



Article

An Object-Based Genetic Programming Approach for Cropland Field Extraction

Caiyun Wen ¹ , Miao Lu ^{1,*}, Ying Bi ² , Shengnan Zhang ¹, Bing Xue ², Mengjie Zhang ², Qingbo Zhou ³ and Wenbin Wu ¹

- ¹ Key Laboratory of Agricultural Remote Sensing, Ministry of Agriculture and Rural Affairs, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China; wencaiyun@caas.cn (C.W.); 82101212293@caas.cn (S.Z.); wuwenbin@caas.cn (W.W.)
- ² School of Engineering and Computer Science, Victoria University of Wellington, Wellington 6140, New Zealand; ying.bi@ecs.vuw.ac.nz (Y.B.); bing.xue@ecs.vuw.ac.nz (B.X.); mengjie.zhang@ecs.vuw.ac.nz (M.Z.)
- ³ Agricultural Information Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China; zhouqingbo@caas.cn
- * Correspondence: lumiao@caas.cn

Abstract: Cropland fields are the basic spatial units for agricultural management, and information about their distribution is critical for analyzing agricultural investments and management. However, the extraction of cropland fields of smallholder farms is a challenging task because of their irregular shapes and diverse spectrum. In this paper, we proposed a new object-based Genetic Programming (GP) approach to extract cropland fields. The proposed approach used the multiresolution segmentation (MRS) method to acquire objects from a very high resolution (VHR) image, and extracted spectral, shape and texture features as inputs for GP. Then GP was used to automatically evolve the optimal classifier to extract cropland fields. The results show that the proposed approach has obtained high accuracy in two areas with different landscape complexities. Further analysis show that the GP approach significantly outperforms five commonly used classifiers, including K-Nearest Neighbor (KNN), Decision Tree (DT), Naïve Bayes (NB), Support Vector Machine (SVM), and Random Forest (RF). By using different numbers of training samples, GP can maintain high accuracy with any volume of samples compared to other classifiers.

Keywords: Genetic Programming; object-based; cropland fields



Citation: Wen, C.; Lu, M.; Bi, Y.; Zhang, S.; Xue, B.; Zhang, M.; Zhou, Q.; Wu, W. An Object-Based Genetic Programming Approach for Cropland Field Extraction. *Remote Sens.* **2022**, *14*, 1275. <https://doi.org/10.3390/rs14051275>

Academic Editors: Feng Gao, Berk Üstündağ, Liying Guo and Yahui Di

Received: 26 January 2022

Accepted: 3 March 2022

Published: 5 March 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

Increasing population and consumption is posing global challenges on agriculture and natural resources [1]. Enhancing food production and sustainability is a promising solution to this grand challenge. The cropland field, as the basic unit for agricultural management, is important for providing critical information on agricultural productivity and the environmental footprint to meet the food security and sustainability needs. Smallholder farms occupy up to 40% of agricultural areas globally [2] and contribute to about 30% of most food commodities [3]. Yet, spatial information of cropland fields of smallholder farms is usually incomplete in developing countries and this hampers addressing food security challenges.

Remote sensing can contribute to agricultural resource monitoring by providing such fundamental cropland field information in a cost-effective way. Graesser et al. [4] developed a method consisting of multi-spectral image edge extraction and multi-scale contrast limited adaptive histogram equalization to detect cropland field parcels from Landsat imagery. Yan et al. [5] integrated a variety of computer vision-based image processing techniques and proposed an automated computational methodology to extract agricultural crop fields from Landsat time series. Based on Landsat satellite data acquired from 9 to 25 years apart,

White et al. [6] employed a multispectral segmentation and nearest neighbor supervised classifier to extract cropland fields and quantified agricultural field size changes. Despite their remarkable contributions, these studies focused on the delineation of huge cropland fields from Landsat images. As cropland fields of smallholder farms are generally less than 2 ha [7], Landsat images are not suitable for accurately detecting field boundaries.

Very High Resolution (VHR) images can be used for mapping cropland fields of smallholder farms in large geographical areas. In general, edge detection, deep learning and object-based image analysis (OBIA) are the main approaches for extracting cropland fields using VHR images. When an object is the basic spatial analysis unit rather than a pixel, salt-and-pepper effects can be avoided. Edge detection is performed by setting filters to detect pixels with significant brightness changes, such as with the Canny operator [8] and the Sobel operator [9]. However, these operators are sensitive to noise and are only suitable for images with less noise and less complexity. The convolutional neural network (CNN) [10,11] is the most commonly used deep learning method for extracting cropland fields. CNN-based methods have achieved promising results in multiscale and multilevel feature learning from remote sensing images [12]. However, these methods typically require large amounts of training data to obtain the desired results. It is typically difficult, time-consuming and expensive to collect, label and process a large volume of training data in many domains. Multiresolution segmentation (MRS) is an object-based segmentation method that has achieved excellent results for delineating cropland fields in different regions [13,14]. It merges adjacent objects into larger objects with minimal spectral and spatial heterogeneity variation. MRS obtains segmentation results in regions of varying complexity without training samples.

Accurately extracting cropland fields of smallholder farms is challenging because of their variability in shapes and cropland reflectance. The boundaries of cropland fields are usually irregular, and the edges tend to be fuzzy because of mixed pixel effects in the satellite images. The surface reflectance of cropland fields is usually diverse because of seasonal variations and different crop patterns, and spectral similarities between croplands and other land cover categories, such as trees, grass and shrub, introduce spectral confusion. Yet, most of the existing literature focuses on shape delineation, not considering the reflectance of cropland fields. Furthermore, studies have often focused on areas where croplands are the only land cover. Therefore, it is important to explore a more comprehensive and accurate method for extracting croplands and simultaneously delineating field boundaries.

Genetic Programming (GP) [15] is a popular Evolutionary Computation paradigm and evolutionary machine learning paradigm [16] under the big umbrella of Artificial Intelligence (AI). Driven by Darwinian selection and biological evolution, GP automatically evolves computer programs to find the best solution to a task. GP has been widely applied to recognition tasks, including cancer diagnosis [17,18], character recognition [19,20] and object recognition [21,22]. In remote sensing, Munoz et al. [23] adopted GP to create a new vegetation index that enables bare ground identification. Cabral et al. [24] applied GP to identify burned areas with a cutoff value of 0.5, showing that standard GP can produce better results than Maximum Likelihood Classification and Regression Tree. The above work shows that GP is a promising approach to handling the challenges of extracting cropland fields of smallholder farms.

The goal of this paper is to propose a new object-based GP approach for extracting cropland fields using high-resolution images. GP constructs a discriminant function for representing classifiers based on spectral, shape and texture features of objects. To the best of our knowledge, the use of GP for cropland field extraction from VHR images has not been explored yet, so this will be the first time using GP to perform this task.

2. The Proposed Approach

The basic idea of the proposed approach is that a series of possible tree classifiers for cropland field extraction are randomly generated based on the object features. Then the qualities of the candidate classifiers are measured by the fitness function that is the

overall accuracy. All classifiers are updated by using genetic operators such as crossover and mutation. The optimal classifier from the last generation is selected to extract all cropland fields.

The overall structure presented in Figure 1 has three steps. In the first step, the high-resolution image is segmented, and object features, i.e., spectral, shape and texture, are extracted as GP inputs. The second step is the evolutionary learning process that constructs the optimal GP classifier, consisting of population initialization, fitness evaluation and population updating. In the third step, the best GP classifier is used to extract cropland fields, and the results are assessed.

