Scholar Publishing Group
Machine Learning Theory and Practice
https://doi.org/10.38007/ML.2025.050103
ISSN 2790-0983 Vol. 5, Issue 1: 27-33

Crack Identification Based on Pytorch and U-Net Neural Network

Qiang Zhou[1,a], Jie Gou[2,3,b,*]

[1]Sichuan Tianmeile Expressway Co., Ltd, Meishan, Sichuan, China 620036
[2]Sichuan Shudao New Energy Technology Development Co., Ltd, , Chengdu, Sichuan, China 610041
[3]Southwest Jiaotong University, Sichuan, China 610031
[a]13076077936@163.com, [b]email:2386159102@qq.com
[*]corresponding author

Keywords: Pavement crack identification; Pytorch; crack; Deep learning; Unet; Python; Semantic segmentation

Abstract: Because the reliability of manual crack detection method is low, the cost is high, and it is greatly affected by subjectivity. This paper studies the image recognition of pavement cracks by using digital image technology. This method has the advantages of high accuracy, high flexibility, high speed and low cost, which is not affected by human experience The image processing of concrete cracks using digital image processing technology includes image grayscale semantic segmentation convolution neural network structural crack recognition and so on In this paper, we use Pytorch platform, Python programming language and U-Net neural network structure to build a model training platform to carry out model training independently, and summarize the processing technology of pavement crack image to provide reference for researchers in this field This method can quickly and effectively identify the type, size, shape and other information of cracks in the image.

Introduction

Currently, China has the largest road network in the world, with the fastest growth rate in road infrastructure. The vast expanse of roadways has made road maintenance a particularly pressing issue. China's road construction has gradually entered a maintenance and repair phase, and the country will face a significant amount of road surface inspections in the future. Traditional manual inspection methods have limitations such as low efficiency, low precision, and high labor costs. As a critical component of national economic development, road transportation plays a crucial role in the growth of key industries such as transportation, manufacturing, tourism, and agriculture. In recent years, with the extension of road service life and the increase in traffic loads, road surface damage detection and maintenance have become primary tasks in the field of transportation. Road cracks are an early manifestation of surface damage, and timely detection and maintenance of these cracks can prevent further deterioration that might pose safety risks, making it of significant practical importance[1].

Early automated detection methods include thresholding techniques[2,3], edge detection methods[4], and machine learning approaches[5]. However, these methods generally suffer from low accuracy and efficiency during the crack recognition process[6].

Fig.1 Cracks in pavement

At present, crack identification and extraction is a critical issue in the field of road surface crack detection, posing significant challenges. On one hand, due to the complexity of road surfaces and the presence of interference factors such as textures, noise, and shadows, there are numerous misjudgments and missed detections in current crack identification tasks. On the other hand, when cracks are known to exist on the surface, how to eliminate the influence of these interference factors and achieve automatic and accurate crack extraction remains another key issue in road crack detection. To improve maintenance effectiveness, the development of fast and efficient automated detection and recognition systems for road surface cracks has become a hot topic. From the perspective of detection methods and equipment development, crack detection based on computer vision (digital image processing) has become the mainstream research direction in this field.

This study utilizes the Pytorch platform combined with the Python programming language and the U-Net neural network to construct a model training platform. The research focuses on the image processing techniques for road surface cracks, resulting in the development of a model capable of accurately identifying cracks in images.

1 Highway Crack Detection

1.1 U-Net Network Architecture

U-Net, proposed in 2015, is a variant of Fully Convolutional Networks (FCN). For more detailed information about FCNs, you can refer to my other paper on FCN, which includes reading and code implementation. The original goal of U-Net was to address issues in biomedical image segmentation. Due to its excellent performance, it has been widely adopted in various domains of semantic segmentation, such as satellite image segmentation, industrial defect detection, and more.

