WATER BODY EXTRACTION FROM REMOTE SENSING IMAGE OF *WULIANGSUHAI* LAKE BASED ON FULLY CONVOLUTIONAL NEURAL NETWORK

Tianming ZHAO[1], Honghui LI[*], Hua HU[2], Xueliang FU[3]

Accurate extraction of water body from remote sensing images is of great significance to monitor the environmental changes of Wuliangsuhai Lake. However, the interference of complex background and surface reed make it difficult to accurately extract the water body of Wuliangsuhai Lake. This paper proposes a novel method (FCN-OPT) of water body extraction based on Fully Convolutional Networks (FCN). The FCN-OPT method first uses FCN to extract and fuse the deep-level features of the image to achieve pixel-level segmentation. It then uses the OpenCV computer vision library to remove the wrongly extracted part and holes of the FCN segmentation result to improve the accuracy of water extraction. Thereby the FCN-OPT method achieves accurate and complete Wuliangsuhai Lake water body extraction. Also, it can extract the bright water area and reed area of Wuliangsuhai Lake simultaneously and eliminate the interference of complex background, significantly reducing the misdivision and leakage of the water body. The experimental results show that, in comparison with Multi-band Spectral Relationship Method and Normalized Difference Water Index (NDWI) method, the maximum Pixel Accuracy (PA) of FCN-OPT is improved by 21.44% and 20.58% respectively, the maximum Mean Pixel Accuracy (MPA) enhanced by 35.03% and 34.56% respectively, the Mean Intersection over Union (MIoU) increased by 17.51% and 28.19% respectively. Moreover, the proposed method FCN-POT achieves good performance when extracting the complete water body of Wuliangsuhai Lake from the remote sensing images of different seasons.

Keywords: Water Body Extraction, Wuliangsuhai Lake, Fully Convolutional Network, Remote Sensing Images

1. Introduction

The Wuliangsuhai Lake is one of the eight freshwater lakes in China. It plays a vital role in maintaining the balance of ecosystems in the surrounding area

[1] M.S., College of computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot, China, e-mail: zhaotm0609@163.com
[*] Prof., Corresponding author, College of computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot, China, e-mail: lihh@imau.edu.cn
[2] College of computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot, China
[3] College of computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot, China

and the middle-up stream of the yellow river. It has many functions, such as irrigating farmland, regulating rivers, regulating the regional ecological environment, etc. However, in recent years, with the development of human society, a large amount of domestic water and industrial sewage has been discharged into Wuliangsuhai Lake, which has caused it to become a heavily eutrophic lake. The inner and boundary parts of the lake were covered by aquatic plants and turned into reed areas, resulting in a decrease in the size of the clear water areas and severe degradation of ecological functions [1]. High-resolution remote sensing images contain multi-segment spectral information and rich spatial feature information [2]. By extracting the lake area in the remote sensing image, can effectively monitor the ecological environment changes in the Wuliangsuhai Lake. Therefore, it is of tremendous research significance to extract the water area of Wuliangsuhai Lake in the remote sensing image wholly and accurately.

At present, the traditional water extraction methods mainly include the single-band threshold analysis method [3], multi-band spectral relation method, water index method and classifier method [4]. The single band threshold analysis method [5] generally uses the reflectivity difference between the water part and the background part in the near-infrared band or mid-infrared band to determine a threshold through experiments. This method is relatively simple, but it usually over extracts and has poor accuracy. The multi-band spectral relationship method [6] analyzes the gray value characteristics of the water body by combining multiple bands, and then extracts the water body. Bi Shuoben et al. [7] improved the Spectral relation method, they selected the 2, 3, 4, and 5 bands of Landsat TM images to establish a new water body extraction model, making it easier to distinguish the water body from the background. Wang Guohua et al. [8] proposed a new method of inter spectral relationship based on the threshold, which can eliminate the influence of shadow areas and has a better extraction effect on small water bodies. The basic principle of the water index method is to select multiple bands for ratio operation to highlight the difference between the water part and the background part, the Normalized Difference Water Index (NDWI) [9] based on the green band and the near-infrared band is commonly used. Xu Hanqiu et al.[10] proposed an improved normalized water index (MNDWI), which selects the green band and mid-infrared band for calculation and can perform accurate water extraction within the range of cities and towns. The classifier method generally selects representative images as training samples to train the classifier [11]. It then generates a classification model to extract the water body according to the difference of spectral features between the water body and the background, such as support vector machine (SVM) [12], decision tree [13] method, object-oriented method, and active learning with a limited number of training samples are widely used. Compared with the single-band threshold method, the multi-band spectral

relation method, and the water index method, the classifier method can achieve more accurate water extraction.

However, part of the water surface of the Wuliangsuhai Lake water area is covered with aquatic plants and algae, and the overall color and reflectivity of the lake water area are inconsistent, showing significant differences. The extraction accuracy of the Wuliangsuhai Lake water body can be greatly improved by using the traditional extraction method.