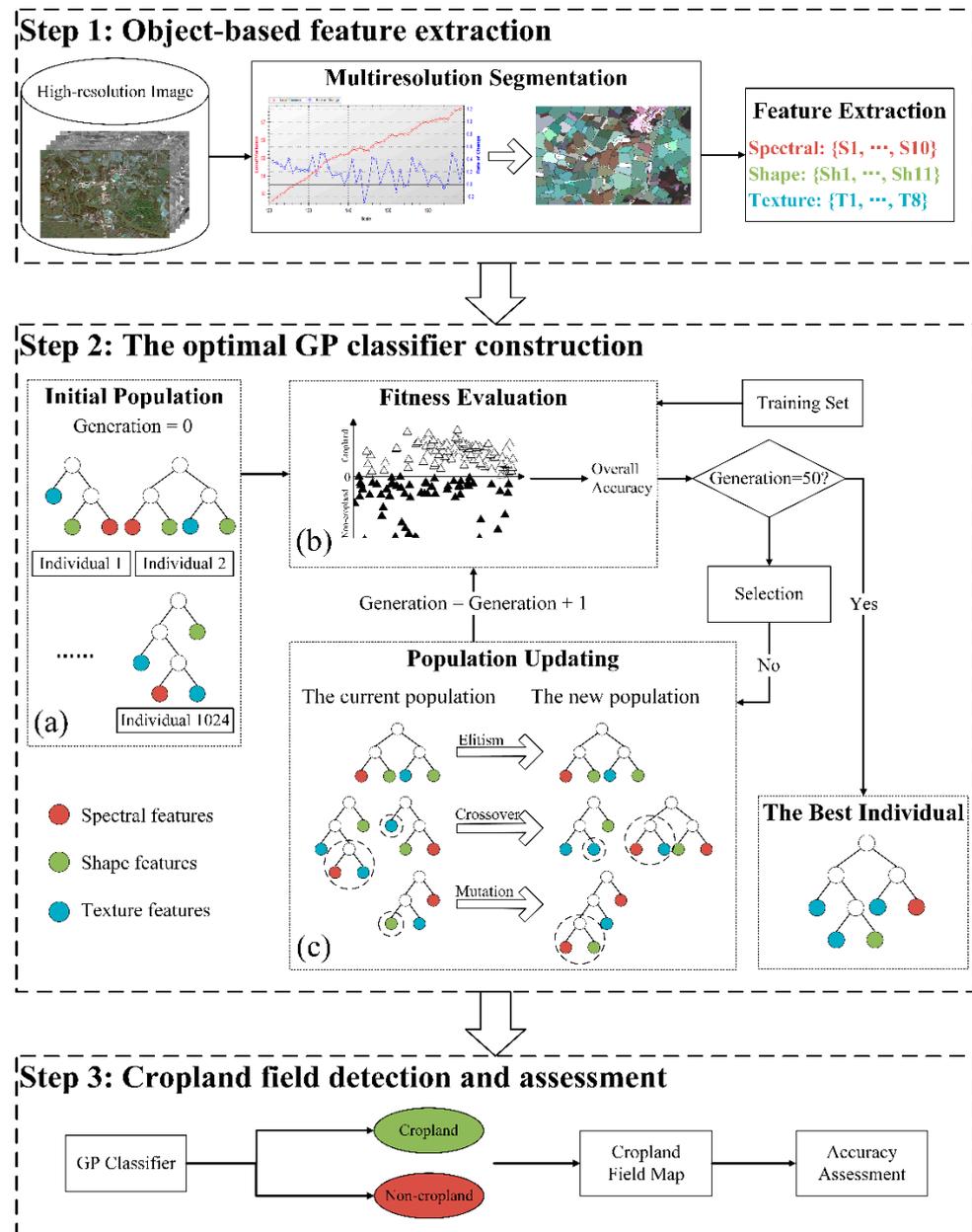


Figure 1. The flowchart of the proposed genetic programming (GP) method. The optimal GP classifier construction includes three parts, initial population (a), fitness evaluation (b) and population updating (c).

2.1. Object-Based Feature Extraction

Segmentation is a bottom-up region-merging process that minimizes the average heterogeneity of pixels and maximizes their respective homogeneity. The segmentation

procedure starts with single image objects of individual pixels and repeatedly merges them in several iterations to larger units until an upper threshold of homogeneity is exceeded locally [25]. We used the Definiens eCognition multiresolution segmentation package [26] to segment the images. This package requires three parameters: object scale, shape, and compactness. With a certain scale, color and shape define the heterogeneity of the objects [27]. The shape parameter value ranges from 0 to 1. The color represents the weight of the object spectral information, so the shape parameter should be less than 0.9. The shape is further divided into compactness and smoothness. These two are related to each other as smoothness equals to 1-compactness. The object stops merging if the object heterogeneity after pre-merging is greater than the scale parameter. Therefore, the scale parameter determines the object size and its determination is crucial to achieving the best segmentation results for cropland fields.

We determined the most appropriate scale using the estimation of the scale parameter (ESP) tool [28]. The basic idea of the ESP tool is to select the scale by calculating the rate of change (ROC) curve for the local variance (LV) of object homogeneity at various scales. The ROC is calculated as:

$$ROC_N = \frac{LV_N - LV_{N-1}}{LV_{N-1}} \times 100 \quad (1)$$

where ROC_N is the rate of change of the scale N , and LV_{N-1} is the local variance of a scale interval smaller than LV_N . The scale corresponding to the peak of the ROC curve is the optimal segmentation scale.

We extracted three types of object features including 10 spectral features, 11 shape features and 8 texture features after image segmentation. The spectral features are the most basic information of remote sensing data and need to be considered thoroughly in the classification task [29]. Cropland fields have a more regular shape than other landscapes, so the shape features are stable and effective in distinguishing cropland objects [30]. Cropland fields have distinctive texture features in high-resolution images, which are commonly extracted by the Gray-Level Co-occurrence Matrix (GLCM) in existing studies [31]. Spectral features (S1–S10), shape features (Sh1–Sh11) and texture features (T1–T8) were calculated based on the image objects. More details are listed in Table S1 in the Supplementary Materials.

The features were normalized to a common scale $[-1, 1]$ because they had different value ranges. Normalization was performed as:

$$x_i^* = \frac{x_i}{\max_{k=1}^C |x_k|} \quad (2)$$

where x_i^* is the normalized feature value, x_i is a feature value, C is the total number of feature values and $\max_{k=1}^C |x_k|$ is the maximum absolute value of the feature. The normalized features were used as the input to GP to construct the optimal classifier for cropland field extraction.

2.2. The Optimal GP Classifier Construction

GP finds the optimal classifier by constructing numerous individuals in the evolutionary process. An individual is composed by terminals and functions, and an example GP classifier is shown in Figure 2. The colored leaf nodes represent terminals (including object features and constants), while the white internal nodes represent functions. In the binary cropland classification, the classifier calculates the probability of the “cropland” class. If the probability of the sample is greater than 0, it is labeled as “cropland”, otherwise it is labeled as “non-cropland”.

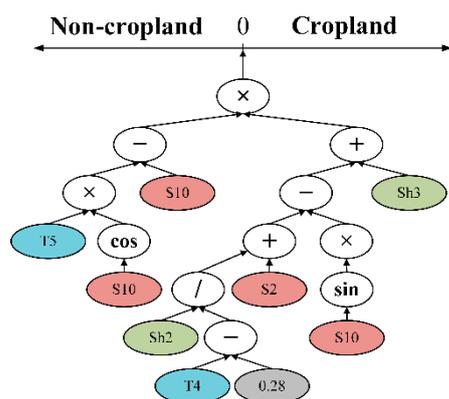


Figure 2. The GP individual represents a classifier. The colored leaf nodes represent terminals, while the white internal nodes represent functions.

2.2.1. Initial Population

A GP population contains 1024 tree-based individuals, and each individual represents a potential classifier to extract cropland fields. An individual is generated by automatically selecting the object features (Section 2.1) or random constants $[-1, 1]$ to form the leaf nodes and functions from $\{+, -, \times, /, \sin, \cos\}$ to form the internal/root nodes.

In this study, the initial population was randomly generated by using the ramped half-and-half method, which generated half of the population using the full method and the other half using the grow method [32]. The minimum tree depth of the initial individual was set to 2 and the maximum depth was defined as 6 [18]. The ramped half-and-half method ensures a diverse range of sizes and shapes of the generated individuals, making it easier to find the optimal classifier to extract cropland fields.

Before the initial population begins to evolve, it is important to determine the necessary evolutionary parameters. The parameter settings for the proposed GP approach followed the commonly used settings in the GP community [18,33] (Table 1).

Table 1. The necessary GP parameter settings.

Parameter	Value	Parameter	Value
Population Size	1024	Elitism Rate	0.01
No. Generation	50	Crossover Rate	0.8
Selection Method	Tournament	Mutation Rate	0.19
Tournament Size	7	Maximum Tree Depth	8

2.2.2. Fitness Evaluation

The fitness evaluation process aims to assess how well each individual/classifier in the current population extracts cropland fields. The individual calculates the probability of the “cropland” class of each sample. If the probability is greater than 0, it is labeled as “cropland”, otherwise it is labeled as “non-cropland”. The fitness function is the overall accuracy (OA), defined as:

$$OA = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \quad (3)$$

where TP refers to true positives, FP refers to false positives, TN refers to true negatives, and FN refers to false negatives. The fitness function is to be maximized, i.e., the higher the fitness value, the better the individual.

