

Article

Automated Detection Method to Extract *Pedicularis* Based on UAV Images

Wuhua Wang ¹, Jiakui Tang ^{1,2,*}, Na Zhang ^{1,2}, Xuefeng Xu ¹, Anan Zhang ¹ and Yanjiao Wang ¹

¹ College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
² Yanshan Earth Key Zone and Surface Flux Observation and Research Station, University of Chinese Academy of Sciences, Beijing 101408, China
* Correspondence: jktang@ucas.ac.cn

Abstract: *Pedicularis* has adverse effects on vegetation growth and ecological functions, causing serious harm to animal husbandry. In this paper, an automated detection method is proposed to extract *Pedicularis* and reveal the spatial distribution. Based on unmanned aerial vehicle (UAV) images, this paper adopts logistic regression, support vector machine (SVM), and random forest classifiers for multi-class classification. One-class SVM (OCSVM), isolation forest, and positive and unlabeled learning (PUL) algorithms are used for one-class classification. The results are as follows: (1) The accuracy of multi-class classifiers is better than that of one-class classifiers, but it requires all classes that occur in the image to be exhaustively assigned labels. Among the one-class classifiers that only need to label positive or positive and labeled data, the PUL has the highest F score of 0.9878. (2) PUL performs the most robustly to change features in one-class classifiers. All one-class classifiers prove that the green band is essential for extracting *Pedicularis*. (3) The parameters of the PUL are easy to tune, and the training time is easy to control. Therefore, PUL is a promising one-class classification method for *Pedicularis* extraction, which can accurately identify the distribution range of *Pedicularis* to promote grassland administration.



Citation: Wang, W.; Tang, J.; Zhang, N.; Xu, X.; Zhang, A.; Wang, Y. Automated Detection Method to Extract *Pedicularis* Based on UAV Images. *Drones* **2022**, *6*, 399. <https://doi.org/10.3390/drones6120399>

Academic Editor: Adam Cross

Received: 17 November 2022

Accepted: 5 December 2022

Published: 6 December 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/>).

Keywords: one-class classification; positive and unlabeled learning (PUL); vegetation extraction; semi-supervised learning (SSL); unmanned aerial vehicle (UAV); *Pedicularis*

1. Introduction

Spatio-temporal variations in grassland species composition are crucial for grassland health evaluation, local ecosystem natural changes analysis, and grassland monitoring strategy formulation [1]. *Pedicularis* plants are most commonly known as lousewort (*Pedicularis* is the Latin word meaning louse). The flowering period is from June to August, and the fruit period is from July to September [2]. It was found that there was a large outbreak of *Pedicularis* in the Bayinbuluke grassland (the second biggest grassland in China), which had been infesting an area of 4.11×10^4 hm². This phenomenon has altered the community structure and composition of local grasslands, decreasing edible forage and affecting grassland animal husbandry development [3]. Therefore, it is necessary to develop a reliable method to identify *Pedicularis*.

Remote sensing serves as a key technology in ground object identification, inversion of surface parameters, and complex attribute analysis [4]. It uses different types of sensors on satellites, aviation aircraft, and unmanned aerial vehicles that provide various data related to the Earth at different spatial scales [5]. The advantage of satellite remote sensing is that it can acquire global data and cover a large area in one imaging, which is a great advantage for large-scale and macroscopic data collection and analysis. However, satellite remote sensing images are affected by the weather, and the spatial resolution is difficult to meet the requirements of refinement. Even though there are high-resolution commercial satellites available, their images are expensive to acquire such as PlanetScope, Maxar, and

Worldview. In recent years, UAVs have become more and more popular in many industries, such as 3D real-world modeling [6], communications [7], transport [8], and remote sensing. UAV remote sensing applications have attracted the interest of scientists because they are not affected by the atmosphere. Due to low flight height, the spatial resolution of images can reach 0.01 m or even higher, and their ability to obtain detailed information regarding specific targets is strong. The disadvantage is that the coverage area per image is very small, and the cost is high. So far, no studies have been carried out on the remote sensing of *Pedicularis* with drones. A few studies have used drones for remote sensing of toxic plants, such as *Hogweed* [9], *Rumex obtusifolius* [10], and *Oxytropis ochrocephala* [11]. However, none of them used the PUL algorithm. UAV-borne multispectral sensors used to obtain images with higher spatial resolution and more spectral bands have outstanding application prospects in *Pedicularis* recognition, especially when distributed in small and scattered patches. To restrain the spread of *Pedicularis* and eliminate its harmful effects, determining how to map the detailed distribution of *Pedicularis* is becoming urgent for grassland administration. This is what attracted us to conduct research on UAV remote sensing.

Remote sensing data have been widely applied in identifying ground objects. An image classification task aims to extract class information from the input. The classification task is especially daunting because most supervised learning projects require a sufficiently large number of training samples. Nevertheless, the definition and acquisition of reference data are often key issues [12]. As a result, classification algorithms for various tasks have been developed and applied in different contexts. Most studies used the multi-class classification method when classifying land cover types, such as support vector machines, decision trees, logistic regression, maximum likelihood, and deep neural networks [13,14]. These models have been used for a long time in land use and land cover classification [15,16], crop classification, and yield estimation [17]. For example, polynomial logistic regression based on semi-supervised learning was proposed for hyperspectral image classification by Shah, S.T.H. et al. [18]. Bo, S.K. et al. [19] proposed a method of dividing data using multi-class classifiers to extract a single land cover type. Muñoz-Marí J. [20] combined geographic weighted regression and logistic regression to analyze spatial changes in classification results of remote sensing images. Li L.H. et al. [21] used an object-oriented random forest algorithm to identify the forest using remote sensing data from the GaoFen-2 satellite that was launched in 2014 by China. Zhao, C. [22] transformed multi-class classification into a binary classification problem in extracting Chinese mangroves, improving classification accuracy by reducing the sample size and feature selection. Numerous studies classified land cover types and selected features based on random forest or the combination of random forest and neural networks [23–25]. Traditional supervised classification methods must mark all land cover types. However, when there is only one land cover type of interest and other types are not considered, marking all the classes is time-consuming and laborious, especially when high-resolution images are used [16,26,27]. Thus, it is of great necessity to cultivate one-class classifiers to extract specific land cover types [28].

One-class classifiers only need to know the feature data of the target of interest to distinguish this specific target from other land cover classes [29,30]. One-class classification is more suitable for extracting specific objects than multi-class classification. In one-class classification, negative classes (non-interested classes) either do not exist or are incorrectly sampled. Numerous one-class classifiers to extract specific ground targets are proposed [19]. The existing one-class classification methods can be divided into the following categories: density estimation method, prior probability estimation method, clustering method, and support domain-based method [31]. All one-class classifiers use the class of interest as the target class and the other class as outliers. One-class SVM (OCSVM) is commonly used and performs well among these classifiers. The OCSVM method was initially introduced by estimating the support of a high-dimensional distribution [20,32]. The classifier has been proven helpful in document classification, land cover classification, and land cover change detection [32]. However, a drawback of OCSVM is that it is sensitive to free parameters that are difficult to adjust. Another is the isolation forest designed for anomaly

detection, which has been developed in recent years. It is used to monitor hyperspectral data for anomalies and changes in land cover types [33,34]. However, isolation forest is rarely applied in remote sensing, particularly in classification. So it is necessary to test the accuracy of isolated forests in one-class classification extraction. These one-class classifiers only require positive class samples, which not only improves efficiency but also lowers sampling labor costs.

In addition to labeled samples, unlabeled samples also provide helpful information for constructing classifiers. Since the accuracy of traditional one-class classifiers could not meet the requirement of reasonable classification, the PUL classifier was developed in recent years. Wan Bo et al. [35] used the PUL method to extract urban areas of the United States from MODIS data, and the mapping accuracy reached 92.9%. R Liu et al. [36] used various PUL classifiers to test the extraction of urban regions in images with different resolutions. Li, W.K. et al. [37] compared the performance of PUL and other one-class classifiers in extracting different land cover types. They found that in urban areas, the extraction accuracy of PUL was 15% higher than OCSVM; in grasslands, the extraction accuracy of PUL was 10% higher than that of OCSVM. Li, W.K. et al. [38] proposed a PUL with a constraints algorithm based on binary classifiers for linear and fractional linear models, which increased the accuracy of PUL extraction of one-class classification. Lei L. et al. [39] proposed a deep one-class crop framework that includes a deep one-class crop extraction module and a one-class crop extraction loss module for large-scale one-class crop mapping. According to the literature review, one-class classification and PUL are primarily applied to urban buildings, large land targets, and monitoring rivers. In contrast, identifying single vegetation is relatively tricky due to the similarity of spectral features between vegetation.

This study explores the effectiveness of using UAV images to identify *Pedicularis* as a single grassland type [40]. We compare the ability of multi-class classification and one-class classification methods to extract *Pedicularis* and carry out performance evaluation of multiple algorithms. Finally, the feature importance ranking based on one-class classification algorithms for extracting *Pedicularis* is provided.