Deep learning technology provides an effective way to solve the above problems. Alex Krizhevsky et al. [14] proposed a convolutional neural network model called Alexnet, which learns more abstract image features through training many data samples to realize image recognition. Anamaria Radoi et al.[15] Using a pre-trained deep neural network architecture based on ResNet50 achieved high accuracy on multi-label classification tasks. But the convolutional neural network has learned more abstract features of the image, which is more suitable for judging the image category, and it is difficult to extract the objects in the image finely. Long et al. [16]proposed Fully Convolutional Network (FCN) structure, which was based on the existing convolutional neural network structure, modified the fully connected layers to the convolutional layers, and used deconvolution to restore the abstract features of the image to the categories of the image pixels, to realize the semantic segmentation of the image. Yang et al. [17] proposed an R-CNN model based on the mask to automatically detect and extract water body in remote sensing images, thus avoiding the complex operation of manual feature extraction. Liang Zeyu et al. [18] used dense connection structure and full convolutional network to extract large and small water bodies accurately. Wang Xue et al. [19] proposed a Fully Convolutional Neural Network model for water object extraction from remote sensing images, which achieved good results in both the degree of automation and accuracy. Yu et al. [20] proposed a novel self-attention capsule feature pyramid network; they enhanced the ability of feature representation through the integration of context-augmentation and self-attention modules, achieved superior performance in extracting water body with different surfaces and backgrounds. Tambe Rishikesh G. et al. [21] proposed an end-to-end multi-feature architecture called W-net; they introduced two refinement modules in the basic neural network model, which can enhance prediction results of water body, reduce blur effect, and detect continuity of boundary pixels.

The remote sensing image of Wuliangsuhai Lake has a complicated background, and the water body shows two parts of bright water area and reed area on the image. It is difficult to extract the water body completely and accurately using traditional methods. In view of this situation, this paper proposes an optimized water extraction method FCN-OPT based on Fully Convolutional Networks. First, we establish a remote sensing image data set of the Wuliangsuhai Lake water body. Then build and train a Fully Convolutional Network to obtain a

water extraction model and use this model to extract water body from the remote sensing image of Wuliangsuhai Lake. Finally, we optimized the extraction results to remove the wrong background and internal holes. The experimental results show that the FCN-OPT method proposed in this paper can reduce the false and missing extraction of the water body, realize the complete and accurate extraction of the water body.

2. Data source

NASA launched the Landsat8 satellite in February 2013 with an altitude of 705km, an orbital inclination of 98.2°, a revisit period of 16 days, and a spatial resolution of 30m. It carries the OLI land imager and TIRS thermal infrared sensor. Unlike the Landsat7 satellite, Band5 in the OLI land imager has a band range of 0.845-0.855um, so the influence of water vapor absorption at 0.825um on the image is excluded, the Band8 panchromatic band is narrower than Landsat7, so the difference between the vegetation area and the non-vegetation area in the image is more prominent. In addition, blue band Band1 [22] and shortwave infrared band Band9 were added to observe coastal zone and water consumption. The data used in this paper are from the United States Geological Survey (USGS) website. We select images with less than 10% cloud cover for high definition. The image of Wuliangsuhai Lake is shown in Fig. 1.

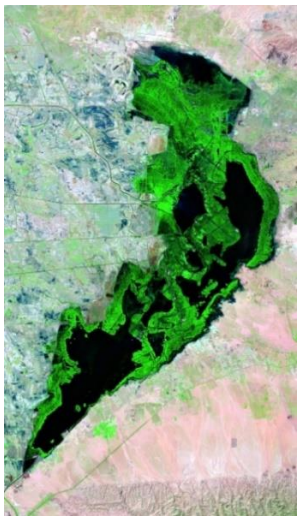


Fig. 1. The Landsat8 OLI image of Wuliangsuhai Lake. Acquired on June 6, 2015. Pseudo-color synthesis of SWIR2, NIR, and Red bands.

3. Water extraction method based on FCN

The workflow of the FCN-OPT method based on the Fully Convolutional

Network (FCN) for extracting lake waters from the Wuliangsuhai Lake remote sensing image is shown in Fig. 2. It mainly includes three steps: Firstly, we obtain the remote sensing image of the Wuliangsuhai area and establish a data set. Then, we construct the FCN network to train the water body extraction model and save the final model. Finally, we use the model to extract the waters of the Wuliangsuhai Lake in the image and optimize the result.

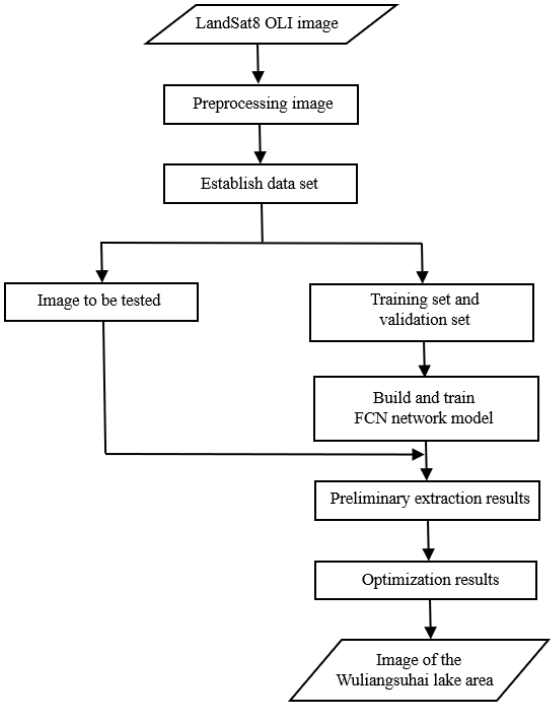


Fig. 2. The workflow of the FCN-OPT method

3.1. Preprocessing images and building data sets

The water body of Wuliangsuhai Lake is seriously eutrophic, and a large number of reed, yellow moss, and other plants grow on the water surface. Therefore, the color of the water body on the remote sensing image of Wuliangsuhai Lake is not uniform, and there is no appropriate public data set, so this paper establishes my own data set.

