

Article

Comparative Analysis of Machine Learning Algorithms in Automatic Identification and Extraction of Water Boundaries

Aimin Li ^{1,*}, Meng Fan ², Guangduo Qin ², Youcheng Xu ² and Hailong Wang ²¹ School of Geo-Science and Technology, Zhengzhou University, Zhengzhou 450001, China² School of Water Conservancy Engineering, Zhengzhou University, Zhengzhou 450001, China; fanm15755409160@163.com (M.F.); gd13223022138@163.com (G.Q.); chengyouxu1998@163.com (Y.X.); whl18307065356@163.com (H.W.)

* Correspondence: aiminli@zzu.edu.cn

Abstract: Monitoring open water bodies accurately is important for assessing the role of ecosystem services in the context of human survival and climate change. There are many methods available for water body extraction based on remote sensing images, such as the normalized difference water index (NDWI), modified NDWI (MNDWI), and machine learning algorithms. Based on Landsat-8 remote sensing images, this study focuses on the effects of six machine learning algorithms and three threshold methods used to extract water bodies, evaluates the transfer performance of models applied to remote sensing images in different periods, and compares the differences among these models. The results are as follows. (1) Various algorithms require different numbers of samples to reach their optimal consequence. The logistic regression algorithm requires a minimum of 110 samples. As the number of samples increases, the order of the optimal model is support vector machine, neural network, random forest, decision tree, and XGBoost. (2) The accuracy evaluation performance of each machine learning on the test set cannot represent the local area performance. (3) When these models are directly applied to remote sensing images in different periods, the AUC indicators of each machine learning algorithm for three regions all show a significant decline, with a decrease range of 0.33–66.52%, and the differences among the different algorithm performances in the three areas are obvious. Generally, the decision tree algorithm has good transfer performance among the machine learning algorithms with area under curve (AUC) indexes of 0.790, 0.518, and 0.697 in the three areas, respectively, and the average value is 0.668. The Otsu threshold algorithm is the optimal among threshold methods, with AUC indexes of 0.970, 0.617, and 0.908 in the three regions respectively and an average AUC of 0.832.

Keywords: water extraction; modified normalized difference water index (MNDWI); remote sensing; machine learning algorithm



Citation: Li, A.; Fan, M.; Qin, G.; Xu, Y.; Wang, H. Comparative Analysis of Machine Learning Algorithms in Automatic Identification and Extraction of Water Boundaries. *Appl. Sci.* **2021**, *11*, 10062. <https://doi.org/10.3390/app112110062>

Academic Editors: Dimitrios S. Paraforos and Anselme Muzirafuti

Received: 31 August 2021

Accepted: 25 October 2021

Published: 27 October 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Water is the source of life: the earth's surface open water body accounts for about 74% of the total earth area, it is an important resource for all life survival, and it is also the most important component of living organisms [1,2]. In China, the distribution of water resources is quite uneven, and the pollution situation is serious. So, how to identify water bodies efficiently and accurately has become a severe issue [3,4].

With the rapid development of aviation and aerospace technology, remote sensing technology has provided advanced support for many fields, including resource survey, environmental monitoring, mapping, and geography [5,6].

The development of remote sensing technology makes it possible to extract water information quickly and accurately, which is substantially different from conventional field survey methods employed in the past [7–10].

Monitoring open water bodies accurately is an important and basic application in remote sensing. Various water body mapping approaches have been developed to extract

water bodies from multispectral images [11–13]. Using remote sensing images to monitor a water body is mainly based on spectral bands and each image's spatial feature, so the identification methods can be categorized into three types from different perspectives.

(1) Water body index method: This method is based on the spectral curves of water bodies, and thresholds are utilized to effectively distinguish water bodies from the background [14]. Different water indexes have already been proposed in the past few decades. Specifically, in 1996, McFeeters [15] introduced the normalized difference water index (NDWI) model to extract water bodies. However, this model is unable to distinguish between dark shadow and water bodies. To overcome the shortcomings of NDWI, in 2006, Xu [16] proposed the modification of normalized difference water index (MNDWI) to enhance open water features in remotely sensed imagery, and this model has better results for urban water bodies extraction. The water body index method has the characteristics of high precision and low computational cost, which has been widely used in practical applications. In the last few decades, the MNDWI of Xu is one of the most widely used water indices for various fields, including surface water mapping, land use/cover change analyses, and ecological research [17–20].

(2) Machine learning methods: These methods feature pixel-based pattern recognition analysis, mainly including supervised and unsupervised classification techniques. The supervised methods mainly include neural network [21–25], support vector machine (SVM) [26–28], logistic regression [29,30], and random forest [31–33], and the unsupervised classification methods mainly include K-means clustering [34] and ISODATA clustering [35,36] methods. The machine learning algorithm has been widely used in remote sensing water extraction due to its high accuracy.

(3) Object-based image analysis methods (OBIA): Due to the limitations of pixel-based classification methods, such as the salt and pepper phenomenon in classification results, object-based classification techniques have been increasingly applied in remote sensing classification in recent years [37,38]. Many successful cases of water body extraction using OBIA methods have been reported [39–43]. Given that urban functional zones (UFZs) are composed of diverse geographic objects, Du et al. [44] presented a novel object-based UFZ mapping method using very-high-resolution (VHR) remote sensing images. Based on object-oriented analysis technology and multi-source data, Guo et al. [45] proposed a multi-level classification scheme based on goals and rules to study the changes of glacier environments.

In addition, some studies also have used synthetic aperture radar (SAR) data to monitor the surface dynamics, because these data are insensitive to clouds [14,46,47]; the area of surface water can be extracted from SAR data based on textural analysis [48], change detection [49], automatic segmentation [50], and classification [51].

At present, machine learning algorithms to extract water bodies mainly include neural networks, support vector machines, and random forest algorithms. The studies carried out in the past have identified the best performing classification algorithm by comparing different classification algorithms. However, none of them provides a comprehensive comparative analysis of some popular classification algorithms [37,52].

There are few studies on the evaluation of the transfer performance of each machine learning algorithm applied to remote sensing images in different periods. Based on Landsat-8 images, this study uses machine learning algorithms such as decision tree, logistic regression, random forest, and neural network to extract water bodies. First of all, the effect of each machine learning algorithm on the test set is discussed. After that, each machine learning algorithm is applied to three different local areas, and its effect on each local area is evaluated. At last, each machine learning algorithm is applied to remote sensing images in different periods to evaluate the model transfer performance of each machine learning algorithm, and three threshold methods are compared. The results could shed light on the future work of water body extraction based on remote sensing.

2. Data and Pre-Processing

2.1. Data

Landsat-8 data from the website (<http://glovis.usgs.gov/> (accessed on 20 October 2021)) of the United States Geological Survey are used. Landsat-8, launched as a collaboration between the United States Geologic Survey (USGS) and National Aeronautics and Space Administration (NASA) on 11 February 2013, carries onboard the OLI push broom multispectral radiometer [53]. As shown in Table 1, the Landsat-8 OLI/TIRS imagery has 11 spectral bands in total, including eight spectral bands (i.e., three visible bands, two bands for describing aerosol, water vapor, and cirrus clouds, two short-wave infrared bands (SWIR) and near infrared (NIR)) with spatial resolution of 30 m, one panchromatic spectral band with a spatial resolution of 15 m, and two thermal spectral bands with a spatial resolution of 100 m [54]. Landsat-8 remote sensing images (path 123; raw 039) of the same area acquired on 4 October 2019 and 20 October 2019 are used in our experiment. Specifically, the data on 20 October 2019 are used to establish the model and compare the effect of each algorithm, and the data on 4 October 2019 are used to examine the performance of model transfer. Three different areas with different surface features are selected from remote sensing images. As shown in Figure 1, Area1 has a large area of water with relatively simple surface object types, while Area2 has a small water area and complex surface environment, and its water extraction is affected by numerous vegetation and mountain shadow. Area3 is located in the urban built-up area and has multiple contiguous water bodies; thus, the water extraction is affected by nearby buildings and roads.

To avoid the effects of too many clouds and aerosol, images with fewer clouds are selected here. All original data are processed by converting the original digital number (DN) value into spectral radiance, through Equation (1) [55]. The formula is given as follows:

$$L_{\lambda} = M_L \cdot Q_{cal} + A_L \quad (1)$$

where:

L_{λ} = spectral radiance ($W/m^2 \cdot sr \cdot \mu m$);

M_L = radiance multiplicative scaling factor for the spectral band (radiance_mult_band_n from the metadata);

A_L = radiance additive scaling factor for the spectral band (radiance_add_band_n from the metadata);

Q_{cal} = raw digital numbers (DN).

Table 1. Spectral band spatial resolution and wavelength of the Landsat-8 image.