2. Material and Methods

2.1. Study Area

In recent years, *Pedicularis* has been out of control in Xinjiang, China, affecting the area's ecology and resulting in financial losses for the animal husbandry industry [41]. As shown in Figure 1, this study was carried out in the Swan Lake area of the Bayinbuluke grassland at an altitude of 2400 m [42]. Here, grassland vegetation is rich in species, and the grass composition is mainly formed of *Kobresia capillifolia*, *Carex* spp., *Stipa purpurea* Griseb, *Agropyron cristatum* (L.) Gaertn, etc. It is a well-established ecosystem with rivers, lakes, snowy mountains, and other water bodies [4].

2.2. Datasets and Pre-Processing

The study used two UAV systems to collect data from the same area on 7 August 2019, namely UAV RGB and UAV multispectral data. The area of interest was approximately 1.4×1.2 km. The DJI M600 (DJI, Shenzhen, China) was equipped with a SONY RX1RII camera with a focal length of 35 mm and a picture size of 35.9×24.0 mm. The UAV flew at a height of 230 m and acquired photos with a resolution of 3 cm. Each RGB image was 7952×5304 pixels. The flight information was planned in DJI GS Pro (DJI, Shenzhen, China), with forward overlap set to 80%, side overlap set to 70%, flight speed set to 12 m/s, and data collection time set to 33 min. SONY RX1RII collected RGB images in the visible spectral band (0.38~0.76 microns). These images directly reflect detailed information such as ground objects' shape, color, and texture.

The multispectral camera is a Swiss Parrot Sequoia multispectral camera, which was mounted on the DJI M600. Each image had a single-band resolution of 1280×960 . The UAV flew at a height of 110 m and acquired data with a resolution of 10 cm. The study area size

was $13,866 \times 11,957$ pixels. We planned the flight information in DJI GS Pro (DJI, Shenzhen, China), set the forward overlap to 75%, side overlap to 70%, flight height to 110 m, flight speed to 12 m/s, and data collection time to 70 min. The research used a calibration panel to calibrate the multispectral images. Furthermore, we used Pix 4D Mapper to mosaic images and ENVI 5.3 for layer stacking. The parameters of the multispectral camera are shown in Table 1.

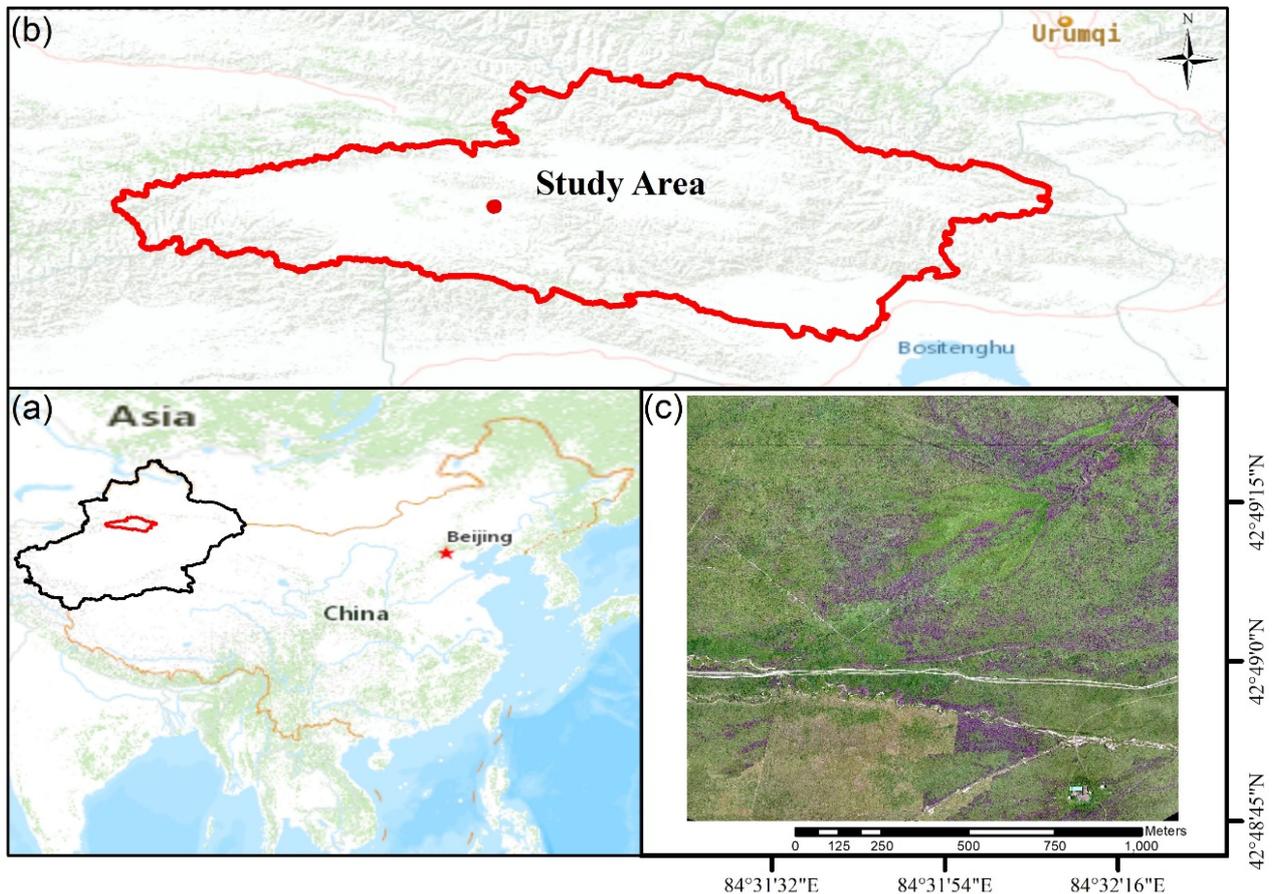


Figure 1. (a) Study site located in Xinjiang (black boundary), China. (b) Hejing County (red boundary), Bayinbuluke grassland (redpoint) located at $42^{\circ} 48'$, $84^{\circ} 31'$, (c) ortho-mosaic image taken by the RGB UAV on 7 August 2019 (purple pixel is *Pedicularis*).

Table 1. Configuration information for Parrot Sequoia.

Band Name	Central Wavelength (nm)	Bandwidth FWHM (nm)
B1(Green)	550	40
B2(Red)	660	40
B3 (Red Edge)	735	10
B4(Nir)	790	40

In this study, three vegetation indices were used as new features to better characterize ground objects. The normalized difference vegetation index (NDVI), ratio vegetation index (RVI), and normalized difference red edge index (NDRE) were derived using Equations (1)–(3). NDVI is recognized as one of the most effective parameters to characterize vegetation change and can reflect vegetation greenness change well [43]. RVI is widely used to estimate and monitor the biomass of green plants [44], and NDRE is mainly used to analyze vegetation health [45]. Numerous experiments have shown that these three

vegetation indices are widely used for land use classification and target extraction and can improve classification accuracy [46–51].

$$\text{NDVI} = \frac{\rho_{\text{nir}} - \rho_{\text{red}}}{\rho_{\text{nir}} + \rho_{\text{red}}}, \quad (1)$$

$$\text{RVI} = \frac{\rho_{\text{nir}}}{\rho_{\text{red}}}, \quad (2)$$

$$\text{NDRE} = \frac{\rho_{\text{nir}} - \rho_{\text{rededge}}}{\rho_{\text{nir}} + \rho_{\text{rededge}}}. \quad (3)$$

The data pre-processing steps were as follows. First, by observing the distribution of land cover types in the study area, we registered UAV multispectral and RGB orthogonal images of the study area. Then, with the help of visible images, we selected samples from multispectral images by visual interpretation methods. We divided each type of sample into training and test sets according to the 7:3 ratio. The sampling points are shown in Figure 2.

