



# A Combination of OBIA and Random Forest Based on Visible UAV Remote Sensing for Accurately Extracted Information about Weeds in Areas with Different Weed Densities in Farmland

Chao Feng<sup>1</sup>, Wenjiang Zhang<sup>1,\*</sup>, Hui Deng<sup>1</sup>, Lei Dong<sup>2</sup>, Houxi Zhang<sup>3</sup>, Ling Tang<sup>1</sup>, Yu Zheng<sup>1</sup> and Zihan Zhao<sup>1</sup>

- <sup>1</sup> College of Earth Science, Chengdu University of Technology, Chengdu 610059, China; nianyi@stu.cdut.edu.cn (C.F.); dengh@cdut.edu.cn (H.D.); 2021050031@stu.cdut.edu.cn (L.T.); 2021050020@stu.cdut.edu.cn (Y.Z.); 2021020032@stu.cdut.edu.cn (Z.Z.)
- <sup>2</sup> No. 2 Geological Team, Tibet Autonomous Region Geological Mining Exploration and Development Bureau, Lhasa 850007, China; qingjiaming@stu.cdut.edu.cn
- <sup>3</sup> College of Forestry, Fujian Agriculture and Forestry University, Fuzhou 350002, China; zhanghouxi@fafu.edu.cn
- \* Correspondence: zhangwenjiang13@cdut.edu.cn

Abstract: Weeds have a significant impact on the growth of rice. Accurate information about weed infestations can provide farmers with important information to facilitate the precise use of chemicals. In this study, we utilized visible light images captured by UAVs to extract information about weeds in areas of two densities on farmland. First, the UAV images were segmented using an optimal segmentation scale, and the spectral, texture, index, and geometric features of each segmented object were extracted. Cross-validation and recursive feature elimination techniques were combined to reduce the dimensionality of all features to obtain a better feature set. Finally, we analyzed the extraction effect of different feature dimensions based on the random forest (RF) algorithm to determine the best feature dimensions, and then we further analyzed the classification result of machine learning algorithms, such as random forest, support vector machine (SVM), decision tree (DT), and K-nearest neighbors (KNN) and compared them based on the best feature dimensions. Using the extraction results of the best classifier, we created a zoning map of the weed infestations in the study area. The results indicated that the best feature subset achieved the highest accuracy, with respective overall accuracies of 95.38% and 91.33% for areas with dense and sparse weed densities, respectively, and F1-scores of 94.20% and 90.57. Random forest provided the best extraction results for each machine learning algorithm in the two experimental areas. When compared to the other algorithms, it improved the overall accuracy by 1.74-12.14% and 7.51-11.56% for areas with dense and sparse weed densities, respectively. The F1-score improved by 1.89–17.40% and 7.85–10.80%. Therefore, the combination of object-based image analysis (OBIA) and random forest based on UAV remote sensing accurately extracted information about weeds in areas with different weed densities for farmland, providing effective information support for weed management.

Keywords: UAV remote sensing; weeds; random forest; object-based image analysis

## 1. Introduction

China is the world's largest producer and consumer of rice, and rice production is tightly linked to the food security of 60% of the country's population. Weeds growing in rice fields compete with rice for water, nutrients, and growing space, resulting in a decline in rice yields and favorable conditions for the growth of rice viruses [1,2]. Asian sprangletop is one of the most harmful weeds, which grows fast and is widely distributed in all the provinces of China, so how to effectively control its growth has become a complex problem. At present, spraying chemicals over large areas has become the primary weed



Article

Citation: Feng, C.; Zhang, W.; Deng, H.; Dong, L.; Zhang, H.; Tang, L.; Zheng, Y.; Zhao, Z. A Combination of OBIA and Random Forest Based on Visible UAV Remote Sensing for Accurately Extracted Information about Weeds in Areas with Different Weed Densities in Farmland. *Remote Sens.* 2023, *15*, 4696. https:// doi.org/10.3390/rs15194696

Academic Editors: Hao Yang, Changping Huang, Dameng Yin, Xiuliang Jin and Zhenhai Li

Received: 16 August 2023 Revised: 8 September 2023 Accepted: 13 September 2023 Published: 25 September 2023



**Copyright:** © 2023 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/). control method, but it is untargeted and inefficient, not only causing much pollution to the ecological environment but also affecting farmers' health. Therefore, it is important to adopt intelligent equipment to automate farmland management.

Obtaining information on weed growth in rice fields is a precondition for automated farmland management. The acquisition of large-scale data has become possible with the development of remote sensing technology and has been widely used in geological disaster monitoring [3,4], environmental monitoring [5], and resource monitoring, such as for crops and forests [6–10]. Although technologies such as satellite and space remote sensing have been the primary methods for observing plant optical features, the images obtained by these technologies can be affected by clouds and have low temporal and spatial resolution. Even high-resolution satellite remote sensing images may not be sufficient for obtaining accurate weed growth information from farmland. On the contrary, as a new low-altitude remote sensing platform, UAVs have the advantages of portability, high resolution, reduced effects from weather, and lower cost when compared to high-altitude remote sensing technology [11–13]. Thus, UAVs are better suited for acquiring information about weed growth. In fact, research results on UAV remote sensing in various fields have been promising, including crop monitoring [14,15], nitrogen content estimation [16], and the extraction of urban impervious surfaces [13].

Two approaches currently exist for traditional remote sensing image processing: pixelbased and object-based. The pixel-based classification method is commonly used for lowand medium-resolution remote sensing images because it can produce excessive noise when applied to high- or ultra-high-resolution remote sensing images. Therefore, many researchers have favored the object-based classification method as an alternative [13,17–19]. The objectbased classification method is based on mutually disjointed regions that exhibit consistent or similar features, such as spectral, geometric, and textural features, after image segmentation has been carried out. This method can extract a wide range of features in greater depth, resulting in improved classification accuracy compared to the pixel-based method. Therefore, the OBIA method is particularly well-suited for the classification of digital orthophoto maps generated from UAV images [19–22]. The traditional machine learning algorithms represented by support vector machines [23], K-nearest neighbors [24], and decision tree [25] have their advantages and limitations for object-oriented image analysis: support vector machines can solve classification problems with small samples, K-nearest neighbors has a simple algorithm and is easy to understand and implement, and the decision tree algorithms require fewer training samples, but the K-nearest neighbors perform poorly in the face of sample imbalance, the decision tree algorithm tends to produce an overly complex model, which is poorly generalized to the data, and the support vector machine algorithm has difficulty in finding a suitable kernel functions for the nonlinear complementarity problem. However, as the feature dimension increases, features of lower importance act as noise, so the traditional OBIA classification algorithm becomes limited; therefore, it is especially important to choose a suitable classification algorithm. The random forest algorithm has been widely used in previous studies for feature extraction and the classification of remote sensing images due to its high accuracy, strong generalization ability, and superior resistance to overfitting [26–28]. Moreover, this algorithm has been shown to perform well in high-dimensional feature spaces, leading to better classification accuracy [29,30].

