International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 2, Number 1, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v2n1.33
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Water Body Information Extraction from Remote Sensing Images based on PSPNet

Jing Zhang*

Beijing University of Civil Engineering and Architecture, Beijing, China

ABSTRACT

Remote sensing image has the characteristics of real-time, periodicity and wide monitoring range. It can quickly and accurately obtain water area, distribution and other information, which is of great significance to the utilization and development of water resources, agricultural irrigation, flood disaster assessment and so on. Since traditional water information extraction methods only use part of image band information, the accuracy of water information extraction is low and has certain limitations. In recent years, convolutional neural network technology has developed rapidly and achieved good results in water information extraction from remote sensing images. Therefore, in this paper, Pyramid Scene Parsing Neural Network (PSPNet) was used to extract water information from Ziyuan-3 multispectral remote sensing images, to make sample sets of water, and train the convolutional neural network model. Compared with the traditional normalized difference water index (NDWI) and support vector machine (SVM), the results show that PSPNet has the highest accuracy and the lowest misclassification rate.

KEYWORDS

PSPNet; SVM; NDWI; remote sensing image; Water body information extraction

1. INTRODUCTION

As a necessary material guarantee for the survival of life on earth, the distribution of water resources has a significant impact on human life, social economy and environmental governance in the region. Remote sensing image has the characteristics of fast update speed, rich information and wide observation range, which can ensure the real-time and dynamic information. Using remote sensing images to extract water information is of great significance for water resources investigation, water conservancy planning, watershed management, flood monitoring and post-disaster assessment [1]. At present, water body information extraction in China is mainly based on medium and high resolution multi-spectral remote sensing image data sources such as SPOT and Landsat. Water body information extraction algorithms mainly include water body index method, spectral relation method, decision tree method, density segmentation method, image classification method, ratio method, threshold method and difference method, among which water body index method and spectral relation method are relatively mature. However, with the increase of remote sensing image resolution, the geographical information expressed is also more and more large, and the traditional water information extraction method has low precision, which can not meet the needs of The Times. In recent years, deep learning technology has achieved good results in target recognition and target detection. As one of the important fields in deep learning, convolutional neural network uses data-driven methods to carry out deep feature learning and finally realize target recognition. Therefore, this paper uses PSPNet network to extract water information in Nanchang city, Jiangxi Province, and carries out a comparative experiment with traditional water extraction methods NDWI and SVM.

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

2. REVIEW ON METHODS OF WATER INFORMATION EXTRACTION

Foreign studies on water extraction using remote sensing images are relatively early: In 1996, McFeeters proposed Normalized Difference Water Index (NDWI) based on the spectral characteristics of Water body. Since the reflectance of Water body in green light band was different from that in near-infrared band, the Difference between them was normalized and Water information was extracted [2]. In 2002, Ryu used band ratio method (TM4-TM3)/(TM4+TM3) to extract water boundary of tidal flat [3]. In 2014, Feyisa proposed AWEI (Automated Water Extraction Index) to extract Water from Landsat images, and the extraction effect is better in areas with more mountain shadow and cloud shadow [4]. In 2016, Sarp extracted water from Burdur Lake in Turkey by Landsat TM/ETM+ images using Support Vector Machine [5].

The technology of water body information extraction using remote sensing images has developed rapidly in China: In 2005, Xu Hanqiu proposed the Modified Normalized differential water index (Modified NDWI), which used the short-wave infrared band to replace the reflectivity of the near-infrared band, effectively achieved a good extraction effect of water around the city[6]. In 2020, Wang Ning took Chaohu Lake basin in Anhui province as the research area and used U-NET model and random forest model to extract water information from Gaofen-1 satellite images. The results showed that the water extraction results of U-NET model were more consistent with the artificial visual interpretation results and could better eliminate the influence of shadows [7]. In 2021, Wang Fan proposed a New Normalized Difference Water Index (NNDWI) to improve the threshold selection error of the traditional index model in detecting water information from remote sensing images. Combining the reflectance of red, green and near-infrared bands, the reflection difference between water body and background ground object is increased, and the results are binarized to obtain water body information quickly, with high extraction accuracy, avoiding artificial threshold selection error, and can better identify the scattered small water body information [8]. In 2022, In order to solve the problem of monitoring changes of surface water resources based on remote sensing images, Zhang Mingfei collected remote sensing images by means of web crawler and constructed data sets through random clipping and data cleaning. High accuracy of water extraction can be achieved through network model training based on high-level features [9].

Although the above methods can extract most of the water body information, the water body spatial information is not fully used, and there is still a large error of leakage grading. In recent years, with the arrival of the era of remote sensing big data, object recognition algorithms based on deep learning have developed rapidly, providing a more effective method for remote sensing water body information extraction.

