Contextual Information Enhancement Network for Crack Segmentation Methods

Subjects: Computer Science, Artificial Intelligence

Contributor: Lili Zhang , Yang Liao , Gaoxu Wang , Jun Chen , Huibin Wang

Convolutional neural-network-based crack segmentation methods have performed excellently. However, existing crack segmentation methods still suffer from background noise interference, such as dirt patches and pitting, as well as the imprecise segmentation of fine-grained spatial structures.

convolutional neural network crack segmentation skip connections

1. Introduction

In recent years, as the urbanization rate of countries around the world increases, a large number of infrastructures, such as bridges, tunnels, and dams, are constructed, providing a solid guarantee for economic development and livelihood security. However, the supervision and maintenance of these facilities has also brought us new challenges. These infrastructures commonly use concrete as the construction material and the surface crack is one of the main symptoms of their damage and destruction ^{[1][2]}. Without timely maintenance, cracks will have a significant impact on the service life and safety of those infrastructures. Other facilities, such as asphalt roads, also need to be checked regularly to ensure that surface cracks can be maintained and repaired in a timely way. Therefore, the automatic identification of surface cracks from optical images of various scenes is of great research importance ^[3]. Due to the development of computer science and image processing technology, it is now possible to partially automate the process of surface crack inspection. However, it is still a difficult task to accurately separate the cracks from the complex image background, as there may be dirt patches, oil stains, pitting, or other noise interferences.

Most of the early crack segmentation techniques rely on traditional digital image processing methods, which often involve multiple pre-processing processes, such as morphological filtering ^{[4][5]}, fuzzy theory methods ^{[6][7]}, and wavelet transform ^{[8][9]}, as well as various crack segmentation methods, such as methods based on the threshold algorithm ^{[10][11]} or the edge detection algorithm ^{[12][13]}. Traditional digital image processing methods are sensitive to interference from external factors, such as light changes and shadow occlusion, making them unusable in complex scenes. Meanwhile, digital image processing methods require manually designed feature operators, which are more difficult and less efficient to implement.

Recently, the application of deep-learning-based convolutional neural networks (CNNs) in the field of computer vision has developed rapidly and has even surpassed human performance in a variety of tasks, such as image classification [14][15], object detection [16][17], and semantic segmentation [18][19]. Compared with traditional digital

image processing methods, CNNs are characterized by their high level of automation and strong feature extraction capability, as CNNs do not rely on manually designed feature operators. In terms of crack recognition applications, some studies localize cracks in images by classification ^{[20][21]} or object detection ^{[22][23]} methods. However, these methods cannot obtain detailed information about the cracks, making them less optimal. Segmentation-based crack recognition methods annotate cracks in images at the pixel level, providing a better level of detailed information, as part of the current mainstream research direction ^[24].

Due to the special morphological characteristics of cracks, the crack segmentation task faces two challenges: the accurate segmentation of fine-grained spatial structures and the ability to adapt to complex background environments. The former requires that the multi-level feature information extracted by the feature extraction network can be fully utilized, while the latter requires the network to possess accurate context awareness. It is shown in [25] that feature maps in different levels explore distinctive information, with shallow feature maps possessing fine spatial information and deep feature maps capturing rich semantic information, while the conversion process from shallow to deep feature maps leads to a loss of detailed spatial information. To recover the lost spatial information in the decoder network, SegNet ^[26] assists the decoder in up-sampling by means of maximum pooling indexing, while U-Net ^[19] feeds the shallow feature information generated in the encoder directly to the decoder network by means of a skip connection. Both of them are based on the symmetric encoder-decoder architecture, and there are some recent studies of crack segmentation which also use similar architectures [27][28]. However, it is demonstrated in ^[29] that the multi-scale feature information in the encoder cannot be fully utilized by delivering information between the same layers of the encoder and decoder networks. Meanwhile, due to the limitation of the empirical receptive field size [30], the plain convolutional neural networks cannot provide sufficient contextual feature information, which is necessary to adapt to complex scenarios. To address these problems, this research proposes a multi-scale contextual information enhancement network (MCIE-Net), which redesigns the connection structure between the encoder and the decoder of the U-Net to capture multi-scale feature information and enhance the decoder's ability to restore fine-grained the spatial structure of cracks; meanwhile, a contextual feature enhancement module, which consists of the pyramid pooling network and channel attention mechanism, is designed to enhance the context awareness of the network.

