4RATFNet: Four-Dimensional Residual-Attention Improved-Transfer Few-Shot Semantic Segmentation Network for Landslide Detection

Shiya Huang[1], Qiang Li[1(✉)], Jiajun Li[2], and Jinzheng Lu[1]

[1] School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, Sichuan, China
liqiangsir@swust.edu.cn

[2] School of Environment and Resource, Southwest University of Science and Technology, Mianyang 621010, Sichuan, China

Abstract. Landslides are hazardous and in many cases can cause enormous economic losses and human casualties. The suddenness of landslides makes it difficult to detect landslides quickly and effectively. Therefore, to address the problem of intelligent analysis of geological landslides, we propose a 4RATFNet network for few-shot semantic segmentation detection in the case of insufficient number of labeled landslide images. First, a residual-attention module is designed to fuse channel features and spatial features for residual fusion. Second, improved transfer learning is used to optimize the parameters of the pre-trained network. Third, the network downscales the four-dimensional convolutional kernel into a pair of two-dimensional convolutional kernels. Finally, the few-shot semantic segmentation network is used to extract support image features and complete the landslide detection for the same features in the query image. The experimental results show that the method performs better when tested on Resnet50 backbone and Resnet101 backbone when the sample size of labeled landslide images is insufficient. Compared with traditional semantic segmentation methods, it can obtain better segmentation results and achieve higher mean intersection over union, indicating that our network has obvious advantages and wider applicability.

Keywords: Four-dimensional convolution kernel · Residual-attention mechanism · Improved transfer learning · Few-shot semantic segmentation · Landslide detection

1 Introduction

Landslide is a highly destructive natural disaster and a derivative of many other natural hazards [1]. Broadly speaking, landslides encompass debris flow and rockfalls, which can be very hazardous [2]. Because of the great danger of landslides

B. Sheng et al. (Eds.): CGI 2023, LNCS 14497, pp. 65–77, 2024.
https://doi.org/10.1007/978-3-031-50075-6_6

and the obvious characteristics of landslides, it is of great significance to detect landslides. It allows ground personnel to promptly respond based on the information obtained.

In recent years, scholars have studied landslide detection methods. Traditional approaches involve field investigations using manual surveys and tools. However, such investigations cover a wide range, and considering the significant hazards of landslides, they can lead to personnel losses and wastage of resources. With the continuous development of computer science, various intelligent technologies have been applied to landslide detection. These include methods based on knowledge prior [3] and machine learning [4]. Tian et al. [5] improve the Transformer model using the temporal convolutional network. This model is sensitive to rapid landslide deformations and has high accuracy. Ding et al. [6] establish a landslide probability model to make effective judgments for landslide detection in specific areas and avoid overfitting issues.

Given that the neural networks currently used for landslide detection require landslide datasets of a certain sample size for training, and that there are few publicly available landslide datasets with annotations. Landslide detection is limited by insufficient training samples and long processing time.

Therefore, we use semantic segmentation [7,8] and attention mechanism [9,10] to achieve landslide detection. We design a few-shot semantic segmentation network based on improved residual attention and optimized transfer learning for landslide detection, as shown in Fig. 1. The residual attention module is designed to focus on both the original images and the key regions of the landslide to capture their feature information. An optimized transfer learning is designed to enhance the underlying feature extraction from few-shot datasets. The mean intersection over union (mIoU) [11] and the foreground-background intersection over union (FB-IoU) [12] are used as performance metrics to evaluate the effectiveness of landslide segmentation detection. The two proposed modules further improve the network response speed and landslide segmentation accuracy.