Many researchers have recently focused on developing accurate methods for identifying and locating weeds in farmland, using both OBIA classification [31–33] and deep learning techniques [34–36]. Although these studies have made progress in the field of agriculture, further research to improve the accuracy and applicability of weed identification results is still necessary. For example, while there have been several studies on weed information extraction using OBIA methods, few have focused on determining the optimal segmentation scale and feature subsets or investigating the relationship between the spatial distribution of weeds and the accuracy of the extraction results [33]. In studies that use deep learning methods to extract information about weeds, most focus on detecting individual weeds in farmland using object detection algorithms. While such studies have demonstrated high levels of accuracy, they may not fully meet the practical requirements for weed management [34–36]. Therefore, in this study, images of areas with different weed densities are segmented according to the optimal segmentation scales, and we manage to obtain the optimal subset of features and analyze the relationship between the extraction effects and the spatial distribution of weeds based on the UAV visible images, and then the extraction results were used to partition the degree of weed infestation so as to obtain a practical and efficient method for weed identification at different weed densities, with the aim of improving weed control efficiency and reducing the use of chemicals.

## 2. Materials and Methods

## 2.1. Study Area

The study area was farmland located in Wenjiang District, Chengdu City, Sichuan Province (103°50′51″E, 30°44′42″N), as shown in Figure 1 (The map is made based on the standard map downloaded from the standard map service website of the National Bureau of Surveying, Mapping and Geographic Information. http://bzdt.ch.mnr.gov.cn/ (accessed on 20 December 2022)). The total area was 150.56 m<sup>2</sup> and flat, was located in the hinterland of the Chengdu Plain, and had a simple geomorphology. Wenjiang District is contained in the subtropical humid climate zone and has four distinct seasons, a mild climate, and an annual average temperature of 16.0 °C, a climate suitable for rice growth.



Figure 1. Location of the study area.

## 2.2. Method

Figure 2 shows the protocol for this study. First, the UAV was used to collect images of the study area, and then the Pix4Dmapper was used to build a full-area Digital Orthophoto Map (DOM). Secondly, weed-dense (Dense) and weed-sparse areas (Sparse) were selected in the DOM of the study area, and then eCognition9.0 combined with ESP2 (Estimation of Scale Parameter 2) (Trimble Germany GmbH, Munich, Germany) [37] was used to determine the best segmentation scale for the two weed-dense areas to segment the images and calculate

the spectral, index, texture, and geometric feature values of each object after segmentation. Schemes S1–S4 were constructed based on these results, and scheme S5 was constructed by combining recursive feature elimination and cross-validation methods. The random forest method was used to classify the five schemes; then, the classification results and classification accuracy were compared to choose the best feature combination scheme. The classification results of the random forest, decision tree, K-nearest neighbor, and support vector machine methods based on the selected best feature combination scheme are shown. Furthermore, we analyzed the applicability of the best classification method in the experimental area for two weed densities.



Figure 2. Protocol for the study.

### 2.2.1. Data Acquisition and Preprocessing

Tillering stage, booting stage, heading stage, and ripening stage are the major stages in the growth of rice, and rice is formed during the heading stage and ripening stage. The tillering stage is the period from the beginning of tillering to the beginning of young spikelets, and rice starts fruiting at the booting stage [38]. Rice in the Chengdu Plain is in the max tillering number stage in mid-May, and farm management can avoid weeds competing with rice for nutrients in the next booting stage and heading stage, which is favorable to the growth of rice nutrient organs in May. In this study, the DJI Air 2S was used to collect visible light images from the study area. The UAV has a 1-inch CMOS visible sensor with a pixel size of 2.4 µm, a pixel value of 20 million, and a camera equivalent focal length of 22 mm. To avoid the shadows affecting detection results, we performed aerial photography on a day when it was cloudy and had good lighting conditions, so we chose to acquire images of the study area on 15 May 2022. When performing aerial photography, the flight altitude was 30 m, the overlap rate in the side direction and the overlap rate in the heading were both 80%, the speed was set to 5 m/s, and the shooting method was hover shooting to avoid blurring the photos due to a slow shutter speed during the flight. A total of 515 images of the study area were collected, and the size of each image was  $5472 \times 3648$  pixels. The original images collected by UAV were stitched together using Pix4Dmapper (Prilly, Switzerland) [39] to generate a DOM with a resolution of 0.55 cm.

## 2.2.2. Image Segmentation

OBIA classification mainly includes image segmentation, feature extraction, and image classification, and image segmentation directly affects classification accuracy. If the segmentation scale is too large, some of the smaller objects will be segmented into other ones. If the segmentation scale is too small, a more "broken segmentation" result will be produced [40]. Therefore, we used eCognition 9.0 and ESP2 to determine the best segmentation scale for the image. After testing, the weights of shape and spectrum were determined to be 0.1 and 0.5, respectively. ESP2 was used to quantitatively evaluate the homogeneity of pixels in the segmented object by calculating a local variance (LV) and rates of change (ROC). The changes in LV and ROC under different scale parameters are shown in Figure 3. The optimal scale is the local optimal scale when the ROC reaches its peak; therefore, in this study, the alternative scales for the dense area are 28, 48, 61, 81, 96, 107, and 118; the alternative scales for the parse area are 31, 44, 64, 86, 95, 105, and 118.



Figure 3. Changes in LV and ROC under different scale parameters.

The different scale segmentation effects are shown in Figure 4. It has been shown that weeds are more likely to be under-segmented when using larger scale parameters (i.e., a segmentation scale larger than 61 in dense areas and larger than 44 in sparse areas), resulting in the detection of more weeds and rice in the same patch. If the scale parameter is small (i.e., a segmentation scale of less than 61 in dense areas and less than 44 in sparse areas), the weeds are more likely to be over-segmented, resulting in the weed object being too fragmented. On the contrary, weeds can be clearly segmented and conform to homogeneity when the scale parameter is 61 for dense areas and 44 for sparse areas.



**Figure 4.** Different scale segmentation effects. The number below each image represents the scale parameter, and the blue ellipses in the picture indicate typical areas with poor segmentation results.

#### 2.2.3. Feature Extraction

The features selected in this study primarily include spectral, index, geometric, and texture features. The details are shown as follows:

- Spectral features (SPEC): The mean and standard deviation of three bands in the visible image (mean\_R, mean\_G, mean\_B; Std\_R, Std\_G, Std\_B), band maximum difference (Max\_diff), and brightness [41];
- (2) Index features (INDE): Difference enhanced vegetation index (DEVI), excess red index (EXR), excess green index (EXG), excess green minus excess red (EXGR), green to blue ratio index (GBRI), greenness vegetation index (GVI), modified green-red vegetation index (MGRVI), normalized green-blue difference index (NGBDI), normalized green-red difference index (NGRDI), red-green-blue vegetation index (RGBVI), and visible-band difference vegetation index (VDVI). The vegetation indices and formulas are shown in Table 1;

Table 1. Vegetation indices and formulas.

