Applications in semantic segmentation

I. Semantic Segmentation Experiment

Semantic segmentation is the pixel-level classification of images. To analyze the improvement effect of spatial resolution and spectral resolution, we input the images with restored spatial resolution and spectral resolution, grayscale image and color image into a semantic segmentation model respectively. The overall experimental framework is shown in Fig. 1. The semantic segmentation model adopted in this paper is the ABCNet model proposed by Li et al. [1]. If the segmentation results obtained by the semantic segmentation model after improving the spatial and spectral resolution are good, it means that the improvement of spatial and spectral resolution in this paper is effective.

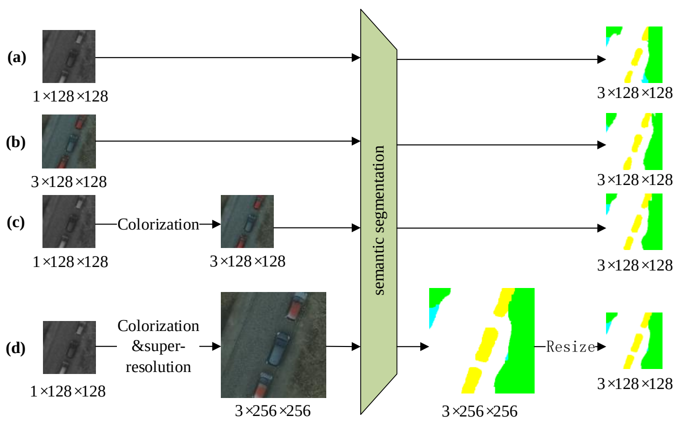

Fig. 1: Framework of semantic segmentation results. (a) segment the grayscale image directly; (b) segment the color image directly; (c) The spectral resolution of the grayscale image is improved before semantic segmentation; (d) The spatial and spectral resolution of the grayscale image is improved before semantic segmentation.

A. Experimental data

The experimental data adopted in this section are from the ISPRS Potsdam dataset, which consists of 38 image patches of the same size, and semantic labels are provided. The semantic labels are surfaces, building, vegetation, tree, and car, which are represented by white, dark blue, lake blue, green, and yellow, respectively, the background is represented by red. We use part of the image blocks in the ISPRS Potsdam dataset for our experiments, while adopting the semantic labels provided by the ISPRS Potsdam dataset. Each image in the ISPRS Potsdam dataset is cut into small image blocks with the size of 128×128, nearly 10,000 small image blocks are selected and randomly divided into test datasets and training datasets according to the ratio of 1:9.

B. Evaluation metrics for semantic segmentation

The semantic segmentation evaluation metrics used in this chapter are Mean Intersection over Union (MIoU), Class Pixel Accuray (CPA), and Mean Pixel Accuracy (MPA), which are calculated based on the confusion matrix. In the confusion matrix, a pixel sample is called True Positive (TP) when its true class is positive and its prediction result is also positive. When the true class of a pixel sample is positive and its prediction result is negative, the sample is called False Negative (FN). When the true class of a pixel sample is negative and its prediction result is positive, the sample is called False Positive (FP). When the true class of a pixel sample is negative, its prediction result is also negative, the sample is called True Negative (TN). CPA is the proportion of pixels with true class i among all pixels predicted as class i, MPA is the average of the CPA of all classes, and the CPA is calculated as follows:

$$\mathrm{CPA} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}, \tag{1}$$

IoU is the ratio of intersection and union between the predicted and true results for class i, and the IoU is calculated as follows:

$$\mathrm{IoU} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}}, \tag{2}$$

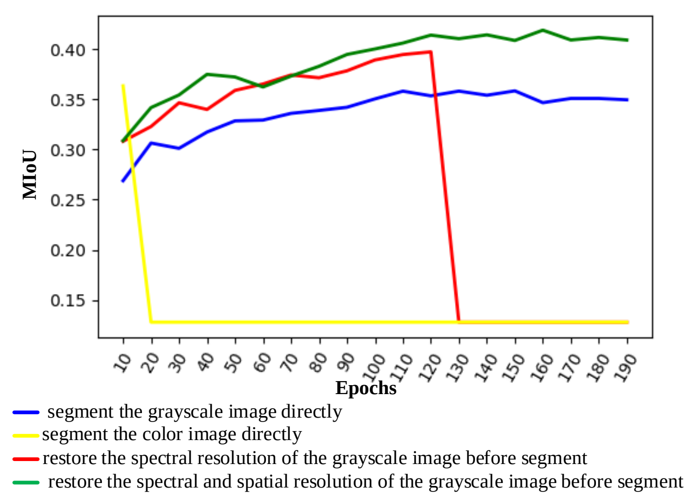

Fig. 2: MIoU scores of semantic segmentation results.

C. Semantic segmentation results

We put four different kinds of images to the ABCNet model for semantic segmentation verification, which are the original grayscale image, the image obtained after spectral resolution restoration of the grayscale image, the original color image, and the image obtained after spectral and spatial resolution

TABLE I: Quantitative results of the semantic segmentation task with different inputs, red text marks the best performance.

MIoU	CPA						MPA
	surfaces	building	vegetation	tree	car	background	
0.3496	0.6043	0.5536	0.3798	0.3502	0.4625	0.2178	0.4860
0.3635	0.5658	0.5104	0.4779	0.3686	0.4966	0.0797	0.4968
0.3974	0.6157	0.6074	0.4889	0.4171	0.4973	0.2967	0.5343
0.4088	0.6342	0.6535	0.4885	0.4212	0.4827	0.3900	0.5427

restoration of the grayscale image. The MIoU scores of the ABCNet model with different inputs are shown in Fig. 2. Compared with the MIoU score of the original grayscale image as the input of the ABCNet model, the MIoU of the original color image with high spectral resolution and the image with improved spectral resolution by our proposed method will soon reach the peak and then fell. This is because the improvement of spectral resolution will make the model ignore the spatial resolution. However, the spatial resolution can reflect the details of ground objects. When the spectral resolution is too high and the spatial resolution is too low relative to the spectral resolution, the details of ground objects will not be identified, resulting in different classes of objects with the same spectrum are divided into the same class. Comparing the green and blue lines in Fig. 2, it can be seen that the trend of MIoU scores of the two lines is consistent, there is no obvious decline. This is because the spectral resolution and spatial resolution of the input image are balanced, the model pays attention to both spatial and spectral resolution when segmenting, which does not ignore the spatial details of the ground objects while classifying the ground objects by color. By comparing the MIoU scores of the four groups of experiments, it can be seen that the maximum values of MIoU of the segmentation after improving the spatial and spectral resolution of the grayscale image are larger than the maximum values of the other groups, that is, the experimental effect of semantic segmentation after improving the spatial and spectral resolution of the grayscale image is the best. In a word, our proposed spatial and spectral resolution restoration scheme can improve the performance of the semantic segmentation model.

