Proceedings of the 7th International Young Geotechnical Engineers Conference – Scott (Ed.)

Soil desiccation crack recognition and quantification using deep learning

Reconnaissance et quantification des fissures de dessiccation du sol à l'aide de l'apprentissage en profondeur

Jin-Jian Xu & Chao-Sheng Tang & Qing Cheng
School of Earth Sciences and Engineering, Nanjing University, China, Xujianjian@smail.nju.edu.cn

Hao Zhang
College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, China

ABSTRACT: Desiccation cracking of soils is of great concern with the advent of global climate change. Accurately obtaining soil crack networks is a significant foundation and premise to study the formation mechanism of shrinkage and desiccation cracking. The current study establishes and compares the performance of a new soil cracks recognition method based on U-Net Convolutional neural network (CNN) architecture with a traditional method for automatic recognition of soil desiccation cracks. A new loss function combing both BCE loss and dice loss is used during training stage to fit imbalance problem. Subsequently, the U-Net with an encoder based on ResNet and a decoder part is trained from end to end on a subset of 524 labeled crack images with 224×224 pixels for semantic segmentation. Experimental results show that the U-Net architecture achieves a satisfactory prediction performance on test sets. This performance is significantly better than the method using Otsu threshold method employed in traditional crack image processing technique. Moreover, the deep learning can obtain higher accuracy than traditional method (binarization by thresholding) in quantifying surface crack ratio, average crack width, total crack length and crack number.

RÉSUMÉ : La fissuration des sols par dessiccation est une préoccupation majeure avec l'avènement du changement climatique mondial. L'obtention précise de réseaux de fissures dans le sol est une base et une prémisse importantes pour étudier le mécanisme de formation des fissures de retrait et de dessiccation. La présente étude établit et compare les performances d'une nouvelle méthode de reconnaissance des fissures du sol basée sur l'architecture de réseau neuronal convolutif (CNN) U-Net avec une méthode traditionnelle de reconnaissance automatique des fissures de dessiccation du sol. Une nouvelle fonction de perte combinant à la fois la perte de BCE et la perte de dés est utilisée pendant l'étape d'apprentissage pour régler le problème de déséquilibre. Par la suite, le U-Net avec un encodeur basé sur ResNet et une partie décodeur est entraîné de bout en bout sur un sous-ensemble de 524 images de fissures étiquetées avec 224×224 pixels pour la segmentation sémantique. Les résultats expérimentaux montrent que l'architecture U-Net atteint des performances de prédiction satisfaisantes sur des ensembles de test. Cette performance est nettement meilleure que la méthode utilisant la méthode de seuil d'Otsu employée dans la technique traditionnelle de traitement d'image de fissure. De plus, l'apprentissage en profondeur peut obtenir une précision plus élevée que la méthode traditionnelle (binarisation par seuillage) pour quantifier le rapport de fissure de surface, la largeur moyenne de fissure, la longueur totale de fissure et le nombre de fissures.

KEYWORDS: desiccation cracking; soil cracks recognition; deep learning; semantic segmentation; crack pattern quantification.

1 INTRODUCTION

The formation of soil desiccation cracks due to moisture evaporation and volumetric shrinkage is a common natural phenomenon in drought climate. Especially for clayey soils, the cracking phenomenon is more common and typical because of the obvious volumetric shrinkage (Tang et al., 2021). The generation of cracks will not only weaken the engineering properties of the soil, but also greatly change its hydraulic properties and weathering degrees, which cause a lot of geotechnical, geological and environmental engineering problems, such as ground subsidence, land slide, embankment/dike collapse, engineering barrier dysfunction (Albrecht and Benson et al., 2001; Vahedifard et al., 2016; Stirling et al., 2020; Xu et al., 2020). Therefore, accurate acquisition of soil crack networks is not only the basis for obtaining the relevant geometrical parameters of the crack networks, but also an important foundation and premise for further study about the formation mechanism of desiccation cracking and the establishment of theoretical models (Wei et al., 2016, Cheng et al., 2021).

