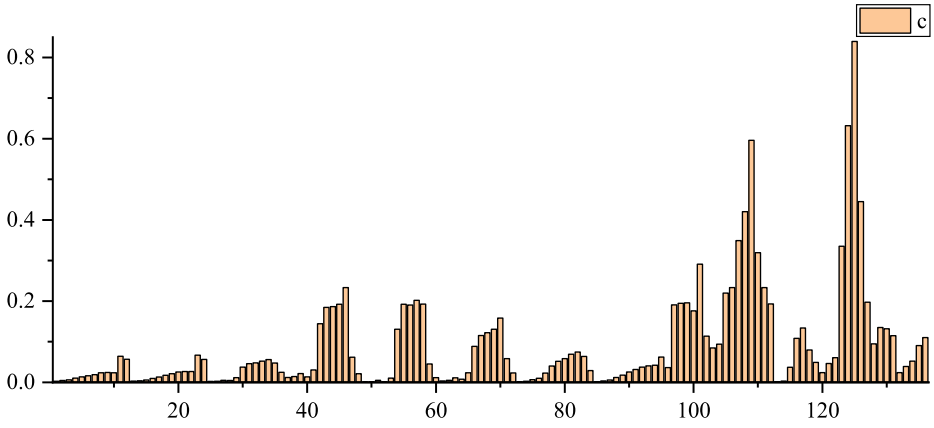

Fig. 9. SVM-based feature importance.

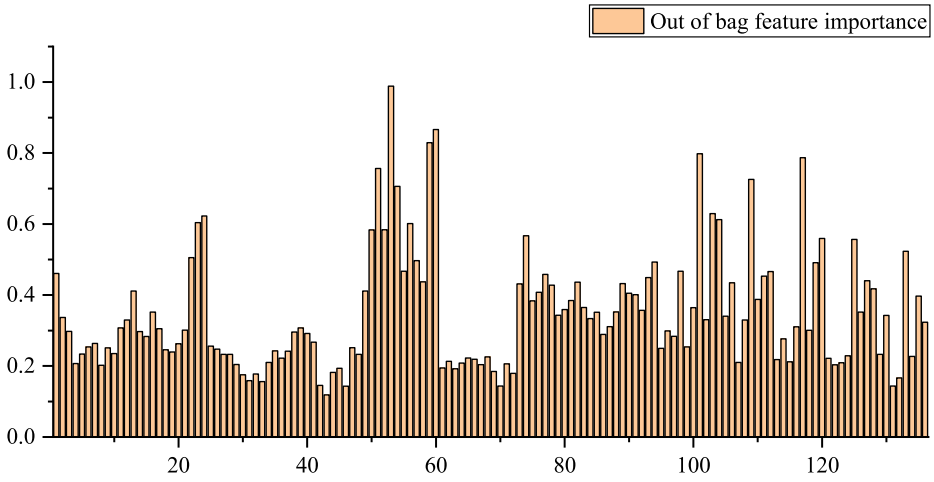

Fig. 10. Feature importance of random forest.

classification accuracy. Features 49 to 60 (July 23rd) have the most significant effect, while features 73 to 136 (spectral features) also have an impact on the accuracy.

4.2. Importance of the temporal features

For the decision tree classifier, Fig. 11 presents the frequency of occurrence of the nodes for the different image dates. The horizontal axe represents the date of images, while the first two numbers represent the month, and the last two numbers represent the day. The vertical axis represents the frequency of occurrence of the nodes.

The image on July 23rd plays a leading role in crop differentiation. This is mainly due to the stems and leaves of the different crops being mature and at the peak stage of growth. Thus, the crops can be distinguished by the different characteristics of the plants. For example, corn is in the jointing stage, and the stems and leaves of the plant are well developed, and the plant height is much higher than that of the other crops. Meanwhile soybean is in the stage of filling seed, and the development of its stem and leaves is also completed, but the plant height is much lower than that of corn. In addition, rice is at the booting stage, beet is in the tuber growing period, and potato is at the end of blooming. All the crops are in different growth stages.

