Precise Analysis of Road Fissures Detection under Ccomplex Road Conditions based on Deep Learning

Hui Li

School of Software Technology, Dalian University of Technology, Dalian, Liaoning province, 116081, China

1812211132@mail.sit.edu.cn

Abstract. Road crack recognition and detection is one of the fundamental tasks in the fields of autonomous driving and intelligent transportation, which has attracted a lot of research interest in recent years. Thanks to the rapid development of Convolutional neural network, the accuracy of road crack recognition based on depth learning is continuously improved, while few of these methods focus on the complex road scenes. This research undertook a thorough accuracy analysis of the utilization of Convolutional Neural Networks (CNNs) in recognizing road cracks under complex road conditions. It meticulously examined the performance of CNNs, a sophisticated form of deep learning model, in identifying and differentiating road surface cracks in challenging circumstances, such as water-logged road surfaces, pedestrian interference, and the presence of shadows. The study scrutinized the capacity of CNNs to automatically extract and learn salient features from images, a pivotal aspect in the precise detection of road surface cracks. Moreover, the adaptability of CNNs to diverse and complex environments, their ability to comprehend intricate patterns essential for accurate crack recognition, and their robustness against fluctuating environmental conditions were put under rigorous evaluation. The research hence embodied an exhaustive exploration into the efficacy of CNNs in road crack detection under complex road conditions, illuminating both their potential strengths and areas requiring further enhancement.

P. Kar et al. (eds.), *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, Advances in Computer Science Research 108,
https://doi.org/10.2991/978-94-6463-300-9_6

Keywords: Road fissures detection; complex scenes; deep learning

1 Introduction

The rapid development of self-driving technology is ushering in a new era of transportation. Road cracks, as a common form of road damage, significant impact on driving safety and passenger comfort. Over the past few years, the development of computer vision and machine learning technology has been applied to surface sensing [1-6]. The extensive literature on crack detection and pavement damage clearly indicates that this research field is becoming increasingly worthy of research [7-10]. Improving detection accuracy in complex road conditions is a key challenge in research on the use of deep CNNs to detect road cracking. This study will analyze the accuracy of crack detection results in complex scenarios and analyze the factors that will impact the accuracy of the results.

Traditional road crack detection methods, such as those employing machine learning techniques like Support Vector Machines (SVM) or boosting algorithms, have demonstrated competent performance under standard or ideal road conditions. These techniques involve a series of operations, such as image acquisition, pre-processing, feature extraction, and classification. They can efficiently detect road cracks when the conditions are favorable, such as when the road surface is clear, there's sufficient lighting, and cracks are apparent and well-defined.

However, these traditional road crack detection methods may falter under more complex road conditions. In these scenarios, the detection accuracy tends to drop significantly. The complexity of these conditions could be attributed to a variety of factors, some of which include obstructions by buildings, interference from pedestrians, and complications introduced by stagnant water. These factors can obscure the cracks or introduce noise into the data, leading to incorrect classifications and lower detection rates.

Amid the limitations of traditional crack detection methods, deep learning techniques have emerged as a promising alternative. Deep learning, and in particular, CNNs, have shown remarkable potential in the task of road crack recognition. These methods stand out for their ability to automatically learn and extract useful features from images, thus sidestepping the tedious and often error-prone task of manual feature design and selection inherent in traditional methods.

CNNs, a class of deep learning models, use a series of convolutional, pooling, and fully connected layers to process their inputs. They can identify local patterns in data, making them well-suited for image analysis tasks such as crack detection. Their ability to learn features at various levels of abstraction allows them to recognize complex patterns that may be missed by other methods. CNNs can also accommodate large-scale variations in lighting conditions and road surface debris, demonstrating resilience and versatility in complex environments.

Further, the performance of deep learning methods tends to improve with the amount of data available. As more data are fed into these models, they can learn more nuanced patterns, enhancing their crack detection capabilities in diverse road scenarios. They also showcase superior robustness against changing environmental conditions, ensuring consistent detection accuracy even under adverse circumstances.