Vegetation Index	Formulas	Reference
DEVI	$(G + R + B) / 3 \times G$	[42]
EXG	$2 \times G - R - B$	[43]
EXGR	$2 \times G - R - B - (1.4 \times R - G)$	[44]
EXR	$1.4 \times R - G$	[45]
GBRI	B / G	[46]
GVI	G / (R + G + B)	[47]
MGRVI	$\left(\mathrm{G}^2-\mathrm{R}^2 ight)$ / $\left(\mathrm{G}^2\mathrm{+}\mathrm{R}^2 ight)$	[48]
NGBDI	(G - B) / (G + B)	[49]
NGRDI	(G - R) / (G + R)	[50]
RGBVI	$\left( \mathrm{G}^{2}-\mathrm{B} imes\mathrm{R} ight)$ / $\left( \mathrm{G}^{2}\mathrm{+B} imes\mathrm{R} ight)$	[51]
RGRI	Ř / G	[52]
VDVI	$(2 \times G - R - B)/(2 \times G + R + B)$	[53]

Note: R, G, and B represent the mean values of the red, green, and blue bands, respectively.

- (3) Geometric features (GEOM): Area, length, length/width, width, border length, number of pixels, volume, asymmetry, border index, compactness, density, elliptic fit, shape, index, roundness, and rectangular fit [54];
- (4) Texture features (GLCM): It contains the mean, standard deviation, entropy, homogeneity, dissimilarity, contract, correlation, and angular second moment [55]. The texture features and formulas are shown in Table 2.

Table 2. Texture features and formulas.

Texture Feature	Formulas		
Mean	$\sum_{i=1}^{N} \sum_{i=1}^{N} iP_{i,j}$		
Standard Deviation	$\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}P_{i,j} \times (i - Mean)^2}$		
Entropy	$\sum_{i=1}^{N} \sum_{j=1}^{N} P_{i,j} \times lnP_{i,j}$		
Homogeneity	$\sum_{i=1}^{N} \sum_{i=1}^{N} P_{i,j} \times \frac{1}{1+(i+j)^2}$		
Dissimilarity	$\sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{P}_{i,j} \times  i-j $		
Contract	$\sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{P}_{i,j} \times (i-j)^2$		
Correlation	$\sum_{i=1}^{N} \sum_{j=1}^{'N} \frac{(i\text{-mean}) \times (j\text{-mean}) \times P_{i,j}^2}{\text{Variance}^2}$		
Angular Second Moment	$\sum_{i=1}^{N} \sum_{j=1}^{N} P_{ij}^2$		

Note: where *I* and *j* are the rank co-ordinates of the element in the image, p(i, j) is the gray joint probability matrix, N is the order of the gray-level co-occurrence matrix, Mean is the mean value, and Variance is the standard deviation.

## 2.2.4. Sample Selection

The experimental sample selection included a training sample and test samples. To accurately assess the extraction performance of the method at different weed densities, a 6:4 ratio of training-to-test samples was selected for both the dense and sparse weed areas using ArcGIS10.8 (ESRI Inc., Redlands, CA, USA) [56]. Firstly, a  $3 \times 3$  fishnet was separately created for each area, then 200 sample points were randomly generated in each fishnet unit (1800 random points were included in each area). Secondly, 173 points were randomly selected as test samples for each area (i.e., 72 weed samples and 101 non-weed samples in the dense area and 80 weed samples and 93 non-weed samples in the sparse area). Finally, 240 sample points were selected as training samples based on the area ratio of weed-to-non-weed areas from the remaining sample points (1370 points) for each area (114 weed samples and 126 non-weed samples in the dense area; 100 weed samples and 140 non-weed samples in the sparse area) to prevent imbalance between the weed and non-weed training samples.

## 2.2.5. Feature Selection

Feature selection, also known as attribute selection, is the process of selecting N features from an existing feature set, M, to optimize specific evaluation metrics. It is also a process of selecting the most effective subset of features from the original set to reduce the dimensionality of the dataset, which is crucial for improving algorithm performance.

Due to the high feature dimensionality, recursive feature elimination (RFE) [57] can be used to grade the importance of all features, and cross-validation (CV) [58] can be used to obtain the optimal number of features, as shown in Figure 5 (here, all the image objects are used as samples for feature dimensionality reduction in order to obtain an accurate feature set in each area). The process of RFE and CV achieve feature selection and are shown below (RFECV): (1) Stages of RFE: (a) Model the current subset of 43 features and calculate the importance of each feature; (b) remove the least important feature or features and update the feature set; (c) repeat steps (a) and (b) until the importance rating of all features is complete. (2) Stages of CV: (a) Select the different number of features by the importance of the features in the RFE stage; (b) cross-validate the selected feature set; (c) determine the highest-rated feature number and complete feature selection.



Figure 5. Relationship between the number of features and accuracy.

The dense area and sparse area feature subsets were reduced from 43 to 17 and 19, respectively, after the feature optimization was completed, as shown in Figure 6. The importance of the preferred features in the different density areas is shown in Figure 7. The common features of the two areas are MGRVI, RGRI, NGRDI, EXR, EXGR, DEVI, GVI, GLCM\_Angular Second Moment (GLCM\_A2M), GLCM\_Entropy, GLCM\_Dissimilarity, Area, Number of pixels, Volume, and Border length.



**Figure 6.** (a) The proportion of the optimal features of different types in the sparse area; (b) the proportion of the optimal features of the different types in the dense area.



Figure 7. Importance of preferred features in areas of different densities.

## 2.2.6. Constructing the Experimental Scheme

Further research on the number of trees in the random forest model can help improve its classification performance. In this study, we tested from 1 to 500 trees under different schemes and evaluated their performance using various scores for each number of trees under the same scheme.

In this study, we designed four additional schemes to validate the effectiveness of the selected feature set. In addition, we created three schemes using the decision tree, K-nearest neighbors, and support vector machine algorithms based on the selected feature set to test the performance of different classifiers in the study areas. The details of the experimental schemes and the number of trees used for each scheme are presented in Table 3. After testing, the depth of both the random forest and the decision tree was 3, the type of the decision tree was "CART", the c-value, gamma, and the kernel in SVM were set to "3", 0.05, and "rbf", respectively, and the K-value in the KNN was set to "2". The relationship between the classification accuracy and the number of trees is illustrated in Figure 8.

Table 3. Experimental scheme and optimal number of trees.

Scheme	Classifier	Features	Number of Features (Sparse/Dense)	Number of Trees (Sparse/Dense)
S1	Random Forest	SPEC	8/8	135/105
S2	Random Forest	SPEC + INDE	20/20	65/106
S3	Random Forest	SPEC + INDE + GLCM	28/28	9/175
S4	Random Forest	SPEC + INDE + GLCM + GEOM	43/43	125/121
S5	Random Forest	RFECV	19/17	26/26
S6	SVM	RFECV	19/17	
S7	Decision Tree	RFECV	19/17	
S8	KNN	RFECV	19/17	—



**Figure 8.** (a) Relationship between accuracy and number of trees in the sparse area; (b) relationship between accuracy and number of trees in the dense area.