The development of computer science and digital image processing technology provided a feasible way to achieve efficient crack networks acquisition. The digital image processing technology is adopted to perform grayscale, binarization and denoising preprocessing operations on the obtained crack images to acquire crack networks (Tang et al., 2008; Cheng et al., 2020). However, this method requires a high quality of the original image and can be significantly influenced by some disruptive factors like uneven lighting environment and uneven surface of the soil or noise. Thus, it is necessary to put forward a new method for crack recognition which is accurate and convenient especially for images with distractions.

Currently, methods based on deep learning have achieved excellent performance on various computer vision tasks, such as image denoising, object detection, image classification and image segmentation (LeCun et al., 2015; Zhang et al., 2019). In the field of geology, deep neural networks can not only be used to accurately identify the groundwater contaminant source and efficiently classify mineral grains, but also be employed to quickly confirm the geologic time of paleontological fossils and radiolarians (Keceli et al., 2017; Mo et al., 2019). It is expected that the deep learning can also be adopted to recognize soil crack networks under bad photographing condition.

The aim of this study is to build an effective U-Net CNN architecture-based method for soil cracks recognition through deep learning. The performance of the proposed method is tested and verified by quantitatively analyze the images obtained from deep learning and traditional method.

2 TRADITIONAL CRACK RECOGNITION METHOD

Traditional crack recognition procedures based on digital image processing technique generally require two steps: (1) Gray-scale. The color value in RGB image (Figure 1 (a)) is affected by the light and other factors, which has an important influence on the later process and analysis. Thus, it is a common operation to get gray-scale image (Figure 1 (b)) in digital image processing; (2) Binarization. Thresholding is used to create binary image due to the distinguishability between the crack and the soil clod in color contrast or pixel value. After the binarizing, the crack and the soil clod are therefore distinguished and illustrated by black and white pixel, respectively, as presented in Figure 1 (c). These operations above can be done by image processing software CIAS, MATLAB, Photoshop, ImageJ and so on (Tang et al., 2008; Liu et al., 2013). However, selecting an appropriate threshold to reflect the true information of the image is always a huge challenge for this subject. In most cases, the selection of the threshold is highly susceptible to subjective judgments, which causes a large error in the subsequent quantification of the crack.

3 DEEP LEARNING METHOD

As shown in Figure 2, an automated semantic segmentation model based on U-Net is utilized to segment soil cracks and clods instead of using thresholding (Ronneberger et al., 2015). The sizes of the feature maps generated during the network forward pass are labelled on the left side, e.g. 224×224 and 14×14. The number under or below the feature maps is the channel. The left part of the pipeline of U-Net is called encoding path to transfer the input image into compact representative features. On the contrary, the right part, called decoding path, binary classification in this task, where means 1 for the crack and 0 for the background. The encoding path and decoding path are constructed by the same number of convolutional layers to a symmetrical architecture. The operation of max-pooling in encoding path is used to cut down the computation cost and enlarge the receptive field, while the up sampling in decoding path is utilized to recover the features lost during encoding. The last 1×1 convolutional layer is used to get the final result. To achieve the combination of both low-level details (from shallow layer) and high-level semantic information (from high layer), U-Net copied low-level representations to the homologous high levels through concatenation ("copy" operation in Figure 2). The left part of the pipeline of U-Net is called encoding path to transfer the input image into compact representations. To train the U-Net for crack image segmentation, the following steps are carried as: data collection, experiment implementation including loss function and post-processing and result.