3. METHODS

3.1. The Normalized Difference Water Index

In the range of visible and near-infrared bands, water recognition is mainly based on spectral reflection differences of ground objects such as water, vegetation and soil. According to the spectral reflection characteristic curve of ground objects, the reflectance of water body is generally lower than 10% compared with vegetation and soil. Because the reflectance of clear water is the highest in the blue-green light band, its reflectance decreases gradually with the increase of wavelength. In the near infrared band after 0.75μm, the reflectance of water is almost zero, so the NDWI value of water is positive. However, for other ground objects such as vegetation and soil, the reflectance in the near-infrared band is greater than that in the visible band, and their NDWI values are all negative. Water bodies can be distinguished from other ground objects according to the rule of the reflectance of ground objects changing with wavelength.

Normalized Difference Water Index (NDWI) is commonly used to extract Water information. The formula of NDWI as Eq. 1:

$$\mathrm{NDWI} = \frac{\mathrm{Green} - \mathrm{NIR}}{\mathrm{Green} + \mathrm{NIR}}. \tag{1}$$

In the formula, Green is the reflectivity of the visible green band of the sensor; NIR is the reflectivity of near infrared band. According to feature spectrum reflection characteristic curve of visible light green band reflection for surface water is stronger, the near infrared wave band of the rule of the absorption of water is stronger, choose the green band and near infrared band images to establish the ratio of the model can well stand out water vegetation and other feature information, and to extract water information for the purpose of the study area.

3.2. Support Vector Machine

SVM is an optimal boundary classification method, which defines the maximum interval classifier in the feature space. The purpose of SVM is to find a compromise value that minimizes the empirical risk and confidence interval, and at the same time, considering the training error and model complexity, it can obtain better classification effect when the number of samples is small [10]. Figure 1 shows the optimal classification line in the linearly separable sample. H represents the maximum classification interval between data sets, and the maximum classification interval is $2/\|\omega\|$ when the sample set is normalized.

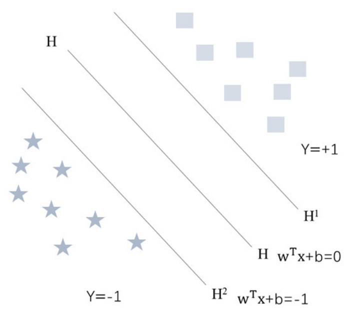

Figure 1. SVM

For nonlinear separable samples, kernel functions need to be introduced. The kernel function can map the sample data in the low-dimensional space to the high-dimensional space, making the samples linearly separable, and its calculation formula is shown in Eq. 2.

$$\mathrm{K(X, Z)} = (\varphi(\mathrm{x}), \varphi(\mathrm{z})). \tag{2}$$

3.3. Pyramid Scene Parsing Neural Network

Hinton proposed deep learning algorithm in 2006, which is a new field in artificial intelligence research [11]. As one of the research fields of deep learning, convolutional neural network has been widely used in remote sensing image classification and target recognition and achieved good results. In this paper, convolutional neural network is used for water object recognition in remote sensing.

Convolutional neural network [12] is a multi-layer feed forward neural network, input layer, convolutional layer, pooling layer, full connection layer, the activation function and output layer. It mainly plays a role in feature extraction layer and feature mapping layer, and is used for visual recognition of images starting from pixel level.

The input layer refers to the input part of the convolutional neural network. For the input of the image, there are a series of pre-processing of the image, including data enhancement, Gaussian blur, etc. The convolutional layer is used to extract the feature information of the input image, and the parameters of the convolutional layer include the preset size, number, moving step, filling size and parameters of the convolutional kernel to be learned. The pooling layer is generally added after the continuous convolution layer to reduce the spatial size of the input feature graph. The full connection layer is usually placed at the end of the network, where the output of each layer of neurons is connected to each neuron in the next layer. The activation function is carried out after the convolution operation in order to activate subsequent neurons only when the weight of signal transmitted by dendrites exceeds the given threshold.

In this study, PSPNet [13] was selected as the network architecture, which was divided into backbone network and enhanced feature extraction network. The backbone network is equipped with MobileNetV2, which has two characteristics. First, the inverted residual structure means 1x1 convolution is deployed before the 3x3 network structure, and 1x1 convolution is deployed after the 3x3 network structure. Second, linear activation function is adopted. Since the inverted residual structure outputs low-dimensional information, linear activation function is used to avoid feature loss, as shown in Figure 2.