2.2.7. Accuracy Evaluation Index

A confusion matrix was used to evaluate the performance of the algorithms and visualize the classification accuracy for supervised classification. In this paper, we combine overall accuracy (OA) and the F1-score coefficient to analyze the classification results due to the presence of dataset imbalances. Here, metrics such as OA (presented in Equation (1)) and the F1-score (presented in Equation (2)) are used to evaluate the classification. The OA is the ratio of test points correctly classified to all test points, and the F1-score can be viewed as a weighted average of the precision and recall of the model, which takes into account both the precision and recall of the classification model [59,60]. By comparing the real category of the samples with the model's predicted results, the results can be classified into the following four cases: true positive (TP), where the predicted value of weeds is consistent with the real value; false positive (FP), where the actual situation is the non-weeds but is incorrectly predicted as weeds; false negative (FN), where the actual situation is the weeds but is incorrectly predicted as non-weeds.

$$OA = \frac{TP + TN}{TP + FP + TN + FN}$$
(1)

$$F1 - Score = 2 \times \frac{\frac{TP}{TP + FP} \times \frac{TP}{TP + FN}}{\frac{TP}{TP + FP} + \frac{TP}{TP + FN}}$$
(2)

#### 3. Results

#### 3.1. Results of the Different Feature Schemes

The density results based on the various schemes and the random forest are shown in Figure 9. There are obvious misclassifications in the classification results obtained by training the model using only scheme S1. This shows that weeds and non-weeds cannot be clearly distinguished using only spectral features, such as Figure 9(c1,c2). Therefore, we randomly counted the distribution of values with the same four index features in different areas (as shown in Figure 10) to determine whether the index features are useful for distinguishing weeds from non-weeds. The weed and non-weed values for each index showed large differences in different areas. So, the index features were applied in the experimental area to construct scheme S2, and the results for scheme S2 showed a more noticeable decrease in misclassifications, such as Figure 9(d1,d2). For the S3 and S4 schemes, we sequentially added texture features and geometric features to investigate whether the classification effect is affected by increasing feature dimensionality. In scheme S3, such as Figure 9(e1,e2), the spectral, index, and texture features were combined, and the extraction effect was further improved compared to schemes S1 and S2. In scheme S4, such as Figure 9(f1,f2), the high-dimensional features achieved a better extraction effect in a dense area but not in a sparse area. The reason is that in dense areas, the spatial distribution of weeds is denser, and the difference in features between the weed patches and their surroundings is more obvious, allowing the weeds to be recognized more easily. However, in areas with a sparse spatial distribution, the weed patches are scattered, and the difference between the features and the surroundings is less obvious, greatly reducing the effectiveness of the weed map extraction. In the S5 scheme, such as Figure 9(b1,b2), we obtained the preferred subset of features by performing feature dimensionality reduction through RFECV; the S5 scheme obtained better classification results than the S4 scheme in different areas, which indicates that high-feature dimensions are not necessarily conducive to improving the weed extraction results.

The classification accuracy of each scheme in areas of different densities is shown in Figure 11. When comparing the classification results of the various feature schemes (i.e., S1–S5), scheme S1 (only SPEC) had the lowest accuracy for both the dense and sparse areas. The feature subset preferred by RFECV in scheme S5 had the highest accuracy (i.e., the overall accuracy in an area of dense weeds was 95.38% and the F1-score was 94.20%; the overall accuracy and F1-score in those areas with dense and sparse weeds increased by 9.83% and 12.72% and 12.72% and 11.71%, respectively, when compared to the S1 scheme. In the study of schemes S1 to S4, features were added to each scheme in turn (e.g., scheme S2 had an index feature added compared to S1; S3 had texture features added compared to scheme S2). The OA and F1-score of most schemes improved in turn; however, the accuracy of the S4

scheme decreased when compared to S3 in two areas; this indicates that the high-dimensional feature sets are not conducive to weed extraction. In the study of scheme S5, all features were inputted into the RFECV algorithm to obtain a subset of preferred features with relatively low dimensionality. The OA and F1-score of scheme S5 improved by 1.74% and 2.00% and 0.58% and 0.33% when compared to scheme S4 in areas with dense and sparse weeds, respectively. Adding all features did not improve the accuracy; the increase in the number of features would have resulted in the phenomenon of feature redundancy, which is not beneficial to the training of the model. According to the visual evaluation and the results of accuracy verification, we concluded that scheme S5, constructed by the subset of features optimized by RFECV, had little misclassification in the sparse area, but the misclassification effect was closest to the true results.



**Figure 9.** The results of different densities based on the schemes and random forest. (**a1**) Dense experiment area of weeds; (**a2**) sparse experimental area containing weeds; (**b1**,**b2**) preferred subset of features by RFECV; (**c1**,**c2**) spectral features; (**d1**,**d2**) spectral features and index features; (**e1**,**e2**) spectral features, index features, and texture features; (**f1**,**f2**) spectral features, index features, texture features, and geometric features. The red circles represent misclassifications in schemes S1 to S4.



**Figure 10.** (**a**) The DEVI values of non-weeds and weeds in different areas; (**b**) the EXGR values of non-weeds and weeds in different areas; (**c**) the EXR values of non-weeds and weeds in different areas; (**d**) the RGRI values of non-weeds and weeds in different areas.



**Figure 11.** Extraction accuracy of different schemes based on random forest. TP and TN indicate the number of test points where the ground objects are correctly classified as weeds and non-weeds, respectively. FP and FN indicate the number of test points where the ground objects are incorrectly classified as weeds and non-weeds, respectively.

## 3.2. Results of the Different Machine Learning Algorithms

In this study, information about weeds was extracted for areas of different densities based on the preferred feature subset using random forest, decision tree, K-nearest neighbor, and support vector machine algorithms and eCognition 9.0 software, and the extraction results for each algorithm are shown in Figure 12. The results show that the extraction effect of random forest was best; the less broken patches are evident in the dense area, and similar broken patches occur in the sparse area due to its spatial distribution, but the weeds and non-weeds can be completely distinguished in the areas of different densities, which is more consistent with the true results. The decision tree algorithm was slightly less effective than random forest for weed information extraction, with a small number of misclassifications in the dense areas and more misclassifications in the sparse areas; the KNN and support vector machine algorithms showed poor extraction results, with many misclassifications.



Figure 12. Extraction results of different algorithms based on the preferred feature subset. (a1,a2) Original image; (b1,b2) random forest; (c1,c2) KNN; (d1,d2) decision tree; (e1,e2) SVM.

The classification accuracy of the various algorithms based on the preferred feature subset is shown in Figure 13. The random forest algorithm achieved the highest extraction accuracy, and the OA and F1-score were as high as 95.38% and 94.20% for the dense area and 91.33% and 90.57% for the sparse area. When compared to the other three algorithms, the OA for dense and sparse areas improved by 1.74–12.14% and 7.51–11.56%, respectively, while the F1-score improved by 1.89–17.40% and 7.85–10.80%, respectively. According to the visual evaluation and accuracy validation results, the random forest algorithm based on the preferred feature subset showed the best weed information extraction results. The classification results for the dense area were better than those for the sparse area for the four different algorithms because the spatial distribution of weeds results in strong or weak differences in the image features (i.e., when the spatial distribution is dense, the image features are strongly different, and when the spatial distribution is sparse, the image features are weak), and the random forest has an optimal performance in the case of an unbalanced dataset classification when compared to the other algorithms. This is because the sample features are preferred in this paper, and the features that are beneficial for extracting weeds are selected to optimize the sample set; secondly, random forests are composed of many decision trees, and the classification results are also voted by the individual trees, which can effectively balance the error in the classification of unbalanced datasets.