As shown in Figure 2, the architecture can be divided into three main parts:

1.The Backbone Feature Extraction: The backbone is used to extract features progressively. The U-Net's backbone is similar to VGG, consisting of stacked convolutional layers and max-pooling layers. This part allows for the extraction of five initial effective feature layers, which will be used in the next step for feature fusion.

2.Enhanced Feature Extraction: The five initial effective feature layers obtained from the backbone are subjected to upsampling and feature fusion (by stacking channels of the upsampled results). This produces a final feature layer that integrates all of the extracted features.

3.Prediction: In this part, the final effective feature layer is used to classify each feature point, essentially performing pixel-wise classification.

This paper primarily focuses on road crack recognition, utilizing a dataset of road surface crack images taken with a mobile phone on the campus of Temple University, USA. The dataset is challenging due to varying lighting conditions, diverse crack shapes, and noisy textures. It serves as an excellent engineering practice case and an algorithm testing case. After processing, the dataset contains 3,368 images of road surface cracks, divided into training (1,896 images), validation (348 images), and testing (1,124 images) subsets. Each image is annotated with pixel-level ground-truth crack labels.

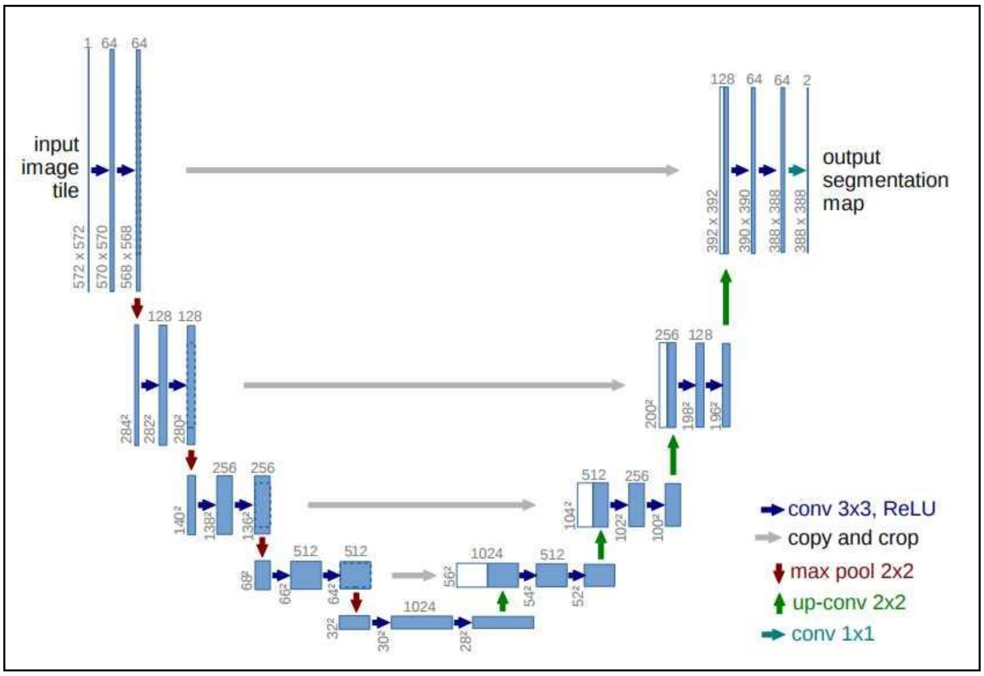

Fig.2 U-Net network framework

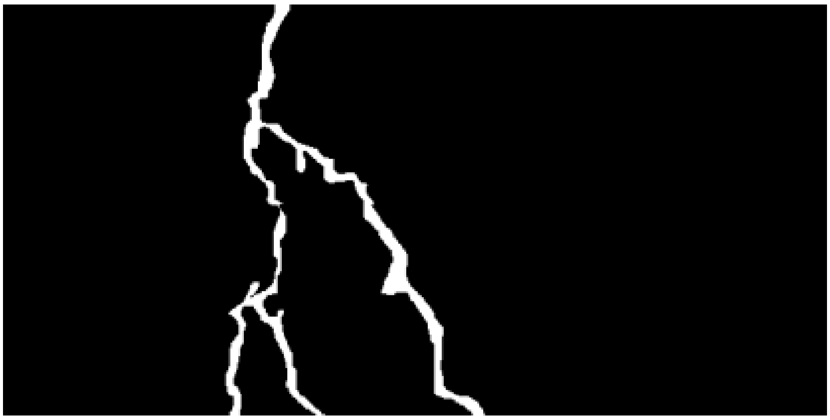

Fig.3 PDS annotation

1.2 Implementation Process

The core hardware environment used for training in this study is shown in Table 1.

Table 1. Core hardware environment

Hardware Component	Specification
Computer Model	Lenovo Legion R9000P Laptop
Graphics Processor	NVIDIA GeForce RTX 3060
Processor	AMD Ryzen 7 5800H
Operating System	Microsoft Windows 11 Home Edition