In the proposed GP approach, the maximum tree size (i.e., the number of nodes) was defined as 70 to avoid the bloating issue. Bloating occurs when the tree keeps growing without any significant improvement in fitness [34]. It often causes a stagnant evolution and ends up with complex individuals. To address this, the fitness value of the GP individuals

is set to 0 when the tree size is greater than 70. By having this constraint, we expect the new GP approach to obtain an interpretable tree with a small size but good performance.

2.2.3. Population Updating

During the evolutionary process, a new population (the next generation) was iteratively updated using the tournament selection method and three genetic operators, i.e., elitism, crossover and mutation. The selection method was used to select individuals with better fitness from the current population for crossover and mutation. The tournament selection method selected the best individual from seven randomly sampled individuals from the population by comparing their fitness values. The selection process was repeated 1014 (population size \times (1 – elitism rate)) times. Each of the selected individuals was operated by the crossover or mutation operator.

The elitism operator copied the 10 (population size \times elitism rate) best individuals with the highest fitness values from the current population to the next generation. Crossover and mutation operations were applied to the selected individuals (parents) to generate new individuals (offspring). The crossover operation swaps the randomly selected subtrees of two individuals to generate two new individuals. The mutation operation randomly selects a subtree and replaces it with a new randomly generated subtree to create a new individual. The processes of elitism, crossover and mutation are shown in Figure 1c. The new populations were reassessed for fitness, and the process was repeated until the number of generations reached 50.

2.3. Cropland Field Extraction and Assessment

The optimal GP individual/classifier calculated the probability and used a threshold of 0 to extract cropland fields. If the probability of the “cropland” class was greater than 0, the object was labeled as “cropland”, otherwise it was labeled as “non-cropland”. We calculated the OA (Equation (3)), kappa coefficient (Equation (4)), commission error (Equation (6)) and omission error (Equation (7)) [35] to evaluate the performance of the GP classifier. The kappa coefficient represents the agreement of the predicted class labels and the ground-truth class labels. The commission measure is the ratio of misclassified cropland fields to the classified cropland fields by the optimal GP classifier. Omission is the ratio of misclassified non-cropland fields to the real number of cropland fields in the test set.

$$Kappa = \frac{OA - P_e}{1 - P_e} \quad (4)$$

where P_e is computed as:

$$\frac{(TP + FN)(TP + FP) + (TN + FP)(TN + FN)}{TP + FP + TN + FN} \quad (5)$$

$$Commission = \frac{FP}{TP + FP} \quad (6)$$

$$Omission = \frac{FN}{TP + FN} \quad (7)$$

3. Experiment Design

The proposed GP approach was examined in two locations with different landscape complexities, and the performances to test the effectiveness and flexibility of GP were compared with five benchmark methods, i.e., K-Nearest Neighbor (KNN), Decision Tree (DT), Naïve Bayes (NB), Support Vector Machine (SVM), and Random Forest (RF).

3.1. Study Areas and Data Source

We selected two study areas in Fucheng Town of Nanning City, Guangxi Province of China (Figure 3). Fucheng is the largest agricultural town in Nanning with 6327 ha of

farmland. Rice, corn, and soybean are the predominant staple crops, and sugarcane and longan are the most important cash crops. The landscapes of the two study areas were selected to be very different to test the flexibility of the GP approach. The cropland fields in Area 1 are larger and more regular than in Area 2. Details of the classifiers' performances were analyzed for Tile 1, Tile 2, Tile 3 and Tile 4.

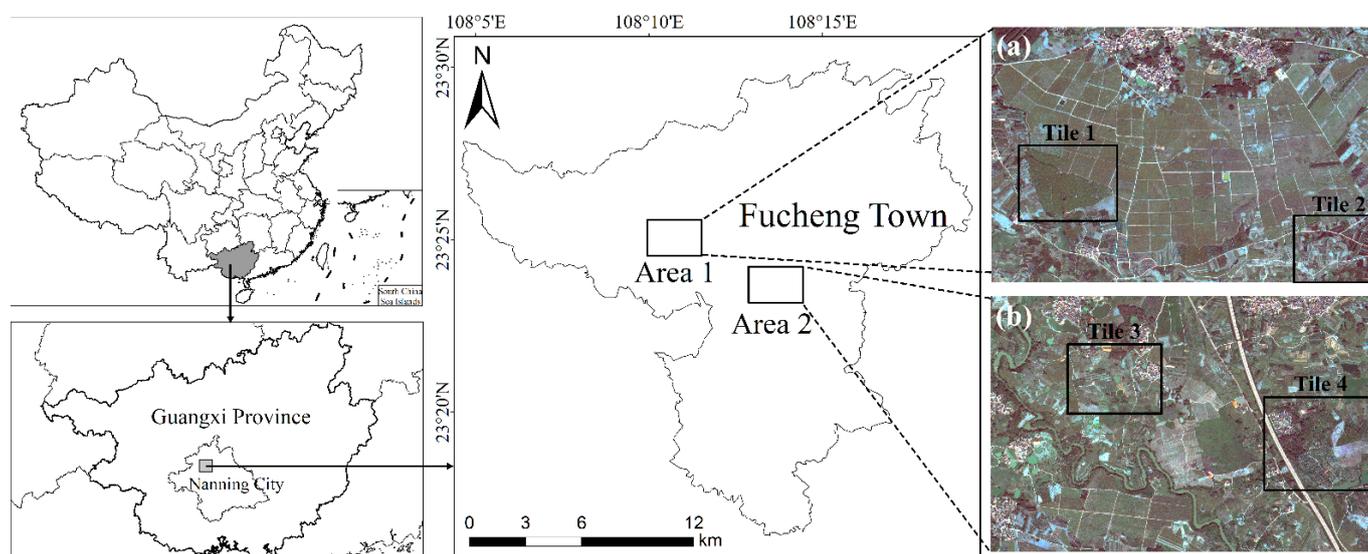


Figure 3. The study areas were located in Fucheng Town of Nanning City, Guangxi Province of China. Area 1 (a) and Area 2 (b) are locations used for assessing the flexibility of the GP approach, and Tiles 1–4 are easily misclassified and are used for additional comparisons and analysis.

We used the VHR image from the Beijing-2 (BJ-2) satellite acquired on 29 October 2018. The BJ-2 satellite constellation is the first Chinese government's civilian commercial satellite project. The constellation is composed of three high-resolution satellites launched in 2015, and it can view the entire Earth's surface every day. The BJ-2 image data contains a panchromatic (PAN) band and four multispectral (MS) bands (Table 2). The orthorectification was conducted for PAN and MS, respectively, after systematic geometric correction. Then the MS bands were resampled to 0.8-m resolution, and image registration was applied to eliminate possible misalignments between the PAN and MS. Lastly, PAN and MS were merged and we obtained the image with 0.8-m resolution.

Table 2. Detailed information of the five spectral bands from the BJ-2 image.

Name	Wavelength (nm)	Resolution (m)
PAN	450–650	0.8
Blue	440–510	3.2
Green	510–590	3.2
Red	600–670	3.2
Near Infrared	760–910	3.2

We selected the training and test samples and labeled them as cropland field or non-cropland field by visual interpretation and field survey. There were 238 samples in Area 1 and 657 samples in Area 2. We selected 70% samples per class for training and 30% samples for testing for each area.

3.2. Comparative Experiments with the GP Classifier

Five representative classifiers, i.e., KNN, DT, NB, SVM and RF, were used as benchmark methods to investigate the effectiveness of the proposed GP approach. These methods used the same training and test samples as GP for performance comparisons.

KNN [36] is one of the simplest machine learning (ML) algorithms. DT [37] consists of a set of hierarchical if–then rules based on a tree structure. NB [38] calculates the posterior probability based on each class’s prior probability and conditional probability and assigns the class label with the maximum probability to a sample. The basic idea of SVM [39] is to automatically find the support vectors that can distinguish classes well, and thus construct a classifier that maximizes the interval between classes. RF [40] is known for its ensemble learning, and the performance of several decision trees is boosted via a voting scheme.

The proposed GP approach and the classification algorithms were implemented in Python using the DEAP (Distributed Evolutionary Algorithm Package) [41] and scikit-learn [42], respectively. In KNN, the number of neighbors k was set to 10 [43]. In DT and NB, all the parameters followed the default settings of the corresponding package. In SVM, the optimal values of C and γ were searched using a systematic grid. In this study, the optimal values of C and γ were 3 and 0.85, and the kernel function in SVM was the Radial Basis Function (RBF). The other parameters in these five algorithms were set following [42]. We used OA, kappa, commission and omission of the validation set to compare the performance of GP with the five classifiers. As a stochastic method, GP was executed 30 independent times using different random seeds in each study area. Each run returned the best GP individual with the highest fitness value, and the best individual with the highest fitness value of 30 individuals was used to extract cropland fields.

