Study on the example segmentation method of remote sensing image based on neural network

Songyi Yu, Guotao Wang*

Heilongjiang University

National Natural Science Foundation of China（51607059）; Heilongjiang Natural Science Fund（QC2017059）

*wanggt@hlju.edu.cn

Abstract. With the rapid development of modern science and technology, people collect remote sensing data from the altitude perspective at the same time, put forward the use of a variety of remote sensing images to solve the military exploration, meteorological analysis, environmental protection, resource exploration and other basic problems. However, since remote sensing images have the characteristics of too large data, high image resolution and extremely low application efficiency, some scholars have used the image features in residual network problems in their research to solve the problem of remote sensing image target segmentation scale difference based on the attention mechanism and single-step case segmentation framework. In this paper, based on the understanding of the research status of neural networks and remote sensing image application, a remote sensing image segmentation model based on multi-level channel attention is proposed according to the model architecture of convolutional neural networks. The final experimental results show that the neural network based remote sensing image case segmentation technology has positive effects.

Keywords: Neural network; Remote sensing image; Instance segmentation; Channel attention.

1. Introducion

From the perspective of practical application, remote sensing image has a huge amount of information data, which has a wide range of application value in the current social construction and development. However, since the image acquisition process will be affected by uncontrollable factors such as air quality, cloud cover and illumination, the problem of block and shadow will appear in the target area. Therefore, remote sensing image in the traditional sense needs manual observation and judgment. With the innovation and development of modern aerospace technology, the acquisition and operation of traditional airborne and space-borne remote sensing images become easier. The activities of aircraft and satellites of various countries have generated a large amount of remote sensing data. However, due to the large size of the image data and the large number of targets, the efficiency of relying only on manual observation is low, so some scholars put forward the automation and intelligent technology theory of remote sensing images in their research. Since the invention of image recording technology, human has been trying to apply it to the remote observation record. [1-3]For example, in the late 1850s, French scholars took a bird 's-eye view of Paris with balloons, which was the earliest aerial photographic record in the world. During the two World Wars, air clapping was used as an important technical means of military observation, and countries gradually realized the important value of aerial reconnaissance and aerial photography. In the late period of the Second World War, various kinds of electromagnetic spectrum remote sensing recording technology has been rapidly developed. The United States, through the interpretation of remote sensing images to plot the military target map, eventually became an important source of information in the Pacific War. After the appearance of satellite aerial remote sensing, the important value of remote sensing makes it gradually develop into an independent subject.

The method based on deep learning neural network plays a positive role in image processing classical problems. Especially after the Agent shows its unique advantages, people gradually realize the unique value of learning-centered technical methods. Neural network is widely used in various

fields to solve classical problems, such as text classification, speech recognition, target detection, etc. In essence, neural network itself is also a research topic. How to obtain more excellent technical results by adjusting network architecture model, training strategy, loss function, etc., is the main problem integrated by scholars in various fields at present. The extraction of medium target of remote sensing image interpretation is a case segmentation problem. Before the deep learning method was proposed, people mainly carried out cluster segmentation by unsupervised K-means, ISODATA and other methods, or realized remote sensing image segmentation by maximum likelihood, minimum distance, support vector machine and other methods, but most of these methods have poor generalization performance. Relevant parameters need to be added manually, so it is difficult to accurately translate data information on a large scale. For example, in the mid-1980s, some scholars proposed the back propagation algorithm of artificial neural network, on which basis, the research upsurge of neural network in machine learning was set off. Due to the problems of over-fitting and long training time of BP neural network, the support vector machine based on statistical learning theory was proposed by scholars in the early 1990s, and its learning effect was significantly better than that of BP neural network. After entering the 21st century, some scholars proposed deep learning in Science. On the one hand, artificial neural networks with multiple hidden layers have excellent feature learning ability, and the learned data can fully show the essential features and help staff visualization or classification processing; on the other hand, the difficulty of deep learning networks in training can be effectively overcome through layer-to-layer unsupervised training. Therefore, on the basis of understanding the research status of neural network and remote sensing image instance segmentation, this paper uses the remote sensing image segmentation model with multistage channel attention as the core to carry out experimental analysis, and verifies the application advantages of remote sensing image instance segmentation method based on neural network according to the final results.[4-6]