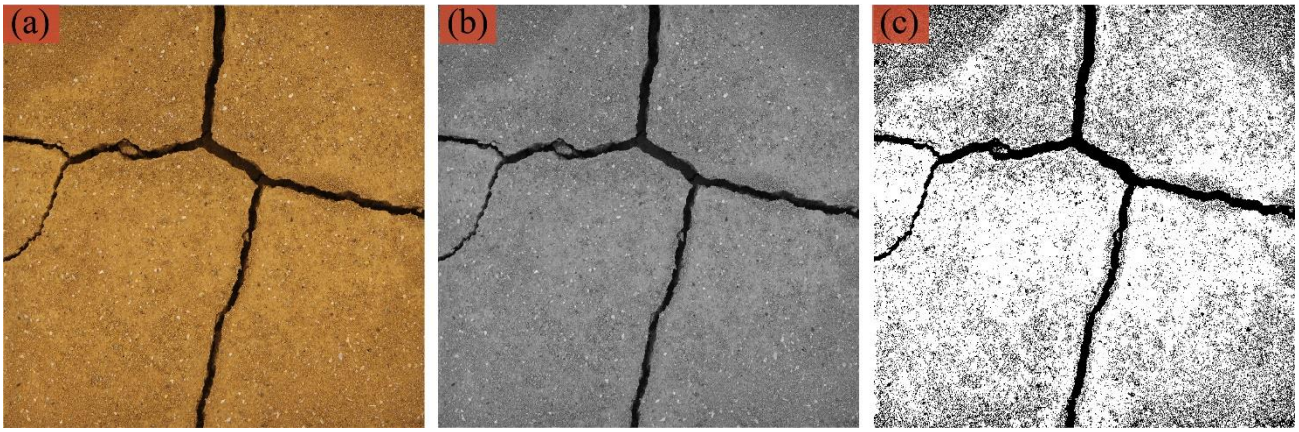


Figure 1. Traditional crack recognition procedures based on digital image processing technique. (a) Original image; (b) Gray-scale image; (c) Binary images obtained by Otsu segmentation method.

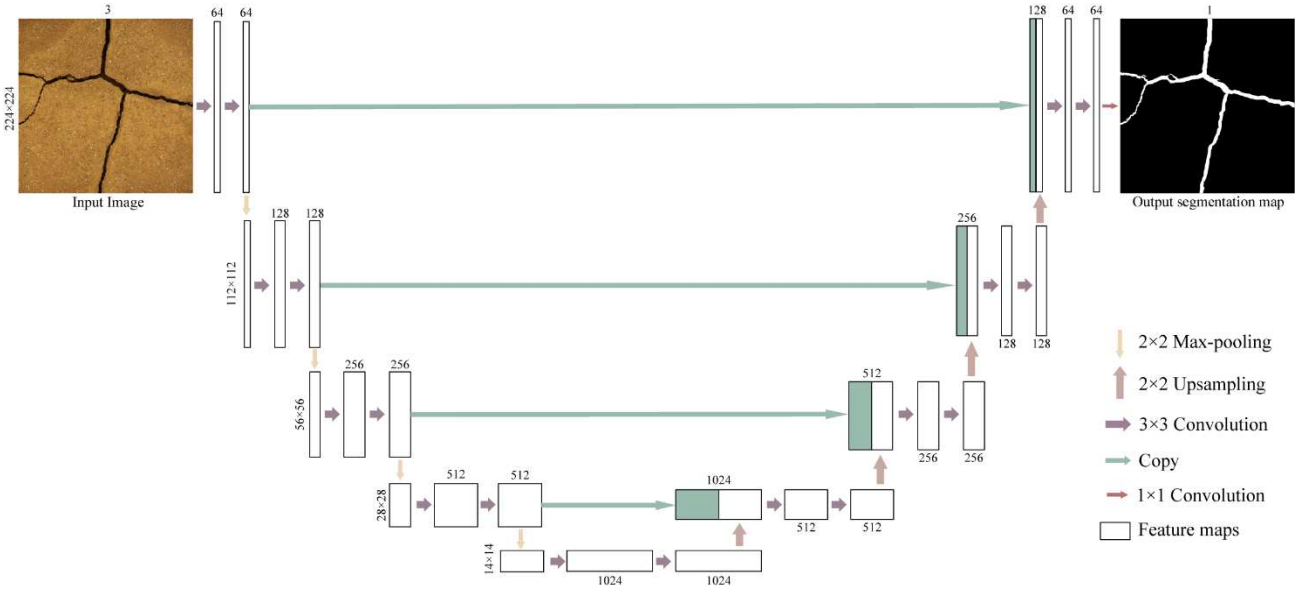


Figure 2. The architecture of U-Net for soil crack image segmentation.