In order to improve the accuracy of remote sensing images, we processed the original remote sensing images obtained from USGS by radiometric calibration and atmospheric correction [23], and the small regional images of Wuliangsuhai Lake are obtained from sizeable remote sensing images through image clipping. In order to clearly distinguish the Wuliangsuhai Lake part from the ground part, band combination processing was carried out on the cropped images. After the band combination processing, the resolution of each image is

897x1569 pixels, and all processed images are divided into test and used for FCN model training. Because the input image of the VGG19 model is 224x224 pixels, if the whole image is directly input to the VGG19 model, the image will be resized. Then the image will lose some features, and the model will not obtain the deeper semantic information of the image. Moreover, it takes up more memory, and the complexity of model training will increase. So, the 897x1569 pixels image for FCN model training is segmented with a 224x224 pixels sliding window to generate sub-images. The step for the sliding window is 224 pixels. This paper divides the remote sensing image into two parts: the lake area part and the background part, mark the lake area part in the sub-images, the rest is background. We get the remote sensing image data set of Wuliangsuhai Lake and divide the data set into the training set and verification set.

3.2. Build FCN model

3.2.1. Introduction to FCN

Convolutional Neural Networks (CNN) is mainly applied in the field of image classification [24]. After the image passes through the convolution layer, ReLU excitation layer, pooling layer, and fully connected layer, the category of the input image can be well judged, but the objects in the image cannot be accurately segmented. Based on the existing CNN model, FCN replaces the fully connected layer of CNN with the convolution layer. After the image passes through the multiple convolutional layers, the output is no longer the probability value of the category to which the image belongs but a rough feature map. Then the deconvolution layer is used for upsampling the feature map to obtain a gray image of the same size as the original image to determine the category of each pixel in the input image. The output is a marked picture, which realizes the extraction of a specific area and completes the semantic segmentation of the image [25].

3.2.2. Build FCN model for water body extraction

The FCN model is usually built on the basis of the VGG neural network model [26]. VGG network model was proposed by the Visual Geometry Group of Oxford University, and it extracts features by stacking multiple small convolution cores (3x3) instead of large convolution cores, which can reduce the amount of parameters and improve the network depth to extract more advanced features of the image. VGG19 model is a variant of the VGG model, including sixteen convolution layers and three fully connected layers. It increases the number of convolution layers to improve the non-linear expression ability and learn more complex image features. At the same time, the problems to be learned at each layer are decomposed into simple problems to perform efficient learning.

Based on the VGG19 network model, this paper first retains the sixteen convolution layers of the VGG19 network structure, discards three full connection layers, adds one pool layer and three convolution layers. We migrate the pre-trained weight parameters of the VGG19 model and then build the FCN network model used to extract the water body of Wuliangsuhai Lake. The constructed network structure is shown in Fig. 3.

In Fig.3, the image represents the input image. The convi represents the convolution layer and the activation function. The pool represents the pooling layer, x*upsampled represents the x-times upsampling layer Softmax represents the loss function and "∑" represents feature fusion.

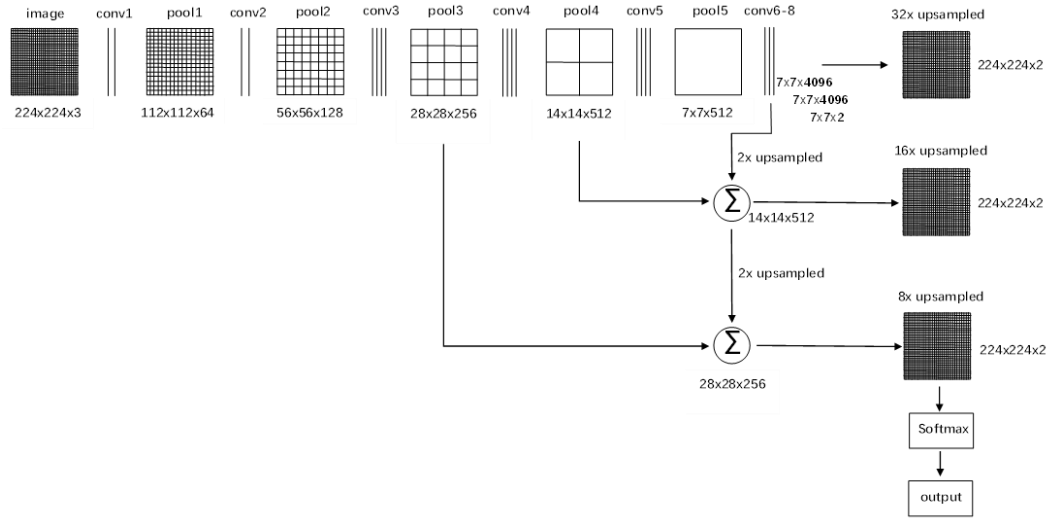


Fig. 3. Structure of FCN network

The FCN network model is mainly composed of the following parts:

1) Convolution layer: Convolutional layer extracts different features of the input image through convolution operations, and then it stores the features in multiple convolution kernels. Convolution kernels optimize the parameters through the back propagation algorithm to get closer to the image features. The convolution kernel details of each convolution layer in the FCN model constructed in this paper are shown in Table 1.

