

XU SHANSHAN

LAKE DETECTION ALGORITHM IN POINT CLOUDS OF THE LIDAR IMAGE BASED ON THREE-DIMENSIONAL CONVOLUTIONAL NEURAL NETWORK

Belarusian National Technical University

An algorithm for detecting lakes in a point cloud of a lidar image based on a three-dimensional convolutional neural network is proposed. The contours of the lakes were extracted from the point clouds of the lidar image and their geometric characteristics were determined using the chain code algorithm. The accuracy of the proposed algorithm for identifying lakes from clouds of laser scanning points was 96.34%. The proposed algorithm can calculate and analyze information about the shape of lakes.

Keywords: laser scanning data; three-dimensional convolutional neural network; lake detection; chain code; outline description

Introduction

Real-time and accurate detection of the evolution trend of lake water surface area is an important means to understand the change law of lakes. Currently, there are several methods to extract lakes^[1-5], for example, by analyzing bands and setting thresholds. These methods require a manual parameter-setting process. An object-oriented method can be used to segment and extract bodies of water by analyzing high-resolution remote sensing images. Thereafter, considering the spectral features, spatial shape, texture, and context of various ground objects, the method classifies the extracted bodies of water based on different parameters^[6-9]. Compared with traditional methods, this method has a high accuracy, can effectively distinguish water and shadow, and suppress the phenomena of salt and pepper. However, this method requires more background information, and the parameters are highly dependent on the scene. The strategy of “global local” distribution iteration has been used to classify plateau lakes by analyzing multi-source remote-sensing data. The advantages of this method are as follows: it does not require any manual intervention operations such as sample collection and parameter input, and it automatically extracts the target water area. However, it is difficult to gather local information with this method, hence, it is difficult to deal with small rivers.

The optimal lake segmentation threshold is determined by assessing the characteristics of lakes in two-dimensional images. Although the above methods can effectively extract lakes, there is no method to analyze the geometric information of lakes. During recent years, laser scanning data have been accumulating. With a high scanning spatial accuracy and short acquisition cycle, three-dimensional (3D) information of objects can be effectively obtained to solve the problem of data occlusion in remote-sensing images. Herein, we propose a new 3D convolutional neural network to extract target objects from laser scanning data.

2 Lake detection in laser scanning data based on the 3D convolution neural network

2.1 Pretreatment of network input

The preprocessing step includes three processes: regional location of candidate lakes, voxel organization of point clouds, and expansion of sample data. In the extraction of candidate lake regions, the plane with the largest area in the point cloud was considered as the candidate lake region. Water is affected by gravity, and the water surface area tends to be horizontal. The problem of lake detection is transformed into the problem of plane extraction in the laser scanning point cloud. Because the normal vector of the water area is known to be (0, 0, 1), the random sample consensus (RANSAC)^[10] algorithm was used to fit the plane in the point cloud. The purpose of RANSAC is to estimate the parameters of

the constructed mathematical model from a set of observation data containing “external points.” RANSAC has the advantage of high robustness. Its disadvantage is that the number of iterations does not have a clear upper limit, but only the target model is calculated with probability. Moreover, the probability is directly proportional to the number of iterations required by the algorithm. The normal vector required by this model was determined, which greatly reduced the number of iterations required by the algorithm. After extracting the candidate lake areas, these areas were divided into voxels, the value of each voxel was 1 or 0; 1 indicates that there is a point in the voxel and 0 indicates that the voxel is empty. The input samples of the depth network must have the same size. In this study, a linear interpolation algorithm was used to normalize the input region to the same size.

2.2 Network structure design

The constructed deep learning network includes the input, convolution, sampling, full connection, and output layers. The sampling and convolution layers appeared alternately in the network. Finally, the identification results were output through a fully connected network. The designed network structure is shown in Figure 1.