We also designed an experiment to test the effect of different sample sizes on the performance of the six methods. The percentage of sample size is defined as the percentage of the present sample size to the original sample size. For example, 40% means that 263 samples ($657 \times 40\%$) in Area 2 were used as inputs of six methods to extract cropland fields, of which 70% were used for training and 30% for testing. The percentage of sample size started at 30% and gradually increased to 100% in 10% step intervals.

4. Results and Analysis

4.1. Object Segmentation

The optimal segmentation scales were obtained by using the ESP tool. In Area 1, the ESP tool calculated the LV of the objects from scale parameters of 120 to 170. The potential optimal scales of Area 1 were selected by the local maxima of ROC, which were 125, 131, 141, 150, 156 and 166 (Figure 4a). We selected 150 as the best scale parameter by visually comparing the segmentation results of these scales. We conducted similar experiments to determine the scale parameter in Area 2, selecting 83 as the optimal parameter (Figure 4b). The scale parameters were different because of the different field sizes in the two regions. The more fragmented the field, the smaller the scale parameter.

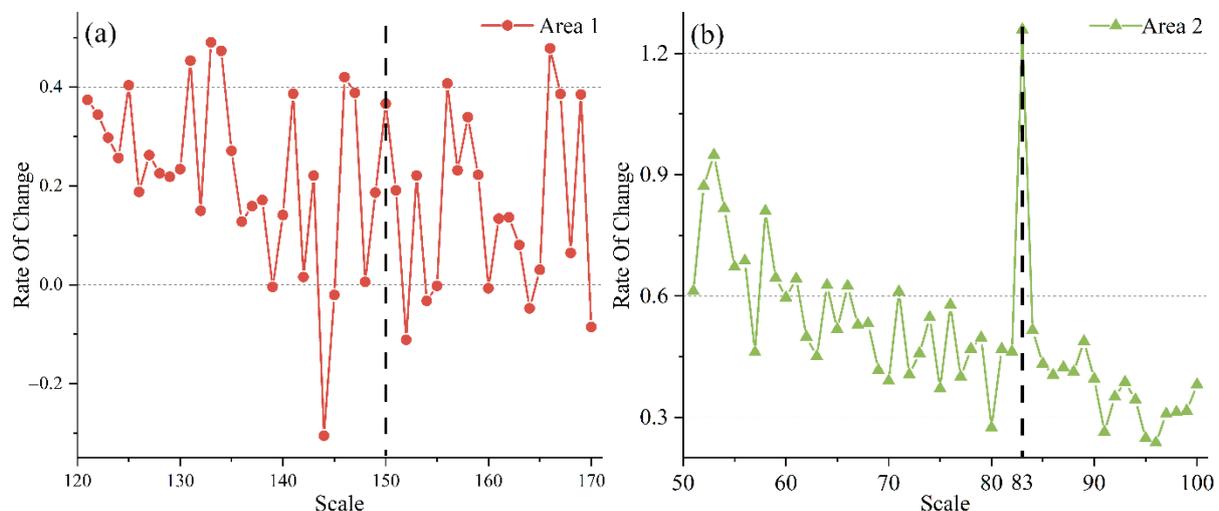


Figure 4. The results of the estimation of scale parameter (ESP) tool in Area 1 (a) and Area 2 (b). The black vertical dashed lines indicate the selected scale parameters of Area 1 and Area 2.

We tested multiple combinations of parameters to determine the optimal one. The shape and compactness parameters of Area 1 were both 0.5, and those of Area 2 were 0.4 and 0.6, respectively. In Area 2, over-segmented objects were merged through an onscreen interactive operation. The segmentation results (Figure 5) illustrate the fragmented and complex geometrical characteristics of smallholder farm fields. The average size of the segmented objects in Area 1 was 38.10 acres, while it was 10.42 acres in Area 2. The spectral, texture and shape features of each object were calculated as inputs for GP.

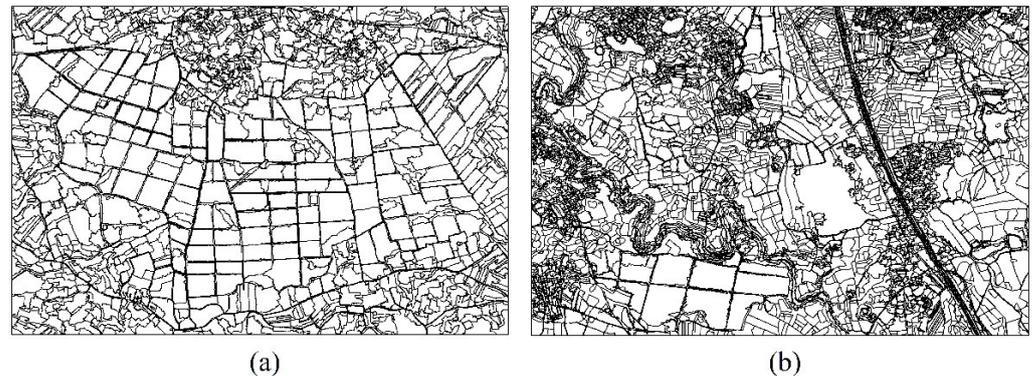


Figure 5. The segmentation results in Area 1 (a) and Area 2 (b).

4.2. Classification Results and Comparison

The results of the cropland field extraction by the object-based GP are visualized in Figure 6. The results show that the classifier evolved by the GP approach effectively distinguishes between cropland and non-cropland in both Area 1 and Area 2. In Area 1, the OA, kappa, commission and omission were 98.6%, 0.97, 0.0% and 3.3%, respectively. Almost all of the cropland and non-cropland fields in the test set were correctly classified. The kappa of GP is greater than 0.95, which means that GP achieves almost perfect agreement between the predicted and the ground-truth class labels [44]. In Area 2, the OA, kappa, commission and omission were 98.0%, 0.95, 1.4% and 3.3%, respectively. The performance is slightly lower than in Area 1, likely because of the more complex landscape.

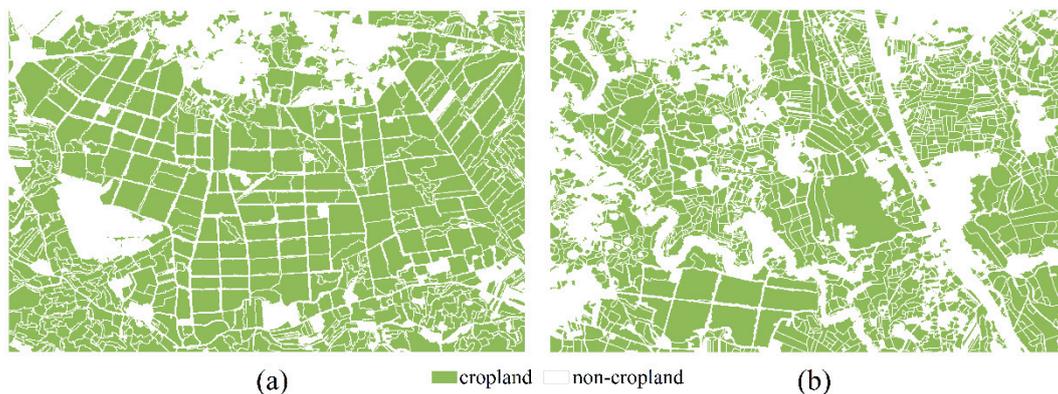


Figure 6. Cropland fields extracted by the object-based GP in Area 1 (a) and Area 2 (b).

The accuracy assessments of the six methods are shown in Table 3. A closer inspection shows that the GP approach in Area 1 and Area 2 present the highest values of OA and kappa. In Area 1, the OA of GP is about 15% higher than that of KNN and NB. SVM has a 12.1% higher commission error than GP. DT and RF have the highest omission error among the six methods. In Area 2, the highest OA obtained by the GP method is 98.0%, which is also higher than for other classifiers. The lowest OA of the six methods is 86.4% of NB, which is 12% lower than for GP. The results show that the proposed GP method works not

only well for the region with larger and regular plots, but also has the potential to produce a good performance for the region with smaller and irregular plots.

Table 3. Classification results of different classifiers in Area 1 and Area 2.