Landsat-8 OLI and TIRS Bands	Wavelength (um)	Spatial Resolution (m)
Coastal/Aerosol	0.435–0.451	30
Blue	0.452–0.512	30
Green	0.533–0.590	30
Red	0.636–0.673	30
NIR	0.851–0.879	30
SWIR-1	1.566–1.651	30
TIR-1	10.60–11.19	100
TIR-2	11.50–12.51	100
SWIR-2	2.107–2.294	30
Pan	0.503–0.676	15
Cirrus	1.363–1.384	30

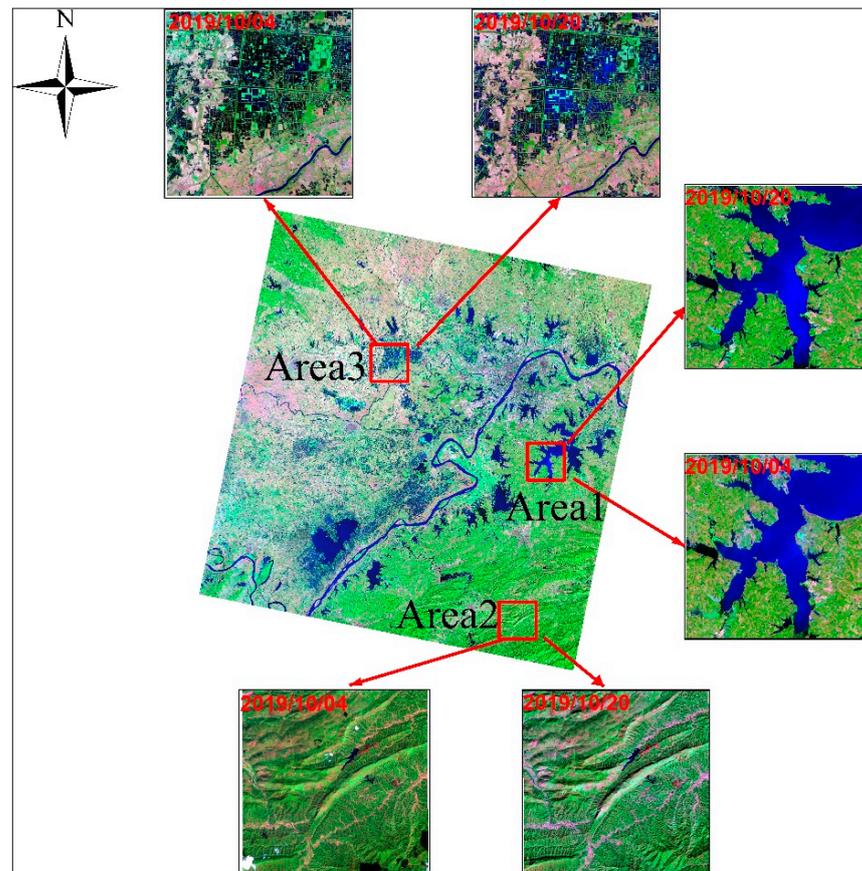


Figure 1. Landsat-8 remote sensing images are displayed in false color in bands 7, 5, and 3. Three local areas are extracted from this image. Area1 has a large area of water distribution with a simple ground environment and is only affected by vegetation; Area2 is affected by mountain shadow and vegetation; Area3 is located in the urban built-up area with scattered water distribution and is affected by roads and buildings.

2.2. Pre-Processing

By adopting spectral band combinations 7/5/4, 7/4/3, 6/5/4, and 4/3/2 combined with visual interpretation, a sample dataset is selected from Landsat images for classification; the sample set contains 340 water samples and 454 non-water samples. To avoid the influences of heterogeneous categories in the subsequent classification, the ratio of other ground object samples to the water body samples remains at 1.3:1.

The characteristics of the data, such as a large correlation between multiple spectral bands in the original images and similar information and structures between different spectral bands, generally bring significant amounts of redundancy. For this reason, principal component analysis (PCA) for dimensionality reduction is applied to remove repetitive and redundant information between various spectral bands [56]. The first and second principal components in the PCA with a cumulative variance contribution of 99% are selected as classification characteristics.

Based on the PCA, four generally used texture features, i.e., contrast, autocorrelation, dissimilarity, and entropy are extracted. The distance is set to be 1 pixel (distance of 30 m), 2 pixels (distance of 60 m), and 3 pixel (distance of 90 m), and 3×3 , 5×5 , 7×7 , and 9×9 are selected as windows with orientations of 0° , 45° , 90° , and 135° . Optimal combined features are selected as the characteristic spectral bands for water body extraction. When the two parameters—i.e., window size and distance—increase, the edges of the images get fuzzy, and the window size shows more effects than distance. Considering the factors of ground objects correlation and image resolution, we set the distance to 1 pixel and select a 3×3 window with four orientations of 0° , 45° , 90° , and 135° .

After the size and window parameters are determined, J-M distance [57,58] and transformed divergence [59] (T-D) in many extracted texture features are used for studying the separability of ground objects; thus, the characteristics ultimately used for classification are determined as well. As shown in Table 2, the separability of the first component (PCA1) and the second component (PCA2) is compared in detail, and the separability of J-M dissimilarity in PCA2 is the optimal. Therefore, in later classifications, a total of six characteristics are selected.

Table 2. Separability of the samples.

	PCA1		PCA2	
	J-M	T-D	J-M	T-D
Contrast	1.723	1.9991	1.825	1.0000
Autocorrelation	1.404	1.6832	1.787	1.989
Dissimilarity	1.562	1.9232	1.84	1.0000
Entropy	1.634	1.9746	1.816	1.999

3. Research Methods

First of all, the performance of machine learning algorithms with a different sample number is discussed. During this process, the optimal parameters of the models are determined and the indices, such as precision and AUC, are used to evaluate the performances of algorithms in the test set. Then, according to spectral characteristics, the water indices are constructed, and on this basis, thresholds are selected; thus, water bodies and other ground objects are classified and identified. Moreover, machine learning methods, such as SVM, decision tree, and random forest, are used to extract water bodies. At last, the accuracy of the test results is verified for the same area at different times.

3.1. MNDWI

In 2006, Xu [16] presented a modification of normalized difference water index (MNDWI) (Equation (2)) by replacing the NIR spectral band used in NDWI with the SWIR spectral band to reduce the influence of building information on water bodies. By using the MNDWI water index method, the MNDWI image is binarized by selecting an appropriate threshold to achieve water bodies extraction. The determination of thresholds affected the accuracy of water body extraction, and different thresholds might be made by subjective judgments of different people. To reduce such influences, three methods for determining thresholds are used for comparison and discussion. The three threshold methods used in this article are as follows: (1) the user-defined threshold method, which is determined according to visual effect through multiple experiments; (2) the Otsu threshold method [60,61]; and (3) the adaptive threshold method, which is used to scan the image through a 3*3 window.

The MNDWI is expressed as follows:

$$\text{MNDWI} = \frac{\text{GREEN} - \text{SWIR}}{\text{GREEN} + \text{SWIR}} \quad (2)$$

where Green is the radiance of the green band, which corresponds to the 3rd Landsat-8 image band; SWIR represents the short-wave infrared band radiance, namely band 6 of the Landsat-8 image.

3.2. Machine Learning Algorithms

In this research, six machine learning algorithms are selected, all of them used the same group of sample set, and the whole samples are divided into a training set and a test set by the ratio of 7:3. Furthermore, in the process of model training, the relevant parameters of the models are further trained by using 10-fold cross-validation with hierarchical sampling

of the training set. Finally, some indices, such as accuracy, recall rate [62], and AUC [63], are utilized to assess the results.

3.2.1. SVM

SVM has a simple structure but a strong generalization ability to solve problems with high-dimensionality, small sample numbers [64,65]. In this study, the Gaussian radial basis function is selected as the kernel function. By using the grid search method in combination with 10-fold cross-validation, the optimal parameters are determined as $C = 3$ and $\gamma = 0.003$.

3.2.2. Decision Tree

The decision tree determines the categories of the samples in the dataset by assigning the sample data to a certain leaf node. There are many methods for constructing the decision tree, but all of them are based on the different purity indices selected and sample attributes for classification [66]. The algorithms ID3, C4.5, C5.0, etc. are generally used. A classification and regression tree (CART) algorithm is used in this study, and pre-pruning is utilized to avoid the overfitting problem. The parameters mainly include the limited depth of the decision tree, the minimum sample number of leaf nodes, and the least sample number of separable leaf nodes. By using the grid search method and 10-fold cross-validation, the final parameters are determined as follows: the entropy is selected as the purity index and the maximum depth is 7. The lowest sample number of separable leaf nodes is 8, and the minimum sample number of leaf nodes is 1.

3.2.3. Multi-Hidden-Layer Neural Network

The neural network uses specific learning algorithms to learn from data through many learning algorithms; however, the network is generally trained by iteratively modifying connection weights and deviations until the error between the output generated by the network and the expected output is smaller than some specified threshold [21]. The input characteristics are passed to the next layer of nerve cells through a non-linear activation function and then continue to be passed down after activation of the nerve cells in this layer. That process is repeated and cycled to the output layer. The repeated superposition of these non-linear functions ensures that the neural network has sufficient non-linear fitting ability, while different activation functions can affect the output of different neural networks. By selecting a sigmoid activation function, it is determined that the neural network structure should have four layers based on multiple tests through cross-validation. Except for input and output layers, the numbers of nerve cells in the two hidden layers are eight and six, respectively.

3.2.4. Random Forest

The random forest is an ensemble method specially designed for a decision tree classifier, and the selection of random attributes is further added to its training process. Using similar parameters to those used for the decision tree, the random forest model is easy to implement and shows good effects [32,33]. In this research, parameters are determined by using cross-validation and grid search methods. The main parameters of random forest are as follows, there are 10 weak estimators in the decision tree, and the maximum depth is 4. Moreover, a Gini function is selected as the purity index.