The model is trained using the open-source framework PyTorch, in a Windows 11 computer environment, along with packages such as numpy, math, argparse, logging, and yaml. The initial training learning rate is set to 0.1, with a batch size of 16 per iteration, and 60 epochs.

To evaluate the model's performance, metrics such as Precision (P), Recall (R), and Mean Average Precision (mAP) are used, with the corresponding calculation formulas as follows::

$$precision = \frac{TP}{TP + FP}$$

$$\text{recall} = \frac{TP}{TP + FN}$$

$$MIOU = \frac{TP}{TP + FP + FN}$$

In the formulas, TP, FP, andFNrepresent True Positives, False Positives, and False Negatives, respectively. AP refers to the area under the Precision-Recall curve, which represents the number of large classes detected. D represents the number of categories, and their specific relationship is shown in Table 2.

Table 2 Deep learning model evaluating indicator

	Predicted Negative	**Predicted Positive**
Actual Negative	True Negative (TN)	False Positive (FP)
Actual Positive	False Negative (FN)	True Positive (TP)

After 10 rounds of training, the growth rate of all metrics slowed down. After 30 rounds of training, the metrics essentially reached a stable state, with no further upward trend. During the training process, there were significant fluctuations in the evaluation metrics when they reached a balanced state, which may have been caused by a learning rate that was too large at the current stage or updates to the model parameters. However, once the model parameters were updated and the learning rate decreased, the evaluation metrics began to stabilize, indicating that the network structure's parameters were able to meet the requirements for the given network structure.

Further training showed more accurate results in the training set, with no signs of overfitting. Among the other evaluation metrics, the Mean Intersection over Union (MIOU) reached 0.86, the Recall reached 0.92, the Precision reached 0.92, and the F1 score reached 0.92. Throughout the training process, it is evident that the model performed well in terms of prediction effectiveness.