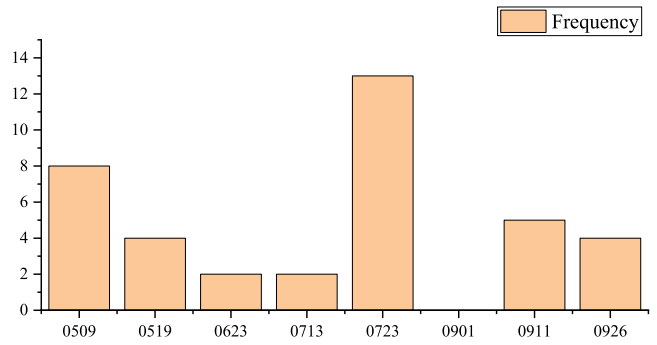

Fig. 11. Frequency of occurrence in the nodes for the different image dates of decision tree.

The frequency of occurrence of images on May 9th was eight and for September 11th it was four. On May 9th, corn, soybean, and potato are in the sowing stage, while rice is at the emergence stage and beet is in the transplanting stage, as shown in Fig. 2. On September 11th, corn, rice, and soybean have matured into the senescence period, while beet and potato are in the period of nutrient accumulation. Last but not least, the images from May 19th, June 23rd, July 13th, and September 26th all contribute to the crop classification. However, the image from September 1st has the worst effect on crop differentiation, compared to the other image dates, so it is not selected as a node in the decision tree.

The added up temporal feature importance results for the different image dates of SVM are shown in Fig. 12. The horizontal axis also represents the date of images, while the vertical axis represents the square of the weight vector.

Not surprisingly, the image from July 23rd plays a leading role in the crop differentiation, the reason for which has been explained above. However, the images from July 13th and September 1st present a more important effect than the other image dates, which is quite different from the result of the decision tree classifier. This difference may have

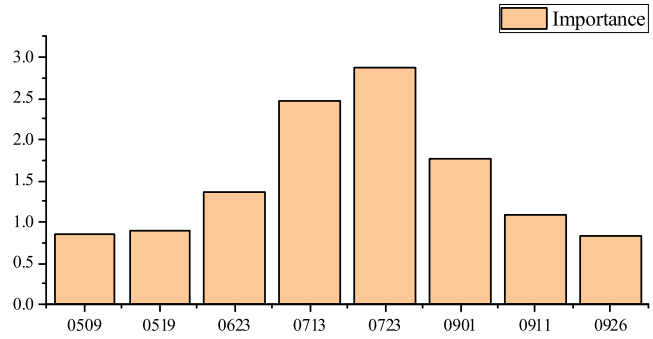

Fig. 12. Cumulative importance of the different image dates in SVM.

been caused by their different mechanisms of model construction, since the model of SVM is only construct by the support vectors, rather than all the samples that were used in the decision tree classifier. Furthermore, the images from May 9th, May 19th, June 23rd, September 11th, and September 26th are also beneficial to the crop type mapping, but are less important.

The added-up results of the temporal feature importance for the different image dates in random forest are shown in Fig. 13. While the horizontal and vertical axes represent the date of images and the average change of prediction error, which is calculated by the out-of-bag data.

Clearly, the image from July 23rd plays a leading role again, which hints at the extraordinary importance of the image from July 23rd. In addition, the images from May 9th, May 19th, September 11th, and September 26th added a certain degree of accuracy, which is similar to the result of the decision tree classifier but differs from that of SVM. Moreover, the remaining images from June 23rd, July 13th, and September 1st also contribute to the crop classification.

4.3. Importance of the spectral features

First of all, from the tree constructed in our study, as shown in Fig. 8, we can see that band 11 on May 19th is effective for distinguishing between water bodies, forest, and other crops. The frequency of the occurrence in the nodes was analyzed according to their spectral and temporal attributes for the decision tree classifier. The results are shown in Fig. 14.

It can be clearly seen that band 1 and NDRI are of great significance for crop type mapping, with the frequency of occurrence being eight and seven. Band 11, band 12, and NDTI also contribute a lot. In addition, band 5, band 6, and band 8A also contribute. However, for the artificially constructed features of NDVI and NDSVI, the influence is not so obvious in the construction of the decision tree. Furthermore, the frequency of occurrence of NDTI, NDRI, band 11, and band 12 indicates that the short-wave infrared bands are very important in fine crop classification. In the same way, from the frequency of occurrence of band 5, band 6, and band 7, we can see that the red edge bands play more important roles in crop type mapping than the traditional red, green, and blue bands.

