coatings

MDPI

Article

Research on Automatic Pavement Crack Recognition Based on the Mask R-CNN Model

Pengcheng Wang [1], Chao Wang [2], Hongwu Liu [3], Ming Liang [4,*], Wenhui Zheng [5,*], Hao Wang [1,6], Shichao Zhu [1], Guoqiang Zhong [2] and Shang Liu [2]

1 Shandong High-Speed Infrastructure Construction Co., Ltd., Jinan 250014, China
2 Shandong Provincial Communications Planning and Design Institute Group Co., Ltd., Jinan 250101, China
3 Shandong High-Speed Group Co., Ltd., Jinan 250014, China
4 School of Qilu Transportation, Shandong University, Jinan 250002, China
5 School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
6 Shandong Hi-Speed Jiwei Expressway Co., Ltd., Jinan 250200, China
* Correspondence: ming.liang@sdu.edu.cn (M.L.); zheng_wenhui@bupt.edu.cn (W.Z.); Tel.: +86-1866-162-8291 (M.L.)

Abstract: Pavement will inevitably be damaged in the process of use; pavement damage detection and assessment are an important part of maintenance management. In order to prevent road diseases, it is necessary to fix the road cracks and implement automatic road crack inspection and monitoring. In this paper, the automatic identification of road cracks is realized by constructing the Mask R-CNN model. The labeled area can be segmented by pixels and positioned at the original data by integrating the image data used for training and the labeled data into a network model. The effect of the training model can be improved by increasing the number of data sets, the pixel of the fracture image, the background of the fracture, and the marking method of the fracture type. The validity and accuracy of the test results were characterized by RPN bounding-box loss, classification loss, mask loss, and total loss.

Keywords: pavement; crack identification; machine learning; mask R-CNN model

check for updates

Citation: Wang, P.; Wang, C.; Liu, H.; Liang, M.; Zheng, W.; Wang, H.; Zhu, S.; Zhong, G.; Liu, S. Research on Automatic Pavement Crack Recognition Based on the Mask R-CNN Model. *Coatings* **2023**, *13*, 430. https://doi.org/10.3390/coatings13020430

Academic Editor: Valeria Vignali

Received: 29 November 2022
Revised: 2 February 2023
Accepted: 10 February 2023
Published: 14 February 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In the service of road operations, the mechanical properties of materials are constantly changing due to the effects of vehicle load, natural factors, and human factors, as well as certain defects in the pavement, which will lead to various types of damage to the pavement as well as the continued decay of the pavement's structure. If the pavement damage is not maintained and repaired on time, the use effect of the road surface as well as the use time of road surface disease will be impacted [1,2]. If not found in time and with proper maintenance and repair, road surface disease as well as traffic transportation will have a great impact on traffic safety and cause a great loss to the economy [3–5].

Pavement will inevitably suffer a variety of damages in the process of use, so damage detection and assessment for pavement is an important part of maintenance management [6]. To better assess the state of the road surface and carry out road surface maintenance management, a timely and effective analysis of the pavement condition is required [7]. But due to the various forms and complex causes of damage, it is difficult to determine the quantitative description of damage conditions. Therefore, how to objectively and scientifically describe the pavement damage condition and how to adopt the detection and evaluation methods to take into account the accuracy of the evaluation results and the simplicity of data collection have become common problems to be solved [8–11].

There are many types of road damage, including cracks, potholes, rutting, looseness, subsidence, etc. In this study, the automatic pavement crack recognition technology based

on digital images is mainly carried out for a single crack. The single crack includes longitudinal cracks and transverse cracks [12]. The longitudinal crack is a single crack that is roughly horizontal to the middle line of the road, while the transverse crack is a single crack that is roughly vertical to the middle line of the road. These two types of cracks are often accompanied by a small number of branches. According to the crack width and fracture edge damage, it can be divided into two grades: mild and severe. A slight crack has a seam width of less than 3 mm, no or few branch joints, and a crack wall that is not scattered or only slightly scattered. The severe crack is wide, the crack wall is scattered and supported, and the main crack is wider than 3 mm. In order to repair the road crack, the first task is the detection of the crack. With the gradual maintenance period of the pavement, the demand for the detection of road cracks is increasing. Artificial detection is still the main method to detect the cracks, but it has some drawbacks, such as being high-risk, subjective, time-consuming, expensive, and sometimes impacting transportation. Because of these problems, there is an urgent need for an efficient, safer, and less costly detection method [13,14]. Using an Unmanned Aerial Vehicle (UAV) for crack detection has the advantages of fast, efficient, permanent certificate retention, economy, and environmental protection. The UAV is lightweight and easy to operate, so that it can carry out external detection on the road through the high-resolution camera mounted on the UAV without affecting the traffic and can quickly obtain high-definition data [15].

Many researchers have focused on this topic. Premachandra et al. [16] conducted variance analysis on the pixels of the collected crack image data and took the highway area containing pavement cracks as the crack image data. The discriminant analysis method was used to extract cracks, and it was found that the algorithm is feasible for pavement crack detection. Oliveira et al. [17] established a set of image processing algorithms to detect pavement cracks with the MATLAB software, in which the algorithm is accurate to pixel points. It lays a valuable foundation for future research, but if this method is used to detect fine fractures, the detection accuracy will be slightly lower. Sha Aimin et al. [18] proposed a detection method for pavement cracks based on CNN. According to the traditional CNN algorithm, this method builds several different models according to the existing picture data and tests the pavement cracks according to multiple models. Although this method reduces the labor cost to a certain extent, it also has obvious shortcomings in that model construction needs a long time and has certain requirements on the capacity of the server. Considering the complexity of the image collection process, images collected from real-word scenes are generally impacted by complex illumination conditions, shadows, the shape and size of cracks, and so on. Palevičius [19] and Pal [20] et al. focused on the influence of shadows on the automatic detection of cracks in pavement. The novel shadow augmentation techniques were developed to increase the accuracy of automatic detection of concrete cracks. They further discussed the limitations of the shadow removal technique for improving crack detection accuracy. From the reported literature, automatic crack recognition dependent on automatic vehicle detection is expensive and not available for the non-road sector, although it is more effective. On the contrary, the lower-priced solutions are less effective and time-consuming. The most common method, which relies on professionals to recognize and assess pavement cracks, is the simplest, but it is not accurate enough. Besides, the manual inspection method requires significant labor and time costs, especially when working in highway areas or in complex weather conditions, which can put inspectors at risk. Therefore, the automated pavement surface inspection and management system can efficiently detect and classify different types of pavement distress, which makes the development and use of computer vision technology in the field of pavement engineering possible. The automatic detection of cracks on pavement by means of computer vision technology provides access to massive amounts of pavement condition data, which contributes data for life cycle analysis (LCA) and life cycle cost analysis (LCCA) [21].