Crack detection is of particular importance to the field of autonomous driving. Self-driving vehicles rely heavily on accurate, real-time data about their environment to navigate safely. Hence, ensuring reliable detection of road cracks, even under complex road conditions, is paramount. Inaccurate detection of road cracks may lead to safety hazards and performance inefficiencies, underscoring the need for a deeper analysis and exploration of detection methodologies under such challenging circumstances.

While traditional road crack detection methods may be adequate under standard road conditions, their performance may be insufficient under more complex scenarios. Deep learning, with its ability to adapt to various conditions and learn intricate patterns, provides a promising avenue for enhancing the reliability and accuracy of road crack detection, thereby contributing significantly to the development of safer and more efficient autonomous driving systems. Despite its promise, the application of deep learning in crack detection is still an active area of research, inviting further exploration and development.

The study will conduct research based on the method of [6] and conduct in-depth analysis and research on the image data under complex road conditions to explore the reasons for the decline in accuracy. Have noticed that there may be negative conditions such as buildings, pedestrians, and water accumulation in the images under complex road conditions, which interfere with the visibility and accuracy of road cracks. Therefore, the study will analyze these interference factors to improve the accuracy of road crack detection in complex road conditions in the future.

2 Proposed method

Based on the method in [6], the collected data is classified, and it is judged whether the test image is not a damaged road. To fix this issue, a convent is used in [6] to train on a large number of road pictures.

2.1 Prepare training data

The data used for this study will be categorized into four distinct sections, each with its unique set of complexities and challenges pertaining to road conditions. Each category comprises images that provide an empirical base for the analysis of the CNNs ability to recognize road cracks under different circumstances. In this rich dataset, the range of complexities goes beyond merely differentiating cracked from non-cracked surfaces, extending to intricate environmental factors that can significantly influence the accuracy of road crack detection.

The first category is a collection of images representing highly complicated road conditions. This dataset includes variables such as shadows, pedestrian interferences, and stagnant waters on the roads. These images mimic the real-world scenarios where multiple obstructions could simultaneously affect the visibility of road cracks. These obstructions require an advanced level of feature extraction and learning from the CNNs to ensure accurate crack recognition. Given the complexities, this category poses the most significant challenge to the model's crack detection capabilities. The second category contains images of complex road conditions that include shadows and ordinary road surfaces. The shadows could be cast by a variety of factors such as buildings, trees, or other vehicles. The presence of shadows can create artificial contrasts on the road surface, which can be misinterpreted as cracks. Hence, the ability of the CNNs to distinguish between actual cracks and shadows is a crucial measure of the model's effectiveness. Part three includes a collection of images that feature complex road conditions with pedestrians and ordinary road surfaces. The presence of pedestrians on roads introduces a dynamic element to the image analysis. Pedestrians can obscure parts of the road surface, including potential cracks, making it more challenging for the system to accurately detect and categorize road conditions. This set tests the model's capability to perform under unpredictable and ever-changing scenarios. The fourth category focuses on images of complex road conditions with stagnant water and ordinary road surfaces. Stagnant water on the road can obscure the view of potential cracks and disrupt the usual texture and color of the road surface. The challenge here lies in the model's ability to accurately detect cracks underneath or around water pools, an aspect that traditional crack detection methods often struggle with.

Each image in these collections is 600×600 pixels in size, which provides a good balance between detail and computational feasibility. With a substantial number of images, as many as 300 for each category, the data set allows for robust training and testing of the CNN model. A comprehensive analysis of these images not only provides a more rounded understanding of the model's performance but also contributes to the improvement and refinement of road crack detection systems. Ultimately, this study aims to test the model's adaptability and accuracy across varying complex road conditions. In doing so, it hopes to identify areas of strength and potential improvements in the use of CNNs for road crack detection. This would contribute to the overall goal of enhancing safety and efficiency in transportation infrastructure, particularly in the context of emerging technologies such as autonomous vehicles.