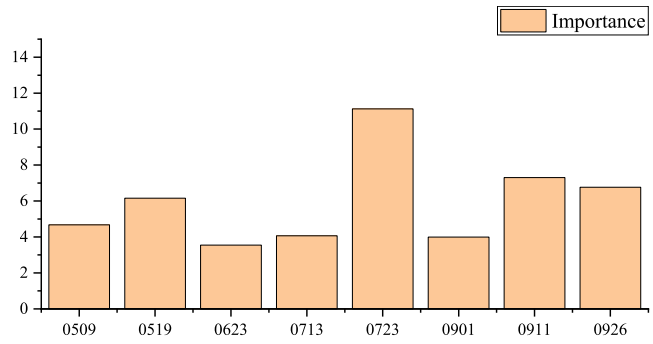

Fig. 13. Cumulative importance of the different image dates in random forest.

Fig. 15 presents the added-up results for the importance of the different spectral features in SVM. The horizontal axe represents different types of feature, while the vertical axis represents cumulative results.

It can be seen that NDRI, NDSVI, and NDVI are more important than the spectral bands, while NDTI and PMI perform well in the classification. This hints that reasonable spectral indices, which are artificially constructed based on expert knowledge, are beneficial to the differentiation of crops. Furthermore, bands 6 to 9, band 11, and band 12 are indispensable in our study, while band 1 to 5 contribute little. Overall, the red edge, near-infrared, and short-wave infrared bands are extremely useful in fine crop classification. Among the different bands, the short-wave infrared bands are the most important, followed by the near-infrared bands. The red edge bands also present a better performance than the traditional red, green, and blue bands in the cumulative importance of the different spectral features of SVM.

Last but not least, the added-up results of the spectral feature importance for random forest are shown in Fig. 16. While the horizontal and vertical axes represent the type of features and the average change of prediction error, which is calculated by the out-of-bag data.

As in the other results, NDRI, NDSVI, and NDTI are more important than the original spectral bands, while NDVI and PMI also perform well in the accurate discrimination of multiple crops. This illustrates that artificially constructing a rational feature space is indeed beneficial to the classification of crops. Moreover, band 5 is rather important for crop differentiation, followed by band 11 and band 12, which further proves that the short-wave infrared bands and the red edge bands are more important than band 2, band 3, and band 4.

4.4. Summary

Although the conclusions with regard to the feature importance for these three methodologies may differ, to some extent, some unanimous conclusions can still be drawn. That is, all the analysis results show that a rational feature space constructed by expert knowledge is important. In detail: 1) July 23rd is the best identification date that can capture the crucial phenological stages efficiently. This is mainly due to the stems and leaves of the different crops being at the peak of growth at that time, as shown in Fig. 2. 2) The short-wave infrared bands are extremely important in fine crop classification since these bands can better characterize the water content and residue cover of the different crop types. 3) The red edge bands of Sentinel-2 play an important role in accurate crop identification, and are usually more important than the traditional red, green, and blue bands. This is mainly due to the red edge bands being better at capturing the biochemical information of vegetation.

4.5. Verification using statistical methods

4.5.1. Verification of the temporal feature importance

The specific analysis results for the 12 spectral bands and the 5 spectral indices are shown in Fig S-1 and Fig S-2 of the supplementary material and the added-up results are shown in Fig. 17.

At this point, it can be seen that the image from July 23rd does play a key role in the classification. In addition, the images from July 13th and September 1st are also extremely important. It can be concluded that images of the vigorous crop growth period are conducive to fine crop classification. Furthermore, it is worth noting that, at an early stage of crop growth, i.e., around May 19th, the different crops can also be differentiated, to some extent. The above analysis of the importance of the image date on the classification result could be used to guide the selection of the image date in classification, so that the distribution patterns and cultivated area information of the different crop types