Table 1

Details of convolution kernel in each convolution layer of FCN model

Convolution layer	kernel_number	kernel_size
conv1_1—conv1_2	64	3x3
conv2_1—conv2_2	128	3x3
conv3_1—conv3_4	256	3x3

Convolution layer	kernel_number	kernel_size
conv4_1—conv4_4	512	3x3
conv5_1—conv5_4	512	3x3
conv6	4096	1*1
conv7	4096	1*1
conv8	2	1*1

2) Activation layer: Non-linear activation function can realize the non-linearization of the network, delete redundant features, and retain useful features. In this paper, the ReLU function [27] is used as an activation function to alleviate network gradient disappearance and accelerate the training speed. Add ReLU function after each convolution layer, and the formula is as shown in (1).

$$F(x) = \max(0, x) \tag{1}$$

In formula (1), x denotes the input of the activation function.

3) Pooling layer: The pooling layer can compress the input image, reduce the dimensionality of the image while retaining main features, speed up the model training, improve the generalization ability of the model, and prevent overfitting. This paper uses the average pooling layer, and the pooling window is 2x2, the pooling step is 2. After the pooling layer processes the image, the size becomes half of the input size. The formula of the pooling layer is as shown in (2).

$$F_{ij} = \frac{1}{c^2}\left(\sum_{i=1}^{c}\sum_{j=1}^{c} G_{ij}\right) + b_1 \tag{2}$$

In formula (2), F_{ij} denotes the output of the pooling layer, G_{ij} denotes the eigengraph matrix obtained by the convolution layer, c denotes the side length and slide step of the pooling window, b_1 denotes the Offset.

4) Upsampling layer: After the convolution and pooling operations of the first five layers, five feature maps are obtained, the resolutions are one half, one fourth, one eighth, one sixteenth, and one thirty-second of the original resolution, and then after the final three-layer convolution operation, the resolution of the obtained feature map is still one thirty-second of the original image. In order to obtain the classification information of pixels in the original image [28], FCN uses the deconvolution upsampling method to restore the feature map to the resolution of the original image. If the feature map output by Conv8 is directly upsampled by 32 times, only the features of the convolution kernel in conv5 are restored, the result is rough and not satisfactory enough. Therefore, FCN introduced a skip structure to integrate the features extracted by conv3, conv4, and conv5 to supplement image details and improve segmentation accuracy. First,

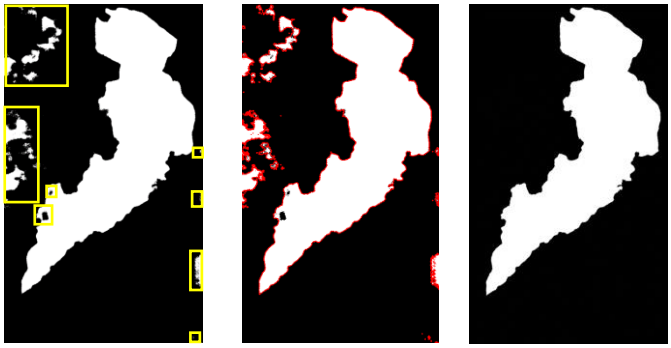
FCN performs 2x deconvolution upsampling two times on the feature map output by conv8. Then FCN performs add operation with the feature map output by pool4 through element-wise addition and perform two times deconvolution upsampling on the result. Finally, FCN adds it with the feature map output by pool3 through element-wise acquisition and then performs 8x deconvolution upsampling to restore the fused feature map to the input image size.

5) Softmax: After upsampling is completed, we use the Softmax function to calculate the probability of each pixel belonging to each category and classify all pixels. The formula is as shown in (3).

$$Soft\max(z_i) = \frac{e^{z_i}}{\sum_{c=1}^{C} e^{z_c}} \tag{3}$$

In formula (3), z_i denotes the output value of the i-th node, C denotes the number of categories to be classified.

3.2.3. Optimization method



(a) (b) (c)

Fig. 4. Optimization process. (a) Initial extraction result of. (b) Detection image of the outermost contour of all regions in (a). (c) Final optimization result of removing interference and holes.

The preliminary extraction results of the water body using the trained FCN model are shown in Fig. 4 (a). It can be seen that there are wrong parts in the background and interior of the water body. For this reason, this paper uses the OpenCV computer vision library to remove the wrong part of the preliminary extraction results. First, we find the outer contours of all areas in Fig. 4 (a), as shown in Fig. 4 (b). Then sort all contours by area size, the contour with the largest area is regarded as the water body of Wuliangsuhai Lake, and other contours are removed. Finally, the contour with the largest area was filled to remove the holes, and the final result was obtained, as shown in Fig. 4 (c).

4. Experiment and discussion

In order to verify the water body extraction method of Wuliangsuhai Lake based on FCN proposed in this paper, a large number of experiments were carried out and compared with the Multi-band Spectral Relationship Method and NDWI method. The evaluation indexes adopted include Pixel Accuracy (PA), Mean Pixel Accuracy (MPA), and Mean Intersection over Union(MIoU). Among them, MIoU is a standard measurement in semantic segmentation [29], so it is used as the main evaluation index in this paper.