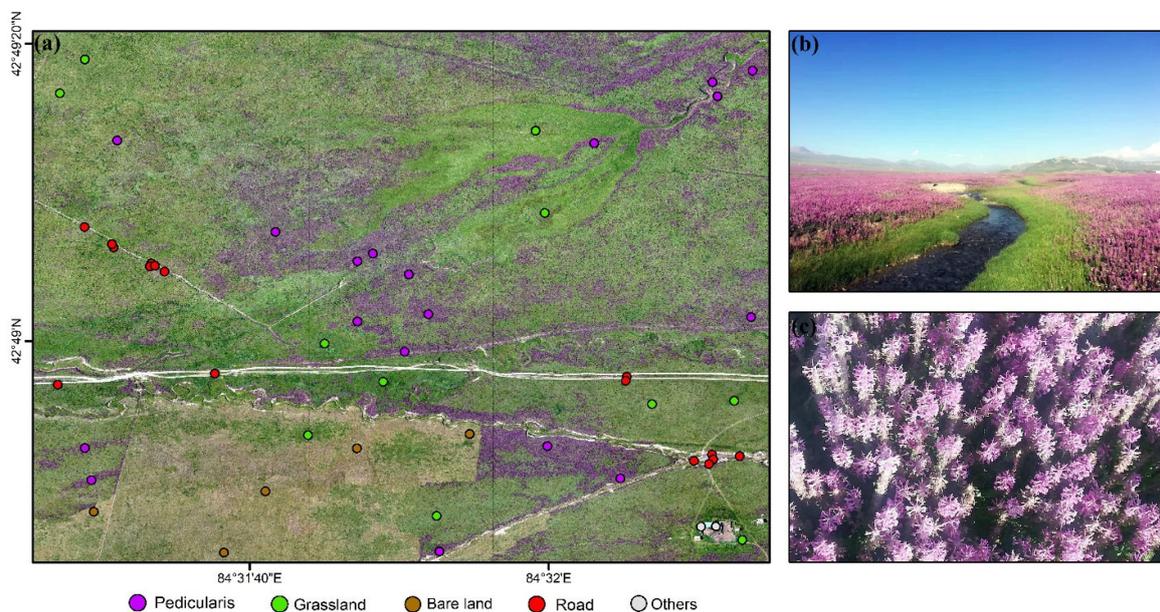


Figure 2. (a) Distribution of samples in RGB image. (b) *Pedicularis* foreground of *Pedicularis*. (c) Close-up photograph of *Pedicularis*.

In this study, sample datasets were selected by visual interpretation. The sample information for the multi-class classification is shown in Table 2. The samples of one-class classification were selected by randomly assigning some samples from training datasets. The test samples were consistent for both one-class and multi-class classifications.

Table 2. Training/testing dataset information for the study regions.

Class	Number of Samples	Number of Pixels	
		Train	Test
<i>Pedicularis</i>	24	157701	67586
Grassland	13	107986	46280
Bare land	5	135141	57918
Road	24	43581	18678
Others	2	3537	1516

Others: buildings in the area of interest.

2.3. Classification Model

This study compared the accuracy of SVM, logistic regression, and random forest algorithms in multi-class classification, OCSVM, isolation forest, and PUL methods in one-class classification algorithms in identifying *Pedicularis*. We used the raw multispectral data (green, red, NIR, rededge) and its derived vegetation indices (NDVI, NDRE, RVI), seven variables in total, as inputs to construct the classifiers. The Scikit-learn framework of python was used to construct the model and adjust the parameters of the model. The construction of these classifiers is summarized below.

Logistic regression is one of the models of multivariate statistical analysis (MSA), which performs maximum likelihood estimation after transforming dependent variables into logit variables. The logistic regression of event probabilities is suitable for describing the relationships between class variables and continuous predictor variables [52]. It mainly tunes the parameter, which is multi_class ('multinomial' and 'ovr').

Cortes and Vapnik proposed SVM based on statistical theory [53]. SVM aims to identify a hyperplane that splits the dataset into predefined discrete classes consistent with the training instance [54]. Three parameters, C, gamma, and kernel, were optimized in this study. Currently, the commonly used kernel functions are linear kernel function, linear kernel, polynomial kernel function, Gaussian kernel function, and sigmoid kernel function. Three values of C (100, 1000, and 10,000) and gamma ('auto' and 'scale') are commonly used.

Random forest uses the bootstrap method, which is used to extract the training dataset from the original sample dataset, and then a decision tree model is trained on each training dataset [55]. Finally, the category with the most votes cast by all base classifiers is the final category [56]. The parameters that determine the classification performance of the random forest classifier include two main categories. One category is the structure of the forest, which is the number of base classifiers. The other category is based on the tree's structure, including the tree's depth, the minimum number of samples required to split the interior nodes, the minimum number of pieces needed for the leaf nodes, and the number of features to find the best split. This experiment adjusted two parameters, n_estimators and max_depth. The n_estimators ranged from 100 to 1000 with a step size of 100, while the range of values for max_depth was from 10 to 100 with a step size of 10.

OCSVM is the most advanced algorithm among one-class classification algorithms. Its principle is to try to obtain a hypersphere as small as possible to contain a positive class training sample. In this study, the OCSVM classifier adopted the radial basis function kernel and optimized the two most essential parameters, nu and γ [57]. The nu ranged from 0.01 to 0.1 with a step size of 0.01, while the values of γ were 0.01, 0.1, and 1.

Isolation forest isolates split the data space with a random hyperplane [58], randomly selecting a feature and then setting a split value between the maximum and minimum value of the selected feature. Then, the classifier continues to split the dataset in the same way until there is only one data point in each subspace. Eventually, the average path length of the samples in all trees is calculated to determine anomalies. Here, isolation forest tuned two parameters. The contamination from 0.01 to 0.2, with a step size of 0.01, was examined, and the max_samples from 0.01 to 0.2 were determined, with a step size of 0.01.

PUL (positive and unlabeled learning) is a semi-supervised learning algorithm based on a positive and unlabeled sample [59]. PUL's goal is to learn dual classifiers from positive samples and mixed unsigned samples. The steps are as follows: (1) determine a set of reliable negation (RN) examples from unlabeled samples (U) based on positive samples (P) and then transform the problem into a binary classification problem and (2) train binary classifiers based on P and RN by iteratively applying existing classification algorithms. This study chose decision trees as the base classifiers and described PUL's classification results in detail. The experiment adjusted two parameters, the number of base classifiers (from 100 to 1000 with a step size of 100) and the number of iterations (from 100 to 2000 with a step size of 100).

2.4. Classification Strategy

We divided sample datasets into training and test datasets according to a ratio of 7:3. We trained and tested the multi-class classifier using all the data in Table 2. The research combined the grid search algorithm and layered 10-fold cross-validation on the training dataset to obtain the best model parameters and ensure the robustness of the model. Ultimately, multi-class and one-class classifiers were tested using the same dataset, and multi-class and one-class classifiers were evaluated in the test dataset using the same metrics.

Both OCSVM and isolation forest are unsupervised classification models, and we used positive data for a semi-supervised classification task in OCSVM and isolation forest, respectively. In OCSVM and isolation forest, the researchers randomly selected 100,000 positive samples from the training datasets. Then, we used two classifiers to fit this training dataset, and the trained model was evaluated for accuracy on test datasets.

In the PUL classifiers, 100,000 positive and 100,000 negative samples were selected from the training dataset in a ratio of 1:1 to train the model. To improve the training efficiency of PUL, 1000 decision trees were trained to fit the training dataset using decision trees as the base classifier. First, we selected 10% of the 100,000 positive class samples as positive ($y = 1$). Second, we marked the remaining 90% of positive class samples and all negative class samples as unlabeled samples ($y = 0$) and randomly selected 10,000 samples from the unlabeled samples as negative ($y = -1$) for training. Finally, 1000 iterations were performed in this manner. The probability of each sample being positive was calculated, and the classification result of this classifier was obtained.

2.5. Accuracy Assessment of Classifiers

Confusion matrices of different classifiers on the test dataset and evaluation metrics are listed in the results section. Recall (R), precision (P), overall accuracy (OA), and F1-Score were used as metrics to evaluate the model. The precision parameter quantifies predictions for a positive class in the collected dataset, which can be achieved using Equation (4). The recall parameter quantifies predictions for all the positive classes, which can be achieved using Equation (5). The F-score considers the P and R parameters, thus indicating model accuracy for the given datasets, as shown in Equation (6). Finally, overall accuracy is the sum of the true positives plus true negatives divided by the total number of individuals tested, as shown in Equation (7). These metrics are sufficient to complete the performance evaluation of each classifier. This research considered the running time of the multi-class and one-class classification models and the difficulty of debugging the optimal parameters, which provided multifaceted judgment criteria for selecting the best classifier.