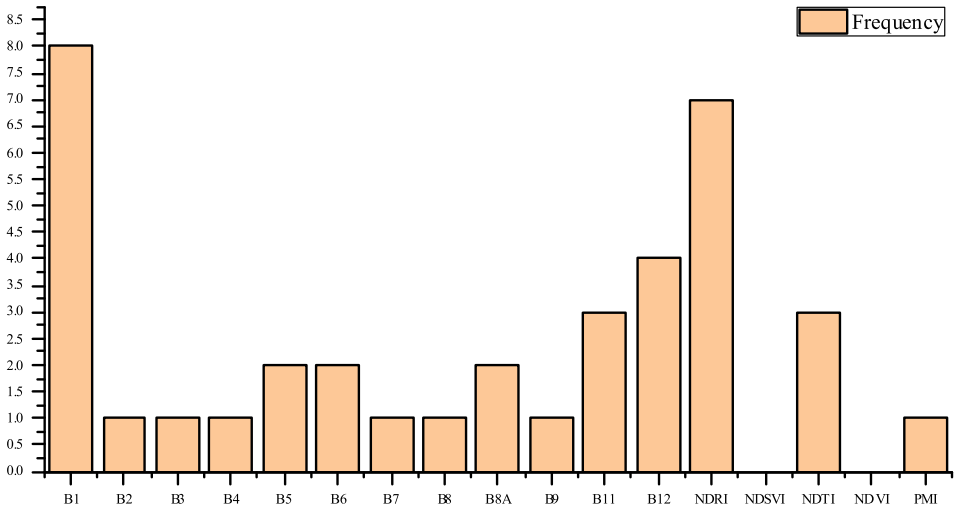

Fig. 14. Frequency of occurrence in the nodes for the different spectral features.

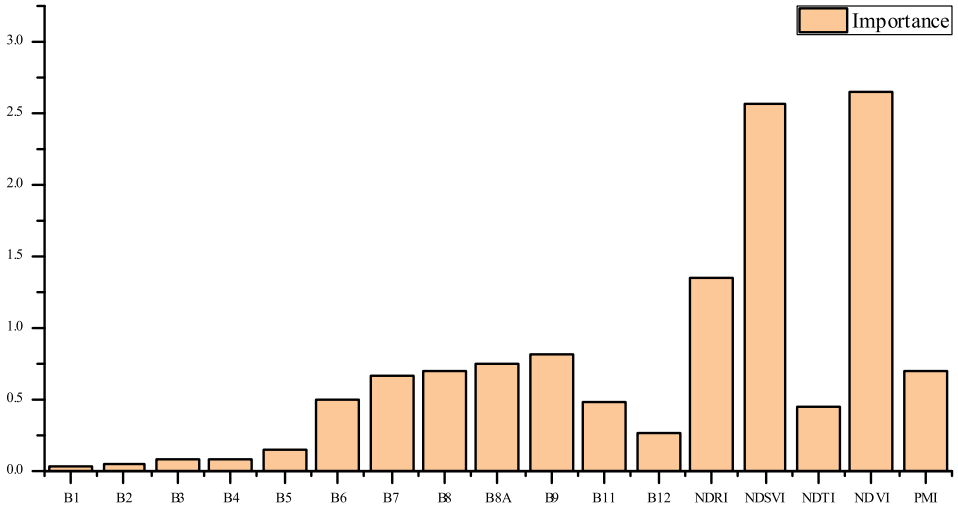

Fig. 15. Cumulative importance of the different spectral features in SVM.

could be obtained at an early stage of crop harvesting, so that the yield could be estimated. In other words, these results indicate that, for the summer crops in our study area, the first estimate could be made at the end of May. A second estimate could then be made at the end of July, which would be more accurate.