Figure 2. MobileNetV2 Network structure

The enhanced feature extraction structure adopted is PSP module, which divides the acquired feature layers into regions of different sizes, and carries out average pooling within each region. Implement aggregation of context information from different regions to improve the ability to obtain global information. The PSPNet network structure is shown in Figure 3.

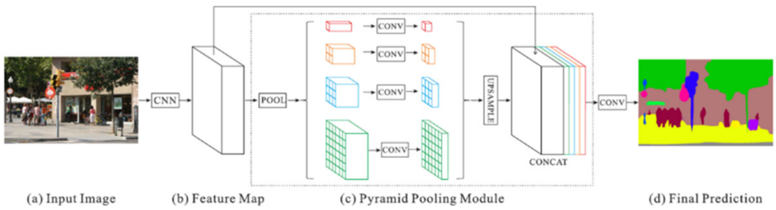

Figure 3. PSPNet network structure

4. EXPERIMENT AND RESULTS

4.1. Data source

Ziyuan-3 multispectral satellites include one face panchromatic TDI CCD camera with a ground resolution of 2.1m, two front and rear panchromatic TDI CCD cameras with a ground resolution of 3.6m, and an face multispectral camera with a ground resolution of 5.8m.

4.2. Experiment Results and Analysis

The experimental environment of this paper is configured as follows: AMD Ryzen 7 5800H with Radeon Graphics CPU, Win10 operating system, pytorch1.2.0 deep learning framework, programming language python, Labelme = 3.16.7 was used to annotate 2/3 data sets of Ziyuan-3 remote sensing images to generate JSON files and convert the JSON files into PNG format, which is the samples used in the training network, as shown in Figure 4.

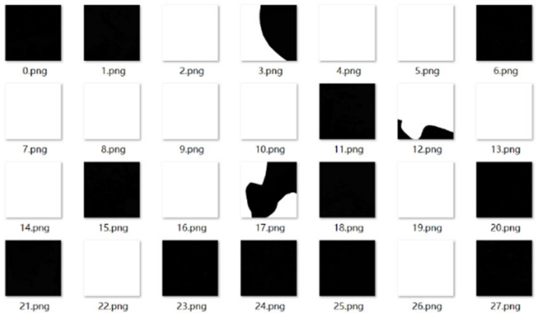

Figure 4. Ziyuan-3 remote sensing image water body training samples

The generated training samples were used to train the model and the training weights were obtained. In order to verify the accuracy of the model, 1/3 data set was used to test the model. The test results are shown in Figure 5:

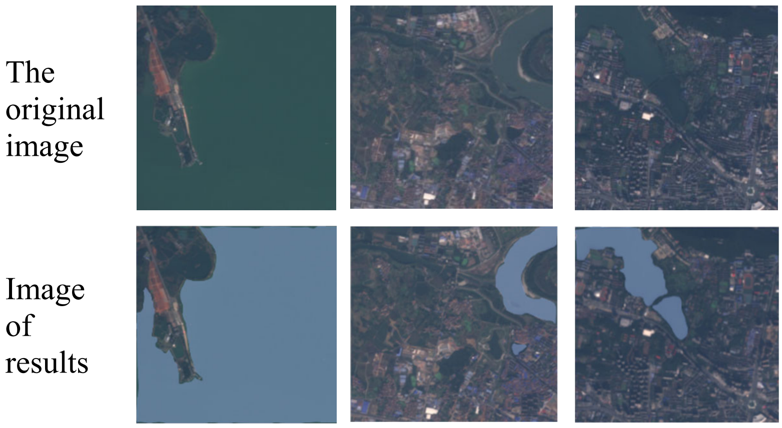

Figure 5. PSPNet model to extract water body information result diagram

Meanwhile, NDWI, SVM and PSPNet were also used for comparative experiments, and the results are shown in Figure 6.

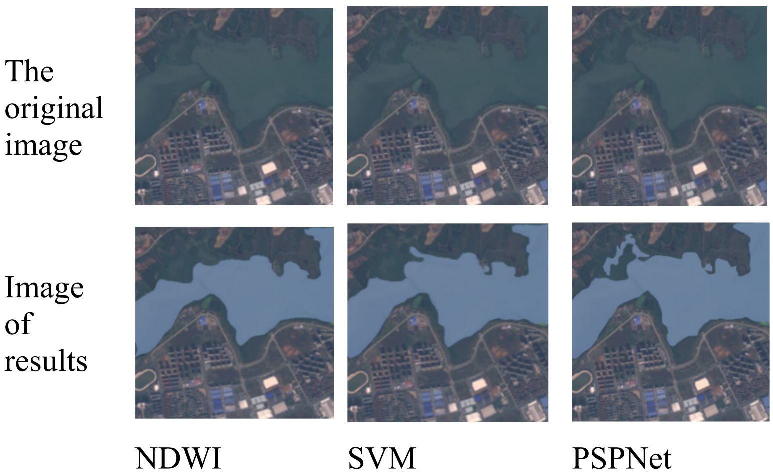

Figure 6. Results of NDWI, SVM and PSPNet were compared

According to the water body results extracted by these three methods, accuracy Eq. 3 and misclassification rate Eq. 4 are used to compare the accuracy of these methods in terms of accuracy and misclassification rate through visual interpretation, as shown in Table 1.

