International Journal of Advanced Engineering and Technology Research
ISSN: 3106-8677 (Print), ISSN: 3106-8685 (Online) | Vol. 4, No. 2, 2026

A Review of Intelligent Recognition of Rock Mass Structures for Inrush Prevention in Tunnels

Zhili Chen

School of Civil Engineering, Henan Polytechnic University, Jiaozuo-Henan-454000, China

Abstract: The scale of tunnel construction in China continues to expand, and a large number of projects are advancing into the complex geological regions of western China, intensifying safety risks such as water inrush and collapse. Traditional manual on-site measurement of rock mass fractures is highly subjective, inefficient and accompanied by operational safety hazards. Focusing on the intelligent recognition of rock mass structures for inrush prevention in tunnels, this paper presents a review of the relevant research: image semantic segmentation is divided into two major categories, namely traditional methods and deep learning methods, and the research status of traditional segmentation algorithms based on thresholding, regions, edges and graph theory, as well as deep segmentation models such as FCN, U-Net, DeepLab, attention mechanisms and CNN-Transformer architectures, is summarized. The achievements related to tunnel face fracture recognition are also reviewed, and the performance of traditional image processing and deep learning in fracture detection is compared; a large number of scholars have demonstrated that deep learning can automatically extract deep-level features of rock mass fractures and achieves better recognition performance than traditional algorithms. Based on a review of the existing research, the paper identifies the shortcomings of current technologies and provides a summary and outlook on the future development directions of intelligent recognition technology for tunnel rock masses.

Keywords: Rock mass structure, Tunnel face fractures, Image semantic segmentation, Deep learning, Intelligent recognition.

1. Introduction

In recent years, along with the steady growth of the Chinese economy, the continuous enhancement of overall national strength and the wide application of high-tech technologies, tunnel and underground engineering has encountered unprecedented development opportunities. Statistics show that by the end of 2024, the number of railway tunnels in operation in China had reached 18,997, with a total length exceeding 24,246 km; in 2024, 424 new railway tunnels were put into operation, with a total length of 738 km, including 10 extra-long tunnels longer than 10 km with a combined length of about 138 km [1]. At the same time, the Outline of the 14th Five-Year Plan explicitly proposes "accelerating the construction of a transportation powerhouse, implementing major projects such as the Sichuan-Tibet Railway, the New Western Land-Sea Corridor and the National Water Network, and advancing a batch of major projects that strengthen foundations, enhance functions and deliver long-term benefits" [2]. This strategic deployment has gradually shifted the focus of transportation tunnel engineering toward the western region, where the geological environment is complex.

However, when a tunnel project passes through a complex geological environment, construction difficulty and safety risk often grow geometrically. Under such high-risk geological conditions, traditional fracture measurement methods still rely mainly on multi-person collaborative work; the results depend heavily on the professional competence and empirical judgment of the operators, are prone to considerable systematic errors, and can hardly guarantee the completeness and consistency of the described information. Especially in large tunnel projects, the workload of fracture measurement is enormous, consuming substantial manpower, financial resources and time, while the data obtained often fail to meet requirements in terms of systematicity and standardization. More alarming is that such measurement work usually has to be carried out in high-risk construction areas, where operators face direct threats from collapse, water inrush and dust, seriously endangering the life and health safety of technicians. To address these problems, scholars have carried out extensive research on the recognition of rock mass structures for inrush prevention in tunnels; these results have not only significantly improved the efficiency of acquiring rock mass structure information, but also effectively reduced subjective interference in data interpretation, while providing a strong guarantee for the safety of construction personnel.

On this basis, this paper focuses on the intelligent recognition of rock mass structures for inrush prevention in tunnels, systematically summarizes the research results of a large number of scholars on image semantic segmentation and intelligent recognition of tunnel face fractures, and further summarizes the existing methods and offers an outlook on future development directions.

2. Image Semantic Segmentation

At present, image semantic segmentation, as one of the core tasks of computer vision, follows two main technical paths: one is based on traditional theories such as digital image processing and graph theory, and the other is represented by deep neural networks, which rely on data-driven end-to-end learning.