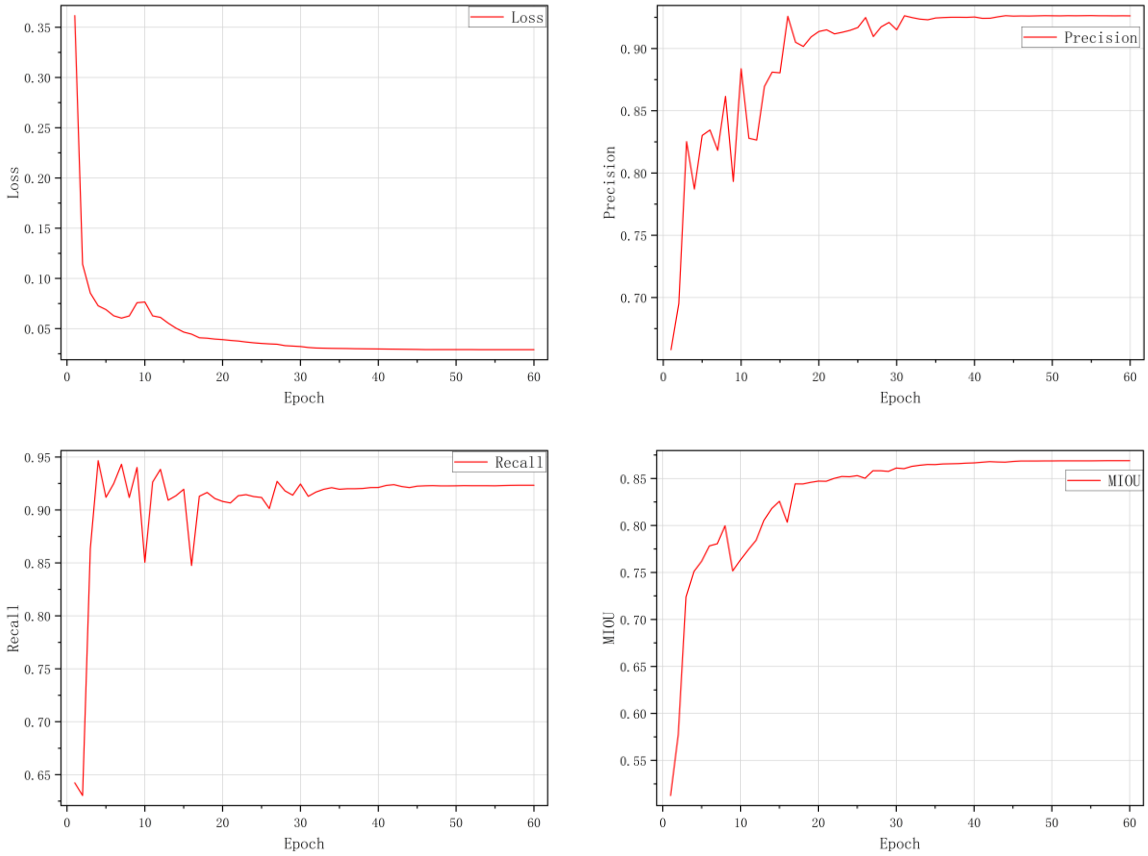

Fig.4 Training information (a) Loss (b) Precision (c) Recall(d) MIOU

1.3 Road Crack Detection Result Analysis

Based on the training results shown in the images, Figure 5 presents the original road surface image, Figure 6 displays the annotated image, and Figure 7 shows the model's predicted results. By comparing these images, it can be observed that the model successfully identifies the main contours of the road cracks in most cases, accurately predicting their locations and shapes, indicating that the model demonstrates strong accuracy in handling road crack detection tasks.

The model performs particularly well in detecting large and prominent cracks, effectively distinguishing crack regions from complex backgrounds. For smaller cracks, although the model exhibits minor edge omissions or discontinuities, these errors do not significantly impact the overall accuracy of crack detection. Furthermore, with subsequent optimization, the model's performance is expected to improve.

Overall, the model shows strong performance in both speed and accuracy when detecting cracks, providing effective support for road maintenance. Despite the presence of minor errors, the model's overall performance is already very close to the requirements for practical applications. Future improvements, such as further optimization and dataset augmentation, will enhance the model's capability to handle a wider variety of crack types and more complex environmental conditions. Therefore, this model holds significant potential for practical applications and widespread use in the field of road crack detection.