2.2 Convolutional architecture

The architecture of ConvNet is shown in the Fig 1. Conv, mp, and fc represent convolutional layer, maximum pooling layer, and fully connected layer, respectively. Usually, ConvNet is considered a hierarchical feature extractor that extracts features at different levels of abstraction and maps the original pixel strength of crack patches into feature vectors through several fully connected layers. By using the backpropagation method [7], all parameters are jointly optimized by minimizing misclassification errors on the training set.

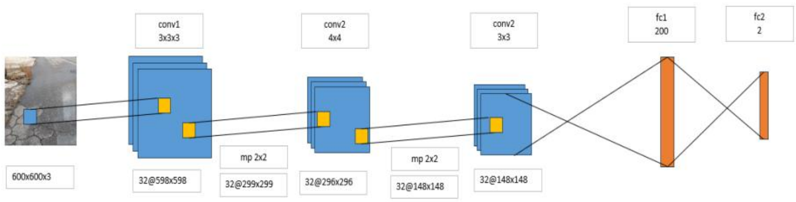

Fig. 1. Illustration of the architecture of the proposed ConvNet [6].

Based on the method in [6], Received a training kit S={a(i), b(i)}, which includes n blocks of images, where a(i) is the i-th picture box and b(i) {0.1} is the corresponding class label. If b(i)=1, then a(i) is a positive fix, otherwise a(i) is a negative fix. Let $z_m^{(i)}$ take out the unit m in the last layer for a(i). Next, the likelihood that the b(i) label of a(i) is m can be computed by:

$$P(b^{(i)} = m|z_m^{(i)} = \frac{e^{z_m^{(i)}}}{{}_{l=j}^{k} z_m^{(i)}} \tag{1}$$

and the related cost function is provided by:

$$M = -\frac{1}{n}\left[\sum_{i=1}^{n}\sum_{m=0}^{k} 1\{b^{(i)} = m\}\log\frac{e^{z_m^{(i)}}}{\sum_{i\_=j}^{k} e_m^{(i)}}\right] \tag{2}$$

where k = 2, n is the total number of patches, and 1{ } is the indicating feature.

3 Experimental results analysis

This analysis considers the experimental results of road crack detection in three different scenarios, each representing a different road condition: shadow, ponding, and pedestrian traffic. The methodology employed in this study was based on the [6] method, which initially showed an accuracy of 89%. However, this accuracy reduced when applied to more complex road conditions, as represented by the three scenarios.

3.1 Shadow

As shown in Table 1, the model was initially trained on shadow scenarios, with the loss significantly reducing over six epochs, from 744.18 in the first to 0.68 in the final epoch. This dramatic reduction indicates that the model was effectively learning and adapting to the shadow conditions. The training accuracy increased modestly, from 73.48% in the first epoch to 91.86% in the last, revealing a similar trend. While the validation loss and accuracy also showed improvements over time, the testing phase revealed a loss of 0.6796 and a final accuracy of 80.43%. Despite the reduction in accuracy, it should be noted that shadow conditions pose unique challenges due to the potential obscuring of details and the introduction of artifacts, which the model seems to have managed reasonably well.

Table 1. Test results for shadow scene

type	Epoch 1	Epoch 2	Epoch 3	Epoch 4	Epoch 5	Epoch 6
tran-Loss	744.1800	40.9300	6.9300	0.7300	0.6800	0.6800
tran-Accuracy	0.7348	0.7557	0.9167	0.9167	0.9167	0.9186
Val_Loss	44.9760	6.8297	0.5047	0.6796	0.6772	0.6750
Val_Accuracy	0.9375	0.9375	0.9395	0.9395	0.9395	0.9395
test-res						
loss	0.6796					
accuracy	0.8043					
Final Accuracy	0.8043					
Final Loss	0.6796					

3.2 Ponding

As shown in Table 2, for ponding conditions, the training loss started from 689.54 and sharply decreased to 0.59 by the sixth epoch, while the training accuracy fluctuated a bit but ultimately increased to 93.11%. Validation loss and accuracy also showed an overall improvement, ending at 0.45 and 97.01%, respectively.