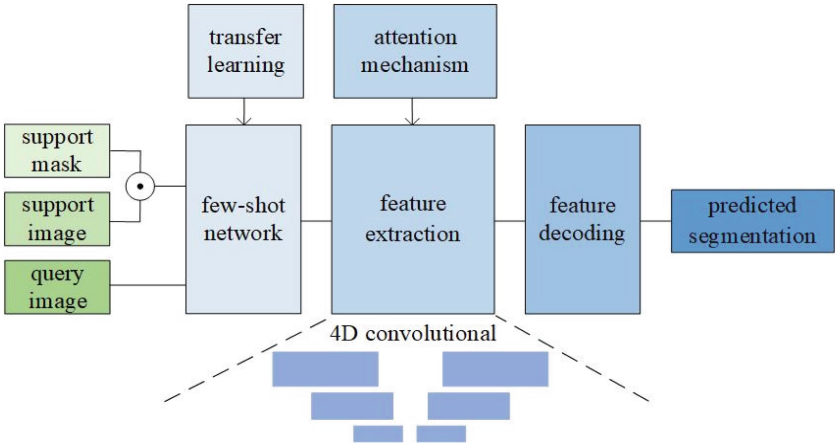


Fig. 1. Structure of 4RATFNet network

2 Related Works

2.1 Few-Shot Semantic Segmentation Network

Few-shot semantic segmentation was first proposed by Shaban et al. [13]. Compared with semantic segmentation based on traditional neural networks, the required datasets are greatly reduced. In the few-shot semantic segmentation network, the support image, support mask image and query image are input in the network that has been trained. And by learning the features of the support image and support mask image, we determine whether there are the same features in the query image and make prediction by mask annotation in the query image. After the few-shot semantic segmentation network was proposed, there have been many more advanced studies for few shots. Zhao et al. [14] achieve the improvement of the performance of the unknown class segmentation by introducing the object module to reduce the interference of the background and to mitigate the overfitting of the network. Min et al. [15] squeeze the network by four-dimensional hyper-correlation and fuse the features at different levels to obtain accurate segmentation results for unknown classes. Fan et al. [16] propose a self-support segmentation network based on the fact that pixels of different objects of the same class are more similar. However, the currently proposed few-shot semantic segmentation networks have not been specifically designed for landslide detection. Landslides have different image features compared to the images in the network training dataset. They occur over large areas and can be easily obscured by forests. When directly using existing networks for landslide detection, the segmentation accuracy is low.

2.2 HSNet Network

The public datasets required for landslide detection are few, and the number of labeled samples is insufficient. The hyper-correlation squeeze few-shot semantic segmentation network (HSNet) [15] achieves high-dimensional feature extraction through multi-level features and 4D convolutions, and segments the same features in the query image. The network structure of HSNet is shown in Fig. 2. First, the support image and support mask image undergo a Hadamard product and are then fed into the hyper-correlation construction module along with the query image. Different hierarchical features are constructed using cosine similarity, forming a hyper-correlation pyramid. Second, the pyramid is input into the 4D convolutional pyramid encoder module, achieving feature downsampling and upsampling and acquiring multi-level features. Finally, the upsampled recovered features are fed back to the 2D convolutional context decoder module to determine the class of each pixel point and get the predicted output of the query mask. The HSNet network can achieve segmentation on the query image when the number of support images is limited. We use this network as the original model and adapt the network in a scenario specific to landslide detection. The network is eventually made capable of segmenting landslide images with high accuracy without relying on a large number of datasets.

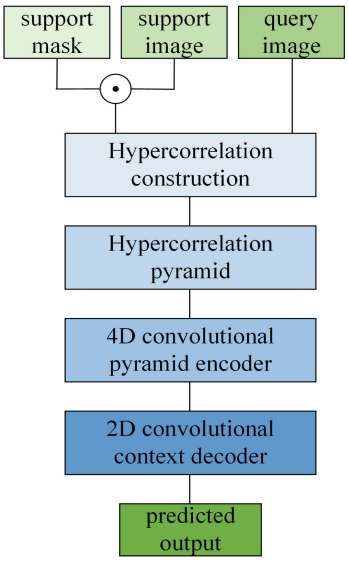


Fig. 2. HSNet network structure

3 Proposed Method

3.1 4D Convolutional

The 4D convolution module in the network is chosen to replace one 4D convolutional kernel with a pair of 2D convolutional kernels [15]. This reduces the number of 4D convolutional kernels used and reduces the loss of network resources. Due to too many parameters, some weight parameters in the neural network are chosen to be sparsely discarded. Focusing only on the parts that have a greater impact on certain weight positions in the network reduces the scale of the network and makes the computation faster.