Classification Algorithm	Area 1				Area 2			
	OA (%)	Kappa	Commission (%)	Omission (%)	OA (%)	Kappa	Commission (%)	Omission (%)
KNN	83.3	0.68	28.6	0.0	90.4	0.78	17.5	13.3
DT	94.4	0.89	6.7	6.7	92.9	0.83	12.9	10.0
NB	83.3	0.67	27.5	3.3	86.4	0.71	29.6	5.0
SVM	93.1	0.86	12.1	3.3	93.9	0.86	11.3	8.3
RF	97.2	0.94	0.0	6.7	96.5	0.92	5.1	6.7
GP	98.6	0.97	0.0	3.3	98.0	0.95	1.4	3.3

We selected four tiles that are easily misclassified from the two areas for additional comparisons and analysis (Figure 7). Tile 1 and Tile 2 are in Area 1, and Tile 3 and Tile 4 are in Area 2. Intuitively, GP has the most accurate classification results. In Tile 1, NB and SVM misclassify almost all of the forest fields as cropland. In contrast, DT, RF and GP correctly classify this forest region as non-cropland. In Tile 2, KNN, DT, NB, SVM and RF have more noticeable omission errors than GP. In Tile 3 and Tile 4, KNN, DT, SVM and RF have commission and omission errors simultaneously, and NB has large commission errors. Compared with these five methods, the effectiveness of GP is clearly shown by visualizing the classification results.

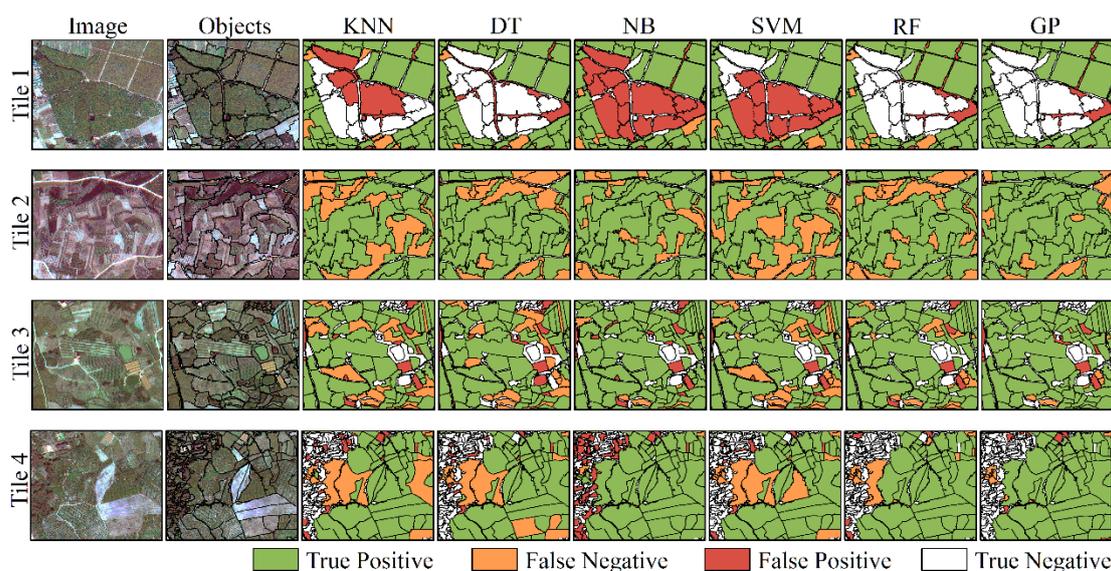


Figure 7. Classification results of six methods in four tiles, i.e., K-Nearest Neighbor (KNN), Decision Tree (DT), Naïve Bayes (NB), Support Vector Machine (SVM), Random Forest (RF) and GP.

The effect of sample size on the performance of the six methods is shown in Figure 8. The GP approach performs with higher accuracy and better stability than the other five classifiers in both study areas. By calculating the mean and standard deviation of the obtained accuracy for different sample sizes (Table S2 in Supplementary Materials), GP obtained the highest accuracy in both areas. In Area 1, although the mean accuracy of RF (92.1%) is second only to that of GP (95.0%) the standard deviation is 0.03 higher than for GP. The standard deviation of NB is lower than that of GP, but the mean accuracy is 12% lower than that of GP. In Area 2, GP gets the highest mean accuracy and the lowest standard deviation. The performance of RF is improved because it has more samples than Area 1. In the two areas, the performance of five ML classifiers fluctuates considerably. This indicates that these ML classifiers are sensitive to the quantity and quality of the sample data. Conversely, GP seems to achieve consistently satisfactory results independently of the volume of samples.

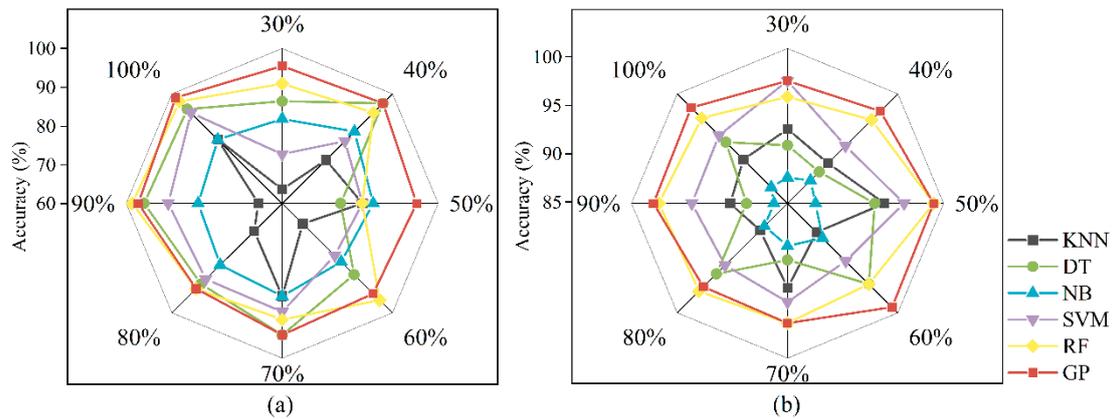


Figure 8. Effect of sample size on six classifiers in Area 1 (a) and Area 2 (b).

4.3. GP Classifier Analysis

The optimal classifier was constructed by GP. The tree-based individuals consist of leaf nodes and internal nodes, and each individual represents a classifier. To better illustrate the superiority of the classifiers generated by GP, the best individuals in the two areas are visualized in Figure 9. These two individuals/classifiers are significantly different in their sizes, structures and used features, suggesting that different levels of geological complexity affect the individuals generated by GP. Spectral features (red nodes) are most frequently used in Area 1 (Figure 9a). However, spectral, shape and texture features are used with about the same frequency in the GP tree found from Area 2 (Figure 9b).

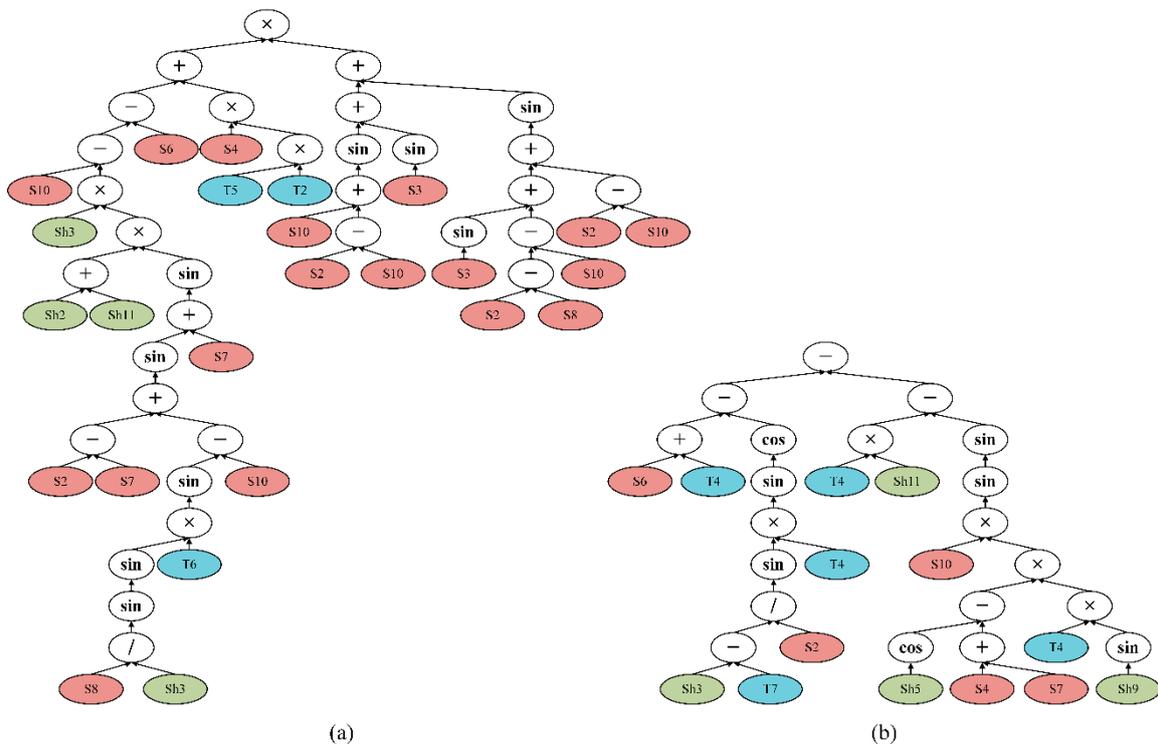


Figure 9. The best GP individuals in Area 1 (a) and Area 2 (b). The colored leaf nodes represent terminals, while the white internal nodes represent functions.