3.2.5. XGBoost

The core of XGBoost is an ensemble algorithm based on the gradient boosting decision tree (GBDT), and it can be used for classification or regression problems. Its modelling process is as follows: a decision tree is built, and one more tree is added upon each iteration to form a strong evaluator integrating many numerical models [67,68]. The accuracy is superior to that of a weak estimator, and its calculation speed and performance are good [69]. The main parameters are set as follows: the maximum depth of each tree is 3,

and a weak classification estimator with 300 decision trees is established. The learning rate is set to be 0.01.

3.2.6. Logistic Regression Algorithm

The logistic regression is a type of classification model. It establishes a regression formula for samples and a sigmoid function is used for classification. For more information, please refer to references [70,71].

4. Experiment and Analysis

4.1. Effects of the Sample Number on Learning Algorithms

For each classification algorithm in machine learning, the basic requirement is that the training and test set are reliable and there are enough samples for training. In this way, a good classifier can be trained. It is assumed that the samples selected by visual interpretation are reliable: namely, the various classes of the sample points are assigned to correct labels. Based on this, a small sample is randomly selected from the training set and divided into a training set and a validation set in the proportion of 7:3. By using the accuracy of the validation set of the small sample as an evaluation index, the effects of the sample number on the classification effects of each algorithm are discussed, so as to judge whether the sample number selected is sufficient to achieve the purpose of the training model.

As demonstrated in Figure 2, the accuracies of the classification algorithms in the validation set of the experiment all tend to increase with the sample number, and they show a smaller error relative to the accuracy in the training set. Moreover, the accuracies gradually tend to be equal. This indicates that there is almost no underfitting of the samples, and the parameters of each algorithm are well adjusted. The accuracy of the logistic regression algorithm is improved rapidly, approximating to the accuracy in the training set when the sample number is small, suggesting that there is almost no overfitting. As the sample number increases, the accuracy stabilizes; however, other classification algorithms need larger samples to achieve this stability, and the accuracy fluctuates (albeit within a small range), therefore, the number of training samples selected in the experiment can meet the needs of model training.

4.2. Analysis of Performance Indices of Machine Learning Algorithms

After testing the performance of the models when using each algorithm on sets of different sample numbers, the effect of each model in the same test set is further evaluated, so as to reflect the predictive abilities of the models to some extent and judge the generalization abilities of the algorithms. As shown in Table 3, the value of the accuracy index and recall index of each model in classifying water bodies and other ground objects are high, the accuracy index is in the range of 0.945–1, and the recall index is in the range of 0.911–1. However, the AUC index can better represent the comprehensive performances of the models and the higher the value, the better the performance [63]. There is little difference in the effect of each machine learning algorithm on the test set, and the AUC index ranges from 0.956 to 0.987; by analyzing AUC data, the logistic regression and XGBoost algorithm are found to perform best on the test set, followed by the SVM, the neural network, then the random forest, while the decision tree has (in general) the worst performance. Whether the evaluation of these algorithms in the test set can accurately represent the generalization abilities of the algorithms for classifying water bodies in the remote sensing images needs to be discussed and studied using remote sensing images acquired under different conditions.

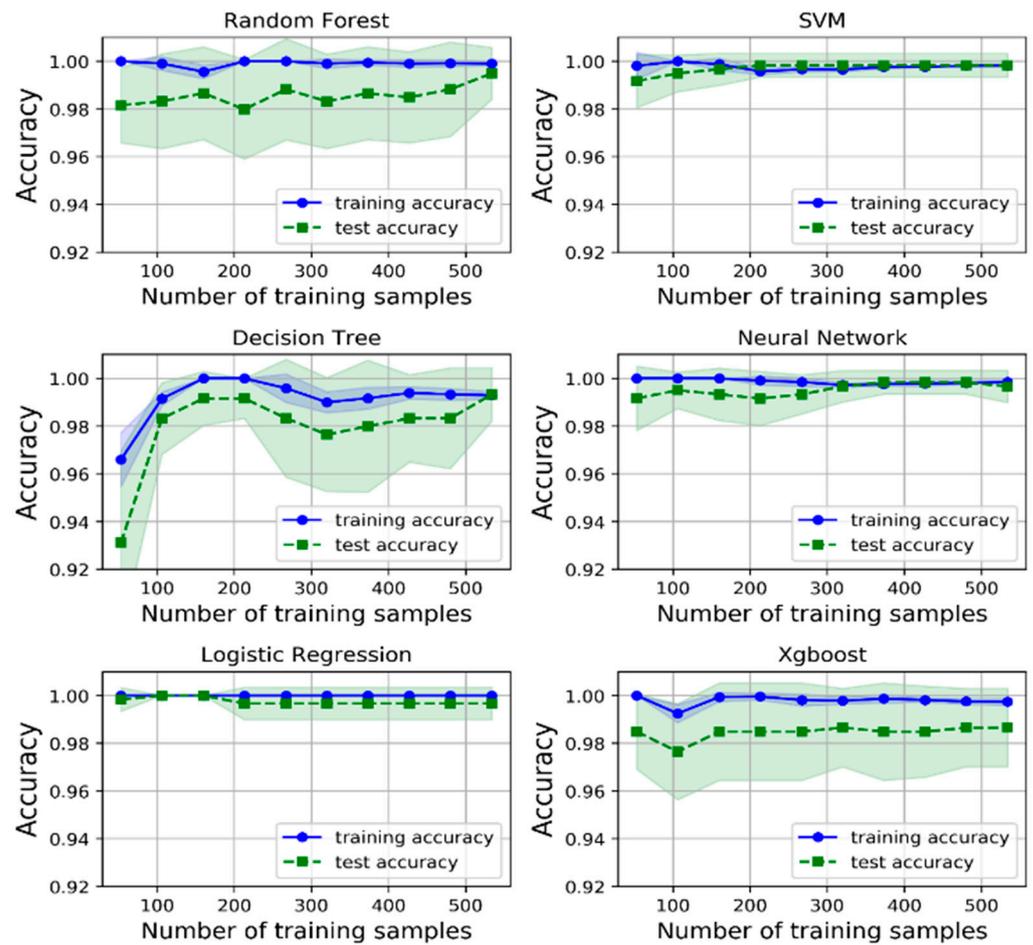


Figure 2. Effects of the sample number on performance of each algorithm.

Table 3. Analysis of performance indices of each algorithm.

		Accuracy	Recall	AUC
SVM	Water	1.000	0.949	0.983
	Other	0.968	1.000	
Random Forest	Water	0.975	0.975	0.979
	Other	0.983	0.983	
Decision Tree	Water	1.000	0.911	0.956
	Other	0.945	1.000	
Neural Network	Water	1.000	0.962	0.981
	Other	0.976	1.000	
Logistic Regression	Water	1.000	0.975	0.987
	Other	0.984	1.000	
XGBoost	Water	1.000	0.975	0.987
	Other	0.984	1.000	

4.3. Comparative Analysis of NDWI and Machine Learning Algorithms

The model established by 2019/10/20 training data is used for water extraction in three areas of 2019/10/20. Statistical results of AUC indicators of each algorithm are shown in Figure 3 (For more details, see Tables A1–A4 in the Appendix A). In general, the XGBoost algorithm has the best accuracy, with an average AUC of 0.966, and the AUC indicators in the three regions are 0.985, 0.972, and 0.941 respectively, which is followed by the random forest algorithm with an average AUC of 0.964, and the AUC indicators in the three regions are 0.985, 0.973, and 0.935; the SVM algorithm has the worst accuracy, the average AUC is 0.898 and the AUC indicators in the three regions are 0.982, 0.789, and

0.923, respectively. When each machine learning algorithm is applied to three different local regions, the average range of AUC index is 0.898–0.966 (for more details, see Table A1 in Appendix A), and the descending order of each machine learning algorithm is XGBoost, random forest, decision tree, logistic regression, neural network, and SVM according to the value of the AUC index. However, this is inconsistent with the conclusion of Section 4.2. In Section 4.2, there is little difference in the accuracy of each machine learning algorithm on the test set, and the AUC index ranges from 0.956 to 0.987. The machine learning algorithms are XGBoost, LR, SVM, NN, RF, and DT in descending order according to the value of the AUC index. It further explains that the evaluation on the test set cannot represent the effect of each algorithm applied in a local area. Among the threshold classification methods, the Otsu threshold algorithm is the best, with an average AUC of 0.957, and the AUC indicators in the three regions are 0.985, 0.922, and 0.964, respectively, followed by the custom threshold algorithm, and the worst performance among all algorithms is adaptive threshold algorithm: the average AUC is only 0.764.