**Figure 13.** Extraction accuracy of different algorithms based on the preferred feature subset. RF represents the random forest algorithm; SVM represents the support vector machines algorithm; DT represents the decision tree algorithm; KNN represents the K-nearest neighbor algorithm.

## 4. Application of Classification Results

Although the random forest algorithm showed excellent results for weed information extraction from farmland, no clear direction about how to use the results to achieve accurate application of weeding was evident. Therefore, we used ArcGIS10.8 to analyze the classification results and calculate the proportion of the number of weed grids in each fishnet, and then the degree of weed infestation in each fishnet was partitioned according to the proportion of the number of weed grids in the fishnet, as shown in Figure 14.



Figure 14. Weed infestation level zoning.

## 5. Discussion

It is difficult to ensure the quality of the images collected by traditional remote sensing by satellites because of the significant effects of weather and the high costs, but UAVs are advantageously unaffected by cloudy weather and can acquire images with higher spatial resolution. With the increasing use of UAV automation, UAV remote sensing also plays an important role in applications such as disaster monitoring, resource investigation, terrain mapping, etc. Recently, machine learning algorithms have appeared in the public eye as representatives of intelligent data processing algorithms, which allow for the processing of large amounts of remote sensing data and solve the problem of the low efficiency of OBIA technology in image processing. At the same time, weed detection in each farmland area is very important, so in this study, we used UAV images to map the weed information extraction and the degree of weed infestation in areas with different weed densities in the same farmland. Farmers can apply herbicides based on a zoning map of weed infestation to reduce expenses and minimize environmental impact.

Weed information extraction in farmland using deep learning algorithms requires many labeled datasets to train the network. However, it can be challenging to obtain high-quality datasets in real life. Additionally, deep learning algorithms can struggle to effectively merge the multiple features of training samples. Therefore, in this study, we used random forest, decision tree, K-nearest neighbor, and support vector machine to extract information about weeds in farmland areas of different weed densities based on the object and optimal feature subsets, respectively. Currently, some researchers have studied combining OBIA with machine learning algorithms [13,61,62]. The high-accuracy extraction of urban impervious surfaces can be achieved by extracting various features such as nDSM, spectral features, index features, geometric features, and texture features [13]. Shrubs are extracted by combining OBIA with different algorithms and different feature sets, and the results show that the random forest algorithm with the best feature subset achieves the best classification accuracy [61]. In the accurate extraction of weeds, weeds in multiple Minnesota wetlands were identified through a combination of OBIA and three machine learning algorithms (i.e., artificial neural networks, random forests, and support vector mechanisms), and each showed an extraction accuracy of 91% [62]. Pena realized the high-accuracy extraction of weeds in maize fields using UAV images and the OBIA method [63]. However, their study did not address the redundancy of high-dimensional features. In our study, we managed to verify that increasing feature dimensionality does not necessarily improve weed extraction accuracy, and RFECV was used to address the redundant features from the feature set; the optimal number of base classifiers in the random forest is obtained experimentally.

For the extraction results of the several machine learning algorithms in the different weed densities we studied, the random forest algorithm achieved the best extraction accuracy. The reasons for this will be discussed below: (1) We used the ESP2 algorithm to achieve the accurate segmentation of the study area image. Systematic evaluations of the segmentation effects at different scales were performed, which enabled the determination of the optimal segmentation scale. This approach ensured that the coverings located in the object were homogeneous and maximized the differences in features within the object of different coverings. (2) We used RFECV to eliminate redundant features. The features are the direct criteria for classification in this paper, while the number of features and the determination of the feature set can directly affect the extraction effect of the algorithms. In view of this, the commonly used RFECV method was used to eliminate redundant features from the feature set; the feature sets in the dense area and sparse area were reduced from 43 to 17 and 19, respectively, and the extraction accuracy of the dense area and sparse area showed that the feature set with its redundant features removed (i.e., scheme S5) performed better than the full feature set (scheme S4). In this study, eliminating redundant features improved the accuracy of the random forest algorithm. (3) We optimized the number of base classifier parameters in the random forest algorithm. The number of base classifiers is the most important parameter in the random forest algorithm. In this study, the number of base classifiers was iterated from 1 to 500 with an optimal feature subset, and the prediction scores of the model were calculated after each training session. Then, the highest prediction score corresponding to the number of base classifiers was used as the final parameter.

While the accuracy achieved in this study was sufficient for practical applications, some limitations and possible errors that hinder the wider application of the research results still exist: (1) The alternative scales were selected based on ESP2; however, manual visual interpretation was still required to select the optimal scale from the alternative scales, which seriously affected the level of automation in removing weeds from farmland. Therefore, the question of how to automatically obtain the optimal scale for image segmentation is an important direction for subsequent research. Additionally, the process of selecting the optimal scale from the alternative scales through manual visual discrimination cannot completely avoid subjectivity, which may bring some errors to selecting the segmentation scale, which may influence the classification result. (2) Weed is an invasive species in farmland, which means that it is not allowed to grow in large numbers, and this limits the selection of weed samples during model training. At the same time, due to the scarcity of weed samples, this may lead to a poorer classification result. (3) The images in this paper were collected in mid-May and were limited to the period when the rice and weed growth in the farmland was in full bloom, and the features differed greatly, whereas in the early stage of weed and rice growth, the features were less different, and it was difficult to determine whether the method was still applicable.

In combination with the tested classifiers and optimized feature sets (by RFECV), the application of OBIA and random forest was able to achieve high OA and F1-scores, and this was used to explore whether the increase in feature dimension is beneficial for the extraction of weeds. Currently, there are several feature downscaling methods that exist. For example, Jing Xia et al. [64] used the max-relevance and min-redundancy algorithms for feature dimensionality reduction; Hongda Li et al. [65] used the unified manifold approximation and projection method and support vector machine to achieve high-precision terrain classification. An improvement in the method may be able to achieve higher extraction accuracy. Another approach that may improve the classification results is the use of boosting algorithms; the core of the boosting algorithm is that the training set remains the same for each round, yet the weight of each sample in the training set changes in the classifier. The representative classifiers of the boosting algorithm are XGBoost, Catboost, gradient boosting decision tree, etc. Among them, the XGBoost algorithm performs well in the application of deforestation trace monitoring [66] and maize lodging detection [67]. However, finding the suitable ensemble classification method for a particular dataset remains a burdensome task because multiple arrangements of classifiers and feature reduction methods can be coupled. Therefore, it is worth exploring this to test the classification effectiveness of the ensemble approach in order to perform weed extraction from UAV images.