1) Pixel Accuracy

PA is the proportion of correctly classified pixel points in all pixel points. The formula is as shown in (4).

$$PA = \frac{\sum_{i=0}^{k} p_{ii}}{\sum_{i=0}^{k}\sum_{j=0}^{k} p_{ij}} \tag{4}$$

In formula (4), k represents the total number of categories, p_{ij} represents the number of pixels that belong to category i but are misclassified into category j, p_{ii} represents the number of pixels that actually belong to category i and are correctly classified into category i.

2) Mean Pixel Accuracy

The MPA is to calculate the pixel accuracy of each category respectively and then calculate the mean value of the pixel accuracy of all categories. The formula is as shown in (5).

$$MPA = \frac{1}{k+1}\sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} pij} \tag{5}$$

In formula (5), the definitions of k, p_{ij} and p_{ii} are the same as formula (4).

3) Mean Intersection over Union

The MIoU is the ratio of intersection and union of real value and predicted value in each category, then sum and average them. The intersection is the set of correctly classified pixels, and the union is the set of correctly classified and incorrectly classified pixels. First, calculate the intersection and union ratio of each class, and then calculate the average value of all classes intersection and union ratio. Ideally, the value of MIoU is 1. The formula is as shown in (6).

$$MIoU = \frac{1}{k+1}\sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij} + \sum_{j=0}^{k} p_{ji} - p_{ii}} \tag{6}$$

In formula (6), p_{ji} represents the number of pixels that belong to category j but are misclassified into category i.

In the experiment, we first establish the data set and divide it into training and verification sets. Then we build the Fully Convolutional Neural Network and train the water extraction model, use the model to extract the images to be tested and optimize the extraction results. Finally, to verify the performance and effectiveness of the FCN-OPT, we compared it with the Multi-band Spectral Relationship Method and NDWI method and extracted the water body from the remote sensing images in different seasons.

4.1. Establish data set

We download Landsat8 OLI satellite images of the Wuliangsuhai Lake area from 2013 to 2020 on the official website of the USGS (http://www.usgs.gov/). We use ENVI5.3 software to perform radiometric calibration and atmospheric correction processing on the image. Then, we select the Wuliangsuhai Lake area by square and crop it from the entire remote sensing image to obtain the original image of the Wuliangsuhai Lake area. The size of the image is 897 x1569 pixels.

We use the band combination function of ENVI software to complete the band combination. Specifically, assign the SWIR2, NIR, and Red bands to the image's R, G, and B bands, respectively, then do the 2% linear stretch. Finally, a clearer original image of the Wuliangsuhai Lake area can be obtained.

The sliding window is used to segment each remote sensing image for training after band combination processing into 28 sub-images. We filter the sub-images and retain the images with both water body and land parts. We use LabelMe software to manually mark the outline of the lake area in each sub-image. The interior of the outline is used as foreground and the rest as background. The marked file is converted into a png format image and processing standard normalization method. We set the pixel value of the background part to 0 and set the pixel value of the lake area part to 1. Finally, 226 pairs of samples were generated for the training model, which was divided into 158 pairs of training samples and 68 pairs of validation samples by a ratio of 7 to 3. There are also 20 unsegmented remote sensing images for testing.

4.2. Training model

The experiment in this paper uses the TensorFlow deep learning framework under the Windows 10 system, uses Python to write programs, the CPU is Intel(R) Core(TM) i5-6500 @ 3.2GHz, the GPU is NVIDIA GeForce GTX 1660 6GB, the RAM is 8GB.

Considering the training speed and prediction accuracy of the model, the training parameters are finally set as follows after several experiments: the batch size is 2, the number of classes is 2, the initial learning rate is 0.0001, the number of iterations is 60,000. We choose Adam optimizer, and the average value of

cross-entropy is used as the loss function. During the first 20,000 training sessions, the loss function value decreased rapidly but fluctuated greatly. After 60,000 times of training, the loss function value gradually converged and stabilized. The variation curve of loss function during training is shown in Fig. 5.

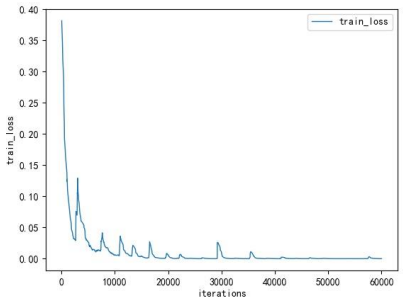


Fig. 5. The change of loss function value during training

4.3. Experimental results and analysis

4.3.1. Comparison of different extraction methods

This section tests the FCN-OPT method's performance by comparing it with the Multi-band Spectral Relationship Method and the NDWI method. Randomly select the images to be tested for experiments, and some extraction results are shown in Fig. 6.

From the visual interpretation point of view, the extraction result of the Multi-band Spectral Relationship Method and the NDWI method are similar, these two methods can extract the water surface area with dark color, but they can not form a complete lake area. They cannot eliminate the interference of complex color background; most of the background outside the lake area is wrongly extracted. And the FCN-OPT method has the best effect and can extract a clear and complete lake area.