4.5.2. *Verification of the spectral feature importance*

The standard deviations corresponding to the different features of

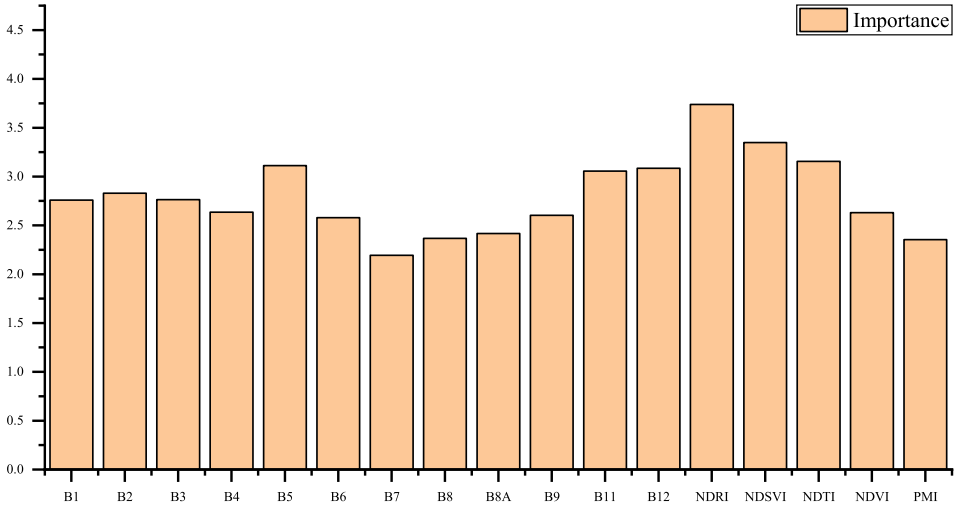

Fig. 16. Cumulative importance of the different spectral features in random forest.

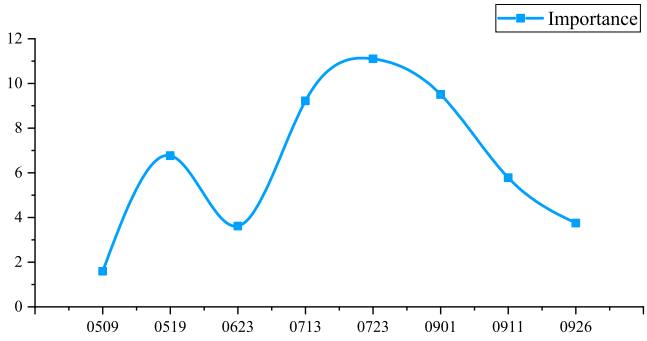

Fig. 17. Importance of the different image dates.

each date are shown in Fig S-3 of the supplementary materials. Normalization and accumulation operations were again performed. The spectral feature importance is shown in Fig. 18.

Clearly, the importance of the artificially constructed spectral indices is significantly higher than that of the spectral bands, with the NDRI, NDSVI, and NDVI indices being the most significant. For the spectral bands, the near-infrared bands (band 8, band 8A, band 9) play important roles in the classification, followed by the short-wave infrared bands (band 11, band 12). Since NDRI, NDSVI, and NDTI all use the short-wave infrared bands, this indicates that the short-wave infrared bands are especially important for crop type mapping. Moreover, the red edge bands (band 6, band 7) of the Sentinel-2 data also assist in the fine classification of crops. This is a unique advantage of the Sentinel-2 data, compared to MODIS and Landsat.

5. Conclusions

In this paper, we first compared the classification results of the three mainstream machine learning classification methods, using a feature space which contained eight temporal phases, with 12 spectral bands and 5 spectral indices for each phase. The spectral indices were artificially constructed, combining the expert knowledge in crop type mapping. Furthermore, the relative importance of the features used in the classification process was analyzed for each classification method, and a statistical analysis was carried out. From the experimental results, we draw the following conclusions, which provide guidance for the optimal temporal window selection and spectral feature space construction for the future crop type mapping application using multi-temporal Sentinel-2 images.

1) Random forest achieved the highest classification accuracy, while SVM came second, and the decision tree algorithm had the poorest classification accuracy. In the case of considering the efficiency of the algorithm, we recommended that the random forest method is used for classification.
2) Images from the vigorous growth period are conducive to the fine classification of crops. A first early season estimate of crop area information could be made at the end of May for summer crops. A second estimate could be made at the end of July.
3) The experiments also showed that a rational feature space can significantly improve the classification accuracy. Specifically, the short-wave infrared bands are extremely important in the fine classification of crops, followed by the near-infrared bands.
4) The unique red edge bands of Sentinel-2 can effectively assist in the identification of different crop types, which is a unique advantage of Sentinel-2 data over MODIS and Landsat.
5) The classification results obtained with the Sentinel-2 data of a 10-m spatial resolution could reflect the phenomena of the fragile farmland in China, and could accurately reflect the spatial distribution patterns of the different crops and their proportions. Compared to MODIS and Landsat data with a coarser spatial resolution, Sentinel-2 data is more suitable for the fine classification of crops when the plots are fragmented.