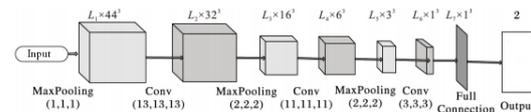


Fig. 1 Structure of the 3D convolutional neural network

There are several feature maps in the convolution and sampling layers. Each feature map is connected to the feature map of the previous layer. The number of feature maps in each layer was recorded as L_i , in which the subscript i is the layer mark, and the number of feature maps is set by the user. In the convolution layer, the characteristic map of the previous layer is convoluted with a learnable kernel. Once the convolution result passes through the activation function, the neurons of this layer are output, and the characteristic map of this layer is formed. The calculation formula for the convolution layer is as follows:

$$X_l^{i,j,k} = B_l + \sum_{i=0}^f \sum_{j=0}^f \sum_{k=0}^f w_l^{i,j,k} X_{l-1}^{i,j,k} \quad (1)$$

where, l represents the layer, w_l is the convolution kernel of this layer, and each characteristic graph can have different convolution kernels with a size of $f \times f \times f$. Each layer has a unique offset B_l . The main function of the convolution layer is to make the feature displacement invariant. If a neuron $X_l^{i,j,k}$ in the feature map on layer l has to be obtained, the neurons near the previous layer

$X_{i,j}^{l,k}$ need to be convoluted with the help of a convolution check.

The sampling layer is relatively simple, and its main function is to reduce the spatial resolution of the network. Generally, the sampling methods include Max-Pooling, Min-Pooling, and Ave-Pooling, which mean that the sampling results are the maximum, minimum, and average values, respectively. In the full connection layer, the conventional Softmax function^[11] is used to output the network results. It should be noted that after each convolution layer, the linear rectification function, also known as the modified linear unit (ReLU)^[12], must be used to enhance noise resistance of the network. The calculation formula is as follows:

$$F(X)_{ReLU} = X^+ = \max(0, X) \quad (2)$$

If the input sample size is $44 \times 44 \times 44$, the processed size of samples in each layer is shown in the upper part of Figure 1. The parameters in the functions maxpooling (f, f, f) and Conv (f, f, f) represent the size of the sampling kernel and convolution kernel, respectively. Finally, the output of the network is lakes or other areas. In summary, a 3D convolutional neural network includes input, sampling, convolution, full connection, and output layers. The input samples are input to the network from the first layer (input layer). After feature extraction through the sampling and convolution layers at intervals, the results are output to the output layer through a full connection layer. In this study, three convolution and sampling operations were used to complete the construction of the network. Although adding more sampling and convolution layers can result in the expression of more complex features, it substantially increases the algorithm complexity and learning time.

The complete process of lake extraction and analysis in this study, it includes the input of the point cloud scene, extraction of all plane regions in the scene based on the flatness information described above, and consider them as candidate lakes. Because of the unorganized nature of the point cloud, before input to the 3D convolutional neural network, in this study, we used the voxelization technology mentioned above, that is, the candidate lake area is divided into $1 \text{ cm} \times 1 \text{ cm} \times 1 \text{ cm}$ voxel, and each region is interpolated to the same size as the input of the network.

The 3D convolutional neural network is a classifier, and the network structure is shown in Figure 1. Both training and test sets were obtained from manually labeled lakes. Because more samples are often needed to complete the training in machine learning and the number of samples in a 3D lake is relatively small, it is necessary to expand the number of training and test sets when training the network. The expansion method is based on the rotation of the X-, Y-, and Z-axes. The range is $[0, \pi]$ and the interval is $\pi/18$. The expansion operation used here is only to translate and rotate the object, and it does not involve scaling. Therefore, the shape features of the original dataset were saved. The output of the network can be divided into two types: the target lake and other areas. w_i and B_i in equation (1) can be obtained by learning samples through 3D convolution neural network. At this time, when a new sample X_i to be tested appears, equation (1) is used for calculation and network transmission until the output layer outputs the category. If it is not a lake, it is filtered out; otherwise, the next step is to extract the contour and analyze the geometric properties of the lake.

3 Experiment and result analysis

The effect and accuracy of this method in laser scanning data will be tested. Firstly, the data to be tested is described; Then, the detection accuracy of Lake based on the proposed convolution neural network is given; Finally, the analysis results of Lake geometry information are displayed.

Six airborne point cloud data are selected to verify the algorithm. Figure 2 shows the satellite image corresponding to the lake in each point cloud data. The airborne point cloud data corresponding to lakes in this paper are downloaded from the public dataset website opentopograph.org (HTTPS: // www.opentopo-

graph.ORG).



