

Research on Water Extraction from Remote Sensing Images based on Ada Boost Algorithm

Dan Wu^{1,2,3,4}, Shenglan Ye^{1,2,3,4}

¹ Institute of Land Engineering and Technology, Shaanxi Provincial Land Engineering Construction Group Co., Ltd, Xi'an 710075, China

² Shaanxi Provincial Land Engineering Construction Group Co., Ltd, Xi'an 710075, China

³ Key Laboratory of Degraded and Unused Land Consolidation Engineering, the Ministry of natural Resources, Xi'an 710075, China

⁴ Shaanxi Provincial Land Consolidation Engineering Technology Research Center, Xi'an 710075, China

Abstract: The extraction of water in remote sensing image is a key step in the application of remote sensing image. Aiming at the existing problems in remote sensing image water extraction, a water extraction method based on Ada Boost algorithm is proposed. In this method, multiple thresholds are set for image segmentation, and the results are voted, and finally the water extraction results are obtained. Using the threshold value of 7 algorithms to synthesize, using the Ada Boost algorithm, the advantages and disadvantages of each algorithm are analyzed and compared. The experimental results show that the water body extracted by a single method and a combination of 7 methods is not ideal, but the water body extracted by a combination of two, three, four and five methods is relatively good, among which the best effect is the water body extracted by a combination of three methods.

Keywords: Remote Sensing Image; Ada Boost Algorithm; Water Extraction.

1. Introduction

With the rapid development of remote sensing technology, more and more scholars begin to pay attention to the application of remote sensing technology in geoscience research [1]. As one of the most important landscape types on the surface, water body plays an important role in ecological environment and urban planning. Water extraction from remote sensing images is a key link in the application of remote sensing images. Therefore, water information extraction from remote sensing images is a hot issue in the research of remote sensing technology [2,3]. Since multi-spectral remote sensing images have rich spectral information and high spatial resolution, it is of great significance to use remote sensing image data to extract ground object information. Water bodies have obvious spectral characteristics, but in traditional supervised classification methods, it is usually necessary to manually select various types of ground objects to train samples, and a lot of manpower, material and financial resources are needed to extract information. In recent decades, ground object information extraction based on remote sensing image data has become a research hotspot.

Most traditional water extraction methods are based on threshold segmentation [4]. Threshold segmentation method is to set multiple thresholds to achieve image segmentation in the process of image segmentation. By comparing the results obtained under different thresholds, it can be found that the results obtained by the single threshold segmentation method have some defects. However, the image segmentation results obtained after the fusion of multiple thresholds can effectively solve the defects of traditional methods [5,6]. At present, water extraction algorithms based on machine learning are increasingly applied in actual production, but most algorithms require manual participation, which will not only increase the running time and complexity of algorithms, but also cause information loss to a certain extent [7-9]. Therefore, in order to reduce manual participation and improve the accuracy of

water extraction, this paper proposes a water extraction method based on the combination of Ada Boost algorithm and machine learning. In this method, multiple thresholds are set for image segmentation, and the results are voted to get the final result, thus reducing the manual participation and improving the accuracy of water extraction.

2. Data Source and Research Method

2.1. Data Sources

In order to verify the method proposed in this paper, Landsat8 OLI images of Lanzhou City were captured and analyzed. The OLI Land Imager consists of nine bands with a spatial resolution of 30 meters, including a panchromatic band of 15 meters. OLI panchromatic band Band8 has a narrow range, which can better distinguish vegetation and non-vegetation features in panchromatic images. In addition, there are two new bands: the blue band (band 1; 0.433-0.453 μ m) mainly used in coastal zone observation, short-wave infrared band (band 9; 1.360-1.390 μ m) including strong water vapor absorption characteristics can be used for cloud detection; Near infrared band5 and short wave infrared band9 are close to the corresponding bands of MODIS. Because the acquisition of remote sensing image is affected by many factors, such as atmosphere, ground shadow and image noise, there is a lot of noise in the image. Therefore, in order to extract clean water image, it is necessary to filter and enhance the image in the pre-processing process. Firstly, the mean filter is used to preprocess the remote sensing image to eliminate the noise in the image. Finally, the image enhancement algorithm based on mathematical morphology processing is used to enhance the pre-processed images. The image for analysis is obtained by atmospheric radiation correction, geometric correction and image segmentation.