The overall workflow of this study is shown in Figure 3 and is divided into the following four parts: data processing, model construction, accuracy assessment, and result prediction. We obtained the geographic distribution of *Pedicularis* in the study area using different algorithms. Based on the comparison results, feature importance analyses were performed on the performing one-class classifiers to compare the contribution of each input feature, thus providing a reference for the selection of input features as auxiliary data in further studies. All the above-mentioned methods were completed with the Python Scikit-learn package.

$$\text{precision} = \frac{TP}{TP + FP}, \quad (4)$$

$$\text{recall} = \frac{TP}{TP + FN}, \quad (5)$$

$$F_{\text{score}} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}, \quad (6)$$

$$\text{OA} = \frac{TP + TN}{TP + TN + FP + FN}. \quad (7)$$

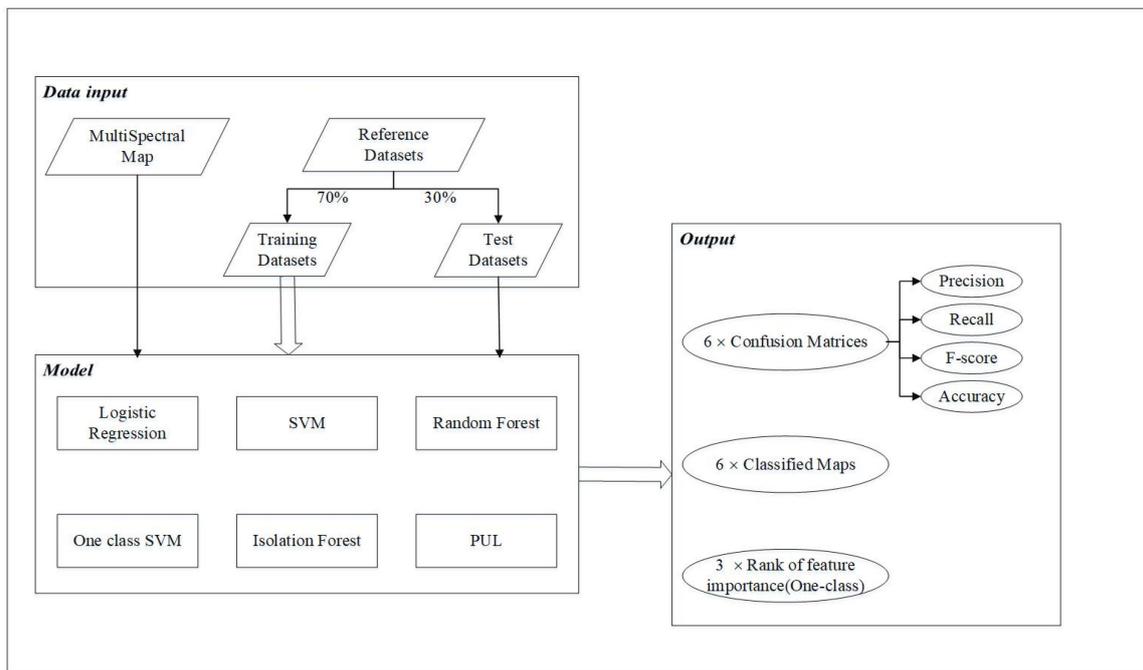


Figure 3. The processing workflow of the study.

TP, FP, FN, and TN are true positive, false positive, false negative, and true negative classifications, respectively.

3. Results

3.1. The Comparisons of the Classification Accuracies of Classifiers

To calculate the best extraction accuracy for *Pedicularis*, we classified the land cover types in the study area using multi-class classifiers that obtained optimal parameters. The land cover types were divided into the following five types: *Pedicularis*, grassland, bare land, road, and others. With the optimal parameter setting, the confusion matrices of three different models on the test dataset were obtained and are presented in Tables 3–5. The performance of each classifier was evaluated using recall, precision, F-Score, and OA.

Table 3. Confusion matrices and evaluation metrics (logistic regression).

	Prediction					Recall (%)
	<i>Pedicularis</i>	Grassland	Bare Land	Road	Others	
<i>Pedicularis</i>	66535	509	540	2	0	98.44
Grassland	340	45359	573	8	0	98.01
Bare land	419	428	57057	14	0	98.51
Road	13	142	809	17690	24	94.71
Others	0	0	52	32	1432	94.45
Precision (%)	98.85	97.67	96.65	99.68	98.35	

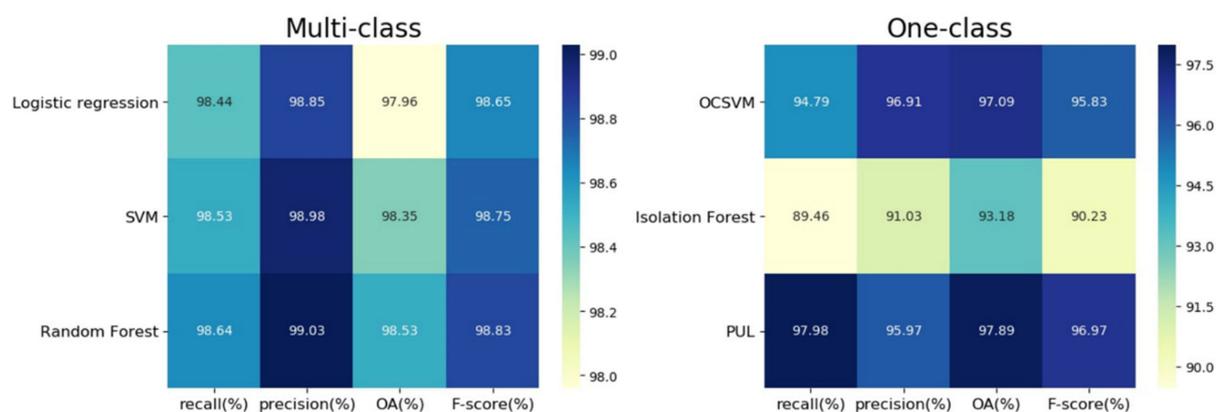
Table 4. Confusion matrices and evaluation metrics (SVM).

	Prediction					Recall (%)
	<i>Pedicularis</i>	Grassland	Bare Land	Road	Others	
<i>Pedicularis</i>	66593	501	491	1	0	98.53
Grassland	300	45753	206	21	0	98.86
Bare land	373	503	57032	10	0	98.47
Road	13	119	553	17984	36	96.14
Others	0	0	0	36	1480	97.62
Precision (%)	98.98	97.60	97.85	99.62	97.62	

Table 5. Confusion matrices and evaluation metrics (random forest).

	Prediction					Recall (%)
	<i>Pedicularis</i>	Grassland	Bare Land	Road	Others	
<i>Pedicularis</i>	66667	418	498	3	0	98.64
Grassland	302	45704	242	32	0	98.75
Bare land	344	403	57111	60	0	98.60
Road	8	91	412	18167	0	97.26
Others	0	0	0	12	1504	99.20
Precision (%)	99.03	98.04	98.02	99.41	1	

As seen in Figure 4, the recall, precision, OA, and F-score of all classifiers were over 97%. Random forest had the best performance with the highest OA and F-score, equal to 98.53% and 98.83%, respectively. Additionally, it was followed by SVM (98.35% and 98.75%) and logistic regression (97.96% and 98.65%). In this experiment, the accuracy of the three classifiers differed very little, and all classifiers could be used as classification methods. As can be evidenced from the precision and recall of other individual land cover types, the performance of the classifiers is consistent with the above results.

**Figure 4.** Evaluation metrics of the different classifiers on test datasets.

When using the one-class classification algorithm, we used *Pedicularis* as the positive class and other samples as the negative class. The confusion matrix and evaluation metrics are shown in Table 6. The evaluation metrics were the same as the multi-class classification. Figure 4 shows that the recall, precision, OA, and F-score of all classifiers were over 90%. The recall, precision, OA, and F-score of OCSVM were 94.79%, 96.91%, 97.09%, and 95.83%, respectively. In the isolation forest, the evaluation metrics were lower, 89.46%, 91.03%, 93.18%, and 90.23%, respectively. PUL had the best performance with recall, precision, OA, and F-score of 97.98%, 95.97%, 97.89%, and 96.97%, respectively.

Compared to the isolation forest, OCSVM had higher recall and precision, which was better than the isolation forest for the identification of *Pedicularis*. In extracting *Pedicularis*, the recalls of OCSVM and isolated forest were slightly lower than the precision. Unlike OCSVM and isolation forest, the PUL was outstanding in all metrics and was relatively close to the result of multi-class classification. Moreover, PUL's precision in extracting *Pedicularis* was slightly lower than its recall. Its overall accuracy and F-score even reached 97.89% and 96.97%.

The result predicted by multi-class and one-class classification shows that the multi-classification algorithm was significantly better than the one-class classification algorithm. As can be seen from Figures 5–7, the accuracy of the multi-class classifiers was higher than the one-class classifiers.

Table 6. Confusion matrices and evaluation metrics (*Pedicularis* is a positive class).