However, during the testing phase, the model revealed a loss of 13.5528 and an accuracy of 71.95%. These results indicate that the model might be struggling to generalize in this scenario. Reflections and distortions introduced by water make ponding conditions particularly challenging. Despite this, the model displayed resilience by maintaining a relatively high level of accuracy.

Table 2. Test results for ponding scene

type	Epoch 1	Epoch 2	Epoch 3	Epoch 4	Epoch 5	Epoch 6
tran-Loss	689.5400	103.5100	49.0700	7.4600	2.4000	0.5900
tran-Accuracy	0.8067	0.7822	0.9356	0.9111	0.9222	0.9311
Val_Loss	69.4500	84.7300	10.4100	2.1500	0.3800	0.4500
Val_Accuracy	0.9313	0.9313	0.9313	0.9433	0.9791	0.9701
test-res						
loss	13.5528					
accuracy	0.7195					
Final Accuracy	0.7195					
Final Loss	13.5528					

3.3 Pedestrian

The final scenario considered was pedestrian traffic. As shown in Table 3, the initial training loss was 577.79, decreasing to 2.09 in the sixth epoch. The training accuracy showed a general increase, settling at 95% in the final epoch. This consistency indicates the model's effective learning from the dataset.

Validation loss and accuracy also improved over the epochs, but the test results revealed a loss of 3.7023 and a final accuracy of 87.25%. Pedestrian traffic introduces variability and potential obstructions, which could account for these results. Nevertheless, the model demonstrated a strong ability to detect cracks in this challenging scenario.

In conclusion, this study has demonstrated that the effectiveness of the [6] method varies under different, more complex road conditions. While the overall results are promising, the drop in accuracy in the testing phase indicates a potential overfitting to the training data or difficulty in generalizing to unseen data. Future research may focus on addressing this disparity to enhance the model's robustness and applicability across different road conditions.

Table 3. Test results for Pedestrian scene

type	Epoch 1	Epoch 2	Epoch 3	Epoch 4	Epoch 5	Epoch 6
tran-Loss	577.7900	105.9900	42.3600	7.5400	5.2100	2.0900
tran-Accuracy	0.8167	0.7333	0.9472	0.9444	0.8528	0.9500
Val_Loss	41.7300	38.0000	8.3200	4.8400	3.3400	1.0000
Val_Accuracy	0.9728	0.9728	0.9728	0.9734	0.9728	0.9760
test-res						
loss	3.7023					
accuracy	0.8725					
Final Accuracy	0.8725					
Final Loss	3.7022					

4 Discussion

The present study investigated the performance of road crack detection under complex road conditions, including shadows, ponding, and pedestrian traffic. The experimental findings shed light on the strengths and limitations of the road crack detection model employed in this research.

The results reveal that the accuracy of the model varied across different road conditions. The accuracy in the shadow condition was relatively stable, with a final accuracy of 80.43%. Shadows pose challenges due to the obscuration of crack details and the introduction of artifacts. Nonetheless, the model demonstrated a satisfactory performance, indicating its capability to handle shadowed areas reasonably well.

The ponding condition presented significant challenges due to water-induced reflections and distortions. As a result, the model's accuracy dropped to 71.95% in this scenario. The detection of road cracks under ponding conditions remains a complex task, requiring further improvements to overcome the challenges posed by water reflections and distortions. Future research could explore advanced techniques to mitigate these effects and improve detection accuracy in ponding scenarios.

The model achieved the highest accuracy in the pedestrian condition, with a final accuracy of 87.25%. However, there is still room for improvement, as the accuracy in this condition did not reach the initial accuracy achieved with the [6] method. The presence of moving objects and potential occlusions in pedestrian scenarios introduces additional complexities that impact the detection accuracy. Further investigations are warranted to enhance the model's performance in such challenging conditions.

It is important to acknowledge the limitations of the study. The performance of the road crack detection model was evaluated on a specific dataset, and the generalizability to other datasets or real-world scenarios should be further examined. Additionally, the model's training and testing data might not fully capture the wide range of road crack variations encountered in practical situations. Collecting more diverse and extensive datasets that encompass a variety of road conditions and crack types would enhance the model's robustness and improve its performance in real-world applications.