In the four-dimensional space, let $(x, x') \in R^4$, $\mathbf{x}$ and $\mathbf{x}'$ respectively represent the positions of two 2D spaces, that is, the surrounding parts of a point on the 2D plane. In the encoder module, 4D convolution is used to obtain the relevant features of the image. The cosine similarity tensor is $c \in R^{H \times W \times H \times W}$, and the 4D convolution kernel is $k \in R^{k' \times k' \times k' \times k'}$. Four-dimensional convolution at position $(x, x') \in R^4$:

$$(c * k)(x, x') = \sum_{(p,p') \in \Psi(x,x')} c(p, p') \times k(p - x, p' - x') \tag{1}$$

where $\Psi(x, x')$ represents a group of neighborhood areas centered on the 4D position (x, x'), which means in the 4D area (x, x'), there is $\Psi(x, x') = \Psi(x) \times \Psi(x')$. 4D convolution operations consume a lot of computing power and occupy a large amount of storage resources.

In order to reduce the application limit of the network to 4D convolution, the method of weight sparseness is used to discard the network weight parameters. The weight sparse process is shown in Fig. 3. In the 2D space represented by $\mathbf{x}$ and $\mathbf{x}'$, relative to each other, the pixels around the center of the two pixels are regarded as having no influence or less influence points. During 4D convolution operation, two pixel centers only have weight values with the other center and the pixel points around the other center. These two centers ignore unimportant weight values for pixels surrounding their own centers, thus accomplishing weight sparsity.

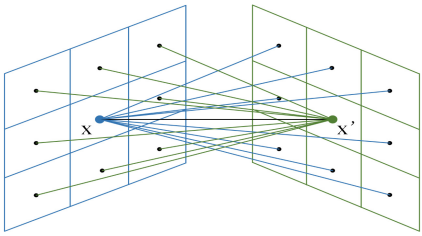


Fig. 3. Weight sparse process

After weight sparsity in the network, when the 4D position (x, x') is adjacent to each other in the corresponding 2D subspaces, the corresponding region set is collected, which is defined as:

$$\Psi_C(x, x') = \{(p, p') \in \Psi(x, x') : p = x\} \tag{2}$$

$$\Psi_{C'}(x, x') = \{(p, p') \in \Psi(x, x') : p' = x'\} \tag{3}$$

The region set of the two centers as a whole is defined as:

$$\Psi_{CA}(x, x') = \Psi_C(x, x') \cup \Psi_{C'}(x, x') \tag{4}$$

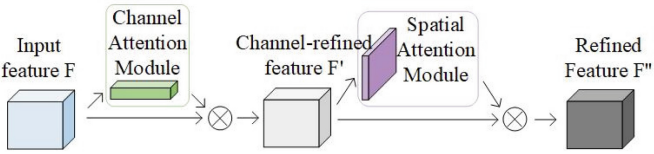
A 4D convolution in the center can be formulated as the sum of two separate 4D convolutions:

$$(c * k_{CA})(x, x') = (c * k_C)(x, x') + (c * k_{C'})(x, x') \tag{5}$$

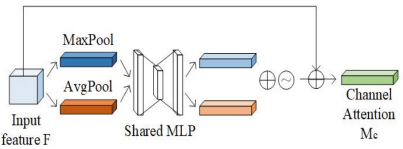
where k_C and $k_{C'}$ are the 4D convolution kernels in the respective central region sets respectively. When computing results for a single center:

$$(c * k_C)(x, x') = \sum_{p' \in \Psi(x')} c(x, p') k_C^{2D}(p' - x') \tag{6}$$

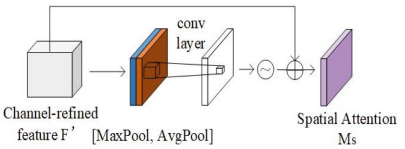
$$(c * k_{C'})(x, x') = \sum_{p \in \Psi(x)} c(p, x') k_{C'}^{2D}(p - x) \tag{7}$$



(a) Original CBAM module structure



(b) Residual-channel attention module structure



(c) Residual-spatial attention module structure

Fig. 4. Attention module

Therefore, the 4D convolution becomes two convolutions on 2D slices through the convolution operation [15]. This further reduces the computational load of convolution and speeds up the operation, which is formulated by Eq. (8):

$$(c * k_{CA})(x, x') = \sum_{p' \in \Psi(x')} c(x, p') k_C^{2D}(p' - x') + \sum_{p \in \Psi(x)} c(p, x') k_{C'}^{2D}(p - x) \quad (8)$$