2. Traditional Image Processing Methods

Most of the traditional crack segmentation methods mainly rely on the color difference between cracks and background or the edge features of cracks to extract cracks from images ^[31]. Kirschke et al. ^[10] used a histogrambased threshold segmentation method to extract road cracks. Cheng et al. ^[11] proposed a threshold segmentation algorithm with reduced sample space and interpolation to optimize the efficiency of crack segmentation. Katakam ^[32] used the method of chunking the image first and then threshold-handling each sub-block separately to improve the accuracy of crack segmentation. Oliveira and Correia ^[33] firstly pre-processed the images using morphological filters and then used dynamic threshold segmentation to segment the cracks. Zhang et al. ^[34] integrated spatial clustering, threshold segmentation, and region-growing methods to obtain a coarse-to-fine segmentation of cracks. In ^{[9][35]}, wavelet transform was used for crack segmentation, while in ^[12], the Canny operator was used to detect the contours of cracks. In addition, there are some studies that identify cracks with the help of machine learning methods. Considering the connectivity of cracks, Fernandes et al. ^[36] used a graph-based (graph-based) approach to extract crack features, and then support vector machines were used to classify the features to obtain a classification of crack types. In ^[37], crack structure features were extracted and learned from annotation data, and, based on this, a crack recognition framework was generated using random structure forest to achieve pixel-level crack segmentation.

3. Deep-Learning-Based Methods

Deep-learning-based crack segmentation methods mostly use semantic segmentation models. In 2015, Long et al. [18] achieved the first end-to-end segmentation of natural images using fully convolutional neural (FCN) networks, which have thus become the most classical network model in the field of semantic segmentation. Liu et al. [25] used a FCN backbone and a deeply supervised approach to upscale and fuse the feature maps from all levels of the backbone, and then applied a guided filter to fuse all feature maps as well as the side outputs to create a segmentation output. Ren et al. [38] used dilated convolution with a different dilation rate in the last four layers of the FCN to expand the receptive field without changing the feature map scale, and used skip connections to deliver shallow feature information, assisting the decoder in generating segmentation results. However, the methods based on FCN networks still suffer from information loss when up-sampling low-resolution feature maps generated in the deep layer of the feature extraction network. To solve the problem, symmetric encoder-decoder-based network structures, such as SegNet ^[26] and U-Net ^[19], have been proposed. In particular, U-Net has had a profound impact on many subsequent studies due to its pioneering concept and excellent performance, and a series of semantic segmentation models such as UNet++ [39] and Unet 3+ [29] have been derived on its basis. Since the detailed spatial information of cracks can be more effectively restored, many recent studies of crack segmentation are based on the SegNet and U-Net structures. Ran et al. [40] introduced a spatial attention mechanism and a channel attention mechanism in SegNet and used spatial pyramidal pooling to capture crack features from different scales. Zou et al. [3] pair-wisely fused the feature maps generated in the encoder and decoder network at the same scale, and generated segmentation results by extracting features from the fused feature maps at multiple scales using a multi-scale fusion component. Lau et al. [27] replaced the plain convolutional neural network of the encoder of U-Net with a residual network and added spatial and channel compression excitation modules to the decoder. Based on U-Net, Han et al. ^[28] designed a skip-level round-trip sampling structure, in which the deep feature maps of the encoder network were up-sampled and aggregated with some shallow feature maps, and then down-sampled and fed into the decoder network. These up- and downsampling actions enhanced the network's memory of transmitting low-level features in the shallow layer, helping the network to pay attention to the distinction between the cracks and the background. Zhao et al. [30] proposed PSPNet, which applies special pyramid pooling to the semantic segmentation task and extracts multi-scale contextual information. Some other studies also explored spatial pyramid pooling, such as the DeepLab series [41] [42][43], although the difference is that DeepLabs use a dilated convolution rather than pooling to obtain contextual information at multiple scales. Sun et al. [44] adopted and enhanced DeepLabv3+, in which a multi-attention module was introduced to dynamically adjust the weights of different feature maps for pavement crack image

segmentation. Yuan et al. ^[45] proposed OCR-Net, which uses object contextual feature representation for contextual information extraction based on object regions, thus explicitly enhancing object information and achieving good results on several mainstream semantic segmentation databases. Zhou et al. ^[46] explored an exemplar-based regime which provides a nonparametric segmentation framework based on non-learnable prototypes, where several typical points in the embedding space are selected for class prototypical representation, and distance to the prototypes determines how a pixel sample is classified. For deep learning models, there has been a bottleneck over the years to acquire sufficient ground-truth supervision, especially for segmentation tasks that require pixel-level annotations. Zhou et al. ^[47] proposed a group-wise learning framework for weakly supervised semantic segmentation that explicitly encodes semantic dependencies in a group of images to discover a rich semantic context for estimating more reliable pseudo ground truths, which are subsequently employed to train more effective segmentation models. König et al. ^[48] proposed a weakly supervised approach for crack segmentation that leverages a CNN classifier to create a rough crack localization map. The map was fused with a thresholding-based approach to segmentation.