CRediT authorship contribution statement

Hongyan Zhang: Conceptualization, Formal analysis, Funding acquisition, Project administration, Resources, Writing - original draft, Writing - review & editing. **Jinzhong Kang:** Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - original draft. **Xiong Xu:** Formal analysis, Project administration, Visualization, Writing - review & editing. **Liangpei Zhang:** Formal analysis, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

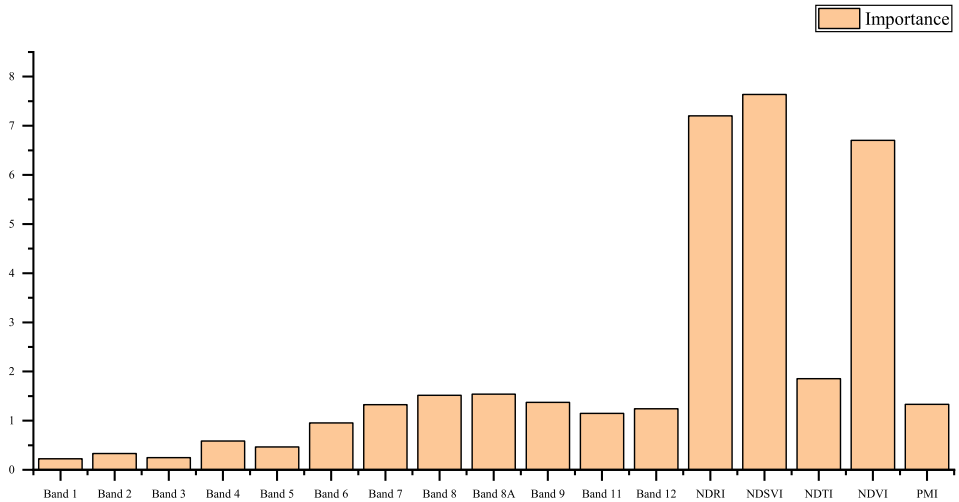

Fig. 18. Importance of the different features.

Acknowledgement

This work was partly supported by the National key research and development program of China under Grant 2018YFB0505400, the National Natural Science Foundation of China under Grant 61871298, 41971299, the Fundamental Research Funds for the Central Universities and Dragon 5 proposal ID. 58817.

References

Asgarian, A., Soffianian, A., Pourmanafi, S., 2016. 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. 127, 531–540.

Awokuse, T.O., 2009. Does agriculture really matter for economic growth in developing countries? Technical Report.

Bolton, D.K., Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agric. For. Meteorol. 173, 74–84.

Breiman, L., Friedman, J., Olshen, R.A., Stone, C.J., 1984. Classification and regression trees. Chapman and Hall/CRC, Boca Raton, FL.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Carrão, H., Gonçalves, P., Caetano, M., 2008. Contribution of multispectral and multi-temporal information from MODIS images to land cover classification. Remote Sens. Environ. 112, 986–997.

Chang, C.C., Lin, C.J., 2011. Libsvm: A library for support vector machines. ACM Trans. Intel. Syst. Technol. (TIST) 2, 27.

Dong, J., Xiao, X., Menarguez, M.A., Zhang, G., Qin, Y., Thau, D., Biradar, C., Moore III, B., 2016. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and google earth engine. Remote Sens. Environ. 185, 142–154.

Foerster, S., Kaden, K., Foerster, M., Itzerott, S., 2012. Crop type mapping using spectral–temporal profiles and phenological information. Comput. Electron. Agric. 89, 30–40.

Gallego, F.J., Kussul, N., Skakun, S., Kravchenko, O., Shelestov, A., Kussul, O., 2014. Efficiency assessment of using satellite data for crop area estimation in Ukraine. Int. J. Appl. Earth Obs. Geoinf. 29, 22–30.

Gelder, B., Kaleita, A., Cruse, R., 2009. Estimating mean field residue cover on mid-western soils using satellite imagery. Agron. J. 101, 635–643.

Gilbertson, J.K., Kemp, J., Van Niekerk, A., 2017. Effect of pan-sharpening multi-temporal Landsat 8 imagery for crop type differentiation using different classification techniques. Comput. Electron. Agric. 134, 151–159.

Hu, Q., Wu, W., Song, Q., Lu, M., Chen, D., Yu, Q., Tang, H., 2017. How do temporal and spectral features matter in crop classification in Heilongjiang province, china? J. Integr. Agric. 16, 324–336.