3.2 Residual-Attention Module

In order to better focus on important areas in the image, human vision is simulated to focus on key content. In the residual module of the HSNet backbone, the CBAM attention module [17] is introduced and improved into the first and last layers of the network. As shown in Fig. 4(a), the preliminary feature **F** obtained after processing is used as input, and the channel attention and the spatial attention are performed in sequence. Finally the improved feature after processing is obtained. The processing is denoted by Eq. 9 and Eq. 10 as:

$$F' = M_c(F) \otimes F \quad (9)$$

$$F'' = M_s(F') \otimes F' \quad (10)$$

where $\otimes$ represents element-wise multiplication, $M_c \in R^{C \times 1 \times 1}$ is the 1D channel attention map obtained after processing by the channel attention module, and $M_s \in R^{1 \times H \times W}$ is the 2D spatial attention map obtained after processing by the spatial attention module.

The residual-channel attention module structure is shown in Fig. 4(b). Firstly, max-pooling and average-pooling are performed on the input feature **F** to aggregate the spatial information of image features. Secondly, the obtained spatial context descriptor is passed into the shared network for processing, and the

channel attention feature vector is obtained by first reducing and then increasing the number of channels. Thirdly, the channel attention feature vectors of max-pooling and average-pooling are summed element-wise. Fourthly, use the sigmoid function for normalization. Finally, the input and output are connected with residuals to obtain a one-dimensional channel attention map M_c.

The residual-spatial attention module structure is shown in Fig. 4(c). Firstly, the channel information of the image is aggregated by max-pooling and average-pooling in the channel direction on the feature $\mathbf{F}'$ improved by the channel. Secondly, dimensionality splicing and merging are performed on the two-layer features. Thirdly, the result is processed by a convolutional layer with a channel number of 1 to adjust the number of channels. Fourthly, use the sigmoid function for normalization. Finally, the input and output are connected with residuals to obtain the two-dimensional spatial attention map M_s.

3.3 Improved-Transfer Learning

Due to the introduction of the residual attention mechanism, the backbone model of the network has changed. If training is performed directly without using a pre-trained model, the network may have slow convergence speed and low segmentation accuracy. Therefore, an improved transfer learning approach is adopted to optimize the network model parameters. It can enhance the learning capability of the network to extract the underlying image features.

The HSNet network after adding the residual attention mechanism selects Resnet50 and Resnet101 as the backbone of the model. After adding channel attention and spatial attention, the network structure changes and the model parameters increase. Therefore, the parameters of the pre-trained model obtained on the ImageNet dataset [18] are used to optimize the network parameters. The ImageNet dataset is a large-scale dataset with more than 14 million images and more than 20,000 categories. Compared with the FSS dataset [19] used in this paper, it has more images and image categories. The FSS dataset contains 1000 categories, each category has 10 images, which is suitable for landslide recognition and segmentation in the case of few samples. When using the pre-trained model of the ImageNet dataset to optimize parameters, the network has more training samples and has stronger generalization.

Improved transfer learning improves the performance of the model on the target domain based on the source domain model. Its structure is shown in Fig. 5. The source dataset is the ImageNet dataset, which is trained on the Resnet50 and Resnet101 backbone to obtain a network model to complete the task. The target dataset is the FSS dataset, which is also trained on the Resnet50 and Resnet101 backbone after adding the residual attention module. The transfer of the network is completed by comparing and adding new weights to the new weight file. The new pre-trained models of Resnet50 and Resnet101 networks after adding the residual attention module on the ImageNet dataset network are obtained through improved transfer learning. And they are used to complete new tasks in the target domain.

4 Experiments

4.1 Datasets

The dataset in this paper uses the FSS dataset [19] for training, and the high-precision aerial imagery and interpretation dataset [20] for testing. We choose the FSS dataset when training the neural network. Compared with the commonly used few-shot datasets PASCAL-5i [13] and COCO-20i [21], the FSS dataset is more effective for model training and have more types. We group the FSS dataset, set 520 of them as the training set, and set 240 of them as the validation set. The types are selected randomly and are not repeated. There is no intersection between the types of the training set and the validation set. To test the segmentation performance of this network in landslide detection, the public dataset of landslides [20] is selected. There are a total of 59 landslide images with pixel-level annotations.