2.1. Research on Traditional Image Segmentation Methods

Before deep learning became the mainstream technology, image semantic segmentation tasks were mainly carried out with traditional image processing methods. Such methods usually follow the general approach of natural image analysis, combine prior knowledge of a specific application domain, build discriminative models suitable for pixel-level

recognition and classification, and rely on manually designed feature extraction procedures to accomplish the segmentation task. From a methodological perspective, traditional image segmentation techniques can be broadly grouped into four typical paradigms, namely threshold-based, edge-based, region-based and graph-theory-based segmentation methods [3].

(1) Threshold-based segmentation methods

Threshold-based image segmentation methods [4-6] mainly rely on the statistical characteristics of gray-level histograms: by setting different gray-level thresholds, pixels are classified so that pixels falling in similar gray-level ranges are treated as the same class, thereby achieving a preliminary partition of regions. For example, Liu et al. [7] proposed automatically determining the optimal segmentation threshold by maximizing the inter-class variance, which realized adaptive binarization of remote sensing images, significantly enhanced the distinction between objects and background, and has been widely applied in vegetation cover segmentation. Kashanipour et al. [8] introduced multidimensional fuzzy rules for threshold optimization, optimizing the histogram segmentation process by fusing spectral and spatial features, and achieved a classification accuracy of 92.3% in vegetation monitoring tasks, although the algorithm complexity is relatively high. Yatendra et al. [9] adopted a histogram-based K-means clustering method for image segmentation, and Wu et al. [10] proposed a multi-threshold segmentation technique for building image extraction. In addition, Li Gengsheng et al. [11] developed an adaptive active contour model that effectively alleviated the segmentation problems caused by intensity inhomogeneity and sensitivity to the initial contour. In terms of algorithmic practicality, classical threshold techniques such as the Otsu algorithm [12] and the Bernsen method [13] have the advantages of simple implementation and high computational efficiency. However, these methods are generally sensitive to noise interference and threshold selection, and their adaptability in complex scenes is limited. Although they can achieve certain results in relatively simple tasks such as object detection and image classification, since they rely only on the single dimension of gray-level information, they tend to produce under-segmentation when facing the diverse objects and backgrounds of complex scenes.

(2) Region-based segmentation methods

Region-based image segmentation methods group pixels with similar attributes into the same class according to the spatial distribution characteristics of the image, thereby forming semantically consistent regions. This class of methods covers a variety of implementation strategies, among which region growing [14] and region merging [15] are two typical representatives. Felzenszwalb and Huttenlocher proposed an influential region segmentation algorithm [16], which dynamically determines the pixel aggregation manner by evaluating the similarity between regions and their connectivity, and the similarity measure adopted comprehensively considers multidimensional features such as color, texture and spatial position. The GrabCut technique developed by Rother et al. [17] takes a different approach: with the initial foreground and background labels provided by the user, pixel classification is achieved by minimizing an energy function, effectively separating the foreground and background of the image. In terms of method improvement and application expansion, Li Yi et al. [18] proposed an improved region growing algorithm for crack identification in metro stations, automatically locating crack seed points by designing a 2:1 pixel matrix and introducing a cyclic algorithm to achieve collaborative growth of multiple seeds, which significantly improved the stability of seed point selection compared with traditional methods. Tu Liping et al. [19] proposed a fast point cloud segmentation method combining K-means clustering and region growing [20], effectively improving processing efficiency. In addition, Z et al. [21] proposed a region growing segmentation method for brain MRI images.

(3) Edge-detection-based segmentation methods

Edge-detection-based image segmentation methods [22] mainly determine object boundaries according to the pixel positions where gray-level or structural features change abruptly; their basic principle is to complete region partitioning by exploiting the edge features at the junctions of different semantic regions. Common edge detection operators include Sobel, Prewitt, Canny and Laplacian [23, 24], which are widely used in various segmentation tasks. In terms of method improvement, Xiong Jinhua et al. [25] used the Canny operator [26] to detect object edges together with subsequent processing, obtaining segmentation results with clear boundaries. Schiewe et al. [27] proposed a contour closure algorithm for the problem of edge fragmentation, so as to enhance edge continuity. Yang et al. [28] constructed a multi-scale Canny optimization model based on geo-information mapping theory, improving sub-pixel localization accuracy while maintaining edge connectivity. Gao et al. [29] designed a Markov-Sobel hybrid architecture, enhancing the adaptability of the algorithm in complex scenes by introducing a probabilistic graphical model. In addition, the weighted guided filtering algorithm developed by Ranjan et al. [30] effectively suppressed noise while significantly improving the preservation of edge details. Although edge-based segmentation methods have made considerable progress, the existing techniques still face major challenges in the edge segmentation of ground objects in tunnel face images.