Huang, J., Sedano, F., Huang, Y., Ma, H., Li, X., Liang, S., Tian, L., Zhang, X., Fan, J., Wu, W., 2016. Assimilating a synthetic Kalman filter leaf area index series into the WOFOST model to estimate regional winter wheat yield. Agric. For. Meteorol. 216, 188–202.

Kang, J., Zhang, H., Yang, H., Zhang, L., 2018. Support vector machine classification of crop lands using sentinel-2 imagery, in: 2018 7th International Conference on Agro-geoinformatics (Agro-geoinformatics), IEEE. pp. 1–6.

King, L., Adusei, B., Stehman, S.V., Potapov, P.V., Song, X.P., Krylov, A., Di Bella, C., Loveland, T.R., Johnson, D.M., Hansen, M.C., 2017. A multi-resolution approach to national-scale cultivated area estimation of soybean. Remote Sens. Environ. 195, 13–29.

Lebourgeois, V., Dupuy, S., Vintrou, É., Ameline, M., Butler, S., Bégué, A., 2017. A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM). Rem. Sens. 9, 259.

Löw, F., Michel, U., Dech, S., Conrad, C., 2013. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines. ISPRS J. Photogramm. Remote Sens. 85, 102–119.

Lu, M., Wu, W., Zhang, L., Liao, A., Peng, S., Tang, H., 2016. A comparative analysis of five global cropland datasets in china. Science China Earth Sci. 59, 2307–2317.

Lupia, F., Antoniou, V., 2018. Copernicus sentinels missions and crowdsourcing as game changers for geospatial information in agriculture. GEOmedia 22.

Massey, R., Sankey, T.T., Congalton, R.G., Yadav, K., Thenkabail, P.S., Ozdogan, M., Meador, A.J.S., 2017. Modis phenology-derived, multiyear distribution of conterminous us crop types. Rem. Sens. Environ. 198, 490–503.

Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E., Six, J., 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. Remote Sens. Environ. 115, 1301–1316.

Piiroinen, R., Heiskanen, J., Mõttus, M., Pellikka, P., 2015. Classification of crops across heterogeneous agricultural landscape in Kenya using AisaEAGLE imaging spectroscopy data. Int. J. Appl. Earth Obs. Geoinf. 39, 1–8.

Qi, J., Marsett, R., Heilman, P., Bieden-bender, S., Moran, S., Goodrich, D., Weltz, M., 2002. Ranges improves satellite-based information and land cover assessments in southwest united states. Eos, Trans. Am. Geophys. Union 83, 601–606.

Quinlan, J.R., 1993. C4. 5: programs for machine learning. Morgan Kaufmann, San Mateo, CA.

Rouse Jr, J., Haas, R., Schell, J., Deering, D., 1974. Monitoring vegetation systems in the great plains with ERTS.

Samberg, L.H., Gerber, J.S., Ramankutty, N., Herrero, M., West, P.C., 2016. Subnational distribution of average farm size and smallholder contributions to global food production. Environ. Res. Lett. 11, 124010.

Song, Q., Zhou, Q.b., WU, W.b., Hu, Q., Miao, L., LIU, S.b., 2017a. Mapping regional cropping patterns by using gf-1 WFV sensor data. Journal of integrative agriculture 16, 337–347.

Song, X.P., Potapov, P.V., Krylov, A., King, L., Di Bella, C.M., Hudson, A., Khan, A., Adusei, B., Stehman, S.V., Hansen, M.C., 2017. National scale soybean mapping and area estimation in the united states using medium resolution satellite imagery and field survey. Remote Sens. Environ. 190, 383–395.

Tatsumi, K., Yamashiki, Y., Torres, M.A.C., Taipe, C.L.R., 2015. Crop classification of upland fields using random forest of time-series Landsat 7 ETM + data. Comput. Electron. Agric. 115, 171–179.

Teluguntla, P., Thenkabail, P., Oliphant, A., Xiong, J., Gumma, M.K., Congalton, R.G., Yadav, K., Huete, A., 2018. A 30-m Landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. ISPRS J. Photogramm. Remote Sens. 144, 325–340.