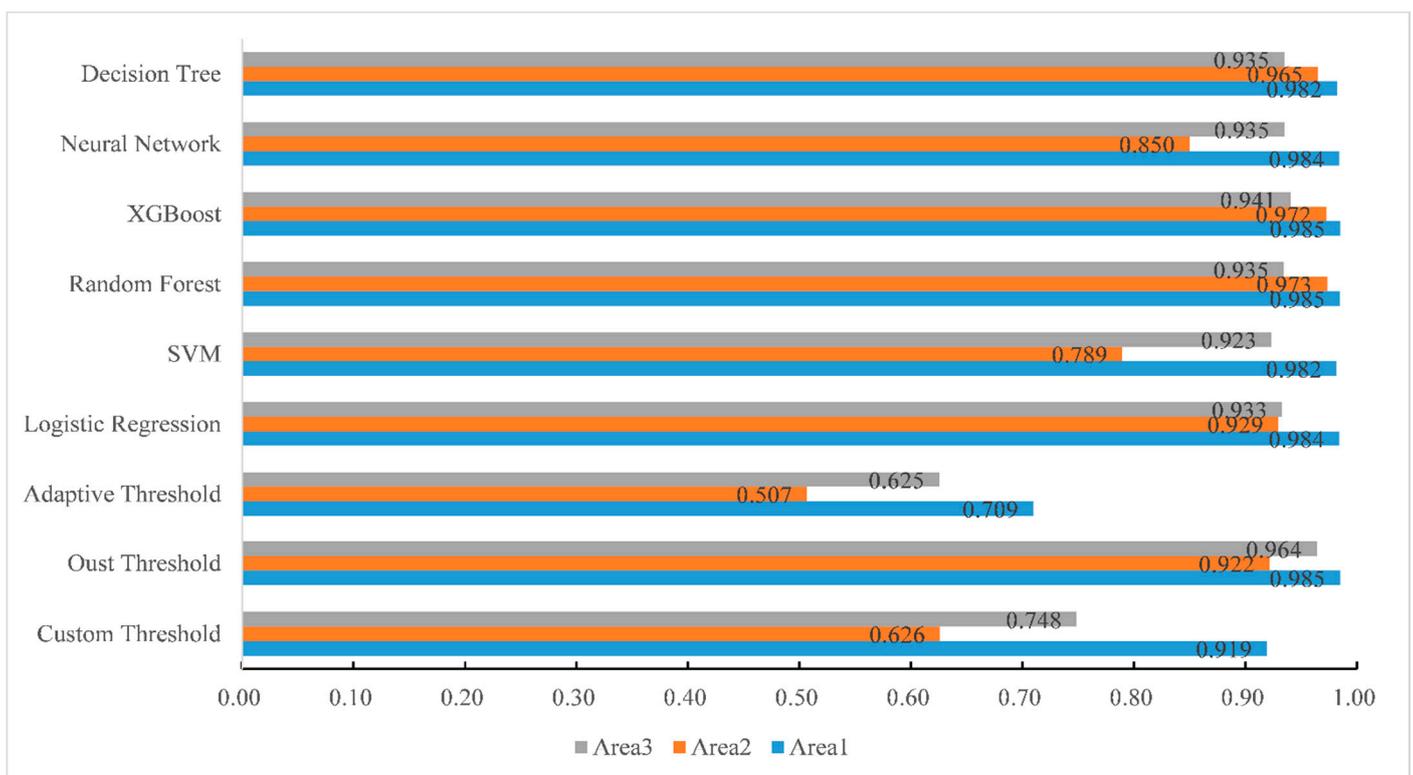


Figure 3. Statistics of the AUC index of each algorithm applied in the three regions.

The image water extraction results of each algorithm were placed in the supplementary materials, as shown in Figure S1: Classification results of each algorithm in Area1 on October 20; Figure S2: Classification results of each algorithm in Area2 on October 20; Figure S3: Classification results of each algorithm in Area3 on October 20. As can be seen from the results graph, compared with other algorithms, the salt and pepper phenomenon for the adaptive threshold and custom threshold is very serious, there is a large number of non-water body “noise”, other algorithms basically have the same visual interpretation effect, and there is no obvious difference, but the edge part is slightly different due to the influence of adjacent features.

4.4. Reliability Test

To discuss the effects of the aforementioned algorithms in water body extraction from remote sensing images in different periods, a remote sensing image captured on 4 October 2019 in the same region is selected. Based on this, the water bodies are classified using the

same algorithms and parameters. The aim is to verify whether the experimental results of each algorithm under different image conditions are reliable and decide whether the models are universal.

The model established by the data of 2019/10/20 is used in the data of 2019/10/04 for water body extraction. The statistical results of the AUC indicators of each algorithm are shown in Figure 4 (for more details, see Tables A5–A8 in Appendix A). As shown in Table 4, the AUC indicators of each machine learning algorithm for three regions all show a significant decline, with a decreased range of 0.33–66.52%. As shown in Figure 4, the differences among the different algorithm performances in the three areas are obvious. In the surface complex Area2, the AUC index of the machine learning algorithms is near 0.5, which means it is difficult to extract water bodies accurately. In Area1 with a simple surface environment, although the accuracy of all machine learning algorithms decreases, the errors are still within an acceptable range. In general, the decision tree algorithm has better transfer performance, with an average AUC of 0.668, and the AUC indexes of the three regions are 0.790, 0.518, and 0.697 respectively. The XGBoost algorithm has an average AUC of 0.631, and its AUC index in the three regions is 0.718, 0.512, and 0.665, respectively. The logistic regression algorithm has the worst accuracy, with an average AUC of 0.392, the AUC index in the three regions is 0.329, 0.489, and 0.357, respectively, which is inconsistent with the conclusion in Sections 4.2 and 4.3. When the model is directly transferred to remote sensing images of different periods for water extraction, the generalization ability of each machine learning algorithm is different. Among the threshold classification methods, the Otsu threshold algorithm is optimal, and its average AUC is 0.832. The AUC indexes in the three regions are 0.970, 0.617, and 0.908, respectively, which exceed the accuracy of the other machine learning algorithms. For the other two threshold algorithm, custom threshold, whose average AUC is 0.700, and the AUC indexes in the three regions are 0.842, 0.549, and 0.708 respectively. The adaptive threshold algorithm has an average AUC of 0.611, and its AUC indicators in the three regions are 0.703, 0.506, and 0.623 respectively. All in all, for different periods of remote sensing images, the threshold method is better than most of the machine learning algorithms, because the sensor imaging is affected by clouds, sun angles, and sensors. Due to the influence of the angle and other factors, the characteristics of remote sensing images will be very different during the adjacent imaging time. Even if there is no major change in the surface features, the pixel value of the remote sensing image could also change significantly. Therefore, the machine learning models trained on the data of 2019/10/20 may not be suitable for different periods.

Table 4. AUC index changes statistics of each machine learning algorithm.

The Method Name		Area1	Area2	Area3	Average
Threshold Method	Custom Threshold	−8.41%	−12.27%	−5.37%	−8.68%
	Otsu Threshold	−1.49%	−33.10%	−5.78%	−13.46%
	Adaptive Threshold	−0.92%	−0.16%	−0.33%	−0.47%
Machine Learning Method	Logistic Regression	−66.52%	−47.37%	−61.74%	−58.54%
	SVM	−66.25%	−32.20%	−60.49%	−52.98%
	Random Forest	−30.12%	−49.73%	−62.68%	−47.51%
	XGBoost	−27.16%	−47.32%	−29.35%	−34.61%
	Neural Network	−30.13%	−42.96%	−62.13%	−45.07%
	Decision Tree	−19.60%	−46.28%	−25.40%	−30.43%

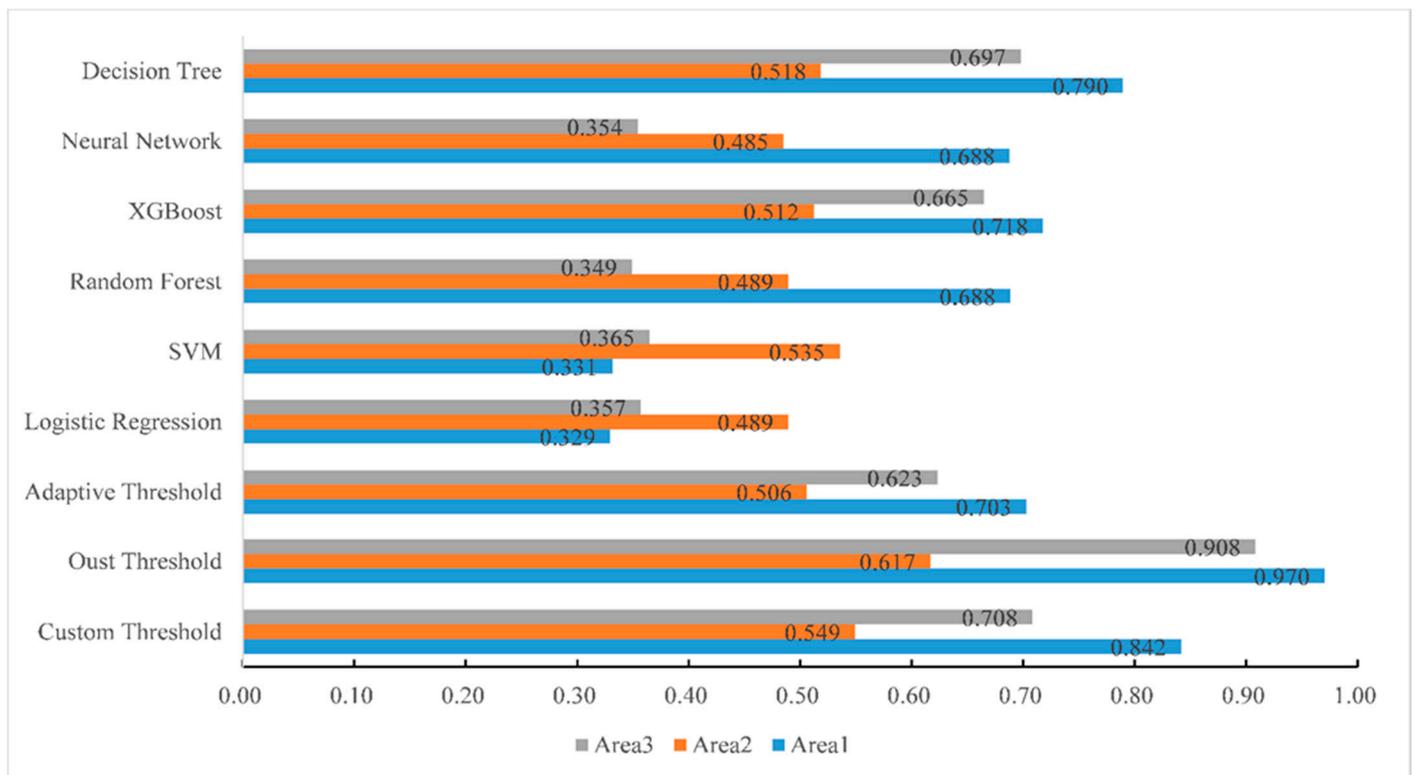


Figure 4. The AUC indexes of the three regions in different periods for each algorithm.