(4) Graph-theory-based segmentation methods

Graph-theory-based image segmentation methods [31, 32] transform the segmentation task into a graph partitioning problem, achieving region partitioning by constructing and optimizing an objective function; representative algorithms include GraphCut [33], GrabCut [34] and OneCut [34]. The GraphCut method proposed by Boykov et al. [33] maps an image into a graph structure and constructs a corresponding energy function, converting the segmentation problem into a combinatorial optimization problem of finding the minimum cut, and thus achieving globally effective segmentation. The GrabCut method further proposed by Rother et al. [17] introduces a Gaussian mixture model on the basis of GraphCut, modeling the object and background regions separately, and is therefore applicable to color image segmentation tasks.

Although graph-theory methods have a rigorous mathematical foundation in theory, they usually require considerable computational overhead and are difficult to apply for efficient segmentation in practice. Overall, although traditional image segmentation methods are rich in research results and relatively simple in algorithm structure, their application scenarios are somewhat limited: they often depend on low-level visual features of images and perform poorly in complex tasks that require high-level semantic understanding. As application demands deepen, such methods can hardly meet the higher requirements for segmentation

accuracy and semantic understanding now raised in various fields.

2.2. Deep Learning-Based Image Semantic Segmentation

With the continuous evolution of deep learning theory and the widespread application of convolutional neural networks (CNNs), significant breakthroughs have been achieved in the field of computer vision. As one of the key tasks in this field, semantic segmentation, after the introduction of CNNs, can extract and exploit the deep semantic information in images more effectively, thus driving the technology from shallow visual feature processing toward deep semantic understanding.

With the continuous evolution of deep learning theory and the extensive penetration of convolutional neural networks (CNNs), remarkable breakthroughs have been made in the field of computer vision. As one of the key tasks in this field, semantic segmentation, after the introduction of CNNs, can extract and utilize the deep semantic information in images more efficiently, driving the technology to leap from shallow visual feature processing toward deep semantic understanding.

Faced with increasingly complex semantic segmentation targets and application scenarios, researchers have successively proposed a series of improved deep-learning-based models, which not only improve segmentation accuracy but also continuously expand the boundaries of practical application. In 2015, Long et al. [35] pioneered the fully convolutional network (FCN), using the VGG network [36] as the backbone architecture and introducing deconvolution layers for upsampling. The model restores feature map resolution through a deconvolution mechanism, strengthens edge details with a multi-level feature fusion strategy, and establishes a pixel-level supervised learning paradigm to enhance semantic consistency. However, FCN still suffers from problems such as high computational cost and a limited receptive field, and is especially prone to fragmentation at the boundaries of large-scale ground objects.

To overcome these limitations, a variety of fully convolutional network variants emerged. In the same year, Ronneberger et al. [37] designed the U-Net structure, which effectively integrates multi-scale contextual information by introducing skip connections between the encoder and the decoder, significantly improving segmentation accuracy. The residual network (ResNet) proposed by He et al. [38] alleviates the vanishing gradient problem in deep networks through residual block structures, enabling models to extend to greater depths. Cai et al. [39] further introduced the atrous spatial pyramid pooling (ASPP) module into U-Net and strengthened the skip connections, enhancing the model's ability to perceive and fuse multi-scale features. In 2017, Badrinarayanan et al. [40] proposed SegNet, whose encoder extracts features through convolution and pooling operations, while the decoder uses pooling indices for upsampling to recover spatial details. Zhao et al. [41] constructed the pyramid scene parsing network (PSPNet), which fuses contextual information at different scales through parallel multi-scale pooling layers, thereby improving segmentation performance in complex scenes. Wu Congzhong et al. [42] designed a dual-module collaborative architecture combining adaptive filtering and residual fusion modules, achieving a significant improvement in mean intersection over union on the Cityscapes dataset. To better balance global and local information, the DeepLab series of models [43-46] emerged, in which techniques such as atrous convolution [47] and spatial pyramid pooling [48] are widely adopted. Atrous convolution expands the receptive field without increasing the number of parameters by adjusting the dilation rate; spatial pyramid pooling integrates contextual information at different scales through multi-scale pooling operations. Chen et al. [49] further combined the two and proposed the ASPP module, effectively strengthening the model's ability to model multi-scale objects.