To evaluate the functions of the pavement service, it is necessary to solve pavement cracks and realize automatic detection and monitoring of pavement cracks. In this study,

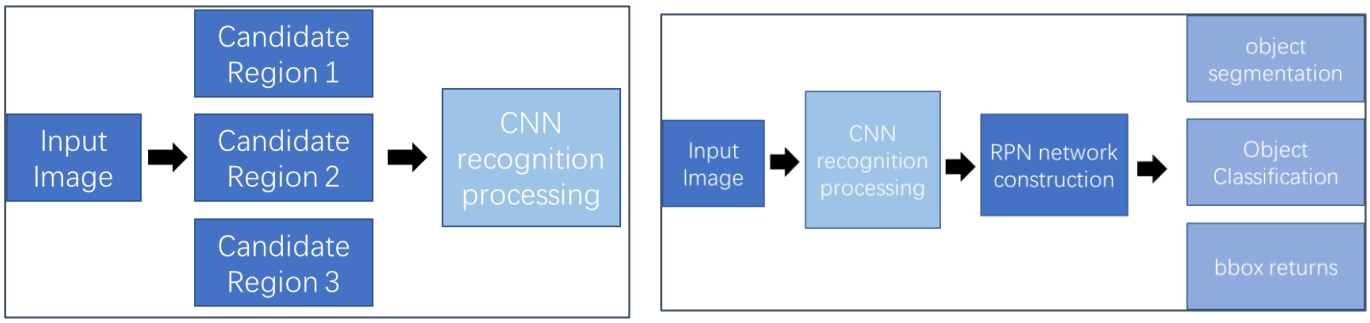
a novel convolutional neural network method in the Mask R-CNN model was proposed. The advantages of the Mask R-CNN model presented in this paper are the use of boundary box information to show breakage and the fast label data sets. The Mask R-CNN model is used to realize automatic identification of pavement cracks. Firstly, the picture data of pavement cracks was collected, sorted out, and processed. Then, the model was built and trained. Finally, the effect was analyzed, and the automatic identification of pavement cracks was realized.

2. Pavement Crack Image: Automatic Recognition Theory

A kind of deep learning algorithm most commonly used for spatial pattern analysis is Convolution Neural Networks, a representative algorithm of deep learning [22]. The research on convolutional neural networks took place from the 1980s to the 1990s, and the time delay network and LeNet-5 were the earliest convolutional neural networks. After the 21st century, with the proposal of deep learning theory and the improvement of numerical computing equipment, convolutional neural networks have developed rapidly and been applied in computer vision, natural language processing, and other fields. The convolutional neural network mimics the biological visual perception mechanism and can carry out supervised and unsupervised learning. The convolution kernel parameter sharing in the hidden layer and the sparsity of inter-layer connections make the convolutional neural network able to calculate grid-like topology features with a small amount of computation, such as pixel and audio learning, stable effects, and no additional feature engineering requirements on the data. The convolutional neural network has the ability of representational learning. By learning the spatial features that can best describe the category or quantity of the target, such as edges, corners, textures, or more abstract shapes, the input data (convolution) is transformed in various aspects and continuously on different spatial scales (such as through pooling operations), so as to classify the input information according to its hierarchical structure by translation invariant. However, with the increase in the number of layers, the quality of the CNN-based model will decline, ultimately leading to the decline of the accuracy of supervised learning. Following the input of image information, R-CNN can generate several candidate regions for each test image. In the candidate regions, the test image will be calculated by CNN. After the CNN calculation of several candidate regions is completed, some results will be obtained, and the final result will be obtained. However, in the R-CNN model, too many candidate areas will lead to an overload of the algorithm server, and it will also consume a lot of time for processing and calculation, and the final processed result will also greatly occupy the server memory space [11,15].

The defect of traditional CNN is that as the number of layers increases, the quality of the model decreases, ultimately leading to a decline in supervised learning accuracy. The reason for the better R-CNN detection effect is that, after the input of the collected image information, the system algorithm will generate several candidate regions for each image tested. In the candidate regions, the test image will be calculated by CNN. After the CNN calculation of several candidate regions is completed, the final result will be obtained after several results are obtained. It can be seen that R-CNN still has some drawbacks. Too many candidate areas will lead to an overload of the algorithm server, and it will also consume a lot of time for processing and calculation, and the final processed result will also greatly occupy the server memory space.

Mask R-CNN is further improved on the basis of RCNN. The optimized algorithm integrates the picture data used for training and the annotated data into a network model. The annotated area can be pixel segmented and positioned at the original data, in the process of model building. Compared with before the improvement, the improved algorithm solves the problem of consuming a lot of time and storage space. Meanwhile, accuracy is higher than before the improvement because there is original data for positioning when the model was built, and the algorithm realizes the results of one-to-one correspondence on pixel points (Figure 1).



(**a**) RCNN (**b**) Mask R-CNN

Figure 1. Working principles of RCNN and Mask R-CNN.

3. Data Collection

Part of the data used in the study was collected by UAV for high-definition images. The main steps include making flight plans, calibrating pixels, and judging the quality of images. First of all, pre-check the site, prepare the flight plan of the UAV, specify the take-off site, shooting angle, flight path, and other information. Considering natural factors, such as wind speed, weather conditions, traffic flow, etc., may cause flight hazards for UAVs. In the process of shooting, it is necessary to consider the light intensity, shooting angle, shooting background, and plan the flight path of the shot so as to achieve high-definition image acquisition.

The rest of the data is collected by means of vehicle-mounted cameras, on-site photos, and Internet search images. The data requirements are that the image has a high contrast in the crack area and that the pixel size is not less than 560*540 (the clearer, the better). All images should be in the same format to facilitate subsequent training operations. The suffix should be.jpg or.png.

4. Data Processing

4.1. Crack Labeling

Image labeling plays a crucial role in computer vision. The target of image labeling is related to the task and belongs to the label of a specific task. Although machine learning is more convenient than the traditional method, it requires more human intervention than we think. In order to make the machine learning effect better, we need to improve the accuracy of the training data and label the images more accurately to get the correct results. Below is the labeling of the collected images with the help of LabelMe software (v4.6.0). The crack labeling pictures are shown in Figure 2.