#### 6. Conclusions

In order to control farmland weeds, the presented method makes full use of the spectral features, index features, texture features, and geometric features of UAV visible images to combine with an OBIA-random forest algorithm to extract information about weeds from experimental farmland areas with different weed densities. In the study, the random forest algorithm was used to test five schemes to select the most accurate one, and based on that scheme, four machine learning algorithms, i.e., random forest, K-nearest neighbor, support vector machine, and decision tree, were used to extract weed information. The extraction results were compared, and weed infestation degree partition mapping was performed based on the best classification results. The following conclusions were drawn: (1) The method using a UAV and the OBIA-random forest algorithm to achieve the accurate and automated management of weeds for farmland is feasible, and the strength of image feature difference is affected by its spatial distribution; the denser the spatial distribution, the worse the extraction effect for farmland. (2) In the study of the S1–S4 schemes, increasing the

dimensionality of the feature can improve the accuracy of weed information extraction, but the high dimensionality of the feature will lead to less accuracy, and using a feature elimination method, such as RFECV, can effectively eliminate features of low importance in the feature set to obtain the optimal subset of features (S5 scheme). The OA and F1-scores of the preferred feature subset reached 95.38% and 94.20% in the dense areas and 91.33% and 90.57% in sparse areas. (3) With the optimized feature subset, the random forest algorithm was clearly superior when compared to machine learning algorithms such as KNN, support vector machine, and decision tree, and the OA of the dense area and sparse area improved by 1.74–12.14%, and 7.51–11.56%, respectively, and the F1-scores improved by 1.89–17.40% and 7.85–10.80%, respectively. The results show that the RFECV method can effectively remove redundant features and improve the accuracy of model weed extraction.

Author Contributions: Conceptualization, C.F., W.Z., H.D. and H.Z.; methodology, C.F., W.Z., H.D. and H.Z.; software, C.F., L.T., Y.Z. and Z.Z.; validation, W.Z. and L.D.; analysis, W.Z. and H.D.; investigation, H.Z., L.D., W.Z. and L.D.; resources, Z.Z.; data curation, C.F., Y.Z., Z.Z., L.T. and H.D.; writing—original draft preparation, C.F., W.Z., H.D. and H.Z.; writing—review and editing, C.F., W.Z., H.Z. and H.D.; visualization, C.F., H.D. and L.D.; supervision, W.Z. and H.D.; project administration, H.D.; funding acquisition, H.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Science and Technology Department of Tibet Key Project (XZ202201ZY0003G), the Science and Technology Department of Tibet Key Project (XZ202001ZY0056G), and the Sichuan Education Department Natural Science Key Project (18ZA0047).

**Data Availability Statement:** Data are available upon request due to restrictions, e.g., privacy or ethics. The data presented in this study are available upon request from the corresponding author.

Acknowledgments: Thanks to scikit-learn for the original RFECV code.

**Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

#### References

- 1. Jin, X.; Liu, T.; McCullough, P.E.; Chen, Y.; Yu, J. Evaluation of convolutional neural networks for herbicide susceptibility-based weed detection in turf. *Front. Plant Sci.* 2023, *14*, 1096802. [CrossRef] [PubMed]
- Zhang, X.; Cui, J.; Liu, H.; Han, Y.; Ai, H.; Dong, C.; Zhang, J.; Chu, Y. Weed Identification in Soybean Seedling Stage Based on Optimized Faster R-CNN Algorithm. *Agriculture* 2023, 13, 175. [CrossRef]
- 3. Rawat, J.S.; Joshi, R.C. Remote-sensing and GIS-based landslide-susceptibility zonation using the landslide index method in Igo River Basin, Eastern Himalaya, India. *Int. J. Remote Sens.* 2012, *33*, 3751–3767. [CrossRef]
- 4. Zou, L.; Wang, C.; Zhang, H.; Wang, D.; Tang, Y.; Dai, H.; Zhang, B.; Wu, F.; Xu, L. Landslide-prone area retrieval and earthquake-inducing hazard probability assessment based on InSAR analysis. *Landslides* **2023**, *20*, 1989–2002. [CrossRef]
- Asadzadeh, S.; Oliveira, W.J.D.; Souza Filho, C.R.D. UAV-based remote sensing for the petroleum industry and environmental monitoring: State-of-the-art and perspectives. J. Pet. Sci. Eng. 2022, 208, 109633. [CrossRef]
- Yang, C.; Shen, R.; Yu, D.; Liu, R.; Chen, J. Forest disturbance monitoring based on the time-series trajectory of remote sensing index. J. Remote Sens. 2013, 17, 1246–1263. [CrossRef]
- He, X.Y.; Ren, C.Y.; Chen, L.; Wang, Z.; Zheng, H. The progress of forest ecosystems monitoring with remote sensing techniques. *Sci. Geogr. Sin.* 2018, 38, 997–1011. [CrossRef]
- 8. Gao, F.; Anderson, M.; Daughtry, C.; Karnieli, A.; Hively, D.; Kustas, W. A within-season approach for detecting early growth stages in corn and soybean using high temporal and spatial resolution imagery. *Remote Sens. Environ.* **2020**, 242, 111752. [CrossRef]
- Guilherme Teixeira Crusiol, L.; Sun, L.; Chen, R.; Sun, Z.; Zhang, D.; Chen, Z.; Wuyun, D.; Rafael Nanni, M.; Lima Nepomuceno, A.; Bouças Farias, J.R. Assessing the potential of using high spatial resolution daily NDVI-time-series from Planet CubeSat images for crop monitoring. *Int. J. Remote Sens.* 2021, 42, 7114–7142. [CrossRef]
- Lu, Y.; Chibarabada, T.P.; Ziliani, M.G.; Onema, J.K.; McCabe, M.F.; Sheffield, J. Assimilation of soil moisture and canopy cover data improves maize simulation using an under-calibrated crop model. Agric. *Water Manag.* 2021, 252, 106884. [CrossRef]
- 11. Li, D.; Li, M. Research advance and application prospect of unmannedaerial vehicle remote sensing system. *Geomat. Inf. Sci. Wuhan Univ.* **2014**, *39*, 505–513. [CrossRef]
- 12. Stroppiana, D.; Villa, P.; Sona, G.; Ronchetti, G.; Candiani, G.; Pepe, M.; Busetto, L.; Migliazzi, M.; Boschetti, M. Early season weed mapping in rice crops using multi-spectral UAV data. *Int. J. Remote Sens.* **2018**, *39*, 5432–5452. [CrossRef]
- 13. Ye, Z.; Guo, Q.; Zhang, J.; Zhang, H.; Deng, H. Extraction of urban impervious surface based on the visible images of UAV and OBIA-RF algorithm. *Trans. Chin. Soc. Agric. Eng.* **2022**, *38*, 225–234. [CrossRef]