References

- 1. Salam, M.; Mathavan, S.; Kamal, K.; Rahman, M. Pavement crack detection using the Gabor filter. In Proceedings of the 16th international IEEE conference on intelligent transportation systems, The Hague, Netherlands, 6–9 October 2013.
- Eisenbach, M.; Stricker, R.; Seichter, D.; Amende, K.; Debes, K.; Sesselmann, M.; Ebersbach, D.; Stoeckert, U.; Gross, H.M. How to get pavement distress detection ready for deep learning? A systematic approach. In Proceedings of the 2017 International Joint Conference on Neural Networks, Anchorage, Alaska, 14–19 May 2017.
- 3. Zou, Q.; Zhang, Z.; Li, Q.; Qi, X.; Wang, Q.; Wang, S. Deepcrack: Learning hierarchical convolutional features for crack detection. IEEE Trans. Image Process. 2019, 28, 1498–1512.
- 4. Li, G. Improved pavement distress detection based on contourlet transform and multi-direction morphological structuring elements. Adv. Mater. Res. 2012, 466, 371–375.
- Su, Z.; Guo, Y. Algorithm on Contourlet Domain in Detection of Road Cracks for Pavement Images. In Proceedings of the 2010 Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science, Hong Kong, China, 10–12 August 2010.
- 6. Das, H.C.; Parhi, D.R. Detection of the Crack in Cantilever Structures Using Fuzzy Gaussian Inference Technique. AIAA J. 2009, 47, 105–115.
- 7. Zhang, D.; Qu, S.; He, L.; Shi, S. Automatic ridgelet image enhancement algorithm for road crack image based on fuzzy entropy and fuzzy divergence. Opt. Lasers Eng. 2009, 47, 1216–1225.

- Zuo, Y.; Wang, G.; Zuo, C. Wavelet Packet Denoising for Pavement Surface Cracks Detection. In Proceedings of the 2008 International Conference on Computational Intelligence and Security, Suzhou, China, 13–17 December 2008.
- 9. Zhou, J.; Huang, P.; Chiang, F.P. Wavelet-Based Pavement Distress Classification. Transp. Res. Rec. J. Transp. Res. Board 2005, 1940, 89–98.
- 10. Kirschke, K.R.; Velinsky, S.A. Histogram-Based Approach for Automated Pavement-Crack Sensing. J. Transp. Eng. 1992, 118, 700–710.
- 11. Cheng, H.D.; Shi, X.J.; Glazier, C. Real-Time Image Thresholding Based on Sample Space Reduction and Interpolation Approach. J. Comput. Civ. Eng. 2003, 17, 264–272.
- Zhao, H.; Qin, G.; Wang, X. Improvement of canny algorithm based on pavement edge detection. In Proceedings of the 2010 3rd International Congress on Image and Signal Processing, Yantai, China, 16–18 October 2010.
- 13. Abdel-Qader, I.; Abudayyeh, O.; Kelly, M.E. Analysis of Edge-Detection Techniques for Crack Identification in Bridges. J. Comput. Civ. Eng. 2003, 17, 255–263.
- Jmour, N.; Zayen, S.; Abdelkrim, A. Convolutional neural networks for image classification. In Proceedings of the 2018 International Conference on Advanced Systems and Electric Technologies, Hammamet, Tunisia, 25 March 2018; pp. 397–402.
- 15. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In Proceedings of the 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015.
- 16. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39, 1137–1149.
- 17. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. arXiv 2018, arXiv:1804.02767.
- Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 8–10 June 2015.
- Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015.
- 20. Cha, Y.J.; Choi, W.; Büyüköztürk, O. Deep learning-based crack damage detection using convolutional neural networks. Comput. Aided Civ. Infrastruct. Eng. 2017, 32, 361–378.
- Pauly, L.; Hogg, D.; Fuentes, R.; Peel, H. Deeper networks for pavement crack detection. In Proceedings of the 34th The International Association for Automation and Robotics in Construction, Taipei, Taiwan, 28 June 2017.