However, the water extraction effect of the threshold method is related to the remote sensing image data, and the water extraction effects of remote sensing images from different periods do not affect each other.

The water extraction results of each algorithm were placed in the supplementary materials, as shown in Figure S4: Classification results of each algorithm in Area1 on October 4; Figure S5: Classification results of each algorithm in Area2 on October 4; Figure S6: Classification results of each algorithm in Area3 on October 4. It can be seen from the classification result diagrams that most of the machine learning pepper and salt phenomenon is very serious, and there is a large number of non-water “noise”. The visual effects of various algorithms are also significantly different.

5. Discussion

This study mainly selects neural network, support vector machine (SVM), logistic regression, random forest, decision tree, and XGBoost from machine learning algorithms, and it selects the MNDWI water index combined with three threshold methods to extract the water bodies. Michael Schmitt [72] pointed out that for a simple surface environment, only the threshold method can achieve satisfactory results, and when the surface environment is slightly more complicated, a supervised classification method, such as SVM, needs to be introduced. However, for the supervised classification method, how to choose the appropriate number of samples is a problem worthy of research. For example, Deepakrishna Somasundaram et al. [73] selected 3765 water samples and 2685 non-water samples from the four-view Landsat-8 OLI image; Wei Jiang et al. [74] selected more than 10,000 water samples and non-water samples in each study area. The choice of these large numbers of training samples brings additional costs. In order to study the influence of sample size on various algorithms, an experiment was designed in this paper, as outlined in Section 4.1. As shown in Figure 2, there are great differences in the number of samples required for various algorithms to reach their optimal. The logistic regression algorithm requires the lowest number of samples, which is close to 110. The SVM algorithm has the best performance when the number of samples reaches 150. As the number of samples

increases, the order of the optimal model is neural network, random forest, decision tree, and XGBoost. The primary task of water body extraction is to select a certain number of samples for the training model. The conclusion of the sample number requirements of each machine learning algorithm in this paper can be used as a reference for other similar applications to reduce the cost of sample selection.

Most studies only use test set samples to evaluate the optimal model and use the selected model for the final classification of images. However, Liu Yang et al. [75] pointed out that in different surface environments, various types of shadows or background noises need to be considered. For example, compared with arid areas, the influence of vegetation on water extraction should be considered in humid areas. In mountainous areas, the extracted water is often mixed with mountain shadow. These types of background information have different influences on different water extraction algorithms [61,76]. For the above reasons, it is worth discussing whether the evaluation effect on the test set can explain the actual generalization performance of the model, that is, whether the evaluation effect on the test set is consistent with the evaluation effect on the local area. For this reason, three local areas with different ground conditions are selected. As shown in Figure 3, in general, the simpler the ground scene, the better the classification accuracy. If the ground scene is complex, the accuracy of various algorithms has a great difference. Generally, three algorithms (decision tree, XGBoost, and Otsu) can perform well in various scenarios. In the case of mountain shadow in the ground background, it is suggested to give priority to the XGBoost algorithm. In the case of roads and buildings in the ground background, besides the XGBoost or decision tree algorithms, a logistic regression algorithm with a relatively simple model can also be tried.

However, when multi-stage extraction research on water bodies is needed, the original model will naturally be directly used to extract water bodies from remote sensing images in other different periods. As shown in Table 4, when various machine learning algorithms are directly used to extract water bodies from remote sensing images in different periods, the AUC indicators of each machine learning algorithm for the three regions all show a significant decline, with a decrease range of 0.33–66.52%. Generally, simple ground scenes have higher accuracy, while complex ground scenes have some effects for different machine learning algorithms. As shown in Table 4, among all the machine learning algorithms, the accuracy of decision tree decreased the least in the three regions on average, and the AUC index decreased 30.43% on average, followed by XGBoost. In the threshold method, although the change of adaptive threshold is small, its accuracy is always very low, while the Otsu algorithm not only has a good accuracy, but also the average decline of the AUC index is small, which is 13.46%. The decision tree algorithm can still achieve better classification results, and the Otsu algorithm also performs well. Experiments show that it is not recommended to directly use the machine learning model to extract water from remote sensing images in different periods. The Otsu classification result can be used as a reference, so that training samples can be selected in other periods quickly and conveniently to extract water bodies using machine learning algorithms.

In summary, for water extraction from remote sensing images, although various algorithms can achieve satisfactory results under certain conditions, none of them can be applied to all remote sensing image and scenes. The factors affecting the classification accuracy of remote sensing images mainly include the complexity of the field landscape, the availability of data, the effectiveness of the processing method, and the experience judgment of the processing personnel [5,76]. Therefore, on the basis of this study, when extracting water from remote sensing images, the water index (MNDWI preferred) can be used first and combined with the Otsu algorithm to classify water bodies. This result is in agreement with the results obtained by Ya'nan Zhou et al. [38], who used the NDWI image to select water samples from the input image. However, if the accuracy does not meet the requirements of the application, on the basis of its classification, researchers can further select the number of samples that meet the requirements of various machine learning algorithms (Figure 2) and select the corresponding machine learning training

model. Among the various machine learning algorithms, XGBoost, decision tree, and logistic regression algorithms are preferentially recommended.

6. Conclusions

Based on Landsat-8 images, decision tree, logistic regression, random forest, neural network, support vector machine, and XGBoost algorithms are used to extract water bodies. Firstly, the effect of each machine learning algorithm on the test set is discussed. Secondly, each machine learning algorithm is applied to three different local areas, and the consistency between the accuracy of each machine learning algorithm on the test set and the accuracy of the local area is evaluated. Finally, each machine learning algorithm is applied to remote sensing images in different periods, the model transfer performance of each machine learning algorithm is examined, and three threshold methods are compared. The following conclusions are drawn:

(1) There are great differences in the numbers of samples required for various algorithms to reach their optimal. The logistic regression algorithm requires a minimum number of samples, about 110. The SVM algorithm has the best performance when the number of samples reaches 150. As the number of samples increases, the optimal order of the model is neural network, random forest, decision tree, and XGBoost.

(2) The accuracy evaluation effect of each machine learning on the test set cannot represent the effect on the local area, because the surface complexity is not same in the three local areas. In Area1 with a single surface type, its AUC range is 0.982–0.985; in Area2 with complex surface environment (numerous vegetation and mountain shadow), its AUC range is 0.789–0.973; in Area3 with wide water distribution, its AUC range is 0.923–0.941 in an urban built-up area.

(3) When the models are directly applied to remote sensing images in different periods, the model accuracy is greatly reduced, the AUC indicators of each machine learning algorithm for three regions all show a significant decline, with a decreasing range of 0.33–66.52%. In general, among the machine learning algorithms, the decision tree algorithm has good transfer performance, with an average AUC of 0.668, and the AUC indexes in the three regions are 0.790, 0.518, and 0.697 respectively. Among the threshold methods, the Otsu threshold algorithm is the optimal, with an average AUC of 0.832 and AUC indexes in the three regions are 0.970, 0.617, and 0.908, respectively.

(4) Owing to the complex distribution of ground objects and many influential factors in the remote sensing image classification, it is difficult to collect small and dispersed water bodies in this research. This limits the performances of these models in the environment with many hill shadows and complex ground objects. The accuracy of these models needs to be further improved; more samples should be collected from images over different areas and periods to train the models in the future.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/app112110062/s1>, Detailed descriptions of Figure S1: Classification results of each algorithm in Area1 on October 20; Figure S2: Classification results of each algorithm in Area2 on October 20; Figure S3: Classification results of each algorithm in Area3 on October 20; Figure S4: Classification results of each algorithm in Area1 on October 4; Figure S5: Classification results of each algorithm in Area2 on October 4; Figure S6: Classification results of each algorithm in Area3 on October 4.

Author Contributions: Supervision, A.L.; Writing—original draft, M.F.; Writing—review and editing, M.F., G.Q., Y.X. and H.W. All authors have read and agreed to the published version of the manuscript.

Funding: The work is supported by the Joint Funds of National Natural Science Foundation of China (Grant no. U1704125).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Appendix A

Table A1. Statistics of the AUC index of each algorithm applied in the three regions.