- Yonah, I.B.; Mourice, S.K.; Tumbo, S.D.; Mbilinyi, B.P.; Dempewolf, J. Unmanned aerial vehicle-based remote sensing in monitoring smallholder, heterogeneous crop fields in Tanzania. *Int. J. Remote Sens.* 2018, 39, 5453–5471. [CrossRef]
- Noguera, M.; Aquino, A.; Ponce, J.M.; Cordeiro, A.; Silvestre, J.; Arias-Calderón, R.; Da Encarnação Marcelo, M.; Jordão, P.; Andújar, J.M. Nutritional status assessment of olive crops by means of the analysis and modelling of multispectral images taken with UAVs. *Biosyst. Eng.* 2021, 211, 1–18. [CrossRef]
- 16. Qin, Z.; Chang, Q.; Xie, B.; Shen, J. Rice leaf nitrogen content estimation based on hysperspectral imagery of UAV in Yellow River diversion irrigation district. *Trans. Chin. Soc. Agric. Eng.* **2016**, *32*, 77–85. [CrossRef]
- 17. Atik, S.O.; Ipbuker, C. Integrating Convolutional Neural Network and Multiresolution Segmentation for Land Cover and Land Use Mapping Using Satellite Imagery. *Appl. Sci.* 2021, *11*, 5551. [CrossRef]
- Guirado, E.; Blanco-Sacristán, J.; Rodríguez-Caballero, E.; Tabik, S.; Alcaraz-Segura, D.; Martínez-Valderrama, J.; Cabello, J. Mask R-CNN and OBIA Fusion Improves the Segmentation of Scattered Vegetation in Very High-Resolution Optical Sensors. 2021, 21, 320. [CrossRef]
- Ye, Z.; Yang, K.; Lin, Y.; Guo, S.; Sun, Y.; Chen, X.; Lai, R.; Zhang, H. A comparison between Pixel-based deep learning and Object-based image analysis (OBIA) for individual detection of cabbage plants based on UAV Visible-light images. *Comput. Electron. Agric.* 2023, 209, 107822. [CrossRef]
- 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]
- 21. Blaschke, T. Object based image analysis for remote sensing. ISPRS-J. Photogramm. Remote Sens. 2010, 65, 2–16. [CrossRef]
- Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; Van der Meer, F.; Van der Werff, H.; Van Coillie, F. Geographic object-based image analysis–towards a new paradigm. *ISPRS-J. Photogramm. Remote Sens.* 2014, 87, 180–191. [CrossRef]
- Liu, T.; Abd-Elrahman, A.; Zare, A.; Dewitt, B.A.; Flory, L.; Smith, S.E. A fully learnable context-driven object-based model for mapping land cover using multi-view data from unmanned aircraft systems. *Remote Sens. Environ.* 2018, 216, 328–344. [CrossRef]
- 24. Tompalski, P.; White, J.C.; Coops, N.C.; Wulder, M.A. Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data. *Remote Sens. Environ.* **2019**, 227, 110–124. [CrossRef]
- 25. Peña-Barragán, J.M.; Ngugi, M.K.; Plant, R.E.; Six, J. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sens. Environ.* **2011**, *115*, 1301–1316. [CrossRef]
- Phiri, D.; Morgenroth, J.; Xu, C.; Hermosilla, T. Effects of pre-processing methods on Landsat OLI-8 land cover classification using OBIA and random forests classifier. *Int. J. Appl. Earth Obs. Geoinf.* 2018, 73, 170–178. [CrossRef]
- 27. Bao, F.; Huang, K.; Wu, S. The retrieval of aerosol optical properties based on a random forest machine learning approach: Exploration of geostationary satellite images. *Remote Sens. Environ.* **2023**, *286*, 113426. [CrossRef]
- Loozen, Y.; Rebel, K.T.; de Jong, S.M.; Lu, M.; Ollinger, S.V.; Wassen, M.J.; Karssenberg, D. Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the random forests method. *Remote Sens. Environ.* 2020, 247, 111933. [CrossRef]
- 29. Wang, L.J.; Kong, Y.R.; Yang, X.D.; Xu, Y.; Liang, L.; Wang, S.G. Classification of land use in farming areas based on feature optimization random forest algorithm. *Trans. CSAE* **2020**, *36*, 244–250. [CrossRef]
- Yang, Y.; Zhang, W. Research on GF-2 Image Classification Based on Feature Optimization Random Forest Algorithm. Spacecr. Recovery Remote Sens. 2022, 43, 115–126. [CrossRef]
- 31. Borra-Serrano, I.; Peña, J.M.; Torres-Sánchez, J.; Mesas-Carrascosa, F.J.; López-Granados, F. Spatial quality evaluation of resampled unmanned aerial vehicle-imagery for weed mapping. *Sensors* 2015, *15*, 19688–19708. [CrossRef] [PubMed]
- 32. Gao, J.; Liao, W.; Nuyttens, D.; Lootens, P.; Vangeyte, J.; Pižurica, A.; He, Y.; Pieters, J.G. Fusion of pixel and object-based features for weed mapping using unmanned aerial vehicle imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, 67, 43–53. [CrossRef]
- Castillejo-González, I.L.; Peña-Barragán, J.M.; Jurado-Expósito, M.; Mesas-Carrascosa, F.J.; López-Granados, F. Evaluation of pixel- and object-based approaches for mapping wild oat (*Avena sterilis*) weed patches in wheat fields using QuickBird imagery for site-specific management. *Eur. J. Agron.* 2014, 59, 57–66. [CrossRef]
- Zhao, H.; Cao, Y.; Yue, Y.; Wang, H. Field weed recognition based on improved DenseNet. *Trans. CSAE* 2021, 37, 136–142. [CrossRef]
- 35. Chen, J.; Wang, H.; Zhang, H.; Luo, T.; Wei, D.; Long, T.; Wang, Z. Weed detection in sesame fields using a YOLO model with an enhanced attention mechanism and feature fusion. *Comput. Electron. Agric.* **2022**, 202, 107412. [CrossRef]
- Wang, C.; Wu, X.; Zhang, Y.; Wang, W. Recognizing weeds in maize fields using shifted window Transformer network. *Trans. Chin. Soc. Agric. Eng.* 2022, 38, 133–142. [CrossRef]
- Rana, M.; Kharel, S. FEATURE Extraction for Urban and Agricultural Domains Using Ecognition Developer. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2019, XLII-3/W6, 609–615. [CrossRef]
- Sheng, R.T.; Huang, Y.; Chan, P.; Bhat, S.A.; Wu, Y.; Huang, N. Rice Growth Stage Classification via RF-Based Machine Learning and Image Processing. *Agriculture* 2022, 12, 2137. [CrossRef]
- 39. Lagogiannis, S.; Dimitriou, E. Discharge Estimation with the Use of Unmanned Aerial Vehicles (UAVs) and Hydraulic Methods in Shallow Rivers. *Water* 2021, *13*, 2808. [CrossRef]
- 40. Yu, H.; Zhang, S.Q.; Kong, B.; Li, X. Optimal segmentation scale selection for object-oriented remote sensing image classification. J. Image Graph. 2010, 15, 352–360.