2.2. Research Methods

AdaBoost algorithm is a voting classification algorithm proposed by Freund et al in 1995, which combines different weak classifiers into strong classifiers through sample

training to improve classification accuracy [11]. In the process of water extraction from remote sensing images, in view of the limitations of a single classifier, the AdaBoost algorithm adopted in this paper can well assign different weights to the water features maximized by each classifier, so that the weak classifier is forced to concentrate on the difficulty of classification. Finally, the learning results are fused with the results of other classifiers by weighted maximum voting, and the accuracy of water extraction is comprehensively improved. In this paper, pixels are classified according to the feature differences in different regions, and the classification results are sorted according to the voting proportion. In this paper, 7 methods, including single-band threshold method, NDWI, MNDWI, 3 different spectral relations and decision tree method, are used as the classifier of AdaBoost algorithm. The specific operations are as follows: Firstly, the image is preprocessed to remove all kinds of noise; Then the optimal segmentation point of each pixel is determined by setting different thresholds. Finally, weighted voting is used to output the segmentation results. This algorithm can not only obtain better segmentation results, but also realize the comparison of segmentation results under different thresholds in the same region. Using MATLAB to implement the Ada Boost algorithm, the implementation code is as follows:

```

I1=imread('danboduananyuzhi.tif');
I2=imread('NDWI.tif');
I3=imread('MNDWI.tif');
I4=imread('pujianguanxi.tif');
I5=imread('wangjinhua.tif');
I6=imread('xuhanqiu.tif');
I7=imread('yanpei.tif');
I=I1+I2+I3+I4+I5+I6+I7;
waterbody1=I>=1;
imwrite(waterbody1,'votealgorithm_oneagree.tif');
waterbody2=I>=2;
imwrite(waterbody2,'votealgorithm_twoagree.tif');
waterbody3=I>=3;
imwrite(waterbody3,'votealgorithm_threeagree.tif');
waterbody4=I>=4;
imwrite(waterbody4,'votealgorithm_fouragree.tif');
waterbody5=I>=5;
imwrite(waterbody5,'votealgorithm_fiveagree.tif');
waterbody6=I>=6;
imwrite(waterbody6,'votealgorithm_sixagree.tif');
waterbody7=(I==7);
imwrite(waterbody7,'votealgorithm_allagree.tif');