(**a**) (**b**)

Figure 2. Crack labeling procedure. (**a**) Crack labeling; (**b**) Completed crack labeling.

After the outline is complete, create labels. Each labeled image should have the same label name. This study made a preliminary plan and used 50 pictures as the initial data, 40 of which were used as training data and 10 as test data. Figure 3 shows the annotation results for some pictures.

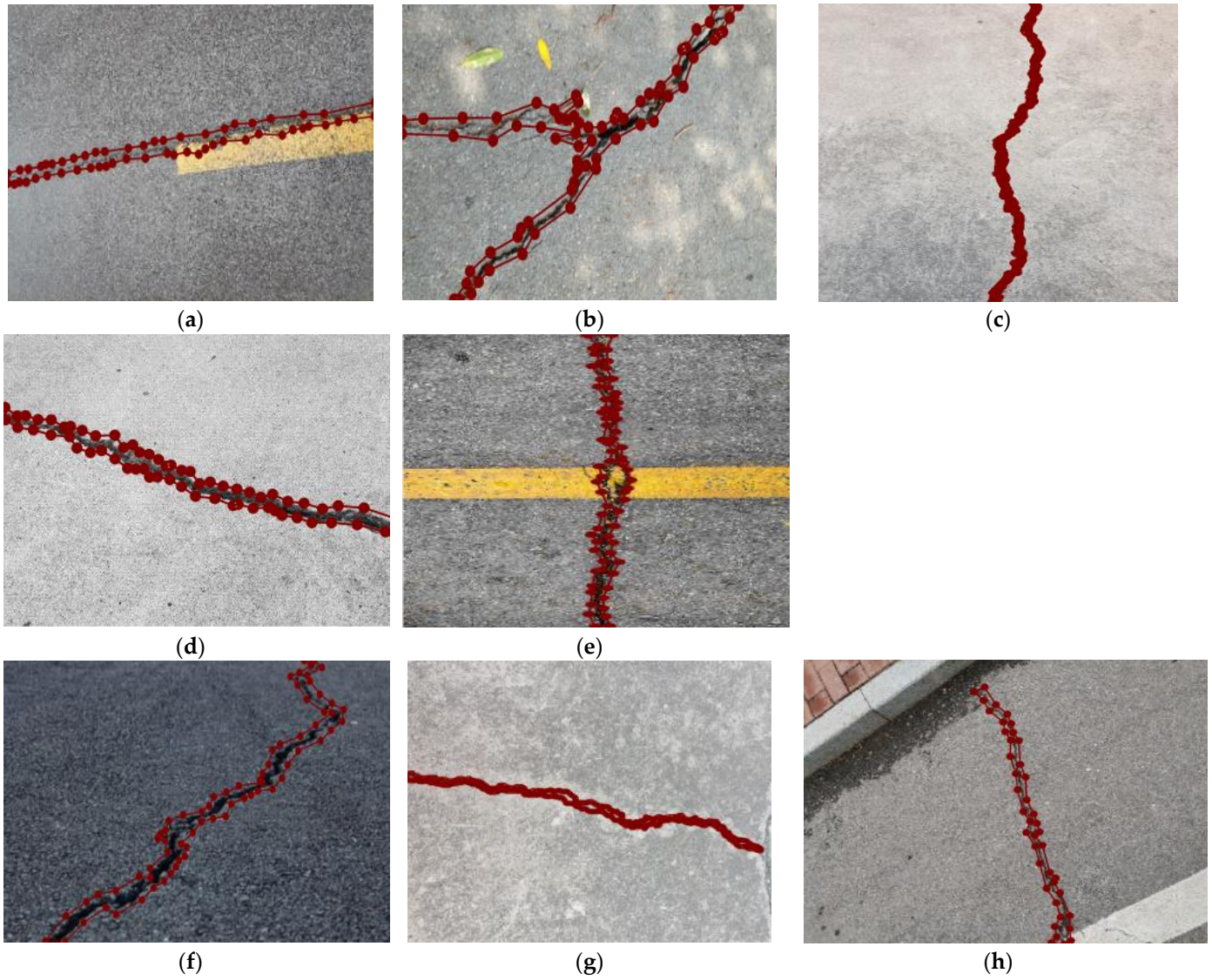


Figure 3. Part of the image data completed with labeling (**a**–**h**).

4.2. Model Building

The Mask-RCNN instance segmentation algorithm is constructed by the backbone, RPN, ROI-Align, and head to achieve "target detection," "target instance segmentation," and "target key point detection" of pavement disease targets such as cracks. Not only can the target be detected in the image, but each pixel can also be labeled to achieve semantic segmentation, as shown in Figure 4.

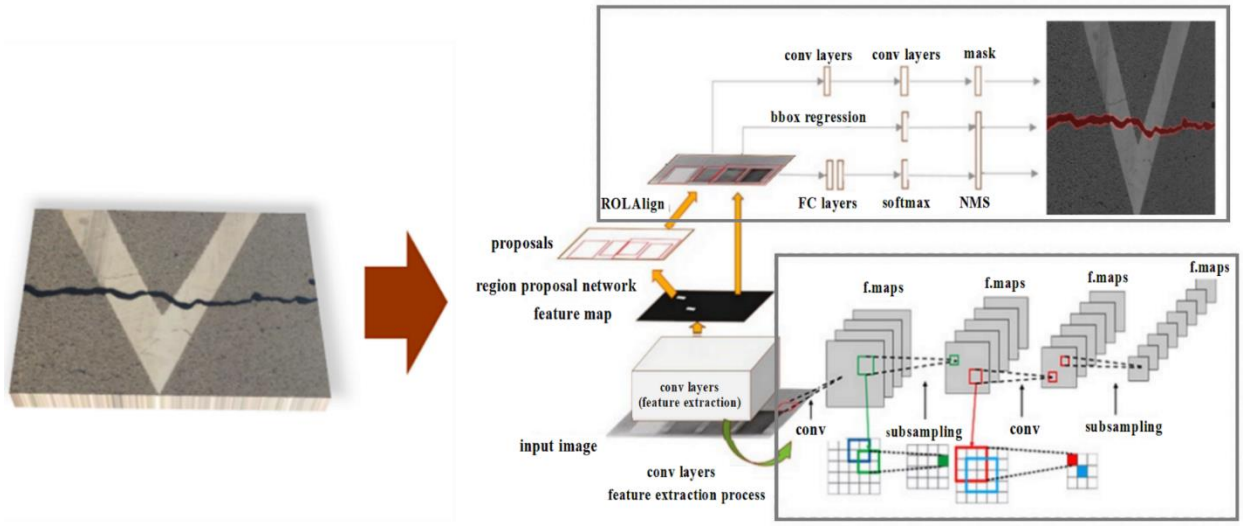


Figure 4. Schematic diagram of the model building process.