Van Deventer, A., Ward, A., Gowda, P., Lyon, J., 1997. Using thematic mapper data to identify contrasting soil plains and tillage practices. Photogramm. Eng. Remote Sens. 63, 87–93.

van der Velde, M., Biavetti, I., El-Aydam, M., Niemeyer, S., Santini, F., van den Berg, M., 2019. Use and relevance of European union crop monitoring and yield forecasts. Agric. Syst. 168, 224–230.

Vogels, M.F., De Jong, S.M., Sterk, G., Addink, E.A., 2019. Mapping irrigated agriculture in complex landscapes using SPOT6 imagery and object based image analysis–a case study in the Central Rift Valley, Ethiopia. Int. J. Appl. Earth Obs. Geoinf. 75, 118–129.

Wang, J., Xiao, X., Qin, Y., Dong, J., Zhang, G., Kou, W., Jin, C., Zhou, Y., Zhang, Y., 2015. Mapping paddy rice planting area in wheat-rice double-cropped areas through integration of Landsat-8 OLI, MODIS, and PALSAR images. Sci. Rep. 5, 10088.

Wang, S., Azzari, G., Lobell, D.B., 2019. Crop type mapping without field-level labels: random forest transfer and unsupervised clustering techniques. Remote Sens. Environ. 222, 303–317.

Wardlow, B.D., Egbert, S.L., 2008. Large-area crop mapping using time series MODIS 250 m NDVI data: an assessment for the us central great plains. Remote Sens. Environ. 112, 1096–1116.

Wei, M., Qiao, B., Zhao, J., Zuo, X., 2018. Application of remote sensing technology in crop estimation, in: 2018 IEEE 4th International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing,(HPSC) and IEEE International Conference on Intelligent Data and Security (IDS), IEEE. pp. 252–257.

Yang, C., Everitt, J.H., Murden, D., 2011. Evaluating high resolution spot 5 satellite imagery for crop identification. Comput. Electron. Agric. 75, 347–354.

Zhai, H., Zhang, H., Zhang, L., Li, P., 2019a. Total variation regularized collaborative representation clustering with a locally adaptive dictionary for hyperspectral imagery. IEEE Trans. Geosci. Remote Sens. 57, 166–180.

Zhai, H., Zhang, H., Zhang, L., Li, P., 2019b. Laplacian-regularized low-rank subspace clustering for hyperspectral image band selection. IEEE Trans. Geosci. Remote Sens. 57, 1723–1740.

Zhang, C., Zhang, H., Du, J., Zhang, L., 2018. Automated paddy rice extent extraction with time stacks of sentinel data: A case study in Jianghan Plain, Hubei, China, in: 2018 7th International Conference on Agrogeoinformatics (Agro-geoinformatics), IEEE. pp. 1–6.

Zhang, H., Liu, L., He, W., Zhang, L., 2020. Hyperspectral image denoising with total variation regularization and nonlocal low-rank tensor decomposition. IEEE Trans. Geosci. Remote Sens. 58, 3071–3084.

Zhai, H., Zhang, H., Zhang, L., Li, P., 2019. Cloud/Shadow Detection based on Spectral Indices for Multi/Hyperspectral Optical Remote Sensing Imagery. ISPRS Journal of Photogrammetry and Remote Sensing 144, 235–253.

Zhang, J., 2004. Risk assessment of drought disaster in the maize-growing region of songliao plain, china. Agric. Ecosyst. Environ. 102, 133–153.

Zhang, J., Zhou, Q., Shen, X., Li, Y., 2019. Cloud detection in high-resolution remote sensing images using multi-features of ground objects. J. Geovisualization Spatial Anal. 3, 14.

Zheng, B., Myint, S.W., Thenkabail, P.S., Aggarwal, R.M., 2015. A support vector machine to identify irrigated crop types using time-series Landsat NDVI data. Int. J. Appl. Earth Obs. Geoinf. 34, 103–112.

Zhong, L., Gong, P., Biging, G.S., 2014. Efficient corn and soybean mapping with temporal extendability: a multi-year experiment using Landsat imagery. Remote Sens. Environ. 140, 1–13.

Zhong, L., Hu, L., Yu, L., Gong, P., Biging, G.S., 2016. Automated mapping of soybean and corn using phenology. ISPRS J. Photogramm. Remote Sens. 119, 151–164.