4 DATA COLLECTION

The model in this paper does not need much more training images (dozens of images are enough), we have collected several images under different environments and sliced these images into small patches for training and testing. To segment crack images obtained from different sources (e.g. laboratory and field), different kinds of samples are collected as illustrated in Table 1. In total, 35 images are collected in this study, where 25 images are used for training and the left 10 images are utilized to test the performance of the U-Net. All these images are labelled with binary mask manually as ground-truth. It should be noted that we annotated a crack image (Figure 3 (a)) with Labelme (a visual image annotation tool written in Python). In Figure 3 (b)-(c), crack area is annotated with polygonal line and the result which is named as label or ground truth is shown in Figure 3 (d), where the white area is represented by 1 and background is 0. For the training and validation set, in order to fit the training process and GPU memory constraint, 26 images for training are random cropped or sliced into patches, e.g. 224×224 resolution. A total 524 patches are generated from the original images and 5% of these patches are used for validation during the training process and the others are used for testing. In terms of the test set, 10 images which are sampled from different sources are employed. These images are resized into a fixed resolution, e.g. 1120×1120, to fit the input constraint of the U-Net and get clearer results.

Table 1. Data collection in different function (training and testing) and sources (laboratory and field)

Function	Image source	Sample ID and resolution
Training	Laboratory	1-6 3200×3200 7-10: 3300×3300
		11-15: 4550×3050 16-17: 2760×4220
		18-19: 1300×1300 20-22: 3500×3500
		23-24: 2240×2240
	Field	25: 3236×3226
Testing	Laboratory	26-28 3200×3200 29-30: 1840×1840
		31-34: 2240×2240
	Field	35: 3741×3598

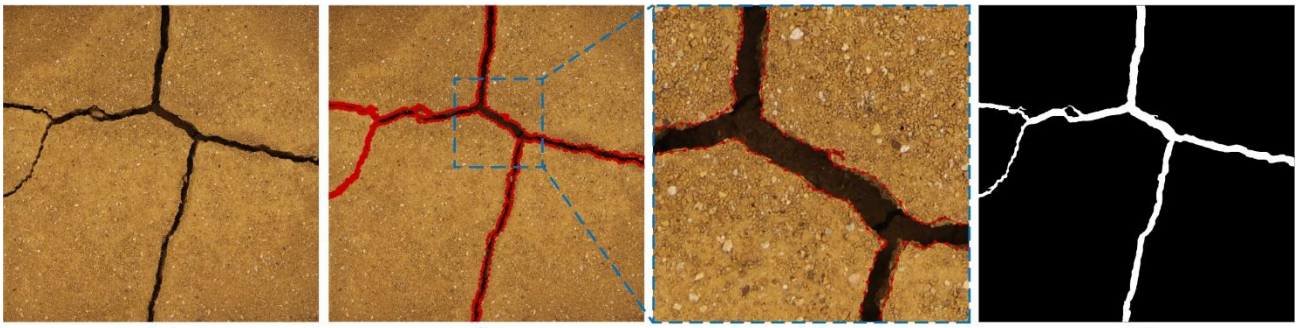


Figure 3. The procedures of annotating crack image. (a) Input image; (b) Image annotated with polygonal line; (c) Detail of annotated image; (d) Label.

5 IMPLEMENTATION

Implementation consists of loss function and post-processing and result analysis. The whole model is initialized randomly and the Stochastic Gradient Descent (SGD) is utilized to optimize the model with initialized learning rate of 0.01 and reduced by a factor of 0.1 at the epoch of 120 and 250 respectively. The model is trained by a batch size of 16 with a Nvidia GTX 1080 GPU with 8 GB memory for 300 epochs. Theoretically, U-Net can take arbitrary size of input images. However, bigger size needs more GPU memory to store the feature maps. To fit the input constraint of the model and match the GPU memory, the original images are randomly cropped and resized into a fixed-sized training image (e.g. 224×224). Loss function measures the "cost" during the training process which means the difference between the true manual annotations and the network predictions. The function is then utilized to update the weights of the network to boost the performance on training set. Specifically, Binary Cross Entropy (BCE) loss is utilized as Eq. 1:

$$L(g,p) = -\sum_i g_i \log p_i \tag{1}$$

where g_i stands for the true label designed manually; p_i for the predicted label; i is the category, specifically $\{1,2\}$ in our task. The dice coefficient can be generalized in 2D segmentation, soft dice loss can be summarized as Eq. 2:

$$D_s(g,p) = 1 - \frac{2\sum_i p_i \times g_i}{\sum_i p_i + \sum_i g_i} \tag{2}$$

where $p_i \in [0,1]$ is the value after sigmoid function; $g_i \in [0,1]$ is the ground truth. BCE loss and the dice coefficient are combined as Eq. 3:

$$L(g,p) = \beta \sum_i g_i \log p_i - (1-\beta)\log\left(\frac{2\sum_i p_i \times g_i + smooth}{\sum_i p_i + \sum_i g_i + smooth}\right) \tag{3}$$

where *smooth* is an extreme small number to avoid division by zero, e.g. 10^{-15}; β is set as 0.5 empirically (Zhang and Hong, 2019).

6 VERIFICATIONS OF THE NEW METHOD BASED ON DEEP LEARNING

6.1 *Training process and test results*

As discussed above, BCE loss and dice coefficient are both used during the training for training set and validation set. Figure 4 shows the training process of U-Net. The training loss drops rapidly in the first 50 epochs and gradually reach the equilibrium at around 200 epochs, at the meanwhile, the dice of validation set gets the contrary results. The model reaches its best performance on test set around the epoch of 230.

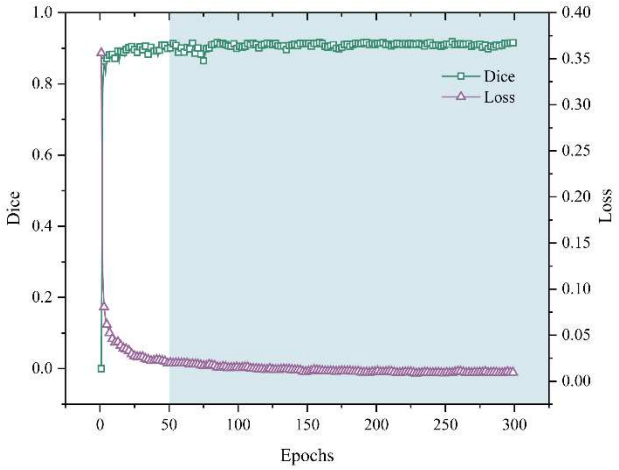
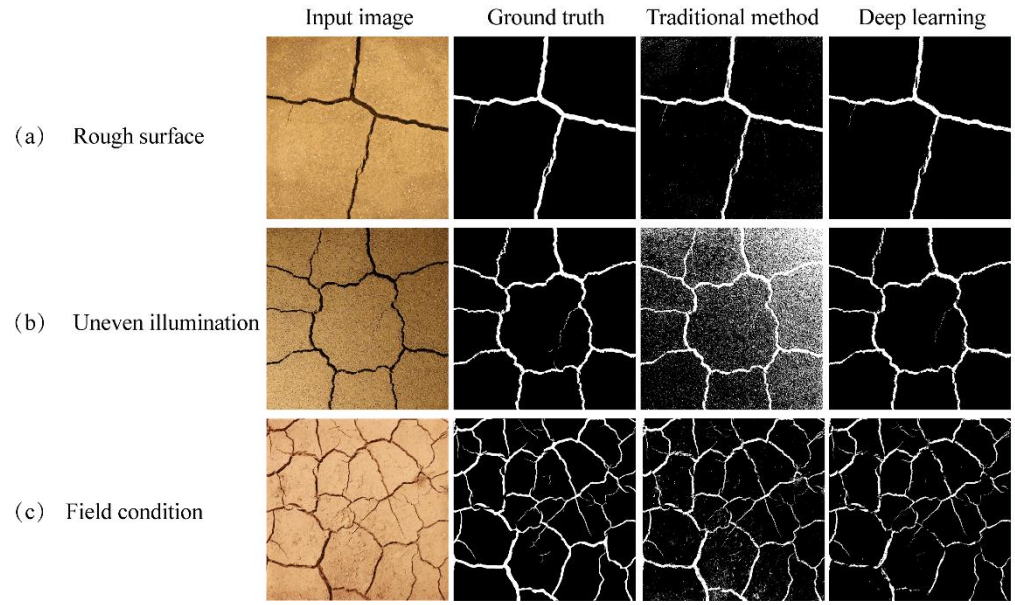