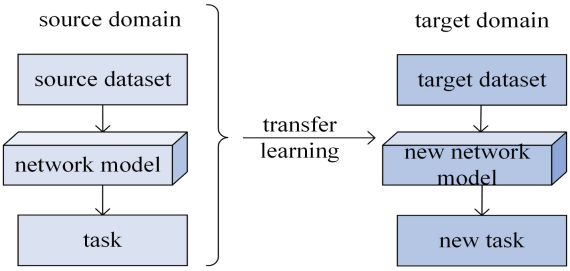


Fig. 5. Improved transfer learning structure

4.2 Evaluation Metrics

In order to verify the performance of the model, we use mIoU [11] and FB-IoU [12] as evaluation metrics to evaluate the error between the predicted segmentation results of the network and the ground truth. The mIoU is to calculate the mean intersection over union of all different types of objects that need to be segmented. The intersection of the pixel points of the predicted segmentation and the ground truth is divided by the pixel point union to obtain the ratios. Then, they are added and divided by the number of types to obtain the mIoU, as shown in Eq. 11:

$$mIoU = \frac{1}{C}\sum_{i=1}^{C} IoU_i = \frac{1}{C}\sum_{i=1}^{C} \frac{pre_i \cap truth_i}{pre_i \cup truth_i} \tag{11}$$

where C is the number of different types that need to be segmented, pre is the prediction result, and truth is the ground truth. FB-IoU calculates the foreground IoU and background IoU respectively, and takes their average value to

obtain the FB-IoU, as shown in Eq. 12:

$$FB - IoU = \frac{1}{2}(IoU_F + IoU_B) \tag{12}$$

where IoU_F is the foreground IoU of the image, and IoU_B is the background IoU of the image.

4.3 Comparative Experiments

Table 1 presents the performance of the currently relevant neural network-based landslide detection methods. Our network is compared with mask region-based convolution neural network(Mask R-CNN) [22] and susceptibility-guided fully convolutional neural network(SG-FCNN) [23]. Mask R-CNN [22] conducts experiments on the Resnet50 and Resnet101 backbones and builds its own landslide dataset. The mIoU values for segmentation of newly formed landslides are listed in Table 1. On the two backbone networks, the mIoU values are 64.95% and 77.94%, respectively. SG-FCNN [23] uses landslide sensitivity as a prior knowledge. Based on the fully convolutional neural network, a boundary and morphological optimization (BMO) module and a mean changing magnitude of objects (MCMO) module are added to effectively improve the accuracy of landslide detection. In Table 1, the value of mIoU for landslide detection using this method in two regions is recorded. The mIoU value is 84.65% on region A and 76.03% on region B. Therefore, the mIoU value of this method is 80.34%.

Table 1. Comparison of the results of the relevant methods

	mIoU	FB-IoU
Resnet50+Mask R-CNN [22]	64.95	–
Resnet101+Mask R-CNN [22]	77.94	–
SG-FCNN+BMO+MCMO [23]	80.34	–
Ours with Resnet50	**82.85**	**89.69**
Ours with Resnet101	**84.19**	**90.63**

The comparison of the data in Table 1 shows that the network in this paper improves 17.9% in mIoU over network [22] and 2.51% over network [23] with ResNet50 as the backbone. In the case of ResNet101 as the backbone, it improves 6.25% over network [22] in mIoU and 3.85% over network [23]. Among these networks, the proposed network in this paper has the highest mIoU, the best effect of segmenting landslides, and requires the least number of landslide datasets, which proves the superiority of the proposed model.

4.4 Ablation Study

In order to verify the effectiveness of the residual-attention module and improved-transfer learning, these two components are studied qualitatively through the ablation study. Eight comparative experiments are carried out, four experiments each with Resnet50 and Resnet101 as the backbone. They are a network without the residual-attention module and not using improved-transfer learning, a network with the residual-attention module but not using improved-transfer learning, a network without the residual-attention module but using improved-transfer learning, and a network with both the residual-attention module and improved-transfer learning. The mIoU and FB-IoU are used as evaluation metrics for the ablation experiments. Since training the network is time-consuming, 160 iteration rounds are set for each ablation experiment to compare their evaluation metrics.