4.3. Model Training

The initial weight of this model's training uses the parameters of the COCO dataset model, and the COCO model is read first at the beginning of the training. Modify the network hyperparameter to match the schema in Table 1, read the prepared data set, and start the feature extraction process. After feature extraction, the training process begins. During the training process, other parts of the network are frozen first and the parameters of the head part of the network are updated. Ten epochs are trained, and each epoch is trained for 100 steps. After the training of the head part, all freezing parameters are unlocked, and the global training is carried out. This process trains 30 epochs and also trains 100 steps for each epoch.

Table 1. Network hyperparameters.

Backbone	Learning Rate	Lr_Momentum	Weight_Decay	Mini_Batch
Resnet101	0.001	0.85	0.0001	2

4.4. Effect Analysis

After building the training model, the test results of the image data are analyzed, and the effect analysis is shown in Figure 5. The loss function L is used to characterize the validity and accuracy of the test results:

$$L = L_{RPNcls} + L_{RPNreg} + L_{cls} + L_{reg} + L_{mask}$$

where L_{RPNcls} and L_{RPNreg} are target classification loss functions and regression box loss functions in the RPN network, respectively. L_{cls}, L_{reg}, and L_{mask} are the classified loss function, regression box loss function, and mask loss function of the output target, respectively. The RPN network loss function is represented as:

$$L = L_{RPN_{cls}} + L_{RPN_{reg}} = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i L_{reg}(t_i, t_i^*)$$

where i is the feature of the anchor point, and p_i* is the prediction probability of the kernel function in machine learning corresponding to the point i. If the anchor point is positive, p_i* is 1; if the anchor point is sensitive, p_i* is 0.

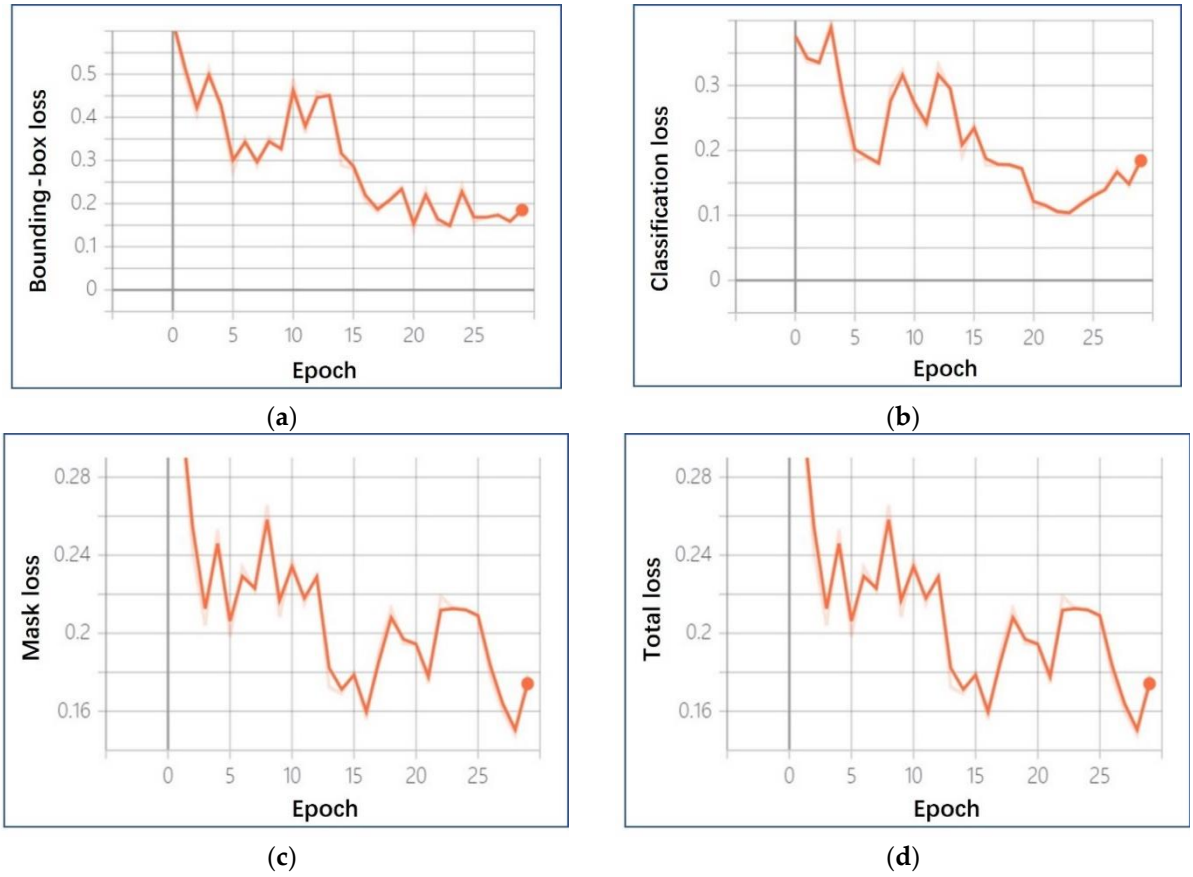


Figure 5. Test effect after model training. (**a**) The bounding-box loss; (**b**) The classification loss; (**c**) The mask loss; (**d**) The total loss.

The output target classification loss function is represented as:

$$L_{cls} = -\sum_i^N p_i \log p_i^*$$

The output regression box loss function is implemented based on the smooth function, which is represented as:

$$L_{reg} = \sum_i smooth(t_i - t_i^*)$$

where the smooth function is 0.5×2 if given$|\times|$given < 1; otherwise, the smooth function given$|\times|$given minus 0.5.