In recent years, attention mechanisms have shown strong application potential in semantic segmentation. This mechanism can screen key semantic information from complex features, thereby improving the accuracy and robustness of models. Oktay et al. [50] introduced an attention gate mechanism in medical image segmentation, enabling the model to adaptively focus on target structures of different shapes. Hu et al. [51] proposed the squeeze-and-excitation module, which adaptively adjusts channel weights by modeling the interdependencies between channels. Zha et al. [41] embedded a pyramid pooling module (PPM) into PSPNet to capture multi-region contextual information and enhance global perception. Nie et al. [52] combined Transformer with CNN, exploiting the former's advantage in long-range dependency modeling to improve global information extraction. The OCNet proposed by Yuan et al. [53] further improved the prediction quality of semantic segmentation by increasing feature map resolution and optimizing the context aggregation mechanism.

3. Current Status of Tunnel Face Fracture Detection Research

Structural planes such as faults, joints, fractures and bedding planes are widely developed within rock masses [54], and their spatial distribution and development degree directly determine the mechanical behavior and overall stability of the rock mass. During the construction of underground projects such as tunnels, under the combined action of in-situ stress and engineering disturbance, the original structural integrity of the rock mass is easily damaged, thereby inducing the initiation, propagation and interconnection of fractures [55]. Therefore, accurately identifying rock mass fractures and implementing early warning is of great significance for effectively preventing engineering disasters such as collapse and rockfall and for ensuring construction safety. The following section systematically reviews the research progress in intelligent recognition and information extraction of rock mass structural fractures.

As the most intuitive morphological manifestation of structural planes, rock mass fractures have traditionally been measured on site and recorded manually by geological engineers, which is inefficient and highly subjective. With the continuous development of image processing technology, vision-based fracture recognition methods have been gradually applied to rock mass fracture analysis [56]. For example, Liu Chun et al. [57] built a fracture image recognition and quantitative analysis system based on digital image processing technology, realizing the automatic extraction of geometric parameters such as fracture length and width; Wang Yang et al. [58] generated digital image models of rock surfaces by means of photogrammetry, on which the digital acquisition of fracture information was completed. Deb et al. [59] used the Canny operator and the Hough

transform to perform edge recognition and structural analysis on binarized fracture images. However, such traditional methods mostly rely on low-order visual features at the pixel level, have relatively simple algorithm structures, and lack the ability to learn autonomously from data, so that they can hardly capture the global semantic information and high-order features of rock mass fractures effectively [60], which greatly restricts recognition accuracy in practical complex scenes.

In recent years, with the continuous penetration of computer vision technology, deep learning has shown broad application prospects in fracture recognition and characterization research in civil engineering. Xue Dongjie et al. [61] constructed a convolutional neural network model that achieved end-to-end intelligent recognition of rock fractures, outperforming traditional edge detection algorithms in both recall and recognition accuracy. Liu Houxiang et al. [62] further applied deep learning to the analysis of tunnel face images, realizing the automatic extraction of multiple types of geological information such as joints, fractures and fragmentation degree, and providing technical support for rapid classification of surrounding rock. Jia Shina [63] improved the YOLOv5 model by adding detection heads and introducing an attention mechanism, enhancing the recognition capability for small-scale fracture targets and the robustness in complex backgrounds. Bhowmick et al. [64] combined UAV imagery with a deep learning model to achieve automatic crack recognition and quantitative analysis of morphological parameters such as area and width. Su Yutong et al. [65] established an intelligent fracture recognition method for coal-rock borehole images based on the YOLOv3 algorithm, verifying the effectiveness of the model in practical borehole scenarios. Lemy and Hadjigeorgiou [66] further improved the continuity of fracture recognition by combining image edge restoration and binarization preprocessing with a neural network. Xu et al. [67] systematically compared the performance of Faster R-CNN and Mask R-CNN in crack detection tasks. Mandal et al. [68] used a supervised deep convolutional neural network to classify image patches, thereby locating and recognizing fracture regions. Zhang et al. [69] employed a BP neural network for crack identification and prediction of development probability, reaching an accuracy of 80%, and constructed a fracture development evaluation system incorporating parameters such as porosity and permeability.