Figure 4. (a) Loss curve and (b) Dice curve during training.

For the images treated using the traditional method, the threshold is determined using Otsu threshold method (Otsu, 1979), as discussed in previous section. Based on the manually labelled ground truth, the results of deep learning method and traditional method on test set are compared. Three metrics of the deep learning and traditional method are counted including overall accuracy for the classification, precision, recall and dice. The results of traditional method in precision, recall and dice scores are 81.74%, 66.56%, 68.48%, respectively, while those of deep learning are 94.38%, 74.43%, 81.13%, respectively. Without bells and whistle, deep learning outperforms with traditional method in a large margin at all evaluation criterions.

To further demonstrate the effectiveness of deep learning, we visualized the output of deep learning compared to traditional method and ground truth. Three typical crack images were selected as examples. When the soil surface is relatively rough (Figure 5 (a)), the binarized image obtained by traditional method results in many miscellaneous spots, indicating that some spots on soil surface are mistaken as cracks, whereas deep learning is a good way to avoid this problem. Figure 5 (b) presents soil surface is unevenly illuminated when photographing, which seriously affects the threshold selection of binarization. However, deep learning can accurately separate soil cracks and clod area under this uneven illumination condition. Figure 5 (c) presents an image obtained in field usually with complex surface condition. As can be seen, the presence of soil aggregates and their shadows on surface can result in numerous noise spots by traditional method but not by deep learning. Additionally, some of the vegetation and branches on soil surface were recognized as cracks by traditional method due to their similar color, but these were not happened by deep learning.



Accuracy metrics (%)

Metrics	TM (a)	DL (a)
Precision	86.09	98.52
Recall	87.16	83.54
Dice	86.62	90.42

Metrics	TM (b)	DL (b)
Precision	17.70	97.05
Recall	95.07	89.36
Dice	29.84	93.05

Metrics	TM (c)	DL (c)
Precision	80.76	97.72
Recall	79.06	73.08
Dice	79.90	79.80

Figure. 5. Example results on the test set of typical soil cracks data set including input image, recognized crack pattern by ground truth, traditional method, and deep learning.

6.2 *Image quantitative analysis*

To quantify soil crack patterns, a professional crack image analysis system CIAS (available at www.climate-engeo.com) developed by Tang et al. (2008) is adopted. In CIAS, a set of geometric parameters for quantifying crack patterns was defined.

In this investigation, the following four geometric parameters are employed: surface crack ratio, average crack width, total crack length, crack number. More details about the four geometrical parameters were introduced by Tang et al. (2008). A total of three typical crack images recognized by ground truth, traditional method and deep learning were quantified by CIAS. To quantify the errors of traditional method and deep learning as compared with ground truth, the following error indicator is defined as Eq. 4:

$$E_r = \frac{\left|G_{t/d} - G_g\right|}{G_g} \times 100\% \qquad (4)$$

where $G_{t/d}$ and G_g are the values geometric parameters of the crack pattern obtained by traditional method/deep learning, and

ground truth; E_r is the quantification error of traditional method or deep learning, respectively.