Model	Reference	Prediction	
		Positive	Negative
OCSVM	Positive	64063	3253
	Negative	2045	122347
	Recall (%)	94.79	98.36
	Precision (%)	96.91	97.41
	OA (%)	97.09	
	F-score (%)	95.83	
Isolation forest	Positive	60463	7123
	Negative	5963	118429
	Recall (%)	89.46	95.20
	Precision (%)	91.03	94.32
	OA (%)	93.18	
	F-score (%)	90.23	
Bagging PU	Positive	64865	2721
	Negative	1339	123053
	Recall (%)	97.98	98.92
	Precision (%)	95.97	97.83
	OA (%)	97.89	
	F-score (%)	96.97	

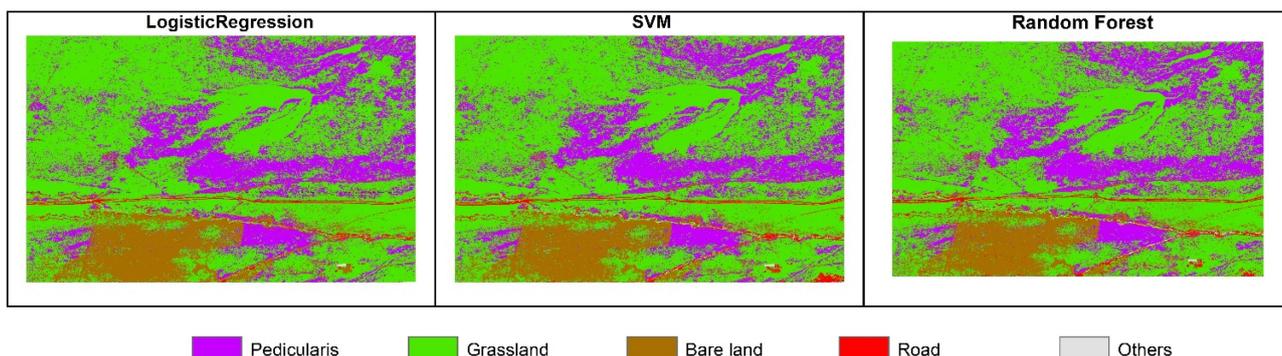


Figure 5. The spatial distribution of each class by multi-class classifiers.

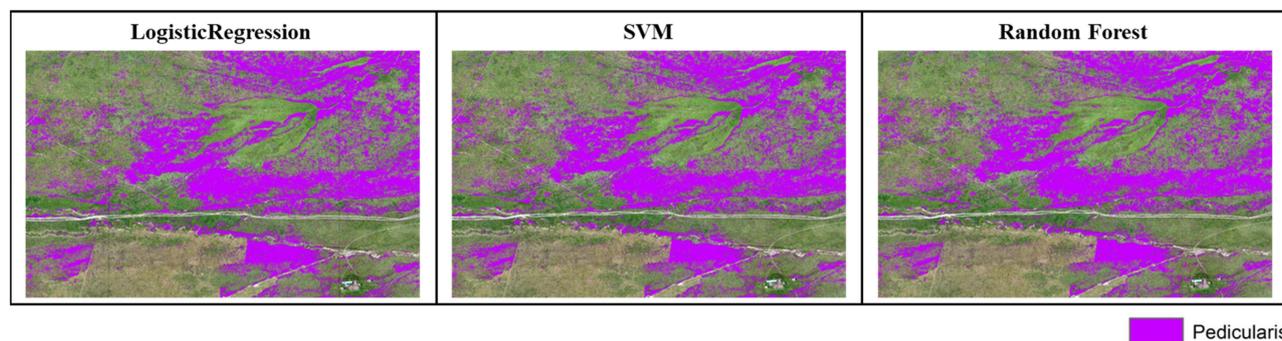


Figure 6. The geographic distribution of *Pedicularis* by multi-class classifiers.

The distribution maps obtained using the OCSVM and isolation forest had slight errors. The map produced by the PUL method is highly consistent with the actual distribution, and its prediction map is very similar to the distribution of *Pedicularis* obtained by the multi-class classifiers. The distribution of *Pedicularis* in the predicted map of OCSVM is more in line with the natural distribution, but the distribution is not dense enough in local areas. The broken speckle phenomenon of the image elements is more prominent. Isolation

forest has low precision, which may cause maps to misclassify other ground objects as *Pedicularis*. In addition, because of low recall, classifiers classify *Pedicularis* as other ground objects. The error of the isolated forest distribution map is the most obvious, and its results can roughly reflect the distribution area of the large area of *Pedicularis*.

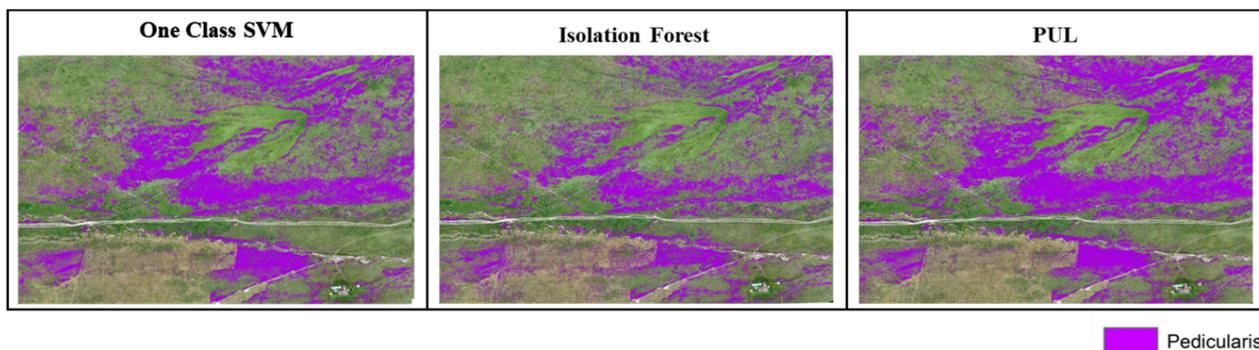


Figure 7. The geographic distribution of *Pedicularis* by one-class classifiers.

3.2. Ranking of Feature Importance for One-Class Classification

To extract ground objects and determine the most important features for identifying *Pedicularis*, we performed feature importance analysis for each one-class classifier. In the first step, on the test dataset, the values of the green bands were randomly ordered, with all other features unchanged. Then, we used the trained model to predict this dataset to obtain the evaluation metrics corresponding to the green band. In the second step, the remaining six features were repeated according to the first step to obtain seven different sets of evaluation metrics. In the third step, to reduce the error of the experiment, this experiment was conducted 10 times, and the average of the 10 results was used as the final result. Finally, we compared the standard evaluation metrics with the seven groups of abnormal metrics. The difference between them was used as a discriminator to determine which feature was the most important, with an enormous difference indicating that the feature was more important.

According to Figure A1 (provided in Appendix A), there may be an improvement in recall or precision by changing the values of individual variables. However, the total OA and F-score were decreased compared with the initial classification results. The classification accuracy of OCSVM was most sensitive to the change in feature values. The isolation forest was the most stable; its OA and F-score were approximately 10%. PUL was insensitive to changes in variables, and its accuracy decreased insignificantly with changes in variables except for the green band.

The trends of F-scores and OA were consistent, using OA and F-scores as the main criteria. The feature importance ranking of OCSVM is Green > NDVI > RVI > Red > Nir > Red Edge > NDRE. Each feature has a pronounced contribution, with Green, NDVI, and RVI being the most important. Isolation forest has an obscure ranking of feature importance, with each variable contributing closely. Green, Nir, and NDVI are most important for isolation forest. PUL's feature importance ranking is Green > Nir > Red > RVI > NDVI > Red Edge > NDRE. Although the three classifiers have different feature importance rankings, green is still the most important feature among the three classifiers.

3.3. Model Evaluation

Table 7 shows the optimal parameters of the model, as well as the training time of the model. Among the multi-class classifiers, logistic regression had the shortest training time, required fewer parameters to be tuned, and had sound performance effects. SVM required the adjustment of two parameters, C and gamma, which had high classification accuracy, but it took more time to fit the model. By contrast, the classification accuracy of the random forest was uppermost, and running speed was faster and more suitable for

Pedicularis prediction. In addition, random forest is insensitive to parameters, does not over-fit, and requires significantly less time to adjust parameters.

Table 7. Parameters and running times of every model.

Model	Parameters	Run Time (s)
Logistic Regression	multi = multinomials	20~30
SVM	C = 1000, gamma = 1	260~290
Random Forest	n_estimator = 100, max_depth = 40	150~170
OCSVM	nu = 0.05, gamma = 0.1	25~35
Isolation forest	Contamination = 0.1, max_samples = 0.1	5~10
PUL	/	500~520

OCSVM took 25~30 s to fit the model and had high accuracy, but its parameters were difficult to adjust. The isolation forest took only 5~10 s to train the model, which took the shortest time, its parameters were easier to adjust, but its accuracy was also the lowest. PUL took approximately 10 min to fit the model because it required 1000 iterations based on the decision tree. Among the one-class classifiers, the model had the highest accuracy and was the most suitable model for extracting *Pedicularis*, and the parameters were adjusted based on the decision tree.

The default parameters are not displayed. Check the documentation for python's Scikit-learn library for a detailed description of the parameters.

4. Discussion

There was a difference between multi-class and one-class classifiers regarding prediction maps. As seen in Figures 6 and 7, the multi-class classifiers had high classification accuracy, and the distribution maps obtained using three different methods were similar. Furthermore, this result was consistent with the distribution of real *Pedicularis*. Among the three classifiers, a stable rank order was found (random forest > SVM > logistic regression). Random forest is widely used in classification tasks because of its excellent generalization performance and is less likely to be overfitted, as can be demonstrated by other studies [23–25]. SVM is not as effective as random forest when classifying land cover types on remote sensing images [26]. Logistic regression is a classical binary classifier. In many studies, it is not used alone for classification but in combination with other methods or as a loss function of neural networks [60].