The Method Name		Area1	Area2	Area3	Average
Threshold Method	Custom Threshold	0.919	0.626	0.748	0.764
	Otsu Threshold	0.985	0.922	0.964	0.957
	Adaptive Threshold	0.709	0.507	0.625	0.614
Machine Learning Method	Logistic Regression	0.984	0.929	0.933	0.949
	SVM	0.982	0.789	0.923	0.898
	Random Forest	0.985	0.973	0.935	0.964
	XGBoost	0.985	0.972	0.941	0.966
	Neural Network	0.984	0.850	0.935	0.923
	Decision Tree	0.982	0.965	0.935	0.961

Table A2. Statistics of various indexes of each algorithm in Area1 on October 20.

Method	Category	Precision	Recall	F1-Score	AUC
Neural Network	water	0.998	0.945	0.971	0.984
	other	0.971	0.999	0.985	
Random Forest	water	0.997	0.950	0.973	0.985
	other	0.973	0.998	0.986	
SVM	water	0.999	0.933	0.965	0.982
	other	0.964	1.000	0.982	
XGBoost	water	0.996	0.952	0.974	0.985
	other	0.974	0.998	0.986	
Logistic Regression	water	0.999	0.942	0.970	0.984
	other	0.969	1.000	0.984	
Decision Tree	water	0.998	0.937	0.966	0.982
	other	0.966	0.999	0.982	
Adaptive Threshold	water	0.508	0.909	0.652	0.709
	other	0.911	0.515	0.658	
Custom Threshold	water	0.838	0.999	0.912	0.919
	other	0.999	0.894	0.944	
Otsu Threshold	water	0.979	0.983	0.981	0.985
	other	0.991	0.988	0.990	

Table A3. Statistics of various indexes of each algorithm in Area2 on October 20.

Method	Category	Precision	Recall	F1-Score	AUC
Neural Network	water	0.708	0.595	0.647	0.850
	other	0.992	0.995	0.993	
Random Forest	water	0.962	0.241	0.385	0.973
	other	0.984	1.000	0.992	
SVM	water	0.585	0.712	0.642	0.789
	other	0.994	0.989	0.992	
XGBoost	water	0.961	0.221	0.360	0.972
	other	0.984	1.000	0.992	
Logistic Regression	water	0.869	0.498	0.633	0.929
	other	0.990	0.998	0.994	
Decision Tree	water	0.948	0.144	0.250	0.965
	other	0.982	1.000	0.991	
Adaptive Threshold	water	0.027	0.670	0.052	0.507
	other	0.986	0.496	0.660	
Custom Threshold	water	0.256	0.825	0.391	0.626
	other	0.996	0.950	0.972	
Otsu Threshold	water	0.856	0.361	0.508	0.922
	other	0.987	0.999	0.993	

Table A4. Statistics of various indexes of each algorithm in Area3 on October 20.

Method	Category	Precision	Recall	F1-Score	AUC
Neural Network	water	0.982	0.690	0.810	0.935
	other	0.889	0.995	0.939	
Random Forest	water	0.999	0.628	0.771	0.935
	other	0.870	1.000	0.930	
SVM	water	0.995	0.567	0.722	0.923
	other	0.852	0.999	0.919	
XGBoost	water	0.998	0.671	0.802	0.941
	other	0.883	0.999	0.938	
Logistic Regression	water	0.997	0.624	0.768	0.933
	other	0.869	0.999	0.929	
Decision Tree	water	0.984	0.682	0.805	0.935
	other	0.886	0.995	0.938	
Adaptive Threshold	water	0.405	0.747	0.525	0.625
	other	0.846	0.558	0.673	
Custom Threshold	water	0.497	1.000	0.664	0.748
	other	1.000	0.593	0.745	
Otsu Threshold	water	0.970	0.893	0.930	0.964
	other	0.958	0.989	0.973	

Table A5. AUC index statistics of each algorithm in three regions on October 4.

The Method Name	Area1	Area2	Area3	Average	
Threshold Method	Custom Threshold	0.842	0.549	0.708	0.700
	Otsu Threshold	0.970	0.617	0.908	0.832
	Adaptive Threshold	0.703	0.506	0.623	0.611
	Logistic Regression	0.329	0.489	0.357	0.392
Machine Learning Method	SVM	0.331	0.535	0.365	0.410
	Random Forest	0.688	0.489	0.349	0.509
	XGBoost	0.718	0.512	0.665	0.631
	Neural Network	0.688	0.485	0.354	0.509
	Decision Tree	0.790	0.518	0.697	0.668

Table A6. Statistics of various indexes of each algorithm in Area1 on October 4.

Method	Category	Precision	Recall	F1-Score	AUC
Neural Network	water	0.379	0.999	0.550	0.688
	other	0.996	0.097	0.177	
Random Forest	water	0.377	1.000	0.547	0.688
	other	1.000	0.087	0.160	
SVM	water	0.020	0.001	0.001	0.331
	other	0.643	0.994	0.781	
XGBoost	water	0.435	1.000	0.607	0.718
	other	1.000	0.284	0.442	
Logistic Regression	water	0.017	0.001	0.001	0.329
	other	0.642	0.990	0.779	
Decision Tree	water	0.579	1.000	0.734	0.790
	other	1.000	0.599	0.749	
Adaptive Threshold	water	0.493	0.918	0.641	0.703
	other	0.913	0.478	0.628	
Custom Threshold	water	0.685	0.998	0.812	0.842
	other	0.998	0.746	0.854	
Otsu Threshold	water	0.949	0.985	0.967	0.970
	other	0.992	0.971	0.981	

Table A7. Statistics of various indexes of each algorithm in Area2 on October 4.

Method	Category	Precision	Recall	F1-Score	AUC
Neural Network	water	0.028	0.999	0.055	0.485
	other	1.000	0.283	0.441	
Random Forest	water	0.022	1.000	0.042	0.489
	other	1.000	0.050	0.095	
SVM	water	0.090	0.039	0.054	0.535
	other	0.980	0.992	0.986	
XGBoost	water	0.024	1.000	0.048	0.512
	other	1.000	0.165	0.283	
Logistic Regression	water	0.001	0.001	0.001	0.489
	other	0.978	0.937	0.957	
Decision Tree	water	0.037	0.992	0.072	0.518
	other	1.000	0.463	0.633	
Adaptive Threshold	water	0.026	0.692	0.050	0.506
	other	0.986	0.451	0.619	
Custom Threshold	water	0.100	0.925	0.181	0.549
	other	0.998	0.826	0.904	
Otsu Threshold	water	0.248	0.314	0.277	0.617
	other	0.986	0.980	0.983	

Table A8. Statistics of various indexes of each algorithm in Area3 on October 4.

Method	Category	Precision	Recall	F1-Score	AUC
Neural Network	water	0.007	0.001	0.001	0.354
	other	0.702	0.944	0.805	
Random Forest	water	0.001	0.002	0.001	0.349
	other	0.697	0.924	0.794	
SVM	water	0.020	0.001	0.002	0.365
	other	0.709	0.978	0.822	
XGBoost	water	0.329	1.000	0.496	0.665
	other	1.000	0.183	0.309	
Logistic Regression	water	0.002	0.001	0.001	0.357
	other	0.712	0.994	0.830	
Decision Tree	water	0.395	1.000	0.567	0.697
	other	1.000	0.386	0.557	
Adaptive Threshold	water	0.399	0.756	0.523	0.623
	other	0.847	0.543	0.662	
Custom Threshold	water	0.419	0.997	0.590	0.708
	other	0.997	0.444	0.615	
Otsu Threshold	water	0.828	0.973	0.895	0.908
	other	0.988	0.919	0.952	