4. Summary and Outlook

This paper reviews two core directions of intelligent recognition of rock mass structures for inrush prevention in tunnels, namely image semantic segmentation and tunnel face fracture detection. The industry is shifting from manual survey toward intelligent recognition. Traditional segmentation algorithms such as threshold-based, edge-based and graph-theory-based methods are computationally simple, but are difficult to adapt to the complex geological environment of tunnels; deep learning models such as FCN, U-Net and DeepLab rely on multi-scale feature fusion and greatly improve segmentation accuracy and adaptability. Fracture detection has evolved from traditional edge operators to deep learning algorithms such as YOLO and Mask R-CNN, which can accurately identify subtle and concealed fractures. Existing research still has shortcomings: most studies use only visible-light images and do not fuse multi-source data such as three-dimensional point clouds, infrared imagery and ground-penetrating radar. In the future, it is necessary to build a multimodal recognition system integrating vision, geometry and geophysical prospecting, fuse multi-dimensional information to identify concealed structures and sudden-disaster hazards, and promote the application of intelligent rock mass recognition technology at tunnel construction sites.

References

[1] Gong, J., Wang, W., Li, X., et al. (2025). Statistics of railway tunnels in China by the end of 2024 and introduction to the tunnels of newly opened projects in 2024. Tunnel Construction, 45(3), 636–653.

[2] Central Committee of the Communist Party of China, & State Council. (2021, March 13). Outline of the 14th Five-Year Plan for National Economic and Social Development of the People's Republic of China and the Long-Range Objectives Through the Year 2035. People's Daily, p. 1.

[3] Kumar, A. (2023). Study and analysis of different segmentation methods for brain tumor MRI application. Multimedia Tools and Applications, 82(5), 7117–7139.

[4] Otsu, N. (1975). A threshold selection method from gray-level histograms. Automatica, 11, 23–27.

[5] Pun, T. (1980). A new method for grey-level picture thresholding using the entropy of the histogram. Signal Processing, 2(3), 223–237.

[6] Yen, J. C., Chang, F. J., & Chang, S. (1995). A new criterion for automatic multilevel thresholding. IEEE Transactions on Image Processing, 4(3), 370–378.

[7] Liu, Y., Mu, X., Wang, L., & Zhang, J. (2012). Automated extraction of vegetation coverage area from UAV imagery using Otsu and morphological processing. Remote Sensing, 4(5), 1440–1460.

[8] Kashanipour, M., et al. (2023). A fuzzy-rule-based multiband thresholding approach for vegetation segmentation using spectral and spatial features. IEEE Transactions on Geoscience and Remote Sensing, 61, 1–15.

[9] Yatendra, S., Abhishek, T., Gupta, R. K., et al. (2023). A CNN-SVM based computer aided diagnosis of breast Cancer using histogram K-means segmentation technique. Multimedia Tools and Applications, 82(9), 14055–14075.

[10] Wu, S., Wu, Y., & Zou, J. (2018). Multi-level thresholding for remote sensing image of urban area based on line intercept histogram. CAAI Transactions on Intelligent Systems, 13(2), 227–235.

[11] Li, G., Liu, G., & Ma, W. (2022). Adaptive image segmentation based on regional information coupling. Laser & Optoelectronics Progress, 59(2), 567–579.

[12] Huang, M., Yu, W., & Zhu, D. (2012). An improved image segmentation algorithm based on the Otsu method. In 2012 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (pp. 135–139). IEEE.

[13] Giannakaki-Zimmermann, H., Huf, W., Schaal, K. B., et al. (2019). Comparison of choroidal thickness measurements using spectral domain optical coherence tomography in six different settings and with customized automated segmentation software. Translational Vision Science & Technology, 8(3), 5.

[14] Hu, J., Zhang, Y., & Xie, S. (2021). Review of the application of domestic remote sensing image classification technology. Computer Engineering and Applications, 57(3), 1–13.

[15] Wang, H., Shen, Z., Zhang, Z., et al. (2021). Improvement of region-merging image segmentation accuracy using multiple merging criteria. Remote Sensing, 13(14), 2782.

[16] Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. International Journal of Computer Vision, 59(1), 167–181.

[17] Rother, C., Kolmogorov, V., & Blake, A. (2004). "GrabCut": Interactive foreground extraction using iterated graph cuts. ACM Transactions on Graphics, 23(3), 309–314.