Analyzing the predicted results of *Pedicularis*, the OCSVM and isolation forest have their limitations due to the absence of the aid of negative class samples. When performing the classification task, OCSVM and isolation forest do not perfectly depict the feature space, and the partitioned interface is not so strict. It was shown that OCSVM had high accuracy in one-class classification, exceeding 90% on high-resolution images. However, it did not perform well in low-resolution images [57]. Isolation forest had never been used for one-class classification of remote sensing images, and the experience results show that its accuracy was lower than that of OCSVM. However, experimental results show that isolation forest had the fastest training speed and easier parameter tuning and could be used as a highly efficient single-class classifier. In addition, OCSVM parameters were difficult to tune and more sensitive to outliers [61]. In contrast, our proposed PUL is not insensitive to outliers and parameters. More importantly, it can use a large number of unlabeled datasets to help the training of the classifier. Therefore, it is reasonable that PUL outperforms OCSVM and isolation forest in this paper. The OA and F-score of PUL were 0.8% and 1.14% higher than OCSVM and 4.71% and 6.74% higher than isolation forest. Similar findings were confirmed in other studies [35].

Adding unlabeled data can enhance classification accuracy when the labeled samples are only positive classes, which is also justified in this paper [59]. Experimental results show that PUL generally provided higher accuracy than OCSVM and isolation forest. The algorithm change borrows the bagging idea, making the model more stable. This research shows that this sampling method produces good classification results, indicating that

this sampling method satisfies the assumption of completely random selection. Further, experimental results show that the tree model is more efficient than the support vector machine in terms of running speed and efficiency. The main advantage of PUL is its ability to obtain higher accuracy, especially when the amount of example data of unlabeled samples is relatively large. However, the algorithm takes the longest time while obtaining the best accuracy. The training time is based on the number of classifiers selected and the number of iterations.

In one-class and multi-class classification algorithms, the change of green band values significantly reduces the model's classification accuracy. The results of this research are consistent with the conclusions drawn from the spectral curve of *Pedicularis*. The spectral curve of *Pedicularis* is different from other grasslands from 550 to 680 nm [62]. The importance of the red band and the vegetation index, calculated based on the red band, are also ranked important. This reason may be due to the difference in the physiological structure of *Pedicularis* from other grasses [63,64]. In addition, we found that object-oriented methods are often used when performing high-resolution image extraction. Object-oriented methods perform better than pixel-based classification methods in extracting regular objects, such as buildings, rivers, or dense grasslands [1]. However, the shape of *Pedicularis* is difficult to obtain, and there is no pattern to its geographical distribution. The government needs to control *Pedicularis* growth early in its emergence. Because of the above reasons, it is necessary to use pix-based classification to extract grassland. This research provides the best practice approach to extracting a single target.

The contribution of this research is to provide a more efficient and accurate algorithm for the extraction of *Pedicularis* based on UAV images. However, extraction experiments were carried out on a small local scale due to the high cost of UAV data collection. More mapping experiments should be conducted in future research since good potential has been shown in our preliminary results. In addition, considering reasonable cost, UAV remote sensing is more suitable for application to a small area. If the task is to extract *Pedicularis* in a vast grassland, it is worthwhile to research migrating the proposed method applied to satellite imagery with higher spatial resolution in the future.

5. Conclusions

This paper shows that the classification accuracy of the multi-class classification method is higher than the one-class classification algorithm when extracting *Pedicularis*. Compared to traditional supervised classification methods, one-class classification can significantly reduce the effort required to assign labels to training samples without loss of prediction accuracy and shows excellent potential for the identification of vegetation. We evaluated the one-class classification method for identifying ground objects, and PUL is the best method to satisfy the condition.

The overall accuracy of *Pedicularis* extracted using the PUL reached 97.89% (F-score = 0.9697). Multispectral imagery can help separate *Pedicularis* from others, with the green band being the most important feature distinguishing *Pedicularis*. In the phase of training models, the training time of PUL depends on the complexity of the base classifier, so the choice of the base classifier is crucial.

The proposed PUL method has great potential in selecting training samples, and it is time-saving and efficient. *Pedicularis* distribution maps obtained in the research showed good potential for detecting and managing harmful plant invasion of grassland-based UAV images. However, UAV remote sensing is more suitable for application to a small area. Practically, we usually have to balance between cost and accuracy when we select remote sensing images of UAVs or satellites. In future research, we will attempt to transfer the proposed PUL method to satellite remote sensing imagery, which is helpful to the application to large grasslands.

Author Contributions: Methodology: W.W. and J.T.; Validation: W.W., J.T. and N.Z.; Formal analysis: W.W.; Investigation: N.Z., Y.W. and J.T.; Writing—original draft preparation: W.W.; Writing—review and editing: W.W., J.T., X.X., A.Z. and N.Z. contributed the same as the corresponding author. All authors have read and agreed to the published version of the manuscript.

Funding: This research was jointly supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA20050103) and the National Key Research and Development Program of China (No. 2020YFC1807102).

Data Availability Statement: Not applicable.

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

Appendix A

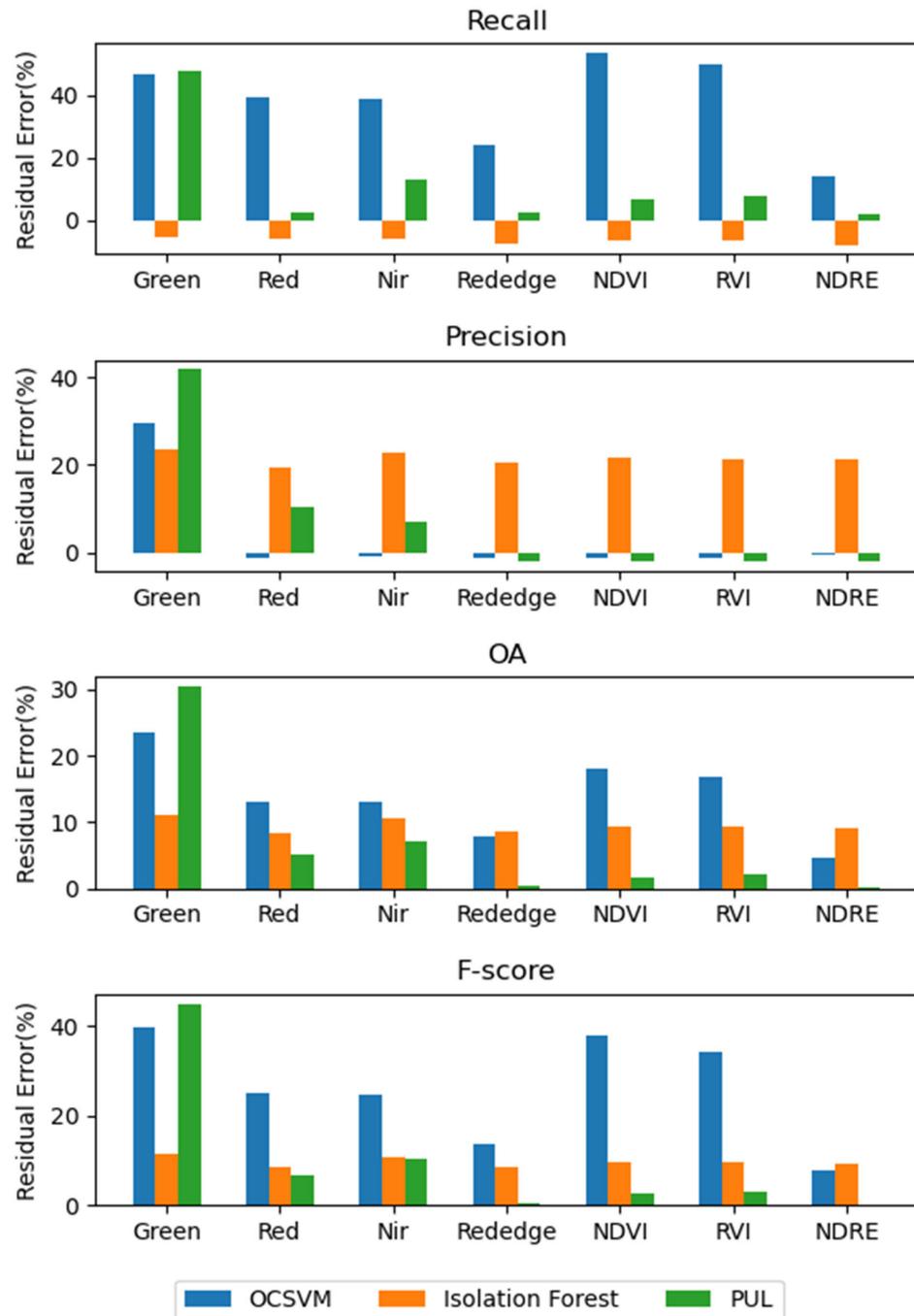


Figure A1. Ranking of feature importance of the one-class classifiers.