The results of the ablation experiments are shown in Table 2, where the mIoU is the main observation. It can be found that the performance of the network improves substantially after adding improved-transfer learning, and improves marginally after adding the residual-attention module. The Resnet50 backbone improves mIoU by 0.7% with the addition of the residual-attention module, by 28.24% with the addition of improved-transfer learning, and by 28.39% when both modules are added. The Resnet101 backbone improves mIoU by 1.03% with the addition of the residual-attention module, by 30.56% with the addition of the improved-transfer learning, and by 30.91% with the addition of both modules.

Table 2. Results of ablation experiments

	mIoU	FB-IoU
Resnet50	54.46	70.71
Resnet50+CBAM	55.16	71.07
Resnet50+transfer learning	82.70	**89.69**
Ours with Resnet50	**82.85**	**89.69**
Resnet101	53.28	69.90
Resnet101+CBAM	54.31	70.26
Resnet101+transfer learning	83.84	90.26
Ours with Resnet101	**84.19**	**90.63**

The results of the ablation experiments show that the model obtained from the original backbone network is not ideal when the number of landslide samples is small. The use of improved-transfer learning can fully exploit the underlying features of the image, so that it can be applied to the detection of landslide features to obtain the underlying features more accurately. Using the residual-attention module can focus the attention of the network on the foreground portion of the image, thus focusing on a relatively small area of landslides occurring in most of the background.

Figure 6 and Fig. 7 respectively show the visualized results of the ablation experiments using ResNet50 and ResNet101 as the backbones. It can be observed that the best segmentation performance, closest to the ground truth, is achieved when both backbones use the improved transfer learning and residual attention modules simultaneously.

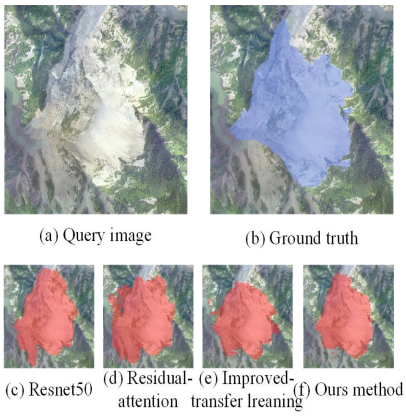


Fig. 6. Visualization results of Resnet50

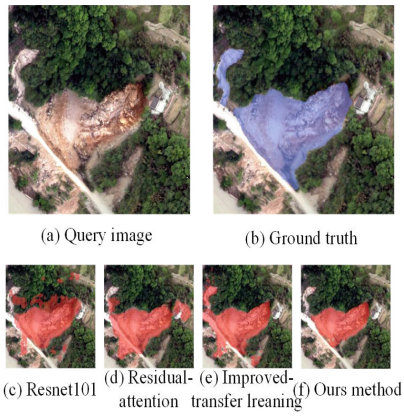


Fig. 7. Visualization results of Resnet101

5 Conclusion and Future Work

In this paper, we propose a few-shot landslide detection method based on residual attention and improved transfer learning. For the problem of few public labeled datasets of landslide, a 4RATFNet network is used to complete the landslide detection. The residual attention mechanism is introduced to focus on the foreground landslide regions in the image from both channel and space aspects, which effectively suppresses the interference of cluttered background. The introduction of improved transfer learning effectively optimizes the pre-trained parameters of the network under the change of the network structure. Our proposed method outperforms existing newer landslide detection methods on the evaluation metric mIoU. And the effectiveness of the two modules is demonstrated by metrics mIoU and FB-IoU. However, our network needs to manually label landslide images in the initial stage, and the more types of landslides are labeled in this stage, the higher the accuracy of final detection will be. In future work, we will preprocess the few-shot datasets to reduce the differences between them and the public large sample datasets, so as to achieve an increase in the similarity between the two datasets and improve the accuracy of landslide detection.

Acknowledgements. This work was supported by the Sichuan Science and Technology Program under Grant No. 2022YFG0148 and the Heilongjiang Science and Technology Program under Grant No. 2022ZX01A16.