[18] Li, Y., Ma, T., Mei, Y., et al. (2025). Crack identification of cover-cut top-down metro stations based on an improved region growing method. Building Structure, 55(1), 128–131, 141.

[19] Tu, L., Hui, Z., Fan, J., et al. (2024). An optimized fast point cloud segmentation method combining K-means clustering and region growing. Bulletin of Surveying and Mapping, (12), 128–131, 154.

[20] Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 3–19).

[21] Weihua, Z., & Ying, S. (2019). A region growing segmentation approach for MRI brain image processing. In 13th International Conference on Anti-counterfeiting, Security, and Identification (ASID) (pp. 188–191).

[22] Guan, X., & Tian, G. (2025). A new digital image edge detection algorithm based on the Canny operator. Journal of Anyang Normal University, 27(2), 11–16.

[23] Yan, H., Zhao, Q., Xie, M., et al. (2021). Edge detection and repair of PCBA components using an adaptive Canny operator. Acta Optica Sinica, 41(5), 97–104.

[24] Zhang, X., Li, M., & Liu, H. (2023). Overlap functions-based fuzzy mathematical morphological operators and their applications in image edge extraction. Fractal and Fractional, 7(6), 465.

[25] Xiong, J., Yu, H., Li, F., et al. (2025). Research on slider edge detection based on an improved Canny algorithm. Information Technology and Informatization, (2), 174–178.

[26] Li, Q., Zhao, F., Xu, Z., et al. (2024). Infrared image edge detection algorithm for electrical equipment based on an improved Canny operator. Engineering Journal of Wuhan University, 57(12), 1759–1765.

[27] Schiewe, J., Segl, K., & Kaufmann, H. (2001). A contour closure algorithm for the analysis of fragmented edges in high-resolution remote sensing imagery. International Journal of Remote Sensing, 22(16), 3223–3235.

[28] Lijun, Y., Mengbo, L., Tongxin, W., et al. (2023). Geo-information mapping improves Canny edge detection method. IET Image Processing, 17(6), 1893–1904.

[29] Hou, J., Guo, Z., Wu, Y., et al. (2022). BSNet: Dynamic hybrid gradient convolution based boundary-sensitive network for remote sensing image segmentation. IEEE Transactions on Geoscience and Remote Sensing, 60, 1–22.

[30] Ranjan, R., & Avasthi, V. (2023). Edge detection using guided Sobel image filtering. Wireless Personal Communications, 132(1), 651–677.

[31] Karthick, P., Mohiuddine, S. A., Tamilvanan, K., et al. (2023). Investigations of color image segmentation based on connectivity measure, shape priority and normalized fuzzy graph cut. Applied Soft Computing, 139, 110239.

[32] Sharma, D., & Singhai, J. (2021). An unsupervised framework to extract the diverse building from the satellite images using Grab-cut method. Earth Science Informatics, 14(2), 777–795.

[33] Boykov, Y. Y., & Jolly, M. P. (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. In Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001 (Vol. 1, pp. 105–112). IEEE.

[34] Tang, M., Gorelick, L., Veksler, O., et al. (2013). GrabCut in one cut. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1769–1776).

[35] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3431–3440).

[36] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[37] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 234–241). Springer.

[38] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770–778).

[39] Cai, J., Shi, J., Leau, Y. B., et al. (2024). Res50-SimAM-ASPP-Unet: A Semantic Segmentation Model for High-resolution Remote Sensing Images. IEEE Access, 12, 2301–2316.

[40] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 2481–2495.

[41] Zhao, H., Shi, J., Qi, X., et al. (2017). Pyramid scene parsing network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2881–2890).

[42] Wu, C., Dong, H., & Fang, J. (2022). Adaptive filtering remote sensing image segmentation network based on attention mechanism. Computer Engineering & Science, 44(11), 2010–2018.

[43] Chen, L. C., Papandreou, G., Kokkinos, I., et al. (2014). Semantic image segmentation with deep convolutional nets and fully connected CRFs. arXiv preprint arXiv:1412.7062.

[44] Chen, L. C., Papandreou, G., Kokkinos, I., et al. (2017). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834–848.

[45] Weber, M., Wang, H., Qiao, S., et al. (2021). DeepLab2: A tensorflow library for deep labeling. arXiv preprint arXiv:2106.09748.

[46] Chen, L. C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587.