References

1. Lu, B.; He, Y. Species classification using Unmanned Aerial Vehicle (UAV)-acquired high spatial resolution imagery in a heterogeneous grassland. *ISPRS-J. Photogramm. Remote Sens.* **2017**, *128*, 73–85. [[CrossRef](#)]
2. Hameed, A.; Zafar, M.; Ahmad, M.; Sultana, S.; Bahadur, S.; Anjum, F.; Shuaib, M.; Taj, S.; Irm, M.; Altaf, M.A. Chemo-taxonomic and biological potential of highly therapeutic plant *Pedicularis groenlandica* Retz. using multiple microscopic techniques. *Microsc. Res. Tech.* **2021**, *84*, 2890–2905. [[CrossRef](#)] [[PubMed](#)]
3. Yanyan, L.I.U.; Yukun, H.U.; Jianmei, Y.U.; Kaihui, L.I.; Guogang, G.A.O.; Xin, W. Study on Harmfulness of *Pedicularis myriophylla* and Its Control Measures. *Arid Zone Res.* **2008**, *25*, 778–782.
4. Liu, Q.; Yang, Z.P.; Han, F.; Shi, H.; Wang, Z.; Chen, X.D. Ecological Environment Assessment in World Natural Heritage Site Based on Remote-Sensing Data. A Case Study from the Bayinbuluke. *Sustainability* **2019**, *11*, 6385. [[CrossRef](#)]
5. Furukawa, F.; Laneng, L.A.; Ando, H.; Yoshimura, N.; Kaneko, M.; Morimoto, J. Comparison of RGB and Multispectral Unmanned Aerial Vehicle for Monitoring Vegetation Coverage Changes on a Landslide Area. *Drones* **2021**, *5*, 97. [[CrossRef](#)]
6. Nakama, J.; Parada, R.; Matos-Carvalho, J.P.; Azevedo, F.; Pedro, D.; Campos, L. Autonomous Environment Generator for UAV-Based Simulation. *Appl. Sci.* **2021**, *11*, 2185. [[CrossRef](#)]
7. Wang, C.-N.; Yang, F.-C.; Vo, N.T.M.; Nguyen, V.T.T. Wireless Communications for Data Security: Efficiency Assessment of Cybersecurity Industry—A Promising Application for UAVs. *Drones* **2022**, *6*, 363. [[CrossRef](#)]
8. Wu, J.; Yu, Y.; Ma, J.; Wu, J.; Han, G.; Shi, J.; Gao, L. Autonomous Cooperative Flocking for Heterogeneous Unmanned Aerial Vehicle Group. *IEEE Trans. Veh. Technol.* **2021**, *70*, 12477–12490. [[CrossRef](#)]
9. Menshchikov, A.; Shadrin, D.; Prutyaynov, V.; Lopatkin, D.; Sosnin, S.; Tsykunov, E.; Iakovlev, E.; Somov, A. Real-Time Detection of Hogweed: UAV Platform Empowered by Deep Learning. *IEEE Trans. Comput.* **2021**, *70*, 1175–1188. [[CrossRef](#)]
10. Valente, J.; Hiremath, S.; Ariza-Sentis, M.; Doldersum, M.; Kooistra, L. Mapping of *Rumex obtusifolius* in nature conservation areas using very high resolution UAV imagery and deep learning. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102864. [[CrossRef](#)]
11. Zhang, X.; Yuan, Y.; Zhu, Z.; Ma, Q.; Yu, H.; Li, M.; Ma, J.; Yi, S.; He, X.; Sun, Y. Predicting the Distribution of *Oxytropis ochrocephala* Bunge in the Source Region of the Yellow River (China) Based on UAV Sampling Data and Species Distribution Model. *Remote Sens.* **2021**, *13*, 5129. [[CrossRef](#)]
12. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS-J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [[CrossRef](#)]
13. Pandey, A.; Jain, K. An intelligent system for crop identification and classification from UAV images using conjugated dense convolutional neural network. *Comput. Electron. Agric.* **2022**, *192*, 106543. [[CrossRef](#)]
14. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [[CrossRef](#)]
15. Ge, G.; Shi, Z.; Zhu, Y.; Yang, X.; Hao, Y. Land use/cover classification in an arid desert-oasis mosaic landscape of China using remote sensed imagery: Performance assessment of four machine learning algorithms. *Glob. Ecol. Conserv.* **2020**, *22*, e00971. [[CrossRef](#)]
16. Hudait, M.; Patel, P.P. Crop-type mapping and acreage estimation in smallholding plots using Sentinel-2 images and machine learning algorithms: Some comparisons. *Egypt. J. Remote Sens. Space Sci.* **2022**, *25*, 147–156. [[CrossRef](#)]
17. Ju, S.; Lim, H.; Ma, J.W.; Kim, S.; Lee, K.; Zhao, S.; Heo, J. Optimal county-level crop yield prediction using MODIS-based variables and weather data: A comparative study on machine learning models. *Agric. For. Meteorol.* **2021**, *307*, 108530. [[CrossRef](#)]
18. Shah, S.T.H.; Qureshi, S.A.; ul Rehman, A.; Shah, S.A.H.; Amjad, A.; Mir, A.A.; Alqahtani, A.; Bradley, D.A.; Khandaker, M.U.; Faruque, M.R.I.; et al. A Novel Hybrid Learning System Using Modified Breaking Ties Algorithm and Multinomial Logistic Regression for Classification and Segmentation of Hyperspectral Images. *Appl. Sci.* **2021**, *11*, 7614. [[CrossRef](#)]
19. Bo, S.K.; Jing, Y.J. Data Distribution Partitioning for One-Class Extraction from Remote Sensing Imagery. *Int. J. Pattern Recognit. Artif. Intell.* **2017**, *31*, 1754018. [[CrossRef](#)]
20. Munoz-Mari, J.; Bovolo, F.; Gomez-Chova, L.; Bruzzone, L.; Camps-Valls, G. Semisupervised One-Class Support Vector Machines for Classification of Remote Sensing Data. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 3188–3197. [[CrossRef](#)]
21. Li, L.H.; Jing, W.P.; Wang, H.H. Extracting the Forest Type From Remote Sensing Images by Random Forest. *IEEE Sens. J.* **2021**, *21*, 17447–17454. [[CrossRef](#)]
22. Zhao, C.; Qin, C.-Z. Identifying large-area mangrove distribution based on remote sensing: A binary classification approach considering subclasses of non-mangroves. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *108*, 102750. [[CrossRef](#)]
23. Dong, L.F.; Du, H.Q.; Mao, F.J.; Han, N.; Li, X.J.; Zhou, G.M.; Zhu, D.; Zheng, J.L.; Zhang, M.; Xing, L.Q.; et al. Very High Resolution Remote Sensing Imagery Classification Using a Fusion of Random Forest and Deep Learning Technique-Subtropical Area for Example. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 113–128. [[CrossRef](#)]
24. Yang, S.T.; Gu, L.J.; Li, X.F.; Jiang, T.; Ren, R.Z. Crop Classification Method Based on Optimal Feature Selection and Hybrid CNN-RF Networks for Multi-Temporal Remote Sensing Imagery. *Remote Sens.* **2020**, *12*, 3119. [[CrossRef](#)]
25. Izquierdo-Verdiguier, E.; Zurita-Milla, R. An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *88*, 102051. [[CrossRef](#)]
26. Kalantar, B.; Pradhan, B.; Naghibi, S.A.; Motevalli, A.; Mansor, S. Assessment of the effects of training data selection on the landslide susceptibility mapping: A comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). *Geomat. Nat. Hazards Risk* **2018**, *9*, 49–69. [[CrossRef](#)]