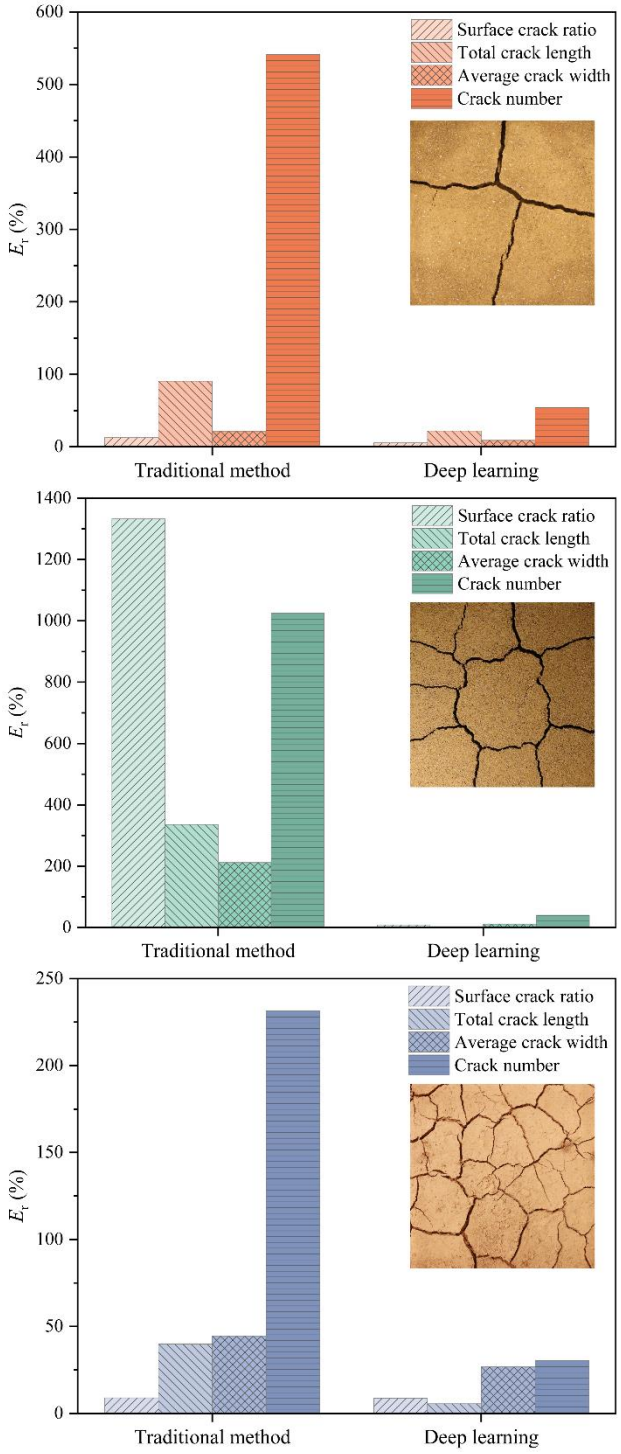


Figure 6. The quantification error of geometric parameters of three typical crack images (Corresponding to three types of soil crack image collection conditions include: rough surface, uneven illumination and field condition) recognized by traditional method and deep learning.

Figure 6 shows the Er of each geometric parameter of the three images recognized by traditional method or deep learning. Generally, deep learning presents higher accuracy than traditional method due to the relatively low E_r of all the four geometric parameters. This is mainly because the concave spots and small soil particles (Figure 5 (a)), relatively darker areas (Figure 5 (b)), soil aggregates (Figure 5 (c)), vegetation and branches (Figure 5 (c)) on soil surface were wrongly recognized as cracks, thereby significantly increasing the quantification errors.

Based on the above analysis, firstly, it is found that the external photographing condition and the internal impurities on the soil surface will seriously affect the quality of image, leading to a large error in cracks recognition by traditional method which is based on color contrast of objectives. However, recognition of soil cracks by deep learning is based on feature extraction of the edge of the crack networks, thereby avoiding both the external and internal influencing factors on image quality.