As can be seen from Figure 5, the effect of automatic crack identification varies from good to bad. The coverage of partial crack identification is not comprehensive enough. The schematic results of pavement crack identification are shown in Figure 6. As can be seen from Figure 6, the confidence level of crack detection for the presented pictures is from 72.5% to 90.0%. Only parts of the whole crack could be recognized in Figure 6d,f. So, the detection results were not ideal. To solve the problem, this study also made improvements to the optimization of the model by proposing solutions and implementing them for the problems that emerged from the above effect analysis. In an analysis of loss function underfitting, the effect needs to be improved, and the mask drawing pixel-level recognition accuracy also needs to be improved. The main problems include: the data set is too small, resulting in weak generalization ability and poor robustness of the model; the distribution of data sets is uneven; the crack data with different characteristics are greatly different; the fracture labeling; and the model hyperparameters have not reached the optimal solution.

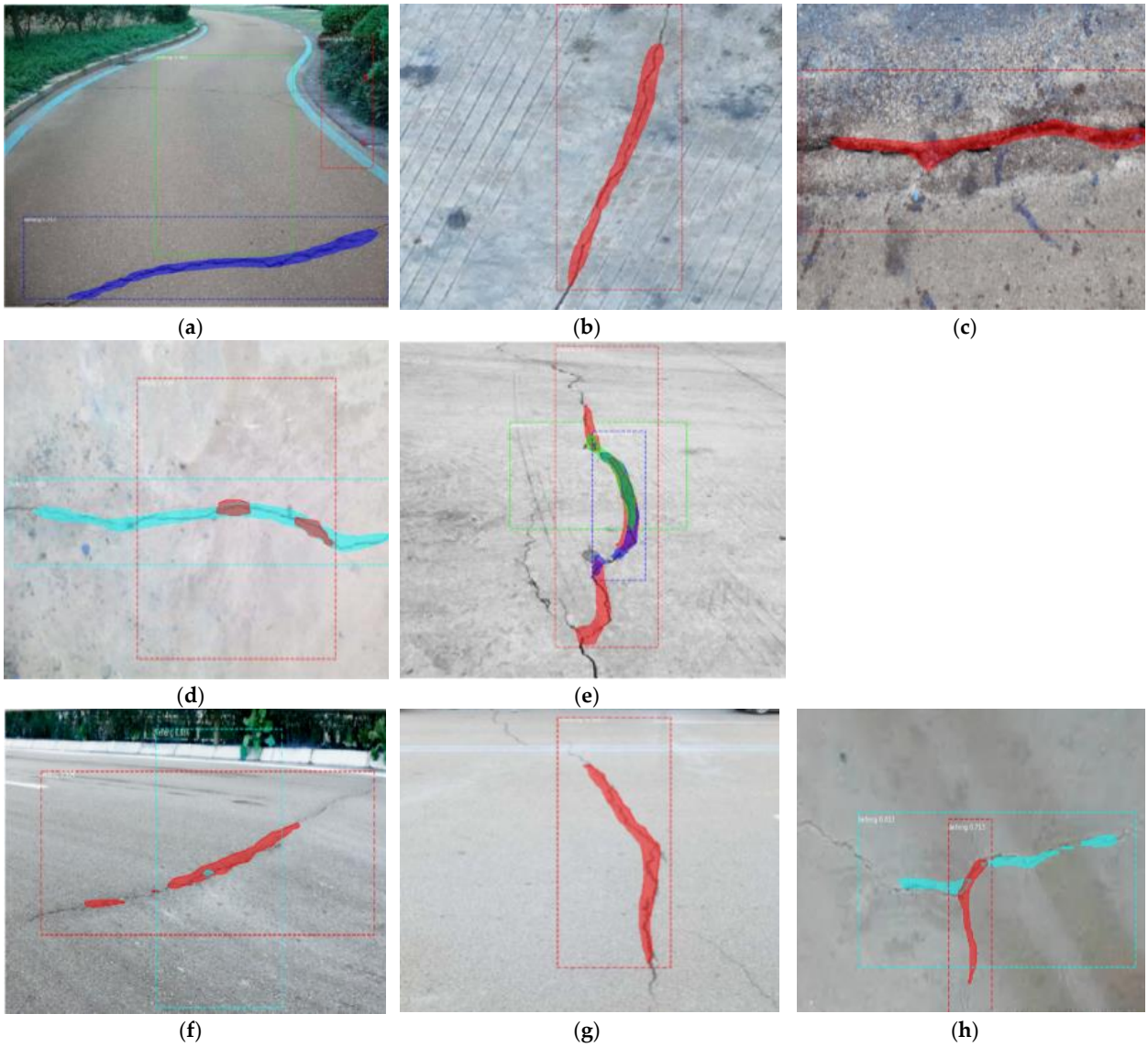


Figure 6. Recognition results of pavement cracks. (**a**) Confidence level: 75.1%; (**b**) Confidence level: 91.1%; (**c**) Confidence level: 75.1%; (**d**) Confidence level: 88.5%; (**e**) Confidence level: 90.0%; (**f**) Confidence level: 72.5%; (**g**) Confidence level: 87.7%; (**h**) Confidence level: 86.9%.

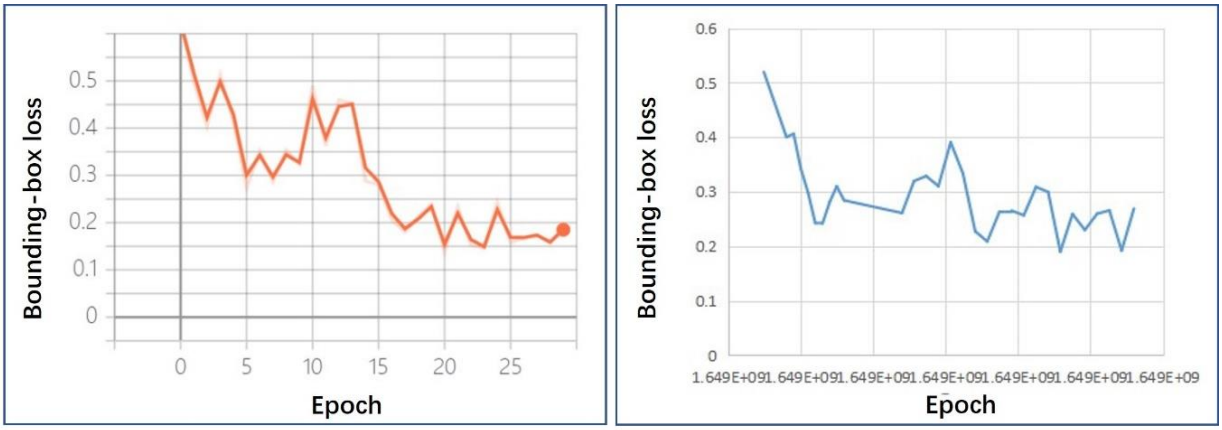
4.5. Research on Model Optimization and Improved Identification Effect

In order to improve the research effect, this paper improved the optimization of the model. Solutions were proposed and implemented to solve the problems in the above effect analysis. Firstly, the data set was too small. In this improvement, 500 new picture data points were added on the basis of 50 picture data points, and the quality of the 500 new picture data points was strictly screened. The quality of the data set was optimized by improving the screening for factors that may affect the final result, such as the pixel of the material, the crack background, and the crack type. In order to address the issue of crack labeling, the labeling operation of 500 new image data was optimized, and the frame selection and crack labeling were refined. Aiming at the problem that the superparameters of the model have not reached the optimal solution, this paper studies the adjustment of the superparameters and optimizes the structure of the model.