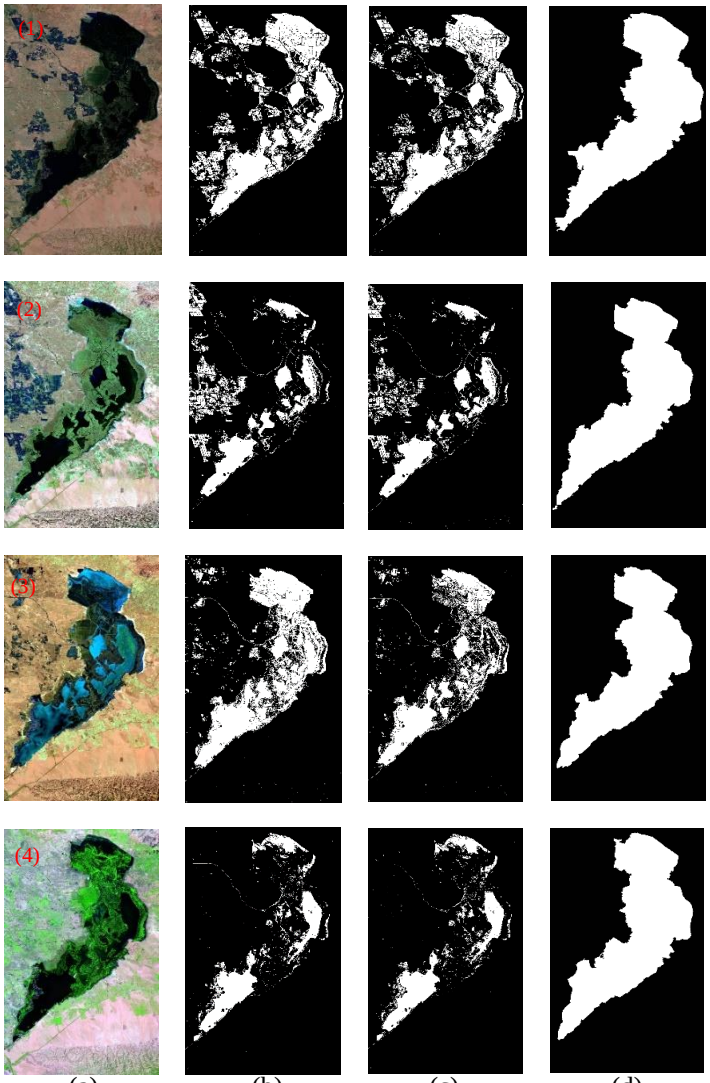


Fig. 6. The extraction results of three kinds of methods. (a) Original image of the test image. (b) The result of water body extraction from column (a) by the Multi-band Spectral Relationship Method. (c) The result of water body extraction from column (a) by the NDWI method. (d) The result of water body extraction from column (a) by the FCN-OPT method.

In order to make a precise comparison, the PA, MPA, and MIoU evaluation indexes of the results by three methods in Fig.6 are shown in Table 2.

Table 2 lists the differences in PA, MPA, and MIoU of the extraction results in Figure 6. Compared with the Multi-band Spectral Relationship Method and the NDWI method, the PA of the FCN-OPT method increased by up to

21.44% and 20.58% most, the MPA of the FCN-OPT method increased by up to 35.53% and 34.56% most.

Table 2

Extraction results of the three methods

Image	Evaluation Indexes	Multi-band Spectral Relationship Method (%)	NDWI Method (%)	FCN-OPT Method (%)
(1)	PA	84.59	85.16	96.48
	MPA	78.25	77.99	97.21
	MIoU	67.36	67.82	91.87
(2)	PA	76.71	77.57	98.15
	MPA	62.85	63.32	97.88
	MIoU	50.96	51.70	95.52
(3)	PA	91.26	86.75	98.69
	MPA	85.55	76.59	98.38
	MIoU	79.25	68.57	96.76
(4)	PA	82.01	81.14	97.69
	MPA	67.15	65.51	97.93
	MIoU	57.23	55.21	94.48

At the same time, it can be seen from Table 2 that the max MIoU of the FCN-OPT method is 96.76%, which is 17.51% higher than the Multi-band Spectral Relationship Method and 28.19% higher than NDWI method, which are much higher than the other two methods. It shows that the FCN-OPT method is better than the other two extraction methods. The reason is that the Multi-band Spectral Relationship Method and NDWI method can effectively extract the bright water area of Wuliangsuhai Lake, but the extraction effect for the reed area covered by aquatic plants is terrible. Moreover, the background interference outside the lake area cannot be excluded. The FCN-OPT method has learned the complex features of the water body on the Wuliangsuhai Lake remote sensing image through model training to extract the bright water area and the reed area together. The boundary of the extracted water body is clear, which greatly reduces the false extraction of the background and the omission of the water body. A complete extraction result can be obtained with higher accuracy. At the same time, the FCN-OPT method realizes automatic extraction without the manual setting of parameters.

4.3.2. Extraction results of images in different seasons

Wuliangsuhai Lake is a typical lake in cold and arid areas, and the spectral characteristics of remote sensing images vary significantly with seasons. In order to further verify the effectiveness of the FCN-OPT method, we selected the images of Wuliangsuhai Lake in spring, summer, autumn, and winter from the

images to be tested and used the FCN-OPT method to extract the lake area. Part of the extraction results are shown in Fig. 7.