The objective evaluation results of the semantic segmentation experiment are shown in Table I, the inputs for the four groups of experiments are grayscale images, color image, images with spectral resolution recovered from grayscale images, images with spectral and spatial resolution recovered from grayscale images. Comparing the first and second group of experiments, it can be seen that the semantic segmentation of color images with higher input spectral resolution has better experimental results. From the overall segmentation results, the MIoU and MPA of the input original grayscale image are higher than that of the input original color image, indicating that the input image with higher spectral resolution can improve the performance of semantic segmentation of remote sensing images. From the perspective of single category, the segmentation effect of ground objects with near-gray color, such as land surfaces and buildings, is not improved after the spectral resolution is restored, which is because these objects do not rely on color information in segmentation. The segmentation effect of other brightly colored objects such as vegetation, trees, and cars will be improved after the spectral resolution is restored. The third group of experiments uses the images generated by the spectral resolution restoration method proposed in this paper as the input of ABCNet model. Compared with the second and third groups of experiments, it can be seen that the MIoU and MPA of the third group of experiments are significantly improved compared with the second group of experiments. It shows that the spectral resolution enhancement method proposed in this paper is better than that of the original color image. Comparing the third group and the fourth group, it can be seen that the improvement of semantic segmentation performance by restoring both spectral and spatial resolution is better than that by only restoring spectral resolution. Comparing the four groups of experiments, the semantic segmentation effect is best when the input image of ABCNet model is the image with restored spectral and spatial resolution at the same time, and the semantic segmentation effect is worst when the input image is the original gray image with low spectral and spatial resolution. Based on the above analysis, the spectral and spatial resolution restoration scheme proposed in this paper are useful for the performance improvement of semantic segmentation models.

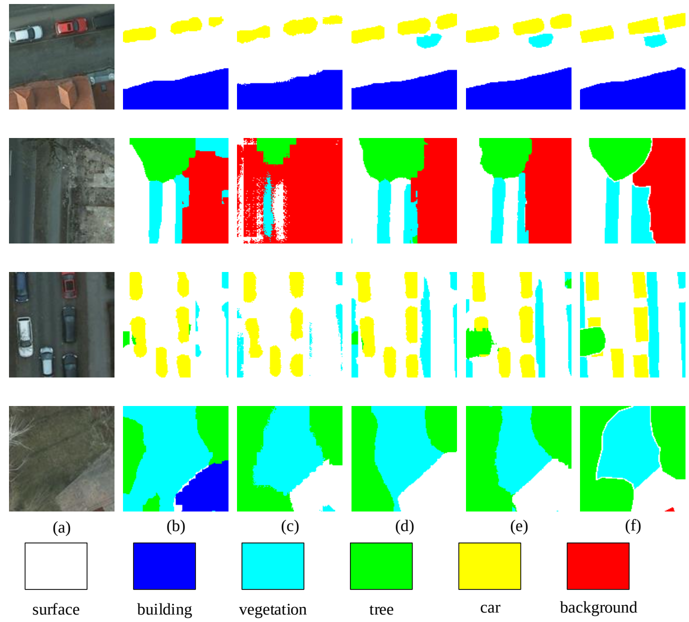

Fig. 3: Semantic segmentation results with different inputs. (a) color images; (b) segmentation results of grayscale images; (b) segmentation results of color images; (c) segmentation results of color images; (d) segmentation results of spectral resolution restored images; (e) segmentation results of spectral and spatial resolution restored images; (f) labels.

The visualization results of the semantic segmentation experiment are shown in Fig. 3, and the closest to the label is the segmentation result of spectral and spatial resolution restored image. For objects with small volume, only input spectral resolution restored image or both spatial and spectral resolution restored image can segment them successfully . For example, in the segmentation results of the first row, the vegetation can only be segmented by the latter two groups of experiments with restored resolution, while the two groups of experiments directly using the grayscale image and color image fail to segment the vegetation. The segmentation effect after restoring the spatial and spectral resolution of the grayscale image is better than the segmentation effect after restoring the spectral resolution of the grayscale image in detail. For example, the green tree on the left of the third row can only be successfully segmented by the group of experiments that restore the spatial and spectral resolution at the same time. The reason is that the spatial resolution of the tree in the image is very low, making it difficult to be segmented without improving the spatial resolution. After increasing the spatial resolution by our spatial resolution restore algorithm, the tree is successfully segmented.

REFERENCES

[1] R. Li, S. Zheng, C. Zhang, C. Duan, L. Wang, and P. M. Atkinson, "Abcnet: Attentive bilateral contextual network for efficient semantic segmentation of fine-resolution remotely sensed imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 181, pp. 84–98, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271621002379