2. Method

2.1 Remote sensing image data set

In this paper, iSAID data set is selected for experimental analysis. As an effective basis for aerial image instance segmentation, it mainly contains 2806 high-resolution images, including Bridges, helicopters, aircraft, sports fields, cars, etc. Each category has a large number of marked instance images. It effectively enhances the learning ability of network for instance and related background. In this data set, the scale changes between categories are very large, and instances of different categories may appear in the same image. Meanwhile, the distribution of data of various objects in the image is not balanced, which truly shows the data distribution in the environment of aerial photography. Since the data of the original aerial remote sensing image is relatively large, the memory and CPU size of the data that is too large should be divided into smaller image blocks according to the conventional processing method to facilitate the test and research.[7-9]

2.2 Attention Mechanism

At present, in visual tasks, researchers mainly use neural network method to learn features in image data, and neural network as a method of learning through computer simulation of biological neuron stimulation and response mechanism, practical ability can enable people to achieve excellent research results in classical visual tasks. Although the current understanding of the cognitive mechanism of human brain is not perfect, some scholars have designed computer algorithms based on the perfect signal processing mechanism, and the most representative one is the attention mechanism. In essence, the attention mechanism as a conscious focus of the human brain, the human visual system can use a very rapid way to find and focus on the important areas of the scene. Under the influence of this technical theory, excellent results have been achieved in design problems in computer vision. The specific calculation formula is as follows:

$$Attention = f\left(g\left(x\right), x\right)$$

In the above formula, g (x) stands for generating attention, which is the focus area processing process; f (g (x), x) represents the process by which the generated attention g (x) processes x.

Existing attention mechanisms in visual tasks include channel attention mechanisms, such as spatial attention mechanism, DANet, etc. In this paper, the spatial attention module will be embedded in the network to improve the performance of the network system. The specific CBAM structure is shown in Figure 1 below:

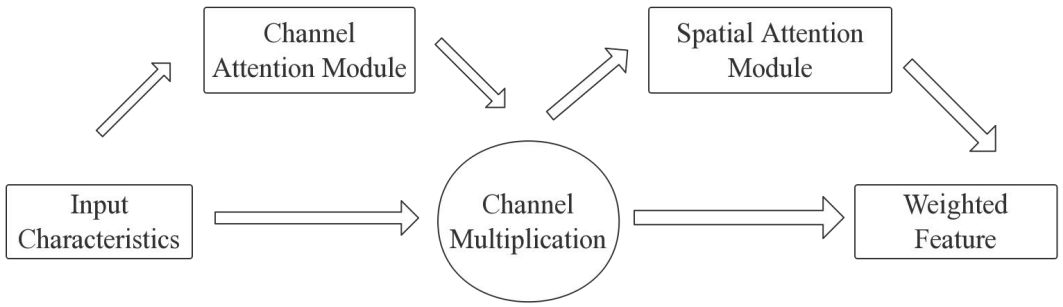

Figure 1 Structure diagram of CBAM

Channel attention mainly guides the network to focus its attention on one or some channels of features through learning, as shown in Figure 2 below:

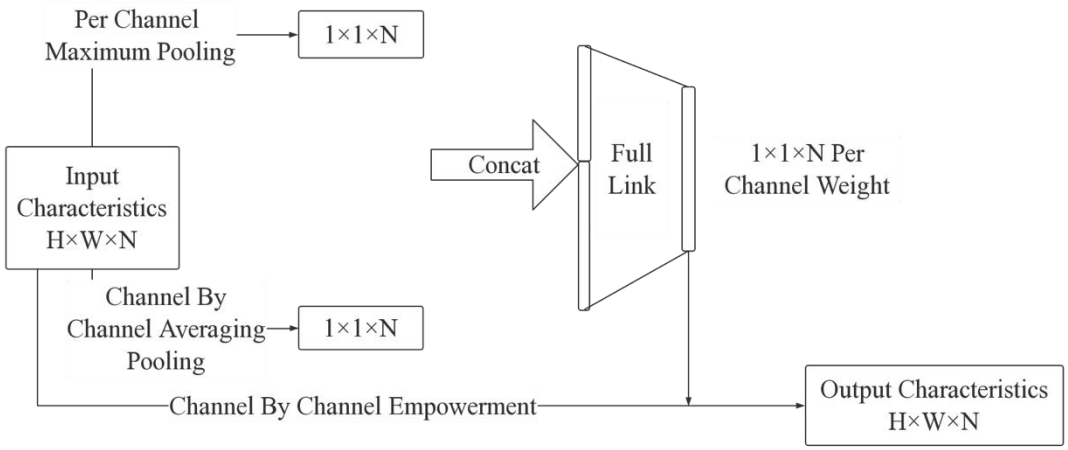

Figure 2. Structure of channel attention

However, the spatial attention mechanism always guides the network to focus its attention on a certain key region of the feature graph, and a greater weight will be assigned to focusing attention in the convolutional neural network, as shown in Figure 3 below:

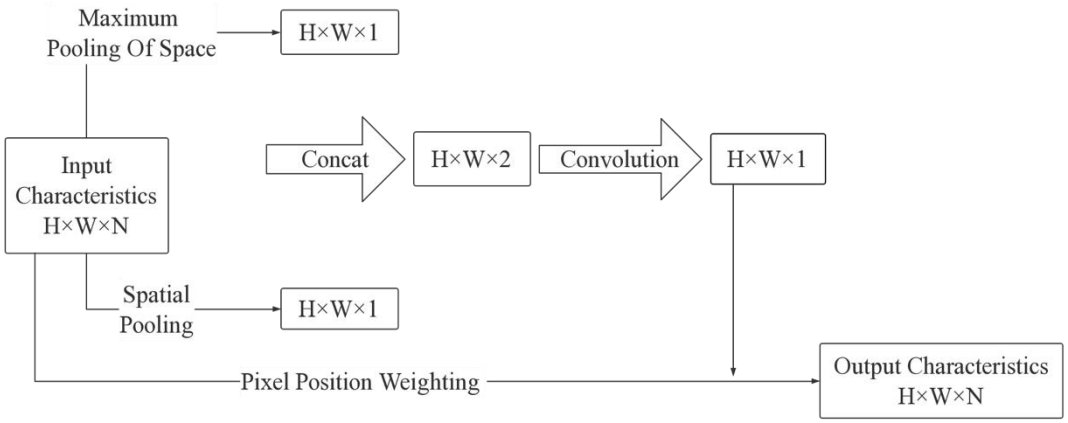

Figure 3. Structure of spatial attention mechanism

2.3 Remote Sensing Image Segmentation Model Based on multi-level Channel Attention (RSIS-MLCA)

The design of RSIS-MLCA model is mainly divided into two components, one is the encoder, the other is the decoder, the former is mainly used to extract features, the latter is mainly used to predict categories. By using convolution-inversion and structure to design encoders and decoders, deep convolutional neural networks are used to extract features in the convolution stage. As the depth increases, the size of feature layer will continue to decrease. This process is called subsampling, which can not only extract higher-level information, but also meet the needs of image segmentation. In order to adapt to the task requirements of remote sensing image segmentation faster, the acquired features are reconstructed in the deconvolution stage, the size of the feature layer is gradually expanded and the depth of the feature layer is reduced, and the input size is recovered after the output image. This process is called upsampling. At the same time, in order to help the system obtain occlusion target information faster, channel attention should be introduced before each step of pooling operation, to provide convolution information for all pixels, enhance the expression ability of useful information, and use multilevel channel attention to achieve feature fusion, so as to provide effective basis for global semantic information and local detail information, and finally obtain better segmentation results.[10-13]