References

- Fletcher, T.D.; Andrieu, H.; Hamel, P. Understanding, management and modelling of urban hydrology and its consequences for receiving waters: A state of the art. *Adv. Water Resour.* **2013**, *51*, 261–279. [\[CrossRef\]](#)
- Papa, F.; Prigent, C.; Rossow, W.B. Monitoring Flood and Discharge Variations in the Large Siberian Rivers From a Multi-Satellite Technique. *Surv. Geophys.* **2008**, *29*, 297–317. [\[CrossRef\]](#)
- Chen, C.; He, X.Y.; Lu, Y.; Chu, Y.L. Application of Landsat Time-Series Data in Island Ecological Environment Monitoring: A Case Study of Zhoushan Islands, China. *J. Coast. Res.* **2020**, *108*, 193–199. [\[CrossRef\]](#)
- Yang, Y.; Yang, Y.; Liu, D.; Nordblom, T.; Wu, B.; Yan, N. Regional Water Balance Based on Remotely Sensed Evapotranspiration and Irrigation: An Assessment of the Haihe Plain, China. *Remote Sens.* **2014**, *6*, 2514–2533. [\[CrossRef\]](#)
- Wang, Y.N.; Huang, F.; Wei, Y.C. Water Body Extraction from LANDSAT ETM plus Image Using MNDWI and K-T Transformation. In Proceedings of the 2013 21st International Conference Geoinformatics, Kaifeng, China, 20–22 June 2013.
- Yang, X.; Zhao, S.; Qin, X.; Zhao, N.; Liang, L. Mapping of Urban Surface Water Bodies from Sentinel-2 MSI Imagery at 10 m Resolution via NDWI-Based Image Sharpening. *Remote Sens.* **2017**, *9*, 596. [\[CrossRef\]](#)
- Durand, M.; Gleason, C.J.; Garambois, P.A.; Bjerklie, D.; Smith, L.C.; Roux, H.; Rodriguez, E.; Bates, P.D.; Pavelsky, T.M.; Monnier, J.; et al. An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope. *Water Resour. Res.* **2016**, *52*, 4527–4549. [\[CrossRef\]](#)
- Yang, X.; Qin, Q.; Grussenmeyer, P.; Koehl, M. Urban surface water body detection with suppressed built-up noise based on water indices from Sentinel-2 MSI imagery. *Remote Sens. Environ.* **2018**, *219*, 259–270. [\[CrossRef\]](#)
- Cui, X.; Guo, X.; Wang, Y.; Wang, X.; Zhu, W.; Shi, J.; Lin, C.; Gao, X. Application of remote sensing to water environmental processes under a changing climate. *J. Hydrol.* **2019**, *574*, 892–902. [\[CrossRef\]](#)
- Li, D.; Wang, G.; Qin, C.; Wu, B. River Extraction under Bankfull Discharge Conditions Based on Sentinel-2 Imagery and DEM Data. *Remote Sens.* **2021**, *13*, 2650. [\[CrossRef\]](#)
- Du, Y.; Zhang, Y.H.; Ling, F.; Wang, Q.M.; Li, W.B.; Li, X.D. Water Bodies' Mapping from Sentinel-2 Imagery with Modified Normalized Difference Water Index at 10-m Spatial Resolution Produced by Sharpening the SWIR Band. *Remote Sens.* **2016**, *8*, 354. [\[CrossRef\]](#)
- Donchyts, G.; Baart, F.; Winsemius, H.; Gorelick, N.; Kwadijk, J.; van de Giesen, N. Earth's surface water change over the past 30 years. *Nat. Clim. Chang.* **2016**, *6*, 810–813. [\[CrossRef\]](#)
- Yang, K.; Yao, F.; Wang, J.; Luo, J.; Shen, Z.; Wang, C.; Song, C. Recent dynamics of alpine lakes on the endorheic Changtang Plateau from multi-mission satellite data. *J. Hydrol.* **2017**, *552*, 633–645. [\[CrossRef\]](#)
- Cao, M.; Mao, K.; Shen, X.; Xu, T.; Yan, Y.; Yuan, Z. Monitoring the Spatial and Temporal Variations in The Water Surface and Floating Algal Bloom Areas in Dongting Lake Using a Long-Term MODIS Image Time Series. *Remote Sens.* **2020**, *12*, 3622. [\[CrossRef\]](#)
- McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [\[CrossRef\]](#)
- Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033. [\[CrossRef\]](#)

17. Feyisa, G.L.; Meilby, H.; Fensholt, R.; Proud, S.R. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 23–35. [[CrossRef](#)]
18. Du, Z.; Li, W.; Zhou, D.; Tian, L.; Ling, F.; Wang, H.; Gui, Y.; Sun, B. Analysis of Landsat-8 OLI imagery for land surface water mapping. *Remote Sens. Lett.* **2014**, *5*, 672–681. [[CrossRef](#)]
19. Singh, K.V.; Setia, R.; Sahoo, S.; Prasad, A.; Pateriya, B. Evaluation of NDWI and MNDWI for assessment of waterlogging by integrating digital elevation model and groundwater level. *Geocarto Int.* **2015**, *30*, 650–661. [[CrossRef](#)]
20. Duan, Z.; Bastiaanssen, W.G.M. Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. *Remote Sens. Environ.* **2013**, *134*, 403–416. [[CrossRef](#)]
21. Zhang, L.; Wang, J.S.; An, Z.Y. Classification method of CO2 hyperspectral remote sensing data based on neural network. *Comput. Commun.* **2020**, *156*, 124–130. [[CrossRef](#)]
22. Li, Y.F.; Liu, C.C.; Zhao, W.P.; Huang, Y.F. Multi-spectral remote sensing images feature coverage classification based on improved convolutional neural network. *Math. Biosci. Eng.* **2020**, *17*, 4443–4456. [[CrossRef](#)]
23. Cui, W.; Zhou, Q.; Zheng, Z.D. Application of a Hybrid Model Based on a Convolutional Auto-Encoder and Convolutional Neural Network in Object-Oriented Remote Sensing Classification. *Algorithms* **2018**, *11*, 9. [[CrossRef](#)]
24. Tian, Y.; Jia, R.S.; Xu, S.H.; Hua, R.; Deng, M.D. Super-resolution reconstruction of remote sensing images based on convolutional neural network. *J. Appl. Remote Sens.* **2019**, *13*, 13. [[CrossRef](#)]
25. Chen, Y.; Fan, R.; Yang, X.; Wang, J.; Latif, A. Extraction of urban water bodies from high-resolution remote-sensing imagery using deep learning. *Water* **2018**, *10*, 585. [[CrossRef](#)]
26. Qi, H.N.; Huang, M.L. Research on SVM ensemble and its application to remote sensing classification. In Proceedings of the International Conference on Intelligent Systems and Knowledge Engineering (ISKE 2007), Chengdu, China, 15–16 October 2007.
27. Alimjan, G.; Sun, T.L.; Jumahun, H.; Guan, Y.; Zhou, W.T.; Sun, H.G. A Hybrid Classification Approach Based on Support Vector Machine and K-Nearest Neighbor for Remote Sensing Data. *Int. J. Pattern Recognit. Artif. Intell.* **2017**, *31*, 1750034. [[CrossRef](#)]
28. Razaque, A.; Frej, M.B.; Almi'ani, M.; Alotaibi, M.; Alotaibi, B. Improved Support Vector Machine Enabled Radial Basis Function and Linear Variants for Remote Sensing Image Classification. *Sensors* **2021**, *21*, 4431. [[CrossRef](#)] [[PubMed](#)]
29. Cheng, Q.; Varshney, P.K.; Arora, M.K. Logistic regression for feature selection and soft classification of remote sensing data. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 491–494. [[CrossRef](#)]
30. Hogland, J.; Billor, N.; Anderson, N. Comparison of standard maximum likelihood classification and polytomous logistic regression used in remote sensing. *Eur. J. Remote Sens.* **2013**, *46*, 623–640. [[CrossRef](#)]
31. Veerabhadraswamy, N.; Devagiri, G.M.; Khaple, A.K. Fusion of complementary information of SAR and optical data for forest cover mapping using random forest algorithm. *Curr. Sci.* **2021**, *120*, 193–199. [[CrossRef](#)]
32. 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](#)]
33. Shetty, S.; Gupta, P.K.; Belgiu, M.; Srivastav, S.K. Assessing the Effect of Training Sampling Design on the Performance of Machine Learning Classifiers for Land Cover Mapping Using Multi-Temporal Remote Sensing Data and Google Earth Engine. *Remote Sens.* **2021**, *13*, 1433. [[CrossRef](#)]
34. Liang, G.; Zhao, X.L.; Zhao, J.H.; Zhou, F.N. Feature Selection and Mislabeled Waveform Correction for Water-Land Discrimination Using Airborne Infrared Laser. *Remote Sens.* **2021**, *13*, 3628. [[CrossRef](#)]
35. Memarsadeghi, N.; Mount, D.M.; Netanyahu, N.S.; Le Moigne, J. A fast implementation of the ISODATA clustering algorithm. *Int. J. Comput. Geom. Appl.* **2007**, *17*, 71–103. [[CrossRef](#)]
36. Mahboob, M.; Genc, B. Evaluation of ISODATA Clustering Algorithm for Surface Gold Mining Using Satellite Data. In Proceedings of the 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), Swat, Pakistan, 24–25 July 2019.
37. Balha, A.; Mallick, J.; Pandey, S.; Gupta, S.; Singh, C.K. A comparative analysis of different pixel and object-based classification algorithms using multi-source high spatial resolution satellite data for LULC mapping. *Earth Sci. Inform.* **2021**, 1–17. [[CrossRef](#)]
38. Zhou, Y.; Luo, J.; Shen, Z.; Hu, X.; Yang, H. Multiscale Water Body Extraction in Urban Environments From Satellite Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4301–4312. [[CrossRef](#)]
39. Janowski, L.; Wroblewski, R.; Dworniczak, J.; Kolakowski, M.; Rogowska, K.; Wojcik, M.; Gajewski, J. Offshore benthic habitat mapping based on object-based image analysis and geomorphometric approach. A case study from the Slupsk Bank, Southern Baltic Sea. *Sci. Total Environ.* **2021**, *801*, 149712. [[CrossRef](#)] [[PubMed](#)]
40. Granger, J.E.; Mahdianpari, M.; Puestow, T.; Warren, S.; Mohammadimanesh, F.; Salehi, B.; Brisco, B. Object-based random forest wetland mapping in Conne River, Newfoundland, Canada. *J. Appl. Remote Sens.* **2021**, *15*, 038506. [[CrossRef](#)]
41. Zhao, B.; Gou, P.; Yang, F.; Tang, P.P. Improving object-oriented land use/cover classification from high resolution imagery by spectral similarity-based post-classification. *Geocarto Int.* **2021**, 1–24. [[CrossRef](#)]
42. Aahlaad, M.; Mozumder, C.; Tripathi, N.; Pal, I. An Object-Based Image Analysis of WorldView-3 Image for Urban Flood Vulnerability Assessment and Dissemination Through ESRI Story Maps. *J. Indian Soc. Remote Sens.* **2021**, 1–16. [[CrossRef](#)]
43. Ma, L.; Zhu, X.X.; Qiu, C.P.; Blaschke, T.; Li, M.C. Advances of Local Climate Zone Mapping and Its Practice Using Object-Based Image Analysis. *Atmosphere* **2021**, *12*, 1146. [[CrossRef](#)]
44. Du, S.H.; Du, S.H.; Liu, B.; Zhang, X.Y. Mapping large-scale and fine-grained urban functional zones from VHR images using a multi-scale semantic segmentation network and object based approach. *Remote Sens. Environ.* **2021**, *261*, 112480. [[CrossRef](#)]