Fig.2 Satellite images in testing data

Corresponding to (a) to (f) in Figure 2 gives the coordinates, area, corresponding airborne point cloud density and acquisition time of the lake.

In the 3D convolutional neural network experiment, a total of 8 different network structures were tested. Among them, the highest accuracy is Test#3, and the corresponding program operation is shown in Figure 3. The program language is Matlab, which is based on the toolbox MatConvNet, the hardware configuration is Intel Core i7 8086k (16 GB memory, 2 NVIDIA geforce GTX 1080 Ti graphics cards), the operating system is windows 10, and the training takes 598.65 seconds. Due to the small number of samples, the training samples need to be expanded, there are 30 Lake training sets in the experiment, and $30 \times 18 \times 3$ samples are obtained after expansion, 80% of which are the training set and the rest are the test set.

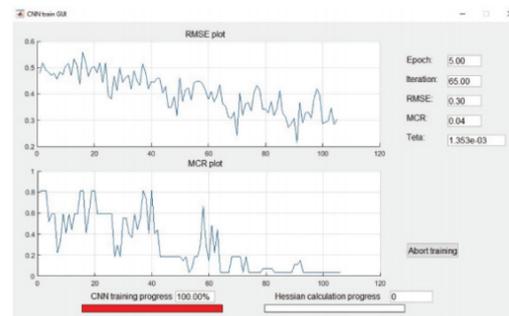


Fig.3 Execution interface results of test#3 (CNN training process, Hessian calculation process)

Through the accuracy evaluation of this paper, it can be seen that the accuracy of the lake recognition algorithm proposed in this paper can meet the requirements of three-dimensional object recognition. Then, based on the proposed chain code method, each boundary extraction result is obtained in the detected Lake area, as shown in Fig. 4. Among them, the size of Barker Lake in Figure 4 (a) is $2158 \text{ M} \times 1209 \text{ m}$, 230000 points in total; In Figure 4 (b), the size of Lake Eleanor is $4807 \text{ M} \times 3199 \text{ m}$, 2.05 million points in total; The size of Cherry Lake in Figure 4 (c) is $6465 \text{ M} \times 2745 \text{ m}$, 2.22 million points in total; In Figure 4 (d), the size of Lake Henry Hagg is $12362 \text{ M} \times 11597 \text{ m}$, 1.07 million points in total; The size of Bullrun Lake in Figure 4 (E) is $9436 \text{ M} \times 7570 \text{ m}$, 1.29 million points in total; The Lake Lost in Figure 4 (f) has a size of $5912 \text{ M} \times 4870 \text{ m}$, 140000 points in total. As can be seen from Figure 4, the algorithm in this paper can effectively extract the boundaries of each lake and accurately describe the complex concave convex areas in the lake. Next, the extraction accuracy will be analyzed.

The lake shape information calculated according to the chain code, the method based on chain code calculation is different from the above point cloud extraction shape analysis method, the length of chain code calculation refers to the longest distance between two points in the horizontal direction, and the width refers to the longest distance between two points in the vertical direction;

The length described in Fig. 4 refers to the horizontal span of the target object, and the width refers to the vertical span, the method adopted is to use an external rectangle to roughly describe the general situation of the object. When calculating the area, this paper uses the chain code method to accurately describe the lake area.

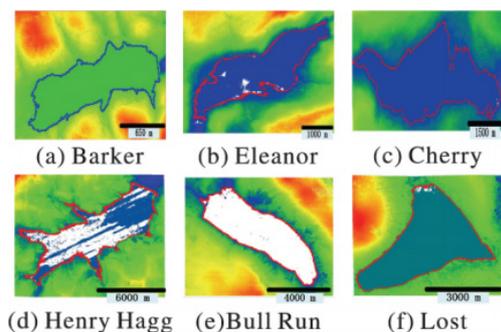


Fig.4 Results of lake detection

Conclusions

Herein, we proposed a 3D convolutional neural network to identify lake areas, and a chain code extraction algorithm was designed to analyze lakes in aerial laser scanning data. The experimental results showed that the shape features of objects in 3D point clouds can be learned using the proposed convolution neural network; the 3D convolution neural network proposed in this paper can effectively filter non-lake areas with an accuracy of 96.34%. In addition, the lake boundary can be accurately described using an eight-direction chain code. The length, width, and area of the lake can be calculated using the chain code. Finally, the area of the lake was accurately calculated by linear fitting. The next step was to extract and analyze other water bodies in the aerial laser scanning point cloud, such as streams, canals, estuaries, and seaports, and detect their changes by extracting their contours.