Based on the analysis of the network structure design shown in Table 1 below, when the RGB three-channel image with 256×256 pixels is input for downsampling operation, the downsampling operation is repeated according to the above rules after two layers of convolution operation with the convolution kernel size of 3×3 and step size of 1, batch normalization layer, attention module and maximum pooling layer, etc. Gradually increase the size of the feature map and reduce the depth of the feature layer, use the attention features contained in the downsampling and the same scale information of the up-sampling to achieve feature stitching, and repeat the up-sampling operation in accordance with the basic provisions, and finally get the 256×256×64 feature map. A 1×1 convolution layer is used to map the feature graphs of 64 channels into a specific number of categories to achieve category analysis.

Table 1 Structure design of RSIS-MLCA network model

Operation	Parameter					
	Input	Output	Kernel-size	Stride	Padding	Image-size
Conv_1＊2	3	64	3	1	1	256×256
AT_Block_1	64	64	-	-	-	256×256
Maxpool_1	64	64	2	2	0	128×128
Conv_2＊2	64	128	3	1	1	128×128
AT_Block_2	128	128	-	-	-	128×128
Maxpool_2	128	128	2	2	0	64×64
Conv_3＊2	128	256	3	1	1	64×64
AT_Block_3	256	256	-	-	-	64×64
Maxpool_3	256	256	2	2	0	32×32
Conv_4＊	256	512	3	1	1	32×32

2						
AT_Block_4	512	512	-	-	-	32×32
Maxpool_4	512	512	2	2	0	16×16
Conv_5＊2	515	1024	3	1	1	16×16
AT_Block_5	1024	1024	-	-	-	16×16
Deconv_6	1024	512	2	2	0	32×32
Conv_6	1024	512	3	1	1	132×32
Deconv_7	512	256	2	2	0	64×64
Conv_7	512	256	3	1	1	64×64
Deconv_8	256	128	2	2	0	128×128
Conv_8	256	128	3	1	1	128×128
Deconv_9	128	64	2	2	0	256×256
Conv_9	128	64	3	1	1	256×256
Conv_10	64	1	1	1	0	256×256

3. Result analysis

In order to verify the segmentation performance of the RSIS-MLCA model studied in this paper, the test was carried out in the data set that has been mastered. The final evaluation results are shown in Table 2 below:

Table 2 Experimental results

	Image1		Image2		Image3		Image4		Image5		Testset	
Metric（%）	P	IOU	P	IOU	P	IOU	P	IOU	P	IOU	P	IOU
U-Net	88.0	70.0	89.9	50.4	87.9	63.0	74.6	41.4	88.5	76.6	83.0	74.6
RSIS-MLCA	89.3	91.2	91.2	83.4	90.1	70.5	81.2	65.6	90.7	80.3	85.2	76.5

Based on the analysis of the above table, it can be seen that all metrics of the RSIS-MLCA model studied in this paper can reach the highest level in each image, among which the accuracy (P) and intersection ratio (IOU) of the average results of the test set exceed 2.2% and 1.9%. From the perspective of training time, the training time of U-Net model is 16 hours and 5 minutes, while the research time of RSIS-MLCA model in this paper is 14 hours and 22 minutes, the latter is only 0.09 times longer than the former. From the perspective of qualitative segmentation map and quantization result, this paper studies that the segmentation result of the RSIS-MLCA model on the road data set is closer to the actual mark map without significantly increasing the calculation cost, and the overall application effect is stronger.[14-15]

In the innovation and development of aerospace technology, we begin to focus on how to extract and apply high resolution remote sensing image target, and use image segmentation technology to segment remote sensing image into areas with actual semantic information, which is one of the

research topics of image processing. Because remote sensing image has such problems as large size, complex type and wide range, the technical means based on remote sensing image object segmentation are still faced with many challenges. In this paper, the remote sensing image segmentation model based on multi-channel attention is proposed. In the network training stage, the channel attention mechanism will be constructed to focus on the target feature information, and the shallow local information and deep semantic information will be fully used to increase the channel attention module, and then the importance of each channel will be automatically obtained to improve the weight of the channel owned by the classification results. To ensure that each scale feature map contains target information, it can be used for more complex remote sensing image target segmentation.