45. Guo, S.C.; Du, P.J.; Xia, J.S.; Tang, P.F.; Wang, X.; Meng, Y.P.; Wang, H. Spatiotemporal changes of glacier and seasonal snow fluctuations over the Namcha Barwa-Gyala Peri massif using object-based classification from Landsat time series. *ISPRS J. Photogramm. Remote Sens.* **2021**, *177*, 21–37. [[CrossRef](#)]
46. Liu, X.; Zhang, Y.; Ling, X.; Huang, X. Automatic and Unsupervised Water Body Extraction Based on Spectral-Spatial Features Using GF-1 Satellite Imagery. *IEEE Geosci. Remote Sens. Lett.* **2018**, *16*, 927–931. [[CrossRef](#)]
47. Bao, L.; Lv, X.; Yao, J. Water Extraction in SAR Images Using Features Analysis and Dual-Threshold Graph Cut Model. *Remote Sens.* **2021**, *13*, 3465. [[CrossRef](#)]
48. Di Baldassarre, G.; Schumann, G.; Brandimarte, L.; Bates, P. Timely Low Resolution SAR Imagery To Support Floodplain Modelling: A Case Study Review. *Surv. Geophys.* **2011**, *32*, 255–269. [[CrossRef](#)]
49. Long, S.; Fatoyinbo, T.E.; Policelli, F. Flood extent mapping for Namibia using change detection and thresholding with SAR. *Environ. Res. Lett.* **2014**, *9*, 35002. [[CrossRef](#)]
50. Pulvirenti, L.; Chini, M.; Pierdicca, N.; Guerriero, L.; Ferrazzoli, P. Flood monitoring using multi-temporal COSMO-SkyMed data: Image segmentation and signature interpretation. *Remote Sens. Environ.* **2011**, *115*, 990–1002. [[CrossRef](#)]
51. Zhao, L.; Yang, J.; Li, P.; Zhang, L. Seasonal inundation monitoring and vegetation pattern mapping of the Erguna floodplain by means of a RADARSAT-2 fully polarimetric time series. *Remote Sens. Environ.* **2014**, *152*, 426–440. [[CrossRef](#)]
52. Srivastava, P.K.; Han, D.; Rico-Ramirez, M.A.; Bray, M.; Islam, T. Selection of classification techniques for land use/land cover change investigation. *Adv. Sp. Res.* **2012**, *50*, 1250–1265. [[CrossRef](#)]
53. Kuhn, C.; de Matos Valerio, A.; Ward, N.; Loken, L.; Sawakuchi, H.O.; Kampel, M.; Richey, J.; Stadler, P.; Crawford, J.; Striegl, R.; et al. Performance of Landsat-8 and Sentinel-2 surface reflectance products for river remote sensing retrievals of chlorophyll-a and turbidity. *Remote Sens. Environ.* **2019**, *224*, 104–118. [[CrossRef](#)]
54. García-Llamas, P.; Suárez-Seoane, S.; Fernández-Guisuraga, J.M.; Fernández-García, V.; Fernández-Manso, A.; Quintano, C.; Taboada, A.; Marcos, E.; Calvo, L. Evaluation and comparison of Landsat 8, Sentinel-2 and Deimos-1 remote sensing indices for assessing burn severity in Mediterranean fire-prone ecosystems. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *80*, 137–144. [[CrossRef](#)]
55. U.S. Geological Survey. Landsat 8 Data Users Handbook. *Nasa* **2016**, *8*, 97.
56. Zhao, G.; Maclean, A.L. A comparison of canonical discriminant analysis and principal component analysis for spectral transformation. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 841–847.
57. Mohajane, M.; Essahlaoui, A.; Oudija, F.; Hafyani, M.E.; Hmaidi, A.E.; Ouali, A.E.; Randazzo, G.; Teodoro, A.C. Land Use/Land Cover (LULC) Using Landsat Data Series (MSS, TM, ETM+ and OLI) in Azrou Forest, in the Central Middle Atlas of Morocco. *Environ.* **2018**, *5*, 131. [[CrossRef](#)]
58. Randazzo, G.; Cascio, M.; Fontana, M.; Gregorio, F.; Lanza, S.; Muzirafuti, A. Mapping of Sicilian Pocket Beaches Land Use/Land Cover with Sentinel-2 Imagery: A Case Study of Messina Province. *Land* **2021**, *10*, 678. [[CrossRef](#)]
59. Forestier, G.; Inglada, J.; Wemmert, C.; Gancarski, P. Comparison of optical sensors discrimination ability using spectral libraries. *Int. J. Remote Sens.* **2013**, *34*, 2327–2349. [[CrossRef](#)]
60. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man. Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
61. Ji, L.; Zhang, L.; Wylie, B. Analysis of Dynamic Thresholds for the Normalized Difference Water Index. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 1307–1317. [[CrossRef](#)]
62. Xie, H.; Zhang, Y.; He, Y.; You, K.; Fan, B.; Yu, D.; Li, M. Automatic and Fast Recognition of On-Road High-Emitting Vehicles Using an Optical Remote Sensing System. *Sens.* **2019**, *19*, 3540. [[CrossRef](#)] [[PubMed](#)]
63. Huo, J.; Gao, Y.; Shi, Y.; Yin, H. Cross-Modal Metric Learning for AUC Optimization. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 4844–4856. [[CrossRef](#)]
64. Pal, M.; Mather, P.M. Support vector machines for classification in remote sensing. *Int. J. Remote Sens.* **2005**, *26*, 1007–1011. [[CrossRef](#)]
65. Min, J.H.; Lee, Y.-C. Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Syst. Appl.* **2005**, *28*, 603–614. [[CrossRef](#)]
66. Zhang, L. Management of offshore oil pollution and logistics transportation based on decision tree. *Arab. J. Geosci.* **2021**, *14*, 1655. [[CrossRef](#)]
67. Ma, M.H.; Zhao, G.; He, B.S.; Li, Q.; Dong, H.Y.; Wang, S.G.; Wang, Z.L. XGBoost-based method for flash flood risk assessment. *J. Hydrol.* **2021**, *598*, 126382. [[CrossRef](#)]
68. Buthelezi, M.N.M.; Lottering, R.T.; Hlatshwayo, S.T.; Peerbhay, K. Comparing rotation forests and extreme gradient boosting for monitoring drought damage on KwaZulu-Natal commercial forests. *Geocarto Int.* **2020**, *1*–24. [[CrossRef](#)]
69. Samat, A.; Li, E.Z.; Wang, W.; Liu, S.C.; Lin, C.; Abuduwaili, J. Meta-XGBoost for Hyperspectral Image Classification Using Extended MSER-Guided Morphological Profiles. *Remote Sens.* **2020**, *12*, 1973. [[CrossRef](#)]
70. Peng, J.; Lee, K.; Ingersoll, G. An Introduction to Logistic Regression Analysis and Reporting. *J. Educ. Res.* **2002**, *96*, 3–14. [[CrossRef](#)]
71. Mishra, V.N.; Kumar, V.; Prasad, R.; Punia, M. Geographically Weighted Method Integrated with Logistic Regression for Analyzing Spatially Varying Accuracy Measures of Remote Sensing Image Classification. *J. Indian Soc. Remote Sens.* **2021**, *49*, 1189–1199. [[CrossRef](#)]
72. Schmitt, M. Potential of Large-Scale Inland Water Body Mapping from Sentinel-1/2 Data on the Example of Bavaria's Lakes and Rivers. *PFG-J. Photogramm. Remote Sens. Geoinf. Sci.* **2020**, *88*, 271–289. [[CrossRef](#)]

73. Somasundaram, D.; Zhang, F.; Wang, S.; Ye, H.; Zhang, Z. Learning vector quantization neural network for surface water extraction from Landsat OLI images. *J. Appl. Remote Sens.* **2020**, *14*, 032605. [[CrossRef](#)]
74. Jiang, W.; He, G.; Long, T.; Ni, Y.; Liu, H.; Peng, Y.; Wang guizhou, W. guizhou Multilayer Perceptron Neural Network for Surface Water Extraction in Landsat 8 OLI Satellite Images. *Remote Sens.* **2018**, *10*, 755. [[CrossRef](#)]
75. Yang, L.; Tian, S.W.; Yu, L.; Ye, F.Y.; Qian, J.; Qian, Y.R. Deep learning for extracting water body from landsat imagery. *Int. J. Innov. Comput. Inf. Control* **2015**, *11*, 1913–1929.
76. Roli, F.; Giacinto, G.; Vernazza, G. Comparison and Combination of Statistical and Neural Network Algorithms for Remote-Sensing Image Classification. In *Neurocomputation in Remote Sensing Data Analysis*; Kanellopoulos, I., Wilkinson, G.G., Roli, F., Austin, J., Eds.; Springer: Berlin/Heidelberg, Germany, 1997; pp. 117–124.