- 22. Tang, J.; Mao, Y.; Wang, J.; Wang, L. Multi-task Enhanced Dam Crack Image Detection Based on Faster R-CNN. In Proceedings of the 2019 IEEE 4th International Conference on Image, Vision and Computing, Xiamen, China, 5–7 July 2019.
- 23. Suh, G.; Cha, Y.J. Deep faster R-CNN-based automated detection and localization of multiple types of damage. In Proceedings of the Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2018, Denver, CO, USA, 5–8 March 2018.
- 24. König, J.; Jenkins, M.D.; Mannion, M.; Barrie, P.; Morison, G. Optimized deep encoder-decoder methods for crack segmentation. Digit. Signal Process. 2021, 108, 102907.
- 25. Liu, Y.; Yao, J.; Lu, X.; Xie, R.; Li, L. DeepCrack: A deep hierarchical feature learning architecture for crack segmentation. Neurocomputing 2019, 338, 139–153.
- Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39, 2481– 2495.
- 27. Lau, S.L.H.; Chong, E.K.P.; Yang, X.; Wang, X. Automated Pavement Crack Segmentation Using U-Net-Based Convolutional Neural Network. IEEE Access 2020, 8, 114892–114899.
- 28. Han, C.; Ma, T.; Huyan, J.; Huang, X.; Zhang, Y. CrackW-Net: A novel pavement crack image segmentation convolutional neural network. IEEE Trans. Intell. Transp. Syst. 2021.
- 29. Huang, H.; Lin, L.; Tong, R.; Hu, H.; Zhang, Q.; Iwamoto, Y.; Han, X.; Chen, Y.W.; Wu, J. Unet 3+: A full-scale connected unet for medical image segmentation. In Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing, Barcelona, Spain, 4–8 May 2020.
- 30. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid scene parsing network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 22–25 July 2017.
- 31. Chambon, S.; Moliard, J.M. Automatic Road Pavement Assessment with Image Processing: Review and Comparison. Int. J. Geophys. 2011, 2011, 989354.
- 32. Katakam, N. Pavement Crack Detection System Through Localized Thresholding. Doctoral Dissertation, University of Toledo, Toledo, OH, USA, 2009.
- Oliveira, H.; Correia, P.L. Automatic road crack segmentation using entropy and image dynamic thresholding. In Proceedings of the 2009 17th European Signal Processing Conference, Glasgow, Scotland, 25 August 2009.
- 34. Zhang, D.; Li, Q.; Chen, Y.; Cao, M.; He, L.; Zhang, B. An efficient and reliable coarse-to-fine approach for asphalt pavement crack detection. Image Vis. Comput. 2017, 57, 130–146.

- 35. Wang, K.C.P.; Li, Q.; Gong, W. Wavelet-Based Pavement Distress Image Edge Detection with À Trous Algorithm. Transp. Res. Rec. J. Transp. Res. Board 2007, 2024, 73–81.
- Fernandes, K.; Ciobanu, L. Pavement pathologies classification using graph-based features. In Proceedings of the 2014 IEEE International Conference on Image Processing, Paris, France, 27– 30 October 2014.
- 37. Shi, Y.; Cui, L.; Qi, Z.; Meng, F.; Chen, Z. Automatic Road Crack Detection Using Random Structured Forests. IEEE Trans. Intell. Transp. Syst. 2016, 17, 3434–3445.
- Ren, Y.; Huang, J.; Hong, Z.; Lu, W.; Yin, J.; Zou, L.; Shen, X. Image-based concrete crack detection in tunnels using deep fully convolutional networks. Constr. Build. Mater. 2020, 234, 117367.
- Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. IEEE Trans. Med. Imaging 2020, 39, 1856– 1867.
- 40. Ran, R.; Xu, X.; Qiu, S.; Cui, X.; Wu, F. Crack-SegNet: Surface Crack Detection in Complex Background Using Encoder-Decoder Architecture. In Proceedings of the 2021 4th International Conference on Sensors, Signal and Image Processing, Nanjing, China, 15–17 October 2021.
- 41. Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. arXiv 2017, arXiv:1706.05587.
- Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 40, 834–848.
- Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. Lect. Notes Comput. Sci. 2018, 833– 851.
- 44. Sun, X.; Xie, Y.; Jiang, L.; Cao, Y.; Liu, B. DMA-Net: DeepLab with Multi-Scale Attention for Pavement Crack Segmentation. IEEE Trans. Intell. Transp. Syst. 2022, 23, 18392–18403.
- 45. Yuan, Y.; Chen, X.; Wang, J. Object-contextual representations for semantic segmentation. In Proceedings of the 16th European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020.
- 46. Zhou, T.; Wang, W.; Konukoglu, E.; Goo, L.V. Rethinking Semantic Segmentation: A Prototype View. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 19–24 June 2022.
- 47. Zhou, T.; Li, L.; Li, X.; Feng, C.M.; Li, J.; Shao, L. Group-Wise Learning for Weakly Supervised Semantic Segmentation. IEEE Trans. Image Process. 2022, 31, 799–811.

48. König, J.; Jenkins, M.D.; Mannion, M.; Barrie, P.; Morison, G. Weakly-Supervised Surface Crack Segmentation by Generating Pseudo-Labels Using Localization with a Classifier and Thresholding. IEEE Trans. Intell. Transp. Syst. 2022, 1–12.

Retrieved from https://encyclopedia.pub/entry/history/show/82277