```

3. Water Extraction Results

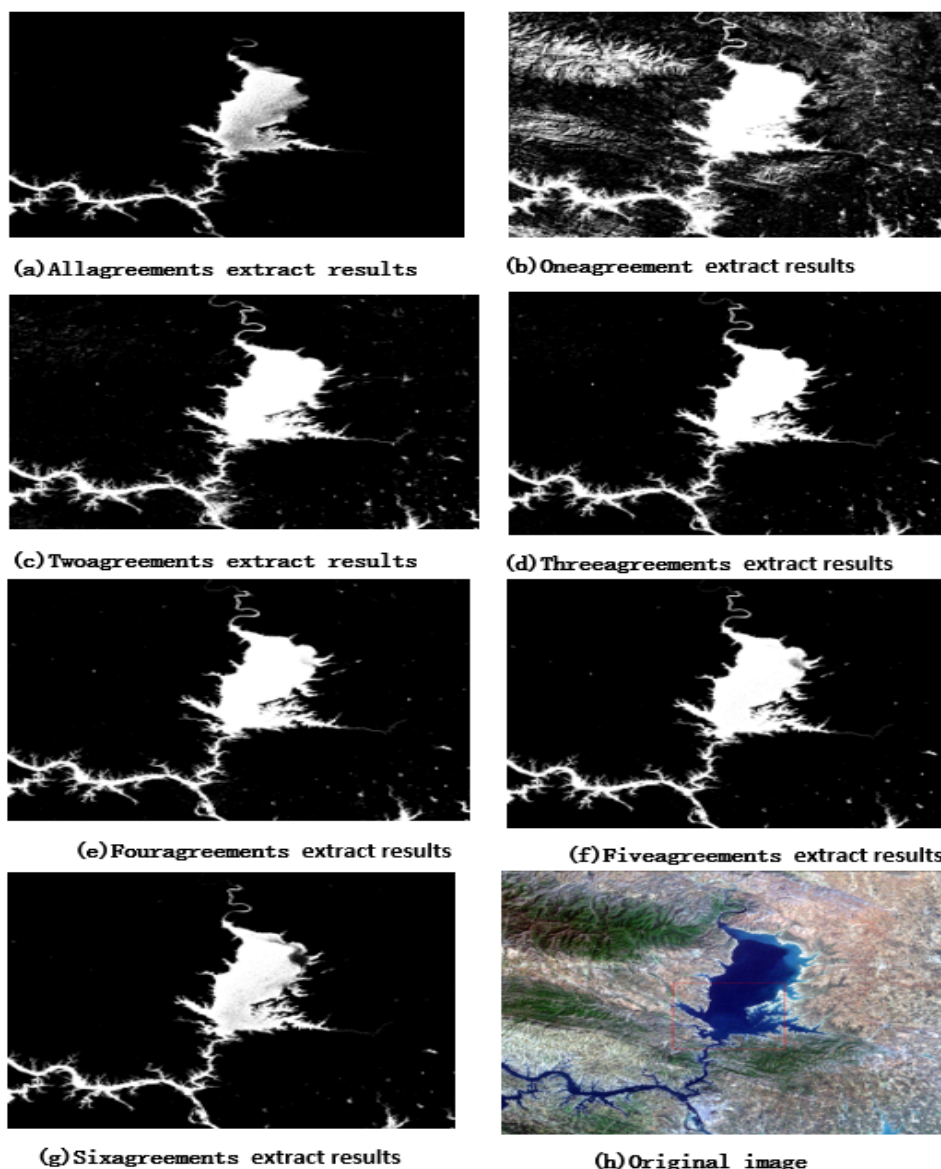


Figure 1. Results of water extraction by Ada Boost algorithm

Through the constructed program, 7 images can be obtained. They are as follows: as long as there is a way to think that the pixel is a water body, it is determined to be a water body, referred to as one agreement; At least two agreements agree that the pixels are water bodies. At least three agreements agree that the pixels are water bodies, referred to as three agreements; There are at least four agreements that identify the pixels as water bodies, referred to as four agreements; At least five agreements agree that pixels are water bodies; There are at least six methods to determine that the pixel is a water body, referred to as six agreements; The pixel must be recognized as a water body by all 7 methods, referred to as all agreements. The Fig 1 is the result.

This Ada Boost algorithm well synthesizes the processing results of different water body search algorithms, and users can choose the processed water body images according to their own needs. The extraction results of one agreement and all agreements are not satisfactory, because they all take extreme values. all agreements treat some bodies of water as other features and one agreement treats some other features as water bodies. Method three agreements, method four agreements and method five agreements have a good extraction effect. These three methods combine the extraction results of most of the methods, so that the image can well distinguish the boundary between water bodies and other features, and extract most of the water bodies. Moreover, this algorithm synthesizes the processing results of different algorithms, so that the image has the advantages of multiple algorithms. However, this Ada Boost algorithm is not specifically proposed according to some special terrain, so its use has limitations, it must exist the results of several existing extraction algorithms.

4. Conclusion

Aiming at the problem of water recognition and extraction in remote sensing images, this paper proposes a method of water extraction in remote sensing images based on Ada Boost algorithm. By synthesizing existing segmentation algorithms, the advantages of different segmentation algorithms are combined, spectral characteristics and spatial characteristics are fully utilized for fusion, and various segmentation algorithms are analyzed and compared. A new segmentation algorithm, Ada Boost algorithm, is proposed. The experimental results show that by synthesizing the existing methods, the water body can be well identified and extracted. However, there are also certain problems. Due to the influence of noise interference, shadow, vegetation and other factors, the phenomenon of misclassification occurs after image segmentation. In the follow-up research, different

methods of image preprocessing will be tried to reduce the influence of experimental data. Therefore, how to eliminate the influence of noise on algorithm accuracy is still the focus of future research.

Acknowledgments

This paper was supported by Shanxi Province Land Engineering Construction Group Project DJNY2022-28.

References

- [1] Rishikeshan, C. A., and H. Ramesh. "An automated mathematical morphology driven algorithm for water body extraction from remotely sensed images." *ISPRS Journal of Photogrammetry and Remote Sensing* (2018).
- [2] Jiang, Wei, et al. "An Effective Water Body Extraction Method with New Water Index for Sentinel-2 Imagery." *Water* (2021).
- [3] Yue, H., et al. "A new accuracy evaluation method for water body extraction." *International Journal of Remote Sensing* (2020).
- [4] Zhuang, Yin, et al. "Harbor Water Area Extraction From Pan-Sharpned Remotely Sensed Images Based on the Definition Circle Model." *IEEE Geoscience and Remote Sensing Letters* (2017).
- [5] Xu X. Research on automatic water extraction scheme of remote sensing image based on image features [J]. *Anhui building*, 2019, 26 (9): 50-51. DOI: 10.16330 / j. carol carroll nki. 1007-7359.2019.09.020.
- [6] Zhao Chenxi, Hu Jingfang, Song Yu et al. Research on water body identification and extraction of image review [J]. *Journal of sensor world*, 2022, 28 (08): 1-9.
- [7] Wang Xin, Xu Mingjun, Li Ke et al. An effective method for water extraction from high-resolution remote sensing images [J]. *Computer Engineering and Applications*, 2019,55(20):145-151+207. (In Chinese)
- [8] Pan, F., X. Xi, and Cheng Wang. "A Comparative Study of Water Indices and Image Classification Algorithms for Mapping Inland Surface Water Bodies Using Landsat Imagery." *Remote. Sens.* (2020).
- [9] Hu, Shunshi, et al. "Automatic Extraction of Water Inundation Areas Using Sentinel-1 Data for Large Plain Areas." *Remote. Sens.* (2020).
- [10] Zhai, K., et al. "Comparison of surface water extraction performances of different classic water indices using OLI and TM imageries in different situations." *Geo-spatial Information Science* (2015).
- [11] Song Yingqiang, Yang Lian 'an, Xu Jingting, Wei Anlei. Water body information extraction based on Landsat-8 OLI remote sensing image and AdaBoost algorithm [J]. *Surveying, Mapping and Geographic Information*,2017,42(03):44-47.