The frequency of the object features in the GP individuals of the two areas is shown in Figure 10. Features with a frequency of 0 in both areas are not listed. NDVI (Normalized Difference Vegetation Index, S10) is the most frequently used feature in GP trees evolved in both areas because NDVI enables the extraction of cropland fields by highlighting

vegetation information. The usage frequency of the shape features is generally lower than that of the other types of features. This is probably because the shape of a field is not an effective feature to identify cropland. In Area 1, Contrast (T2) is only less frequently used than NDVI because texture information based on high-resolution imagery can enhance the ability to identify cropland. In Area 2, Entropy (T4) and Contrast (T2) rank second and third among all 29 features. This suggests that different levels of geological complexity affect feature selection for GP.

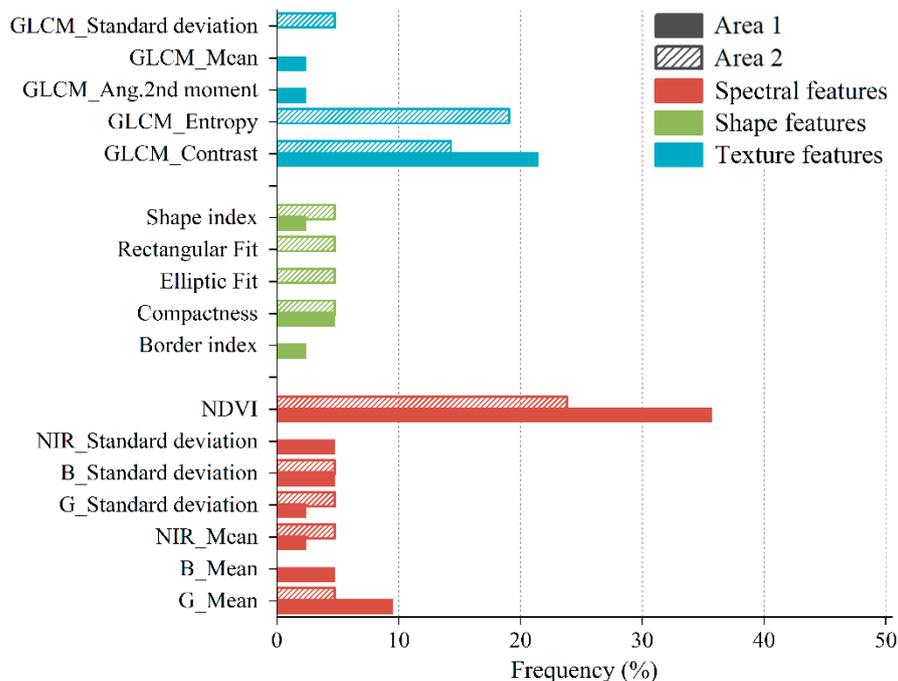


Figure 10. The frequency of the original features in the GP individuals/classifiers. Features with a frequency of 0 in both areas are not listed.

The two classes of the original features and the probability of GP classifiers are visualized in scatter plots (Figure 11). The original features were reduced to two dimensions using t-Distributed Stochastic Neighbor Embedding (t-SNE) [45]. The t-SNE method is a well-known and commonly used technique for dimensionality reduction and visualization. It is used to transform high-dimensional datasets into low-dimensional spaces and typically achieves better visualization results than other methods [45]. The best individual obtained by GP calculated the probability of samples and split it by a threshold of 0. In the two areas, Figure 11a,c show different classes that are highly mixed, making it difficult to distinguish a cluster. In contrast, Figure 11b,d show clearer clusters of the GP classifiers, and fewer points are clustered into the wrong classes. This indicates that the new classifiers constructed by GP eliminate the noise in the original features and make it easier to extract cropland fields.

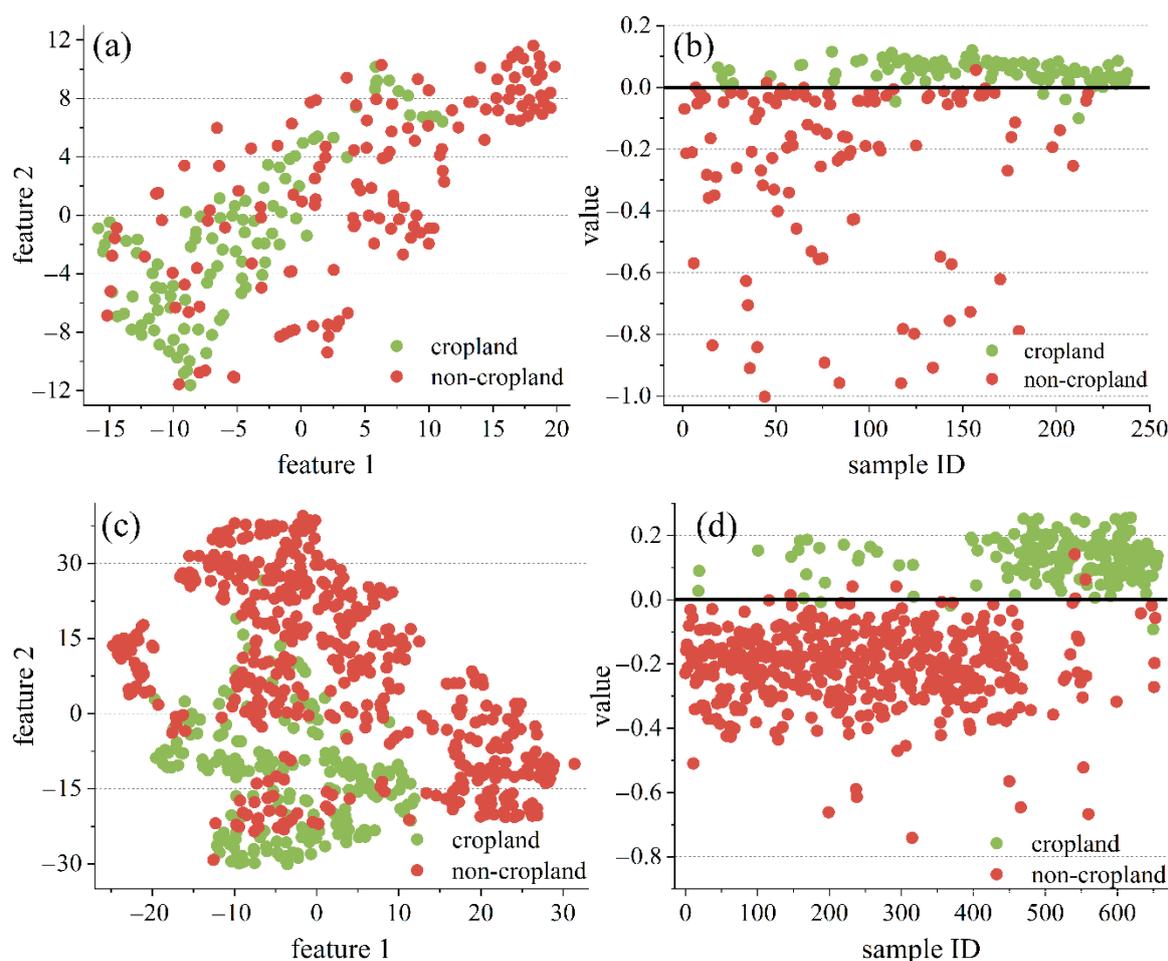


Figure 11. The visualization results in Area 1 (a,b) and Area 2 (c,d). The original features are reduced to two dimensions using t-Distributed Stochastic Neighbor Embedding (t-SNE) (a,c). The GP classifiers divide the samples using a threshold of 0 (b,d).