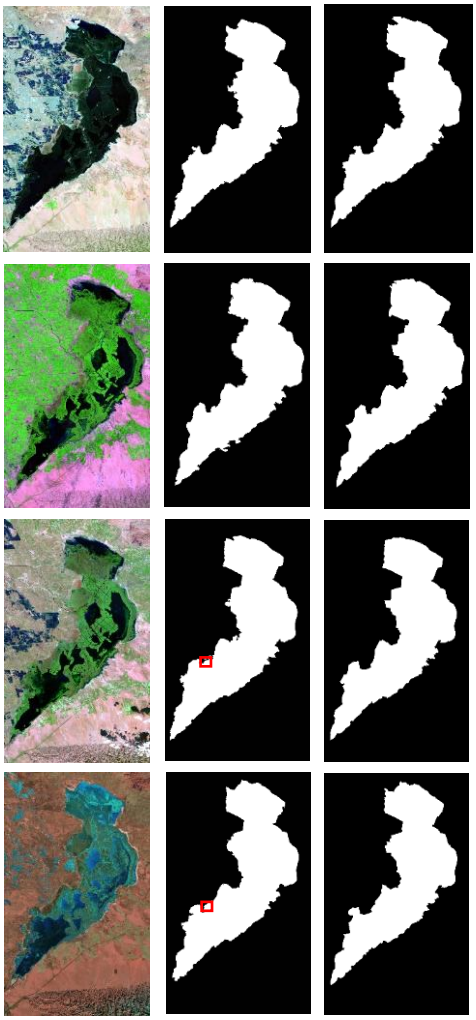


Fig. 7. Extraction results in different seasons. (a) Four images of spring, summer, autumn, and winter from top to bottom, (b) Extraction results using FCN-OPT method, (c) reference standard for extraction results.

From Fig. 7, one can see that, compared with the reference standard, the images of Wuliangsuhai Lake in spring and summer can be extracted entirely by the FCN-OPT method. But there is a small area of miss extraction in the boundary part of the image extraction results in autumn and winter, as shown in the red box in column (b) of Fig.7. Reason analysis found that in the training samples of some years, this area is part of the lake water area so the FCN-OPT method will have the wrong extraction phenomenon in this area.

We Use PA, MPA, and MIoU to evaluate the extraction results of the FCN-OPT method, as shown in Table 3. There are few wrong extraction areas in the spring and summer images of Wuliangsuhai Lake, so the three evaluation indexes are higher than the extraction results in autumn and winter. The PA reaches 98.68%, and the MPA reaches 98.28%. The extraction result of summer images is the best, and the MIoU is 96.84%. The extraction result of the winter image is the worst, and the MIoU is 95.36%, which is 1.48% lower than the summer image. However, the evaluation indexes of Wuliangsuhai Lake in four seasons are all above 95%, and the overall extraction accuracy is high, indicating that the FCN-OPT method can obtain more accurate extraction results.

Table 3

Comparison of image extraction effects in different seasons

Image	PA (%)	MPA (%)	MIoU (%)
spring	98.65	98.28	96.68
summer	98.68	98.15	96.84
autumn	98.48	97.67	96.32
winter	98.10	97.17	95.36

5. Conclusions

This paper proposes an optimized extraction method FCN-OPT based on the Fully Convolutional Network to extract the water body of Wuliangsuhai Lake. The FCN-OPT method establishes remote sensing image data set of Wuliangsuhai Lake in the beginning. Then it constructs the Fully Convolution Network and trains the model to classify each pixel of the Wuliangsuhai Lake remote sensing image. After that, the lake area is extracted preliminarily, and the interference part is removed. A clear and complete lake area can be obtained finally. In comparison with other water body extraction methods, the experimental results show that the FCN-OPT method is superior to the Multi-band Spectral Relationship Method and the NDWI method, the PA increased by up to 21.44% and 20.58%, the MPA increased by up to 35.03% and 34.56%, the MIoU reached 96.76%, increased by up to 17.51% and 28.19% respectively. For Wuliangsuhai images in different seasons, the FCN-OPT method can obtain ideal results. PA, MPA, and MIoU indexes are all above 95%, although there are small area mislift defects.

At present, the study on water body extraction has only been carried out in Wuliangsuhai Lake, not in other lakes. The next step is to increase the data set according to other lakes' characteristics to achieve high-precision water body extraction for different types of lakes.

Acknowledgments

This research was financially supported by the National Natural Science Foundation of China (62041211, 61962047), the National Key Research and Development Program of China (2019YFC049205), the Natural Science Foundation of Inner Mongolia (2020MS06011, 2019MS06015).

R E F E R E N C E S

[1]. *D. Ba,* "Characterization of spatial and temporal distribution of environmental factors of Wuliangsuhai Lake and evaluation of its eutrophication", Transactions of Oceanology and Limnology, no. 04, 2019, pp. 108-114.

[2]. *Y. Chen, R. Fan, X. Yang, J. Wang, A. Latif,* "Extraction of Urban Water Bodies from High-Resolution Remote-Sensing Imagery Using Deep Learning", Water, **vol.** 10, no. 5, 2018.

[3]. *U. Nguyen, L. Pham, T. D. Dang,* "An automatic water detection approach using Landsat 8 OLI and Google Earth Engine cloud computing to map lakes and reservoirs in New Zealand, Environmental Monitoring and Assessment", **vol.** 191, no. 4, 2019, pp. 1-12.

[4]. *D. Li, B. Wu, B. Chen, Y. Xue, Y. Zhang,* "Review of water body information extraction based on satellite remote sensing", Journal of Tsinghua University (Science and Technology), **vol.** 60, no. 02, 2020, pp. 147-161.

[5]. *L. Su, Z. Li, F. Gao, M. Yu,* "A review of remote sensing image water extraction", Remote Sensing for Land & Resources, **vol.** 33, no. 01, 2021, pp. 9-19.