References

1. Wang, X., Fan, X., Xu, Q., Du, P.: Change detection-based co-seismic landslide mapping through extended morphological profiles and ensemble strategy. ISPRS J. Photogramm. Remote. Sens. **187**, 225–239 (2022)
2. Ghorbanzadeh, O., Shahabi, H., Crivellari, A., Homayouni, S., Blaschke, T., Ghamisi, P.: Landslide detection using deep learning and object-based image analysis. Landslides **19**(4), 929–939 (2022)
3. Feizizadeh, B., Ghorbanzadeh, O.: GIS-based interval pairwise comparison matrices as a novel approach for optimizing an analytical hierarchy process and multiple criteria weighting. GI_Forum **1**, 27–35 (2017)
4. Ghorbanzadeh, O., Blaschke, T., Aryal, J., Gholaminia, K.: A new GIS-based technique using an adaptive neuro-fuzzy inference system for land subsidence susceptibility mapping. J. Spat. Sci. **65**(3), 401–418 (2020)
5. Tian, Y., et al.: A transformer-based model for short-term landslide displacement prediction. Acta Sci. Naturalium Univ. Pekinensis **59**(2), 197–210 (2023)
6. Ding, X., Zhao, X., Wu, X., Zhang, T., Xu, Z.: Landslide susceptibility assessment model based on multi-class SVM with RBF kernel. China Saf. Sci. J. **32**(3), 194–200 (2022)
7. Huang, Y., Shi, P., He, H., He, H., Zhao, B.: Senet: spatial information enhancement for semantic segmentation neural networks. Vis. Comput. 1–14 (2023)
8. Ma, Z., Yuan, M., Gu, J., Meng, W., Xu, S., Zhang, X.: Triple-strip attention mechanism-based natural disaster images classification and segmentation. Vis. Comput. **38**(9–10), 3163–3173 (2022)
9. Qin, Y., Chi, X., Sheng, B., Lau, R.W.: GuideRender: large-scale scene navigation based on multi-modal view frustum movement prediction. Vis. Comput. 1–11 (2023)
10. Yang, P., Wang, M., Yuan, H., He, C., Cong, L.: Using contour loss constraining residual attention U-net on optical remote sensing interpretation. Vis. Comput. 1–13 (2022)
11. Chen, Q., Yang, Y., Huang, T., Feng, Y.: A survey on few-shot image semantic segmentation. Front. Data Comput. **3**(6), 17–34 (2021)
12. Rakelly, K., Shelhamer, E., Darrell, T., Efros, A., Levine, S.: Conditional networks for few-shot semantic segmentation. In: International Conference on Learning Representations (2018)
13. Shaban, A., Bansal, S., Liu, Z., Essa, I., Boots, B.: One-shot learning for semantic segmentation. In: British Machine Vision Conference 2017, pp. 167.1–167.13 (2017)
14. Zhao, Y., Price, B., Cohen, S., Gurari, D.: Objectness-aware few-shot semantic segmentation. arXiv preprint arXiv:2004.02945 (2020)
15. Min, J., Kang, D., Cho, M.: Hypercorrelation squeeze for few-shot segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 6941–6952 (2021)
16. Fan, Q., Pei, W., Tai, Y.W., Tang, C.K.: Self-support few-shot semantic segmentation. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds.) ECCV 2022, Part XIX. LNCS, vol. 13679, pp. 701–719. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-19800-7_41

17. Woo, S., Park, J., Lee, J.Y., Kweon, I.S.: CBAM: convolutional block attention module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 3–19. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_1
18. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE (2009)
19. Li, X., Wei, T., Chen, Y.P., Tai, Y.W., Tang, C.K.: FSS-1000: a 1000-class dataset for few-shot segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2869–2878 (2020)
20. Zeng, C., Cao, Z., Su, F., Zeng, Z., Yu, C.: High-precision aerial imagery and interpretation dataset of landslide and debris flow disaster in Sichuan and surrounding areas. China Sci. Data **7**(2), 195–205 (2022)
21. Nguyen, K., Todorovic, S.: Feature weighting and boosting for few-shot segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 622–631 (2019)
22. Jiang, W., et al.: Deep learning for landslide detection and segmentation in high-resolution optical images along the Sichuan-Tibet transportation corridor. Remote Sens. **14**(21), 5490 (2022)
23. Chen, Y., et al.: Susceptibility-guided landslide detection using fully convolutional neural network. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. **16**, 998–1018 (2022)