After the marking and processing of 500 image data points, the Mask-R CNN model was built, among which 450 image data points were used for model training and the other 50 image data points were tested. The analysis and comparison of loss function

effects before and after optimization are shown in Figure 7. The main purpose of the RPN bounding-box loss function (Figure 7a) is to determine the accuracy of accurate target positioning, or the positioning of the minimum external rectangle generated outside the target. The loss is mainly calculated by the difference between the predicted external frame and the marked external frame to characterize the prediction effect, the smaller the value, the more accurately the target can be framed. The classification loss function (Figure 7b) is mainly used to judge the accuracy of the target category in the ROI extraction area, and the smaller the value, the higher the accuracy of recognition. The mask loss function (Figure 7c) is mainly a function of whether the individual pixels in the picture are correctly classified, i.e., the loss is an evaluation of the segmentation effect, and its lower value means that the individual pixels are correctly classified. The total loss function (Figure 7d) is the sum of all loss functions and is used to evaluate the overall effectiveness of the training process. It can be seen that the mask loss function and total loss function decrease significantly. Minimizing the loss function is our training objective. For the results of training in the study, the bigger data sets lead to more accurate detection results. However, the larger data sets are more time-consuming. According to the loss function, the data sets selected in the study are moderate. It is possible to obtain derisible accuracy, and efficient accuracy is appropriate.

After the comparison of the above functions, the optimized function is relatively smoother, indicating that the optimized effect is better than that before optimization and proving that the improvement is effective. Both the slight crack and the severe crack can be detected by the model. Because the width of the crack is smaller, the detection accuracy is lower; however, the more severe the crack, the better the detection. The model has drawbacks and cannot give the length and width of the cracks, but it can solve the problem of consuming a lot of time and storage space. At the same time, the accuracy is higher than before the improvement. In order to obtain a better recognition effect, it is necessary to further expand the data set. The crack identification results of the newly added 500 image data set are shown in Figure 8. As can be seen, the confidence level of crack detection for the presented pictures is all higher than 90%. The highest confidence level is 98.5%. Compared to Figure 6, the confidence level is significantly improved. This indicates that the model proposed in the study shows a good detection effect on pavement cracks. Comparing to previous studies [23], the Mask R-CNN model with the new convolutional neural network method presented in the work shows the advantage of fast labeling data sets followed by higher accuracy. Furthermore, because images collected from real-word are generally impacted by noises and other image obstructions such as shadows, they may decrease the accuracy for detecting cracks. The impact of these image obstructions needs to be further evaluated in further study to check the model's applicability.



(**a**)