$$\text{Accuracy} = \frac{\text{number of correctly extracted water targets}}{\text{number of real targets}} \times 100\%. \quad (3)$$

$$\text{Misclassification rate} = \frac{\text{number of non water objects extracted}}{\text{number of all objects extracted}} \times 100\%. \quad (4)$$

Table 1. The accuracy comparison of NDWI, SVM and PSPNet

Methods	Accuracy(%)	Misclassification rate(%)
NDWI	88.37	30.62
SVM	82.17	9.69
PSPNet	90.19	6.81

As can be seen from Table 1, NDWI threshold method can be used to extract most water bodies, but some mountain shadows are easily misidentified as water bodies with high accuracy, but the misclassification rate is as high as 30.62%, and the misclassification rate of SVM is relatively small. The accuracy of PSPNet is 90.19%, while the misclassification rate is only 6.81%. The accuracy of PSPNet is significantly better than that of NDWI and SVM, and the better image spot boundary can be obtained.

5. CONCLUSIONS

The main purpose of this study is to accurately extract water body information and provide necessary data support for water resources development, investigation and water conservancy planning, which is of great significance for watershed management, flood monitoring and post-disaster assessment. In view of the difficulty in ensuring the reliability and accuracy of traditional water area extraction methods, this study used PSPNet to extract water information from the multispectral images of domestic Ziyuan-3 remote sensing images. The accuracy of water information extraction was 90.91% and the misclassification rate was 6.81%. The experimental results show that the proposed method is superior to the NDWI and SVM extraction methods. However, this study still has some limitations. First, the resolution of Ziyuan-3 remote sensing image is high, and there is a background similar to the target water body in the image, which leads to the model misclassification and low accuracy. Secondly, due to the problems of target occlusion and similarity between target and background in remote sensing images, appropriate feature extraction modules should be selected according to these problems to improve the model accuracy. Finally, under certain conditions, the training batch is increased to improve the model accuracy.

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

[1] Zhu Changming, et al. " Adaptive Iterative Extraction of Small Linear Water Body in River Channel Assisted by DEMs. Journal of Surveying and Mapping 42.02 (2013): 277-283.

[2] McFeeters, and K. S. . "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features." International Journal of Remote Sensing17.7(1996):1425-1432.

[3] Ryu, J. H., J. S. Won, and K. D. Min. "Waterline extraction from Landsat TM data in a tidal flat: A case study in Gomso Bay, Korea." Remote Sensing of Environment 83.3(2002):442-456.

[4] Feyisa, G. L. , et al. "Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery." Remote Sensing of Environment 140.1(2014):23-35.

[5] Sarp G, Ozcelik M. Water body extraction and change detection using time series: A case study of Lake Burdur, Turkey. Journal of Taibah University for Science 11.3(2017): 381-391.

[6] Han-Qiu, X. U. . "A Study on Information Extraction of Water Body with the Modified Normalized Difference Water Index (MNDWI)." Journal of Remote Sensing 05(2005):589-595.

[7] Wang Ning, et al. " Application of U-net model in water extraction from high-resolution remote sensing images. "Land and Resources Remote Sensing 32.01 (2020): 35-42.

[8] Wang Fan, et al. "An Improved Method for Fast Extraction of Water Information from Remote Sensing Images." People's Yangtze River 52.06 (2021): 223-228. doi:10.16232/j.cnki.1001-4179.2021.06.038.

[9] Zhang Mingfei, et al. "Water Extraction from Remote Sensing Images Based on Convolutional Neural Network." Sensors and Microsystems 41.01 (2022): 72-74 + 88. doi:10.13873/J.1000-9787(2022)01-0072-03.

[10] Ak, J. , SuykensJ., and Vandewalle. "Least Squares Support Vector Machine Classifiers." Neural Processing Letters (1999).

[11] Hinton, G. E., et al. "Improving neural networks by preventing co-adaptation of feature detectors." Computer Science 3.4(2012):pages. 212-223.

[12] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012).

[13] Zhao, H., et al. "Pyramid Scene Parsing Network." IEEE Computer Society IEEE Computer Society, 2017.