Fig.5 Real Images

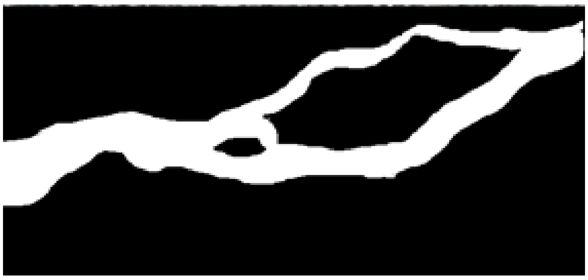

Fig.5 Annotated Images

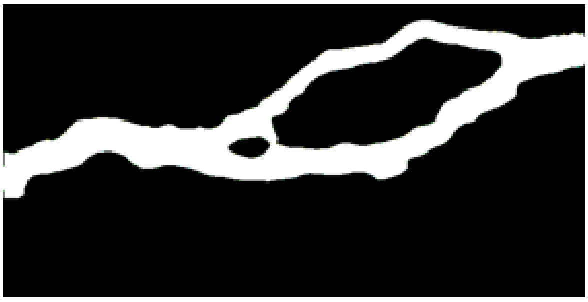

Fig.7 Predicted Images

2 Conclusion

This paper provides a brief overview of the current research status in road damage detection, with a particular emphasis on road cracks as an early manifestation of road deterioration and their importance for road maintenance. In addressing road damage identification, this paper investigates various methods and, based on the current research status, selects a deep learning-based image recognition approach for road crack detection, achieving fast and accurate identification of road surface cracks.

Utilizing the U-Net network architecture, a model for handling road crack detection tasks was constructed. The model was optimized using the training and testing datasets, and the evaluation parameter trends during the training process were analyzed. After testing, the model demonstrated excellent performance in terms of accuracy and speed in crack detection. It showed feasibility on the validation dataset and exhibited strong performance across Precision, Recall, MIOU metrics.

Throughout the research process, we found that while the U-Net network achieved satisfactory results in crack detection, there is still room for improvement in metrics such as Mean Intersection over Union (MIOU). This indicates that the network structure or feature extraction methods may need further optimization to achieve more precise object feature recognition.

3 Prospects

At present, there have been many research in the field of image recognition and classification using deep learning, with various neural network architectures being employed to handle such tasks. This paper focuses solely on the use of the U-Net network architecture for detection. However, due to the simplicity of the network structure, it may not fully capture the feature information of objects, which could limit further improvements in detection accuracy. Many improved methods based on the U-Net architecture already exist, or other network structures are used to process image data, reducing the loss of image information during processing, and enhancing both speed and accuracy for rapid road condition recognition.

In addition to modifying the network structure, the techniques for capturing images have become more advanced. Vehicle-mounted cameras can now quickly capture road conditions, efficiently obtaining clearer images, allowing for the timely detection of road issues that can be addressed promptly. Moreover, with the help of deep learning, more comprehensive datasets can be utilized for training, enabling the classification or grading of various road defects. This provides quicker and more effective information for constructing road damage treatment workflows, ultimately improving the efficiency of road damage remediation and enhancing resource utilization.

References

[1] Niu, K. (2023). Analysis of the causes and prevention measures of highway pavement cracks. Transportation World, Z1, 77-79+82.

[2] Li, H., Wu, Z., Nie, J., et al. (2020). Airport pavement crack automatic detection algorithm based on depth images. Journal of Traffic and Transportation Engineering, 20(06), 250-260. DOI:10.19818/j.cnki.1671-1637.2020.06.022.

[3] Chen, T., Cai, Z., Zhao, X., et al. (2020). Pavement crack detection and recognition using the architecture of SegNet. Journal of Industrial Information Integration, 18, 100144. DOI:10.1016/j.jii.2020.100144.

[4] Guo, C., Gao, W., Zhou, D. (2024). Research on road surface crack detection based on SegNet network. Journal of Engineering and Applied Science, 71(1), 54. DOI:10.1186/s44147-024-00391-0.

[5] Tokhmechi, B., Memarian, H., Noubari, H. A., et al. (2009). A novel approach proposed for fractured zone detection using petrophysical logs. Journal of Geophysics and Engineering, 6(4), 365-373.

[6] Gao, L., Rao, F., Yang, Z., et al. (2024). Research on intelligent recognition technology for road surface cracks based on semantic segmentation. China & Foreign Highway, 44(5), 241-247. DOI:10.14048/j.issn.1671-2579.2024.05.027.