5. Discussions

This study proposed an object-based GP approach to cropland extraction. Compared with the five non-GP methods, i.e., KNN, DT, NB, SVM and RF, the proposed GP approach achieved the highest overall classification accuracy in two areas. The results demonstrated that the classifiers constructed by GP were effective for cropland extraction.

The two individuals in Area 1 and Area 2 show that different levels of landscape complexity affect the individuals generated by GP, and GP can automatically find the best (most fit) classifiers from different areas. The flexibility of GP is shown in two aspects. Firstly, based on the input features and defined functions, the tree structure of GP makes it to produce a large number of different individuals in the population. This provides a large solution space for the target problem and allows GP to have potentially good global searchability. Secondly, the fitness function determines the evolutionary direction of the population and guides the GP search towards the best solution via the evolutionary learning process. Furthermore, GP can be easily applied to different tasks with a flexible tree-based representation and find the best solutions with various lengths, functions, features, and shapes that achieve good performance. By visualizing the optimal GP classifiers, we can better understand how GP constructs classifiers from the two areas for effective classification. It is clear what types of features are selected by GP and how the features are combined by functions, which suggests that GP is highly interpretable and comprehensible. The analysis of feature frequency suggests that GP can automatically select useful features to construct classifiers for effective cropland extraction. We designed

comparative experiments with five classifiers (KNN, DT, NB, SVM and RF) to verify the effectiveness of GP for cropland field extraction—in particular, we examined how sample sizes affect the classification accuracy—and GP did show a very good performance. In several similar studies, Qian et al. [46] found that KNN was more sensitive to the sample sizes compared with SVM and DT using an object-based classification procedure. Thanh Noi and Kappas [47] found that SVM produced the highest OA with the least sensitivity to the training sample sizes compared with KNN. Modica et al. [48] found that SVM and RF resulted as the most stable classifiers compared with KNN and NB. These conclusions are consistent with the results of our study. SVM and RF show good performances in these articles. However, our results show that GP significantly outperforms SVM and RF in both average accuracy and stability. This shows that GP can play a greater role than other ML methods in cropland field extraction with different sample sizes.

Compared with DT and RF with the tree structure, GP also has obvious advantages. GP can generate a large number of different individuals, even with few samples, and selects the optimal individual from them. DT is also a tree-structured classifier. Its leaf nodes are different classes, and internal nodes are logical expressions. DT selects features and determines thresholds by the Gini index. Therefore, if the training sample is fixed, the resulting decision tree is fixed. However, GP can automatically select the leaf features and internal functions to construct the classifiers/individuals that can find the optimal global solution. Its higher classification accuracy compared to that of decision trees also attests to its better performance. Random Forest (RF) is an ensemble classifier that constructs many decision trees and uses the majority voting method to obtain the classification results. It uses the bootstrap sampling method to randomly select a training set for building a decision tree. Therefore, if the sample size is small, the samples taken for each decision tree are similar, resulting in similar decision trees. RF cannot take advantage of ensemble learning when the number of samples is too small, but GP can build a large number of potential classifiers and still yield highly accurate classification results. Therefore, GP is more applicable to real-world tasks with small field survey samples resulting from cost or privacy constraints.

Although one of the GP advantages is that it does not require extensive domain knowledge [49], it can cope with domain knowledge easily—this paper actually finds that adding expert knowledge for high-resolution images works better. We tried to use GP with only four spectral features (without using domain knowledge), but the results were not satisfactory. As VHR images usually have a low spectral resolution, the original feature space needs to be enriched with subjectively selected features. In this study, we extracted shape and texture features and used them together with the original spectral bands as inputs of GP. The performance of GP was significantly improved, which means that shape and texture were the most effective features.

6. Conclusions

This study explored the potential of the object-based GP approach for cropland field extraction in a binary classification task. A VHR image was segmented by the MRS method and spectral, shape, and textural features were extracted from the segments to be used as inputs to the GP approach. Then, the GP method automatically selected useful features to evolve the optimal classifier. The performance of GP was examined in two representative areas with different landscape complexities. Compared to five well-known classification methods, i.e., KNN, DT, NB, SVM and RF, GP performed significantly better. The results confirmed that GP is an effective approach for cropland extraction.

The advantages of GP classifiers include three aspects: interpretability, flexibility, and effectiveness, particularly when the number of training samples is small. The classifiers evolved by GP can be easily visualized as trees to analyze which features are selected and how they are constructed into tree-based classifiers. Our analysis shows that NDVI and Contrast were used most frequently, and shape features were generally used less frequently than the other two types of features due to the variable shape of crop fields in the two areas.

In different landscape complexity areas, GP can construct many potential classifiers and find the optimal classifier based on the fitness function. Compared with KNN, DT, NB, SVM and RF, GP can achieve the desired results with any volume of training samples, even with small samples sizes.

In the future, owing to a flexible variable-length representation and high interpretability, GP will be applied to automatically construct high-level features in multiclass classification tasks.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs14051275/s1>, Table S1: The descriptions and value ranges of the spectral, shape and texture features; Table S2: The means and standard deviations corresponding to the accuracy obtained by the six methods using different sample size in two areas, i.e., K-Nearest Neighbor (KNN), Decision Tree (DT), Naïve Bayes (NB), Support Vector Machine (SVM), Random Forest (RF) and Genetic Programming (GP).

Author Contributions: C.W., M.L. and Y.B. conceived of the original design of this paper. C.W., M.L., Y.B. and W.W. improved the structure of the paper. S.Z., B.X., M.Z. and Q.Z. provided comments on this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work is financially supported by the National Natural Science Foundation of China (41921001, 42071419 and 41801371), the Project of Special Investigation on Basic Resources of Science and Technology (No. 2019FY202501), and the National Key Research and Development Program of China (2019YFA0607400), the Fundamental Research Funds for Central Non-profit Scientific Institution (1610132020016).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Thanks are given to Agricultural Science and Technology Information Research Institute, Guangxi Academy of Agricultural Science, China, for providing the remote sensing data. We are especially grateful to the anonymous reviewers and editors for appraising our manuscript and for offering instructive comments.

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

References

1. Foley, J.A.; Ramankutty, N.; Brauman, K.A.; Cassidy, E.S.; Gerber, J.S.; Johnston, M.; Mueller, N.D.; O'Connell, C.; Ray, D.K.; West, P.C.; et al. Solutions for a cultivated planet. *Nature* **2011**, *478*, 337–342. [[CrossRef](#)] [[PubMed](#)]
2. Lesiv, M.; Laso Bayas, J.C.; See, L.; Duerauer, M.; Dahlia, D.; Durando, N.; Hazarika, R.; Kumar Sahariah, P.; Vakolyuk, M.; Blyshchyk, V.; et al. Estimating the global distribution of field size using crowdsourcing. *Glob. Chang. Biol.* **2019**, *25*, 174–186. [[CrossRef](#)] [[PubMed](#)]
3. Herrero, M.; Thornton, P.K.; Power, B.; Bogard, J.R.; Remans, R.; Fritz, S.; Gerber, J.S.; Nelson, G.; See, L.; Waha, K.; et al. Farming and the geography of nutrient production for human use: A TRANSDISCIPLINARY analysis. *Lancet Planet. Health* **2017**, *1*, E33–E42. [[CrossRef](#)]
4. Graesser, J.; Ramankutty, N. Detection of cropland field parcels from Landsat imagery. *Remote Sens. Environ.* **2017**, *201*, 165–180. [[CrossRef](#)]
5. Yan, L.; Roy, D.P. Conterminous United States crop field size quantification from multi-temporal Landsat data. *Remote Sens. Environ.* **2016**, *172*, 67–86. [[CrossRef](#)]
6. White, E.V.; Roy, D.P. A contemporary decennial examination of changing agricultural field sizes using Landsat time series data. *Geo* **2015**, *2*, 33–54. [[CrossRef](#)]
7. Zhang, X. *Study on Farm Size and Smallholders Productivity in China*; Chinese Academy of Agricultural Sciences: Beijing, China, 2018.
8. Watkins, B.; van Niekerk, A. A comparison of object-based image analysis approaches for field boundary delineation using multi-temporal Sentinel-2 imagery. *Comput. Electron. Agric.* **2019**, *158*, 294–302. [[CrossRef](#)]
9. Lavreniuk, M.; Kussul, N.; Shelestov, A.; Dubovyk, O.; Low, F. Object-Based Postprocessing Method for Crop Classification Maps. In Proceedings of the 38th IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 22–27 July 2018; pp. 7058–7061.