7 SUMMARY AND CONCLUSIONS

In the present study, a new deep learning-based automatic soil cracks recognition method using U-Net architecture is proposed for segmentation on soil desiccation crack images. The backbone of the U-Net encoder is selected as ResNet and a new loss function combing both BCE loss and dice loss is utilized during training stage to fit imbalance problem. Then, the whole encoder-decoder U-Net architecture is trained from end to end on a specific subset of annotated soil crack images of the same dataset. The results showed that all results of the proposed deep learning are better than traditional method in the three metrics. In addition, purely from a methodological perspective, the deep learning-based crack recognition method proposed in this study is not limited to soils. It may also be applicable in other materials with cracks. Once the model was trained well, it can be used in cracks recognition under various situations.

8 ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant No. 41925012, 41902271, 41772280, 41572246), National Key Research and Development Program of China (2019YFC1509902), Natural Science Foundation of Jiangsu Province (Grant No. BK20171228, BK20170394), and the Fundamental Research Funds for the Central Universities.

9 REFERENCES

Albrecht, B.A. and Benson, C.H., 2001. Effect of desiccation on compacted natural clays. *Journal of Geotechnical and Geoenvironmental Engineering* 127(1), 67-75.

Cheng, Q., Tang, C.S., Xu, D., Zeng, H. and Shi, B., 2021. Water infiltration in a cracked soil considering effect of drying-wetting cycles. *Journal of Hydrology* 593, 125640.

Cheng, Q., Tang, C.S., Zeng, H., Zhu, C., An, N. and Shi, B., 2020. Effects of microstructure on desiccation cracking of a compacted soil. *Engineering Geology* 265, 105418.

Keceli, A.S., Kaya, A. and Keceli, S.U., 2017. Classification of radiolarian images with hand-crafted and deep features. *Computers & Geosciences* 109, 67-74.

LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. *Nature* 521(7553), 436-444.

Liu, C., Tang, C.S., Shi, B. and Suo, W.B., 2013. Automatic quantification of crack patterns by image processing. *Computers & Geosciences* 57, 77-80.

Mo, S.X., Zabaras, N., Shi, X.Q. and Wu, J.C., 2019. Deep autoregressive neural networks for high - dimensional inverse problems in groundwater contaminant source identification. *Water Resources Research* 55(5), 3856-3881.

Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9(1), 62-66.

Ronneberger, O., Fischer, P. and Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation *Medical Image Computing and Computer-Assisted Intervention, Pt Iii* 9351, 234-241.

Stirling, R.A., Toll, D.G., Glendinning, S., Helm, P.R., Yildiz, A., Hughes, P.N. and Asquith, J.D., 2020. Weather-driven deterioration

processes affecting the performance of embankment slopes. *Géotechnique*, 1-13.
Tang, C.S., Shi, B., Liu, C., Zhao, L.Z. and Wang, B.J., 2008. Influencing factors of geometrical structure of surface shrinkage cracks in clayey soils. *Engineering Geology* 101(3-4), 204-217.
Tang, C.S., Zhu, C., Cheng, Q., Zeng, H., Xu, J.J., Tian, B.G. and Shi, B., 2021. Desiccation cracking of soils: A review of investigation approaches, underlying mechanisms, and influencing factors. *Earth-Science Reviews* 216, 103586.
Vahedifard, F., Robinson, J.D. and AghaKouchak, A., 2016. Can protracted drought undermine the structural integrity of California's earthen levees? *Journal of Geotechnical and Geoenvironmental Engineering* 142(6), 02516001.
Wei, X., Hattab, M., Bompard, P. and Fleureau, J.M., 2016. Highlighting some mechanisms of crack formation and propagation in clays on drying path. *Geotechnique* 66(4), 287-300.
Xu, J.J., Zhang, H., Tang, C.S., Cheng, Q., Liu, B. and Shi, B., 2020. Automatic soil desiccation crack recognition using deep learning. *Géotechnique*, 1-53.
Zhang, H. and Hong, X.G., 2019. Recent progresses on object detection: a brief review. *Multimedia Tools and Applications* 78(19), 27809-27847.
Zhang, H., Zhou, F., Wu, Q., Wu, W. and Hu, R.Q., 2021. A novel automatic modulation classification scheme based on multi-scale networks. arXiv preprint arXiv: 2105. 15037.