[6]. *H. Wang, F. Qin,* "Summary of the research on water body extraction and application from remote sensing image", Science of Surveying and Mapping, **vol.** 43, no. 05, 2018, pp. 23-32.

[7]. *S. Bi, Y. Qian, Q. Wang, Y. Guo,* "Research on water information extraction algorithm based on TM images", Science Technology and Engineering, **vol.** 14, no. 03, 2014, pp. 222-226.

[8]. *G. Wang, L. Pei, Q. Du, X. Li,* "Spectral relation method for water body extraction from ZY-3 imagery", Remote Sensing Information, **vol.** 35, no. 03, 2020, pp. 117-121.

[9]. *S. K. McFEETERS,* "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features", International Journal of Remote Sensing, **vol.** 17, no. 7, 1996, pp. 1425-1432.

[10]. *H. Xu,* "A study on information extraction of water body with the Modified Normalized Difference Water Index (MNDWI)", National Remote Sensing Bulletin, no. 05, 2005, pp. 589- 595.

[11]. *W. Wang, X. Chen, W. Wu, X. Gao,* "Method of automatically extracting urban water bodies from high-resolution images with complex background", Computer Science, **vol.** 46 , no. 11, 2019, pp. 277-283.

[12]. *G. Sarp, M. Ozcelik,* "Water body extraction and change detection using time series: A case study of Lake Burdur, Turkey", Journal of Taibah University for Science, **vol.** 11, no. 3, 2016.

[13]. *T. Acharya, A. Subedi, D. Lee,* "Evaluation of Machine Learning Algorithms for Surface Water Extraction in a Landsat 8 Scene of Nepal", Sensors (Basel, Switzerland), **vol.** 19, no. 12, 2019.

[14]. *A. Krizhevsky, I. Sutskever, G. Hinton,* "ImageNet Classification with Deep Convolutional Neural Networks", Advances in neural information processing systems, **vol.** 25, no. 2, 2012.

[15]. *A. Radoi , M. Datcu* "Multilabel Annotation of Multispectral Remote Sensing Images using Error-Correcting Output Codes and Most Ambiguous Examples", Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, 2019, 12(7): 2121-2134.

[16]. *J. Long, E. Shelhamer, T. Darrell,* "Fully Convolutional Networks for Semantic Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, **vol.** 39, no. 4, 2015, pp. 640-651.
[17]. *F. Yang, T. Feng, G. Xu, Y. Chen*, "Applied method for water-body segmentation based on mask R-CNN", Journal of Applied Remote Sensing, **vol.** 14, no. 1, 2020, pp. 014502-014502.
[18]. *Z. Liang, Y. Wu, H. Yang, X. Yao*, "Full-automatic water extraction method for remote sensing imagery based on Densely Connected Fully Convolutional Neural Network", Remote Sensing Information, **vol.** 35, no. 04, 2020, pp. 68-77.
[19]. *X. Wang, L. Sui, M. Zhong, D. Li, L. Dang,* "Fully Convolution Neural Networks for water extraction of remote sensing images", Bulletin of Surveying and Mapping, no. 06, 2018, pp. 41-45.
[20]. *Y. Yu, Y. Yao,H. Guan, et al.* "A self-attention capsule feature pyramid network for water body extraction from remote sensing imagery", International Journal of Remote Sensing, 2021, 42(5):1801-1822.
[21]. *G. Tambe Rishikesh, N. Talbar Sanjay and S. Chavan Satishkumar*, "Deep multi-feature learning architecture for water body segmentation from satellite images", Journal of Visual Communication and Image Representation, 2021, 77.
[22]. *D. Qiao, J. Zheng, H. Lu, L. Deng*, "Application of water extraction methods from Landsat imagery for different environmental background", Journal of Geo-information Science, **vol.** 23, no. 04, 2021, pp. 710-722.
[23]. *D. K. Bolton, J. M. Gray, E. K. Melaas, M. Moon, L. Eklundh*, "Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery", Remote Sensing of Environment, **vol.** 240, 2020, pp. 111685.
[24]. *Z. Miao, K. Fu, H. Sun, X. Sun, M. Yan,* "Automatic Water-Body Segmentation from High-Resolution Satellite Images via Deep Networks", IEEE Geoscience and Remote Sensing Letters, 2018, pp. 1-5.
[25]. *S. Zheng, S. Jayasumana, B. Romera-Paredes, et al.,* "Conditional Random Fields as Recurrent Neural Networks", 2015 IEEE International Conference on Computer Vision (ICCV). IEEE, 2015.
[26]. *M. Pai, V. Mehrotra, U. Verma, R. Pai*, "Improved Semantic Segmentation of Water Bodies and Land in SAR Images Using Generative Adversarial Networks", International Journal of Semantic Computing, **vol.** 14, no. 01, 2020, pp. 15.
[27]. *X. Liu, G. Wang, H. Yang, Y. Liu, Y. Wang*, "Road extraction from remote sensing image based on Fully Convolutional Networks", Remote Sensing Information, **vol.** 33, no. 01, 2018, pp. 69-75.
[28]. *T. Zheng, Q. Wang, J. Li, F. Zheng, Y. Zhang, N. Zhang*, "Automatic water extraction from GF-6 image based on Deep Learing", **vol.** 21, no. 04, 2021, pp. 1459-1470.
[29]. *A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, J. Garcia-Rodriguez*, "A Review on Deep Learning Techniques Applied to Semantic Segmentation", 2017.