[47] Yu, F., & Koltun, V. (2015). Multi-scale context aggregation by dilated convolutions. arXiv preprint arXiv:1511.07122.

[48] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(9), 1904–1916.

[49] Chen, L. C., Zhu, Y., Papandreou, G., et al. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 801–818).

[50] Oktay, O., Schlemper, J., Folgoc, L. L., et al. (2018). Attention U-Net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.

[51] Zhang, C., Jiang, W., Zhang, Y., et al. (2022). Transformer and CNN hybrid deep neural network for semantic segmentation of very-high-resolution remote sensing imagery. IEEE Transactions on Geoscience and Remote Sensing, 60, 1–20.

[52] Nie, T., Zhao, Y., & Yao, S. (2024). Multi-path Segmentation Network Based on CNN and Transformer for Skin Lesion Image. In Proceedings of the Asian Conference on Computer Vision (pp. 3137–3153).

[53] Yuan, Y., Huang, L., Guo, J., et al. (2021). OCNet: Object Context Network for Scene Parsing. arXiv preprint arXiv:2101.04041.

[54] De Silva, V. R. S., Konietzki, H., Märten, H., et al. (2023). Grain-scale numerical simulation of crystalline rock fracturing using Soundless Cracking Demolition Agents for in-situ preconditioning. Computers and Geotechnics, 155, 105187.

[55] Leng, B., Zhang, Y., Yang, H., et al. (2021). Rapid Recognition of Rock Mass Fractures in Tunnel Faces. Journal of Southwest Jiaotong University, 56(2), 246–252, 322.

[56] Chen, J., Yang, T., Zhang, D., et al. (2021). Deep learning based classification of rock structure of tunnel face. Geoscience Frontiers, 12(1), 395–404.

[57] Liu, C., Wang, B., Shi, B., et al. (2008). An analytical method for morphological parameters of rock and soil fractures based on digital image recognition. Chinese Journal of Geotechnical Engineering, (9), 1383–1388.

[58] Wang, Y., Wang, S., Guo, M., et al. (2011). Rapid digital identification of joint information on tunnel faces and stability analysis. Chinese Journal of Geotechnical Engineering, 33(11), 1734–1739.

[59] Deb, D., Hariharan, S., Rao, U. M., et al. (2008). Automatic detection and analysis of discontinuity geometry of rock mass from digital images. Computers & Geosciences, 34(2), 115–126.

[60] You, Q. (2018). Image semantic segmentation based on deep convolutional encoder-decoder and adversarial learning [Master's thesis]. Xiamen University.

[61] Xue, D., Tang, Q., Wang, A., et al. (2019). Intelligent recognition of fracture geometry in rock and concrete based on FCN. Chinese Journal of Rock Mechanics and Engineering, 38(S2), 3393–3403.

[62] Liu, H., Li, W., Zha, H., et al. (2018). Classification method of highway tunnel surrounding rock based on deep learning. Chinese Journal of Geotechnical Engineering, 40(10), 1809–1817.

[63] Jia, S. (2023). Research on small object detection algorithm based on improved YOLOv5 [Master's thesis]. Nanchang University.

[64] Sutanu, B., Satish, N., & Ashok, V. (2020). Vision and deep learning-based algorithms to detect and quantify cracks on concrete surfaces from UAV videos. Sensors, 20(21), 6299.

[65] Su, Y., Yang, W., & Li, J. (2021). Intelligent fracture recognition method for coal-rock borehole images based on YOLO v3. Coal Mine Safety, 52(4), 156–161.

[66] Lemy, F., & Hadjigeorgiou, J. (2003). Discontinuity trace map construction using photographs of rock exposures. International Journal of Rock Mechanics and Mining Sciences, 40(6), 903–917.

[67] Xu, X., Zhao, M., Shi, P., et al. (2022). Crack detection and comparison study based on faster R-CNN and mask R-CNN. Sensors, 22(3), 1215.

[68] V. M., L. U., & Y. A. (2018). Automated Road Crack Detection Using Deep Convolutional Neural Networks. In 2018 IEEE International Conference on Big Data (Big Data).

[69] Zhang, G., Xia, Y., Zhao, Y., et al. (2024). Tight carbonate reservoir evaluation case study based on neural network assisted fracture identification and analytic hierarchy process. Journal of Petroleum Exploration and Production Technology, 14(7), 1825–1842.