- 41. Tian, J.; Wang, L.; Yin, D.; Li, X.; Diao, C.; Gong, H.; Shi, C.; Menenti, M.; Ge, Y.; Nie, S.; et al. Development of spectral-phenological features for deep learning to understand Spartina alterniflora invasion. *Remote Sens. Environ.* **2020**, 242, 111745. [CrossRef]
- Zhou, T.; Hu, Z.; Han, J.; Zhang, H. Green vegetation extraction based on visible light image of UAV. *China Environ. Sci.* 2021, 41, 2380–2390. [CrossRef]
- Sánchez-Sastre, L.F.; Alte Da Veiga, N.M.; Ruiz-Potosme, N.M.; Carrión-Prieto, P.; Marcos-Robles, J.L.; Navas-Gracia, L.M.; Martín-Ramos, P. Assessment of RGB vegetation indices to estimate chlorophyll content in sugar beet leaves in the final cultivation stage. *AgriEngineering* 2020, 2, 128–149. [CrossRef]
- Yang, B.; Wang, M.; Sha, Z.; Wang, B.; Chen, J.; Yao, X.; Cheng, T.; Cao, W.; Zhu, Y. Evaluation of aboveground nitrogen content of winter wheat using digital imagery of unmanned aerial vehicles. *Sensors* 2019, 19, 4416. [CrossRef] [PubMed]
- 45. Jiang, J.; Zhang, Z.; Cao, Q.; Tian, Y.; Zhu, Y.; Cao, W.; Liu, X. Use of a digital camera mounted on a consumer-grade unmanned aerial vehicle to monitor the growth status of wheat. *J. Nanjing Agric. Univ.* **2019**, *42*, 622–631. [CrossRef]
- Sellaro, R.; Crepy, M.; Trupkin, S.A.; Karayekov, E.; Buchovsky, A.S.; Rossi, C.; Casal, J.J. Cryptochrome as a sensor of the blue/green ratio of natural radiation in Arabidopsis. *Plant Physiol.* 2010, 154, 401–409. [CrossRef]
- Liu, Y.; Chen, Y.; Yue, D.; Feng, Z. Information extraction of urban green space based on UAV remote sensing image. *Sci. Surv. Mapp.* 2017, 42, 59–64. [CrossRef]
- Bendig, J.; Yu, K.; Aasen, H.; Bolten, A.; Bennertz, S.; Broscheit, J.; Gnyp, M.L.; Bareth, G. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Obs. Geoinf.* 2015, 39, 79–87. [CrossRef]
- Zhao, J.; Yang, H.B.; Lan, Y.B.; Lu, L.Q.; Jia, P.; Li, Z.M. Extraction method of summer corn vegetation coverage based on visible light image of unmanned aerial vehicle. J. Agric. Mach. 2019, 50, 232–240. [CrossRef]
- Elazab, A.; Bort, J.; Zhou, B.; Serret, M.D.; Nieto-Taladriz, M.T.; Araus, J.L. The combined use of vegetation indices and stable isotopes to predict durum wheat grain yield under contrasting water conditions. *Agric. Water Manag.* 2015, 158, 196–208. [CrossRef]
- 51. Li, C.C.; Niu, Q.L.; Yang, G.J.; Feng, H.; Liu, J.; Wang, Y. Estimation of leaf area index of soybean breeding materials based on UAV digital images. *Trans. Chin. Soc. Agric. Mach.* **2017**, *48*, 147. [CrossRef]
- 52. Gamon, J.A.; Surfus, J.S. Assessing leaf pigment content and activity with a reflectometer. *New Phytol.* **1999**, *143*, 105–117. [CrossRef]
- Wang, X.; Wang, M.; Wang, S.; Wu, Y. Extraction of vegetation information from visible unmanned aerial vehicle images. *Trans. Chin. Soc. Agric. Eng.* 2015, 31, 152–159. [CrossRef]
- 54. Yu, N.; Li, L.; Schmitz, N.; Tian, L.F.; Greenberg, J.A.; Diers, B.W. Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. *Remote Sens. Environ.* **2016**, *187*, 91–101. [CrossRef]
- Batool, F.E.; Attique, M.; Sharif, M.; Javed, K.; Nazir, M.; Abbasi, A.A.; Iqbal, Z.; Riaz, N. Offline signature verification system: A novel technique of fusion of GLCM and geometric features using SVM. *Multimed. Tools Appl.* 2020. [CrossRef]
- Mirzahossein, H.; Sedghi, M.; Motevalli Habibi, H.; Jalali, F. Site selection methodology for emergency centers in Silk Road based on compatibility with Asian Highway network using the AHP and ArcGIS (case study: I. R. Iran). *Innov. Infrastruct. Solut.* 2020, 5, 113. [CrossRef]
- Lin, X.; Yang, F.; Zhou, L.; Yin, P.; Kong, H.; Xing, W.; Lu, X.; Jia, L.; Wang, Q.; Xu, G. A support vector machine-recursive feature elimination feature selection method based on artificial contrast variables and mutual information. *J. Chromatogr. B* 2012, 910, 149–155. [CrossRef]
- Soper, D.S. Greed Is Good: Rapid Hyperparameter Optimization and Model Selection Using Greedy k-Fold Cross Validation. *Electronics* 2021, 10, 1973. [CrossRef]
- Alcantara, L.M.; Schenkel, F.S.; Lynch, C.; Oliveira Junior, G.A.; Baes, C.F.; Tulpan, D. Machine learning classification of breeding protocol descriptions from Canadian Holsteins. J. Dairy Sci. 2022, 105, 8177–8188. [CrossRef]
- 60. Ye, S.; Pontius, R.G.; Rakshit, R. A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. *ISPRS-J. Photogramm. Remote Sens.* **2018**, *141*, 137–147. [CrossRef]
- Li, Z.; Ding, J.; Zhang, H.; Feng, Y. Classifying individual shrub species in UAV images-a case study of the gobi region of Northwest China. *Remote Sens.* 2021, 13, 4995. [CrossRef]
- Anderson, C.J.; Heins, D.; Pelletier, K.C.; Knight, J.F. Improving Machine Learning Classifications of Phragmites australis Using Object-Based Image Analysis. *Remote Sens.* 2023, 15, 989. [CrossRef]
- 63. Pena, J.M.; Torres-Sanchez, J.; de Castro, A.I.; Kelly, M.; Lopez-Granados, F. Weed mapping in early-season maize fields using object-based analysis of unmanned aerial vehicle (UAV) images. *PLoS ONE* **2013**, *8*, e77151. [CrossRef] [PubMed]
- Jing, X.; Zou, Q.; Yan, J.; Dong, Y.; Li, B. Remote Sensing Monitoring of Winter Wheat Stripe Rust Based on mRMR-XGBoost Algorithm. *Remote Sens.* 2022, 14, 756. [CrossRef]
- Li, H.; Cui, J.; Zhang, X.; Han, Y.; Cao, L. Dimensionality Reduction and Classification of Hyperspectral Remote Sensing Image Feature Extraction. *Remote Sens.* 2022, 14, 4579. [CrossRef]

- Bhagwat, R.U.; Uma Shankar, B. A novel multilabel classification of remote sensing images using XGBoost. In Proceedings of the 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India, 29–31 March 2019; pp. 1–5. [CrossRef]
- 67. Han, L.; Yang, G.; Yang, X.; Song, X.; Xu, B.; Li, Z.; Wu, J.; Yang, H.; Wu, J. An explainable XGBoost model improved by SMOTE-ENN technique for maize lodging detection based on multi-source unmanned aerial vehicle images. *Comput. Electron. Agric.* **2022**, *194*, 106804. [CrossRef]

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