27. Eddy, P.R.; Smith, A.M.; Hill, B.D.; Peddle, D.R.; Coburn, C.A.; Blackshaw, R.E. Comparison of neural network and maximum likelihood high resolution image classification for weed detection in crops: Applications in precision agriculture. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Denver, CO, USA, 31 July–4 August 2006; pp. 116–119.
28. Pi, W.Q.; Bi, Y.G.; Du, J.M.; Zhang, X.P.; Kang, Y.C.; Yang, H.Y. Classification of Grassland Desertification in China Based on vis-NIR UAV Hyperspectral Remote Sensing. *Spectroscopy* **2020**, *35*, 39–50.
29. Yue, J.; Fang, L.Y.; Ghamisi, P.; Xie, W.Y.; Li, J.; Chanussot, J.; Plaza, A. Optical Remote Sensing Image Understanding with Weak Supervision: Concepts, Methods, and Perspectives. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 250–269. [[CrossRef](#)]
30. Li, W.K.; Guo, Q.H. A maximum entropy approach to one-class classification of remote sensing imagery. *Int. J. Remote Sens.* **2010**, *31*, 2227–2235. [[CrossRef](#)]
31. Dambros, C.S.; Morais, J.W.; Azevedo, R.A.; Gotelli, N.J. Isolation by distance, not rivers, control the distribution of termite species in the Amazonian rain forest. *Ecography* **2017**, *40*, 1242–1250. [[CrossRef](#)]
32. Mack, B.; Roscher, R.; Waske, B. Can I Trust My One-Class Classification? *Remote Sens.* **2014**, *6*, 8779–8802. [[CrossRef](#)]
33. Zhang, C.; Wu, R.; Li, G.; Cui, W.; Jiang, Y. Change detection method based on vector data and isolation forest algorithm. *J. Appl. Remote Sens.* **2020**, *14*, 024516. [[CrossRef](#)]
34. Alonso-Sarria, F.; Valdivieso-Ros, C.; Gomariz-Castillo, F. Isolation Forests to Evaluate Class Separability and the Representativeness of Training and Validation Areas in Land Cover Classification. *Remote Sens.* **2019**, *11*, 3000. [[CrossRef](#)]
35. Wan, B.; Guo, Q.H.; Fang, F.; Su, Y.J.; Wang, R. Mapping US Urban Extents from MODIS Data Using One-Class Classification Method. *Remote Sens.* **2015**, *7*, 10143–10163. [[CrossRef](#)]
36. Liu, R.; Li, W.K.; Liu, X.P.; Lu, X.C.; Li, T.H.; Guo, Q.H. An Ensemble of Classifiers Based on Positive and Unlabeled Data in One-Class Remote Sensing Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 572–584. [[CrossRef](#)]
37. Li, W.K.; Guo, Q.H.; Elkan, C. A Positive and Unlabeled Learning Algorithm for One-Class Classification of Remote-Sensing Data. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 717–725. [[CrossRef](#)]
38. Li, W.K.; Guo, Q.H.; Elkan, C. One-Class Remote Sensing Classification From Positive and Unlabeled Background Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 730–746. [[CrossRef](#)]
39. Lei, L.; Wang, X.; Zhong, Y.; Zhao, H.; Hu, X.; Luo, C. DOCC: Deep one-class crop classification via positive and unlabeled learning for multi-modal satellite imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102598. [[CrossRef](#)]
40. Lyu, X.; Li, X.B.; Dang, D.L.; Dou, H.S.; Wang, K.; Lou, A.R. Unmanned Aerial Vehicle (UAV) Remote Sensing in Grassland Ecosystem Monitoring: A Systematic Review. *Remote Sens.* **2022**, *14*, 1096. [[CrossRef](#)]
41. Nie, H.; Gao, J. Research Progress on the Ecological Impact and Spreading Mechanism of Weeds on Degraded Grassland. *Chin. J. Grassl.* **2022**, *44*, 101–113.
42. Bao, A.; Cao, X.; Chen, X.; Xia, Y. Study on Models for Monitoring of Aboveground Biomass about Bayinbuluke grassland Assisted by Remote Sensing. In Proceedings of the Conference on Remote Sensing and Modeling of Ecosystems for Sustainability, San Diego, CA, USA, 13 August 2008.
43. Carlson, T.N.; Ripley, D.A. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sens. Environ.* **1997**, *62*, 241–252. [[CrossRef](#)]
44. Gnyp, M.L.; Miao, Y.X.; Yuan, F.; Ustin, S.L.; Yu, K.; Yao, Y.K.; Huang, S.Y.; Bareth, G. Hyperspectral canopy sensing of paddy rice aboveground biomass at different growth stages. *Field Crops Res.* **2014**, *155*, 42–55. [[CrossRef](#)]
45. Jorge, J.; Vallbe, M.; Soler, J.A. Detection of irrigation inhomogeneities in an olive grove using the NDRE vegetation index obtained from UAV images. *Eur. J. Remote Sens.* **2019**, *52*, 169–177. [[CrossRef](#)]
46. Chen, J.; Li, X.; Wang, K.; Zhang, S.; Li, J.; Zhang, J.; Gao, W. Variable Optimization of Seaweed Spectral Response Characteristics and Species Identification in Gouqi Island. *Sensors* **2022**, *22*, 4656. [[CrossRef](#)] [[PubMed](#)]
47. Guo, J.-b.; Huang, C.; Wang, H.-g.; Sun, Z.-Y.; Ma, Z.-H. Disease Index Inversion of Wheat Stripe Rust on Different Wheat Varieties with Hyperspectral Remote Sensing. *Spectrosc. Spectr. Anal.* **2009**, *29*, 3353–3357. [[CrossRef](#)]
48. Huete, A.R.; Jackson, R.D. Soil and atmosphere influences on the spectra of partial canopies. *Remote Sens. Environ.* **1988**, *25*, 89–105. [[CrossRef](#)]
49. Liu, Q.; Zhang, T.; Li, Y.; Li, Y.; Bu, C.; Zhang, Q. Comparative Analysis of Fractional Vegetation Cover Estimation Based on Multi-sensor Data in a Semi-arid Sandy Area. *Chin. Geogr. Sci.* **2019**, *29*, 166–180. [[CrossRef](#)]
50. Su, J.; Liu, C.; Coombes, M.; Hu, X.; Wang, C.; Xu, X.; Li, Q.; Guo, L.; Chen, W.-H. Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery. *Comput. Electron. Agric.* **2018**, *155*, 157–166. [[CrossRef](#)]
51. Zhou, X.-X.; Li, Y.-Y.; Luo, Y.-K.; Sun, Y.-W.; Su, Y.-J.; Tan, C.-W.; Liu, Y.-J. Research on remote sensing classification of fruit trees based on Sentinel-2 multi-temporal imageries. *Sci. Rep.* **2022**, *12*, 11549. [[CrossRef](#)]
52. Hossain, M.A.; Jia, X.P.; Benediktsson, J.A. One-Class Oriented Feature Selection and Classification of Heterogeneous Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 1606–1612. [[CrossRef](#)]
53. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
54. Chang, C.-C.; Lin, C.-J. LIBSVM: A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 27. [[CrossRef](#)]
55. Xu, M.; Watanachaturaporn, P.; Varshney, P.; Arora, M. Decision tree regression for soft classification of remote sensing data. *Remote Sens. Environ.* **2005**, *97*, 322–336. [[CrossRef](#)]

56. Pott, L.P.; Amado, T.J.C.; Schwalbert, R.A.; Corassa, G.M.; Ciampitti, I.A. Satellite-based data fusion crop type classification and mapping in Rio Grande do Sul, Brazil. *ISPRS-J. Photogramm. Remote Sens.* **2021**, *176*, 196–210. [[CrossRef](#)]
57. Zhao, L.; Li, Q.; Zhang, Y.; Du, X.; Wang, H.; Shen, Y. Study on the potential of whitening transformation in improving single crop mapping accuracy. *J. Appl. Remote Sens.* **2019**, *13*, 034512. [[CrossRef](#)]
58. Liu, F.T.; Ting, K.M.; Zhou, Z.-H. Isolation-Based Anomaly Detection. *ACM Trans. Knowl. Discov. Data* **2012**, *6*, 1–39. [[CrossRef](#)]
59. Mordelet, F.; Vert, J.P. A bagging SVM to learn from positive and unlabeled examples. *Pattern Recognit. Lett.* **2014**, *37*, 201–209. [[CrossRef](#)]
60. Duie Tien, B.; Khosravi, K.; Shahabi, H.; Daggupati, P.; Adamowski, J.F.; Melesse, A.M.; Binh Thai, P.; Pourghasemi, H.R.; Mahmoudi, M.; Bahrami, S.; et al. Flood Spatial Modeling in Northern Iran Using Remote Sensing and GIS: A Comparison between Evidential Belief Functions and Its Ensemble with a Multivariate Logistic Regression Model. *Remote Sens.* **2019**, *11*, 1589. [[CrossRef](#)]
61. Scholkopf, B.; Platt, J.C.; Shawe-Taylor, J.; Smola, A.J.; Williamson, R.C. Estimating the support of a high-dimensional distribution. *Neural Comput.* **2001**, *13*, 1443–1471. [[CrossRef](#)]
62. Gao, S.; Lin, J.; Ma, T.; Wu, J.; Zheng, J. Extraction and Analysis of Hyperspectral Data and Characteristics from Pedicularis on Bayanbulak Grassland in Xinjiang. *Remote Sens. Technol. Appl.* **2018**, *33*, 908–914.
63. Hu, J.D.; Li, K.H.; Deng, C.J.; Gong, Y.M.; Liu, Y.Y.; Wang, L. Seed Germination Ecology of Semiparasitic Weed Pedicularis kansuensis in Alpine Grasslands. *Plants* **2022**, *11*, 1777. [[CrossRef](#)] [[PubMed](#)]
64. Sui, X.L.; Kuss, P.; Li, W.J.; Yang, M.Q.; Guan, K.Y.; Li, A.R. Identity and distribution of weedy Pedicularis kansuensis Maxim. (Orobanchaceae) in Tianshan Mountains of Xinjiang: Morphological, anatomical and molecular evidence. *J. Arid Land* **2016**, *8*, 453–461. [[CrossRef](#)]