10. Liu, H.; Luo, J.; Sun, Y.; Xia, L.; Wu, W.; Yang, H.; Hu, X.; Gao, L. Contour-oriented Cropland Extraction from High Resolution Remote Sensing Imagery Using Richer Convolution Features Network. In Proceedings of the 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Istanbul, Turkey, 16–19 July 2019; pp. 1–6.
11. Persello, C.; Tolpekin, V.A.; Bergado, J.R.; de By, R.A. Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. *Remote Sens. Environ.* **2019**, *231*, 111253. [[CrossRef](#)]
12. Zhang, L.P.; Zhang, L.F.; Du, B. Deep Learning for Remote Sensing Data A technical tutorial on the state of the art. *Ieee Geosci. Remote Sens. Mag.* **2016**, *4*, 22–40. [[CrossRef](#)]
13. Conrad, C.; Fritsch, S.; Zeidler, J.; Rucker, G.; Dech, S. Per-Field Irrigated Crop Classification in Arid Central Asia Using SPOT and ASTER Data. *Remote Sens.* **2010**, *2*, 1035–1056. [[CrossRef](#)]
14. Vogels, M.F.A.; de Jong, S.M.; Sterk, G.; Douma, H.; Addink, E.A. Spatio-Temporal Patterns of Smallholder Irrigated Agriculture in the Horn of Africa Using GEOBIA and Sentinel-2 Imagery. *Remote Sens.* **2019**, *11*, 143. [[CrossRef](#)]
15. Koza, J.R. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*; MIT Press: Cambridge, MA, USA, 1992.
16. Al-Sahaf, H.; Bi, Y.; Chen, Q.; Lensen, A.; Mei, Y.; Sun, Y.N.; Tran, B.; Xue, B.; Zhang, M.J. A survey on evolutionary machine learning. *J. R. Soc. N. Z.* **2019**, *49*, 205–228. [[CrossRef](#)]
17. Hong, G.; Nandi, A.K. Breast Cancer Diagnosis Using Genetic Programming Generated Feature. In Proceedings of the 2005 IEEE Workshop on Machine Learning for Signal Processing, Mystic, CT, USA, 28–30 September 2005; pp. 215–220.
18. Ain, Q.U.; Xue, B.; Al-Sahaf, H.; Zhang, M. Genetic Programming for Multiple Feature Construction in Skin Cancer Image Classification. In Proceedings of the 2019 International Conference on Image and Vision Computing New Zealand (IVCNZ), Dunedin, New Zealand, 2–4 December 2019; pp. 1–6.
19. Bi, Y.; Xue, B.; Zhang, M. Genetic Programming With a New Representation to Automatically Learn Features and Evolve Ensembles for Image Classification. *IEEE Trans. Cybern.* **2021**, *51*, 1769–1783. [[CrossRef](#)] [[PubMed](#)]
20. Teredesai, A.M.; Govindaraju, V. Issues in evolving GP based classifiers for a pattern recognition task. In Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No. 04TH8753), Portland, OH, USA, 19–23 June 2004; pp. 509–515.
21. Liang, Y.; Zhang, M.; Browne, W. *Wrapper Feature Construction for Figure-Ground Image Segmentation Using Genetic Programming*; Springer: Cham, Switzerland, 2017; pp. 111–123.
22. Bi, Y.; Xue, B.; Zhang, M.J. An Effective Feature Learning Approach Using Genetic Programming With Image Descriptors for Image Classification [Research Frontier]. *IEEE Comput. Intell. Mag.* **2020**, *15*, 65–77. [[CrossRef](#)]
23. Muñoz, J.; Cobos, C.; Mendoza, M. Vegetation Index Based on Genetic Programming for Bare Ground Detection in the Amazon. In *Advances in Computational Intelligence*; Castro, F., MirandaJimenez, S., GonzalezMendoza, M., Eds.; Lecture Notes in Computer Science; Springer International Publishing AG: Cham, Switzerland, 2018; Volume 10633, pp. 259–271.
24. Cabral, A.I.R.; Silva, S.; Silva, P.C.; Vanneschi, L.; Vasconcelos, M.J. Burned area estimations derived from Landsat ETM plus and OLI data: Comparing Genetic Programming with Maximum Likelihood and Classification and Regression Trees. *ISPRS J. Photogramm. Remote Sens.* **2018**, *142*, 94–105. [[CrossRef](#)]
25. Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm. Remote Sens.* **2004**, *58*, 239–258. [[CrossRef](#)]
26. Baatz, M.; Schape, A. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. In Proceedings of the Angewandte Geographische Informations-Verarbeitung XII, Wichmann Verlag, Karlsruhe, Germany; 2000; pp. 12–23.
27. Yan, G.; Mas, J.F.; Maathuis, B.H.P.; Xiangmin, Z.; Van Dijk, P.M. Comparison of pixel-based and object-oriented image classification approaches—a case study in a coal fire area, Wuda, Inner Mongolia, China. *Int. J. Remote Sens.* **2007**, *27*, 4039–4055. [[CrossRef](#)]
28. Drăguț, L.; Tiede, D.; Levick, S.R. ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 859–871. [[CrossRef](#)]
29. Zhong, L.H.; Gong, P.; Biging, G.S. Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 1–13. [[CrossRef](#)]
30. Jones, C. Operating department assistants: Leave it to the specialists. *Nurs. Mirror.* **1982**, *154*, 16.
31. Kwak, G.H.; Park, N.W. Impact of Texture Information on Crop Classification with Machine Learning and UAV Images. *Appl. Sci.* **2019**, *9*, 643. [[CrossRef](#)]
32. Poli, R.; Langdon, W.B.; McPhee, N.F. *A Field Guide to Genetic Programming*. 2018. Available online: <http://www.gp-field-guide.org.uk> (accessed on 1 January 2022).
33. Tran, B.; Xue, B.; Zhang, M. Genetic programming for multiple-feature construction on high-dimensional classification. *Pattern Recognit.* **2019**, *93*, 404–417. [[CrossRef](#)]
34. Luke, S. Modification point depth and genome growth in genetic programming. *Evol. Comput.* **2003**, *11*, 67–106. [[CrossRef](#)]
35. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* **2002**, *80*, 185–201. [[CrossRef](#)]
36. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [[CrossRef](#)]
37. Quinlan, J.R. Induction of Decision Trees. *Mach. Learn.* **1986**, *1*, 81–106. [[CrossRef](#)]
38. Maron, M.E.; Kuhns, J.L. On Relevance, Probabilistic Indexing and Information Retrieval. *J. ACM* **1960**, *7*, 216–244. [[CrossRef](#)]
39. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]

40. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
41. Fortin, F.A.; De Rainville, F.M.; Gardner, M.A.; Parizeau, M.; Gagne, C. DEAP: Evolutionary Algorithms Made Easy. *J. Mach. Learn. Res.* **2012**, *13*, 2171–2175.
42. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
43. Georganos, S.; Grippa, T.; Vanhuysse, S.; Lennert, M.; Shimoni, M.; Kalogirou, S.; Wolff, E. Less is more: Optimizing classification performance through feature selection in a very-high-resolution remote sensing object-based urban application. *GISci. Remote Sens.* **2017**, *55*, 221–242. [[CrossRef](#)]
44. Munoz, S.R.; Bangdiwala, S.I. Interpretation of Kappa and B statistics measures of agreement. *J. Appl. Stat.* **1997**, *24*, 105–111. [[CrossRef](#)]
45. Van Der Maaten, L.; Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
46. Qian, Y.G.; Zhou, W.Q.; Yan, J.L.; Li, W.F.; Han, L.J. Comparing Machine Learning Classifiers for Object-Based Land Cover Classification Using Very High Resolution Imagery. *Remote Sens.* **2015**, *7*, 153–168. [[CrossRef](#)]
47. Thanh Noi, P.; Kappas, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors* **2018**, *18*, 18. [[CrossRef](#)]
48. Modica, G.; De Luca, G.; Messina, G.; Pratico, S. Comparison and assessment of different object-based classifications using machine learning algorithms and UAVs multispectral imagery: A case study in a citrus orchard and an onion crop. *Eur. J. Remote Sens.* **2021**, *54*, 431–460. [[CrossRef](#)]
49. Bi, Y.; Xue, B.; Zhang, M.J. Genetic Programming With Image-Related Operators and a Flexible Program Structure for Feature Learning in Image Classification. *IEEE Trans. Evol. Comput.* **2021**, *25*, 87–101. [[CrossRef](#)]