Conclusion

To sum up, the remote sensing image segmentation method based on multilevel channel attention studied in this paper meets the needs of image processing in the new era, but it can only enhance the edge information of the target by using feature fusion, and cannot completely restore the edge information of the target. Therefore, future researchers should continue to explore the remote sensing image case segmentation method with neural network as the core. In order to improve the continuity of target spatial information and fully show the integrity of target edge, more countermeasures training methods should be considered.

Reference

[1] Jinyang Zhang. Remote sensing image case segmentation based on deep neural network [J]. Modern Computer, 2022, 28(9):85-90.

[2] Junzhi Yang,Jinliang Wu, Zhijun. Research on Aircraft Target Detection in Remote Sensing Images Based on Multi-scale circumferential Frequency Filtering and Convolutional Neural Network [J]. Journal of Electronics and Information Technology, 2021, 043(005):1397-1404.

[3] Guohua Yue,Xiaoli Xing. Remote sensing image registration Method Based on Convolutional Neural Network and Correction network [J]. Computer Applications and Software, 2021, 38(11):6.

[4] Panpan Zhu,Shuaipeng Li,Liqiang Zhang, et al. High resolution Remote sensing image building extraction based on multi-task learning [J]. Journal of Geo-Information Science, 2021, 23(3):10.

[5] Zheng Zhang,Yubo Ma,Chang 'an Liu, et al. Research on bidirectional attenuation loss method for rotating object detection in Remote sensing images [J]. Journal of Electronics & Information Technology, 2022, 44(CQ):1-9.

[6] Shuangling Zhu,Linzi Gu,Mujiang Ai Li, et al. Research on Semantic segmentation method of Remote sensing image based on convolutional neural network [J]. Computer Science and Applications, 2021, 11(2):14.

[7] Junzhi Yang,Jinliang Wu,Jun Zhi. Research on Aircraft Target Detection in Remote Sensing Images Based on Multi-scale circumferential Frequency Filtering and Convolutional Neural Network [J]. Journal of Electronics and Information Technology, 2021, 43(5):8.

[8] Lianwei Ma,Weiyuan Ning, Liwei Jiao,Shuaidong Xue. Research on Remote Sensing image Change Detection Method Based on U-Net Convolutional Neural Network [J]. Energy and Environmental Protection, 2022, 44(11):102-106.

[9] Zheyi Zhang,Weihua Cao,Rui Zhu, et al. High resolution Remote Sensing image Scene Classification Method Based on Sparse Representation of pulsed Convolutional Neural Network [J]. Control and Decision, 2022, 37(9):9.

[10] Fei Wen,Jiawei Mo,Yuqi Hu, et al. Evaluation of litchi flowering growth based on convolutional neural network remote sensing image [J]. Journal of South China Agricultural University, 2023, 44(1):11.

[11]Zitong Ma, Guodong Wang. Human case segmentation based on double-flow convolutional neural network [J]. Laser & Optoelectronics Progress, 2022, 59(16):7.

[12] Zhao Liu,Tong Zhao,Feifan Liao. Research on gap repair algorithm based on convolutional neural network prediction results. Remote Sensing Technology and Application, 2021, 36(2):9.

[13] Yanan Zhao,Hanbing Deng, Liu Ting, et al. Case segmentation method of maize seedling plant image based on weak supervised learning [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(19):10.

[14] Jiantao Ye,Wei Yang,Yanguo Ke, Hengyang Zhao,Di Hu,Changwei Zhao. Intelligent Diagnosis Method for Thermal Defects of Open Switchgear Based on infrared image Instance Segmentation [J]. High Voltage Electrical Apparatus, 2021, 57(12):201-208.

[15] Wei Li,Guixiong Liu,Chenggang Zeng. Research on Drainage pipe defect detection Method based on case segmentation +CCTV [J]. Electronic Measurement Technique, 2022(003):045.