REFERENCES

1. Zhou Y, Tuzel O. Voxelnet: end-to-end learning for point cloud based 3D object detection/ Zhou Y, Tuzel O. - Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018 - p. 4490-4499.
2. Uy M. A., Lee G. H. PointNetVLAD: deep point cloud based re-retrieval for large-scale place recognition / Uy M.A., Lee G.H. - Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018 – p. 4470-4479.
3. Qi C.R., Su H., Mo K., et al. PointNet: deep learning on point sets for 3D classification and segmentation/ Qi C.R., Su H., Mo K. - Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017 – p. 77-85.
4. Qi C. R., Yi L., Su H., et al. PointNet++: deep hierarchical feature learning on point sets in a metric space / Qi C. R., Yi L., Su H. - Proceedings of the 2017 International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2017 – p. 5099-5108.
5. Wang W., Yu R., Huang Q., et al. SGPNet: similarity group proposal network for 3D point cloud instance segmentation/ Wang W., Yu R., Huang Q. - Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018 – p.2569-2578.
6. Shen Y., Feng C., Yang Y., et al. Mining point cloud local structures by kernel correlation and graphpooling / Shen Y., Feng C., Yang Y. - Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018 p. 4548-4557.
7. Landrieu L., Simonovsky M. Large-scale point cloud semantic segmentation with superpoint graphs / Landrieu L., Simonovsky M. - Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018 – p. 4558-4567.
8. Huang Q., Wang W., Neumann U. Recurrent slice networks for 3D segmentation of point clouds/ Huang Q., Wang W., Neumann U. - Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018 – p. 2626-2635.
9. Huang J., You S. Point cloud labeling using 3D convolutional neural network / Huang J., You S. - Proceedings of the 3rd International Conference on Pattern Recognition. Piscataway: IEEE, 2016 p.2670-2675.
10. Schnabel R., Wahl R., Klein R. Efficient RANSAC for point-cloud shape detection. / Schnabel R., Wahl R., Klein R - Computer Graphics Forum, 2007, No. 26(2) – p. 214-226.
11. Whiteson S., Stone P. Evolutionary function approximation for reinforcement learning. / Whiteson S., Stone P. - Journal of Machine Learning Research, 2006, No.7 – p. 877-917
12. Nair V., Hinton G.E. Rectified linear units improve restricted Boltzmann machines / Nair V., Hinton G.E. - Proceedings of the 27th International Conference on Machine Learning. Madison: Omnipress, 2010 – p.807-814.

СЮ ШАНЬШАНЬ

АЛГОРИТМ ОБНАРУЖЕНИЯ ОЗЕР В ОБЛАКАХ ТОЧЕК ЛИДАРНОГО ИЗОБРАЖЕНИЯ НА ОСНОВЕ ТРЕХМЕРНОЙ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Белорусский национальный технический университет

Предлагается алгоритм обнаружения озер в облаке точек лидарного изображения на основе трехмерной сверточной нейронной сети. Контуры озер были извлечены из облаков точек лидарного изображения, а их геометрические характеристики определены с использованием алгоритма цепного кода. Точность предложенного алгоритма идентификации озер по облакам точек лазерного сканирования составила 96,34%. Предлагаемый алгоритм позволяет рассчитывать и анализировать информацию о форме озер.

Ключевые слова: данные лазерного сканирования; трехмерная сверточная нейронная сеть; обнаружение озер; цепной код; наброски описание.

Сю Шаньшань, аспирантка кафедры программного обеспечения информационных технологий БНТУ. Научные интересы: Нейронные сети, обработка облаков точек лидарных изображений земной поверхности.

E-mail: 529031585@qq.com