Figure 7. *Cont.*

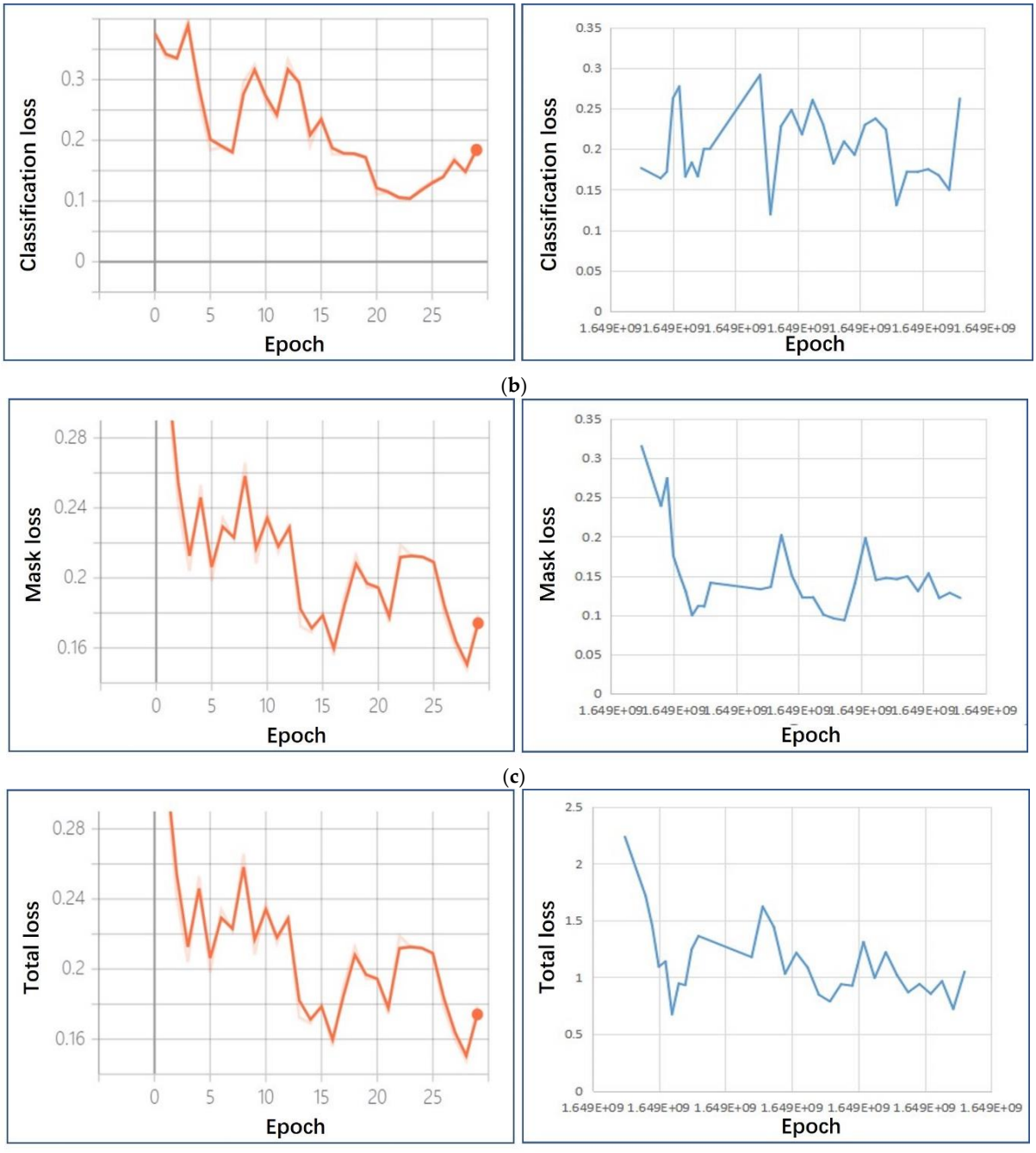


Figure 7. Analysis and comparison of loss function effects before and after optimization. (**a**) The comparison of bounding-box loss; (**b**) The classification loss comparison of the output target; (**c**) The mask loss comparison of the output target; (**d**) The comparison of total loss.

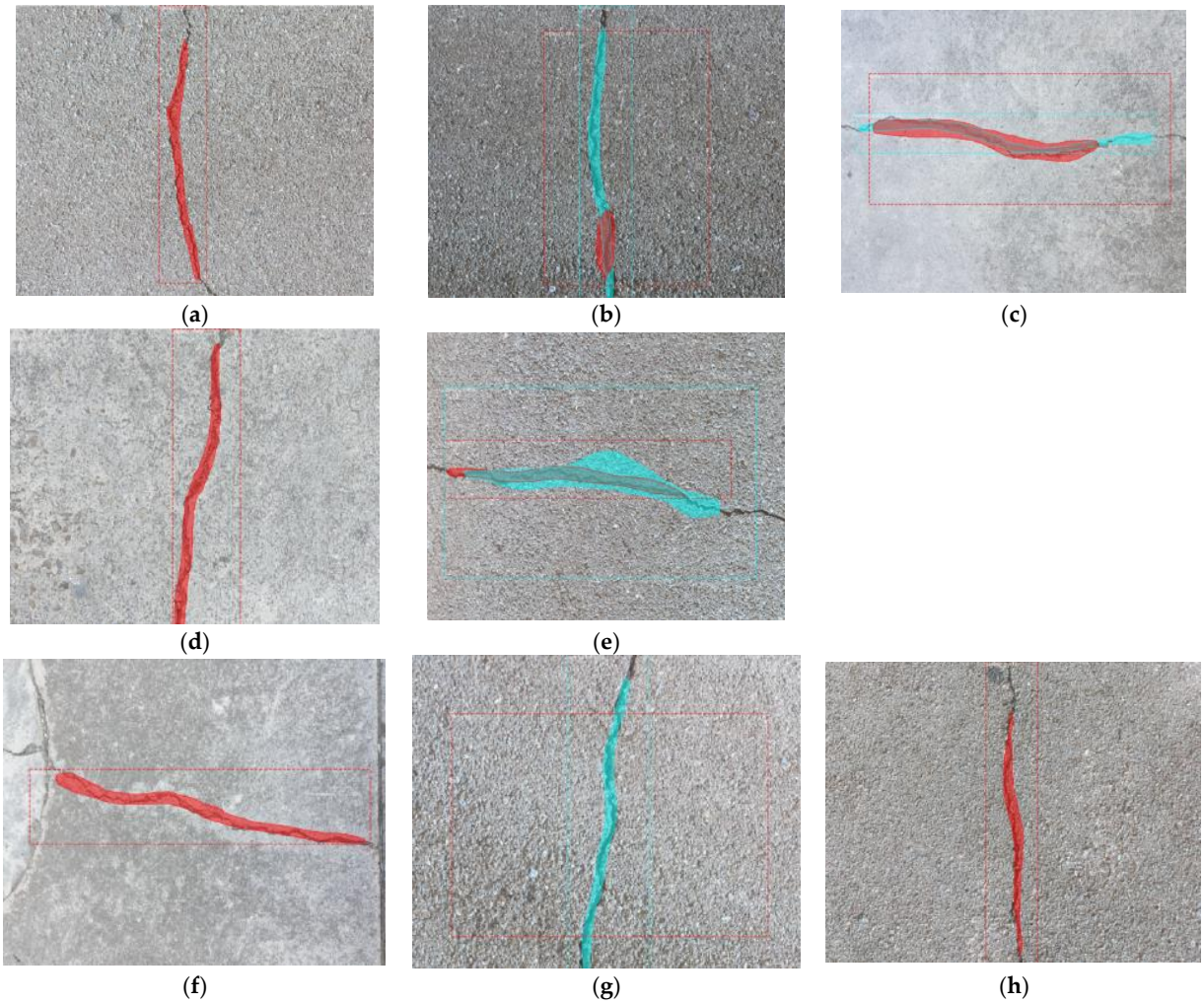


Figure 8. Recognition results for pavement cracks after optimization. (**a**) Confidence level: 95.0%; (**b**) Confidence level: 98.5%; (**c**) Confidence level: 93.2%; (**d**) Confidence level: 97.4%; (**e**) Confidence level: 92.9%; (**f**) Confidence level: 94.3%; (**g**) Confidence level: 91.7%; (**h**) Confidence level: 90.3%.

5. Conclusions

In this paper, the novel convolutional neural network method in the Mask R-CNN model was constructed to realize the automatic identification of pavement cracks. The pavement image data was collected by UAVs and vehicle-mounted cameras. The Mask-RCNN algorithm is constructed by the backbone, RPN, ROI-Align, and head, and the effect of the training model was effectively improved by increasing the number of data sets, pixels of crack pictures, crack background, and crack type labeling methods. After optimizing the model, the confidence levels of crack detection for pictures are all higher than 90% (the highest value is 98.5%), compared to other results below 90%. The RPN bounding-box loss function, the output target classification loss, the mask loss, and the total function were further used to prove the validity and accuracy of the test results. The model proposed in the study shows a good detection effect on pavement cracks. As a result, the advantage of the presented Mask R-CNN model is fast labeling of data sets followed by higher accuracy, which solves the problem of time consumption and storage space consumption. The presented Mask R-CNN model is beneficial for crack detection and pavement condition inspection to provide support for pavement maintenance. Besides, the impact of the image obstructions, like shadows, needs to be further evaluated in the future to check the model's applicability.