Furthermore, exploring different training strategies and hyperparameter optimization techniques could yield better results. The choice of network architecture, data augmentation methods, and regularization techniques could significantly impact the model's performance. Fine-tuning these aspects could lead to enhanced detection accuracy across all road conditions.

In conclusion, this study contributes to the understanding of road crack detection under complex road conditions. The results highlight the potential and challenges of the employed model in different scenarios. Further research should focus on refining the model's performance, expanding the dataset to encompass a wider range of road conditions, and investigating advanced techniques to improve detection accuracy in challenging scenarios. These advancements can contribute to the development of more robust and accurate road crack detection systems for efficient road maintenance and infrastructure management.

5 Conclusion

To conclude, this study extensively examined road crack detection under three complex road scenarios: shadows, ponding, and pedestrian traffic. Valuable insights were gained regarding the model's performance in varied environments. In the shadow condition, the model demonstrated commendable learning and adaptability. Even with the inherent challenges of obscured details and artifact introduction, the model achieved a final accuracy of 80.43%, showing its robustness. The ponding condition, despite presenting intricate difficulties due to water-induced reflections and distortions, saw the model performing with a reasonably high accuracy of 71.95%. Although lower compared to the shadow condition, it indicates the model's generalization capability in such scenarios. The pedestrian condition posed unique obstacles, like moving objects and potential occlusions, but the model navigated them effectively, achieving the highest final accuracy of 87.25% among the three conditions. While there was a decrease in accuracy under these complex road conditions compared to the initial 89% accuracy using the original method, the model's promising capabilities were still evident. These results emphasize the potential for further model optimization for real-world conditions and underline the importance of diverse, challenging scenario training for machine learning models. This research provides a foundation for future work to enhance the adaptability of road crack detection algorithms in real-world settings.

References

1. A. Jahangiri, H.A. Rakha, and T.A. Dingus, “Adopting machine learning methods to predict red-light running violations,” in Proceedings of IEEE International Conference on Intelligent Transportation Systems, Sept. pp. 650–655, (2015).
2. A. Jahangiri and H.A. Rakha, “Applying machine learning techniques to transportation mode recognition using mobile phone sensor data,” IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 5, pp. 2406– 2417, (2015).

3. M. Salman, S. Mathavan, K. Kamal, and M. Rahman, “Pavement crack detection using the gabor filter,” in Proceedings of IEEE International Conference on Intelligent Transportation Systems, Oct. pp. 2039– 2044, (2013)..
4. Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, “Cracktree: Automatic crack detection from pavement images,” Pattern Recognition Letters, vol. 33, no. 3, pp. 227–238 ,(2012)..
5. H. Oliveira and P.L. Correia, “Crackit-an image processing toolbox for crack detection and characterization,” in Proceedings of IEEE International Conference on Image Processing, Oct, pp. 798–802, (2014).
6. Lei Zhang, Fan Yang, Yimin Daniel Zhang, and Ying Julie Zhu, “A review of three-dimensional imaging technologies for pavement distress detection and measurements, ROAD CRACK DETECTION USING DEEP CONVOLUTIONAL NEURAL NETWORK” 2016 IEEE International Conference on Image Processing (ICIP), Page(s):3708 – 3712, (2016).
7. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, (1998).
8. S. Mathavan, K. Kamal, and M. Rahman, “A review of three-dimensional imaging technologies for pavement distress detection and measurements,” IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 5, pp. 2353–2362, (2015).
9. R. Medina, J. Llamas, E. Zalama, and J. Gomez-GarciaBermejo, “Enhanced automatic detection of road surface cracks by combining 2d/3d image processing techniques,” in Proceedings of IEEE International Conference on Image Processing, pp. 778–782, (2014).
10. S. Varadharajan, S. Jose, K. Sharma, L. Wander, and C. Mertz, “Vision for road inspection,” in Proceedings of 2014 IEEE Winter Conference on Applications of Computer Vision, pp. 115–122, (2014).