Author Contributions: P.W., C.W. and M.L.: Data curation; Writing—original draft; Writing—review and editing. H.L., H.W. and W.Z.: Data curation. S.Z. and G.Z.: Investigation. S.L.: Investigation. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the Qilu Young Scholars Program of Shandong University, the Natural Science Foundation of Shandong Province (CN) (No. ZR2020ME244), the Jinan Research Leader Studio (202228101), and the National Key Research and Development Plan Project (2022 YFB2603300).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declared that they have no conflicts of interest in this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

1. Xin, X.; Jiang, H.; Liang, M.; Yao, Z.; Zhang, J.; Luo, W.; Zhang, X. Chemical, rheological properties and microstructure of road asphalt prepared from deoiled asphalt, slurry oil and polymers. *Constr. Build. Mater.* **2020**, *257*, 119571. [CrossRef]
2. Xin, X.; Luan, X.; Su, L.; Ma, C.; Liang, M.; Ding, X.; Yao, Z. The innovative self-sensing strain sensor for asphalt pavement structure: Substitutability and synergy effects of graphene platelets with carbon nanotubes in epoxy composites. *Front. Mater.* **2022**, *9*, 4. [CrossRef]
3. Yang, Y. *Research on Pavement Crack Image Segmentation Method Based on Fractal Dimension*; Changan University: Xi'an, China, 2014.
4. Majidifard, H.; Jin, P.; Adu-Gyamfi, Y.; Buttlar, W.G. Pavement image datasets: A new benchmark dataset to classify and densify pavement distresses. *Transp. Res. Rec.* **2020**, *2674*, 328–339. [CrossRef]
5. Ma, D.; Fang, H.; Wang, N.; Xue, B.; Dong, J.; Wang, F. A real-time crack detection algorithm for pavement based on CNN with multiple feature layers. *Road Mater. Pavement Des.* **2022**, *23*, 2115–2131. [CrossRef]
6. Ibragimov, E.; Lee, H.J.; Lee, J.J.; Kim, N. Automated pavement distress detection using region based convolutional neural networks. *Int. J. Pavement Eng.* **2022**, *23*, 1981–1992. [CrossRef]
7. Peng, B.; Jiang, Y.; Chen, C.; Wang, K.C.P. Automatic parallel cracking detection algorithm based on 1 mm resolution 3D pavement images. *J. Southeast Univ. Nat. Sci. Ed.* **2015**, *45*, 1190–1196.
8. Feng, X. *Research on Algorithm of Asphalt Pavement Cracks Detection Based on Range Image*; Hubei University of Technology: Wuhan, China, 2016.
9. Ke, W.; Chen, H.; Lei, Y.; Zhang, T. Prediction method for asphalt pavement crack based on GRNN neural network. *J. Shenzhen Univ. Sci. Technol.* **2017**, *34*, 378–384. [CrossRef]
10. Chen, C.; Seo, H.; Zhao, Y. A novel pavement transverse cracks detection model using WT-CNN and STFT-CNN for smartphone data analysis. *Int. J. Pavement Eng.* **2021**, *23*, 4372–4384. [CrossRef]
11. Aslan, O.D.; Gultepe, E.; Ramaji, I.J.; Kermanshachi, S. Using Artifical Intelligence for Automating Pavement Condition Assessment. In *International Conference on Smart Infrastructure and Construction 2019 (ICSIC) Driving Data-Informed Decision-Making*; ICE Publishing: London, UK, 2019; pp. 337–341.
12. Piao, W. *Research on Segmentation Algorithm of Pavement Crack in Complex Environment*; Zhengzhou University: Zhengzhou, China, 2019.
13. Lv, J. *Research and Implementation of Pavement Crack Detection Method Based on Deep Learning*; Southeast University: Nanjing, China, 2019.
14. Pan, Y.; Chen, X.; Sun, Q.; Zhang, X. Monitoring Asphalt Pavement Aging and Damage Conditions from Low-Altitude UAV Imagery Based on a CNN Approach. *Can. J. Remote Sens.* **2021**, *47*, 432–449. [CrossRef]
15. Li, J.; Liu, T.; Wang, X.; Yu, J. Automated asphalt pavement damage rate detection based on optimized GA-CNN. *Autom. Constr.* **2022**, *136*, 104180. [CrossRef]
16. Premachandra, C.; Waruna, H.; Premachandra, H.; Parape, C.D. Image Based Automatic Road Surface Crack Detection for Achieving Smooth Driving on Deformed Roads. In Proceedings of the 2013 IEEE International Conference on Systems, Man, and Cybernetics, Manchester, UK, 13–16 October 2013; pp. 4018–4023.
17. Oliveira, H.; Correia, P.L. Crack-IT an image processing toolbox for crack detection and characterization. In Proceedings of the IEEE International Conference on Image Processing 2014, Paris, France, 27–30 October 2014; pp. 798–802.
18. Sha, A.; Zheng, T.; Gao, J. Recognition and measurement of pavement disasters based on convolutional neural networks. *China J. Highw. Transp.* **2018**, *31*, 1–10.
19. Palevičius, P.; Pal, M.; Landauskas, M.; Orinaitė, U.; Timofejeva, I.; Ragulskis, M. Automatic detection of cracks on concrete surfaces in the presence of shadows. *Sensors* **2022**, *22*, 3662. [CrossRef] [PubMed]

20. Pal, M.; Palevičius, P.; Landauskas, M.; Orinaitė, U.; Timofejeva, I.; Ragulskis, M. An overview of challenges associated with automatic detection of concrete cracks in the presence of shadows. *Appl. Sci.* **2021**, *11*, 11396. [CrossRef]
21. Riekstins, A.; Haritonovs, V.; Straupe, V. Life cycle cost analysis and life cycle assessment for road pavement materials and reconstruction technologies. *Balt. J. Road Bridge Eng.* **2020**, *15*, 118–135. [CrossRef]
22. Nguyen, N.T.H.; Le, T.H.; Perry, S.; Nguyen, T.T. Pavement crack detection using convolutional neural network. In Proceedings of the Ninth International Symposium on Information and Communication Technology, Da Nang City, Viet Nam, 6–7 December 2018; pp. 251–256.
23. Xu, X.; Zhao, M.; Shi, P.; Ren, R.; He, X.; Wei, X.; Yang, H. Crack detection and comparison study based on faster R-CNN and mask R-CNN. *Sensors* **2022**, *22*, 1215. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.