

# Automated Asphalt Crack Detection Using Convolutional Neural Networks

Dr. Kanmani P<sup>1</sup>, Anushka S<sup>2</sup>, Anushka Agarwal<sup>3</sup>

<sup>1</sup>Department of Data Science And Business Systems, SRM Institute of Science and Technology, Chennai, India

<sup>2</sup>Department of Data Science And Business Systems, SRM Institute of Science and Technology, Chennai, India

<sup>3</sup>Department of Data Science And Business Systems, SRM Institute of Science and Technology, Chennai, India

**Abstract** - Road damage caused by asphalt cracks is a significant issue in the civil engineering industry as it poses a threat to road and highway safety. Detecting and classifying cracks is a difficult undertaking because of the intricate pavement conditions created by various factors such as shadows, oil stains, and water spots. These factors can create challenges in differentiating cracks from the surrounding pavement. The focus of our study was to put forward a architecture of a deep convolutional neural network (DCNN) that can automatically detect and categorize pavement cracks. To train DCNN, we utilized RGB images of pavement cracks that were captured manually with a resolution of 1024x768 pixels. These images were then segmented into patches measuring 32x32 pixels. During the training of the DCNN, we employed two filter sizes, which were 3x3 and 5x5. Our presented approach achieved a recall of 98%, precision of 99%, and accuracy of 99%, successfully detecting the presence of cracks in the images. The DCNN was also capable of classifying With fair classification accuracy for both filter sizes and no noticeable difference in accuracy between the two filter sizes, pavement cracks into no cracks, transverse, longitudinal, and alligator. In contrast to bigger filter sizes, smaller filter sizes required greater processing time during training. Overall, 94.5% accuracy was achieved while using our suggested method to classify different kinds of cracks.

**Key Words:** Crack detection, DCNN, Faster R-CNN, Unet, YOLO, ResNet-50

## 1. INTRODUCTION

The construction and thoughtful design of roads have made them the most important component for bridging urban and rural regions, which benefits both social and economic growth [1]. The research on automated crack detection systems has garnered considerable interest from both the university and the industry since they are safer, less expensive, more effective, and more objective [2], [3]. Early research' approaches typically combine or enhance well-known edge detection [6], mathematical morphology [5], and thresholding [4] methods for processing digital images. In a typical road network, many environmental factors are present on thousands of kilometres of pavement, including traffic volumes, temperature variations, and moisture variations, all of which can result in the development of road defects. Asphalt road surface fissures are the most problematic of these flaws. According to Figure 1, the three main divisions of typical forms of fractures are transverse, longitudinal, and alligator cracks. If these warning signals of deterioration are disregarded, potholes may form, increasing

the risk on the road. Early identification and upkeep techniques has been created to track information regarding pavement cracks to be able to improve the functionality and longevity of pavement. The conventional technique for assessing pavement in poor nations is crack analysis by human inspection, or so-called non-computer vision.

Nevertheless, this approach requires a lot of manual labour, is sensitive to human subjectivity, and depends on specialised knowledge. For pavement inspection, automatic crack identification technologies based on computer vision and image processing have taken the lead. Recently published works have proposed deep learning-based methods for autonomous pavement crack identification. For instance, Huang et al. [7] presented a Faster R-CNN and U-Net model for the automatic identification of asphalt pavement fractures. Similarly, Wu et al. [8] proposed a deep learning-based technique for autonomous pavement crack detection and quantification, using photos taken from unmanned aerial vehicles. Additionally, Ravanbakhsh and Saffar [9] introduced a convolutional neural network-based approach for automatic identification of road cracks. These studies demonstrate the efficacy of deep learning-based techniques for autonomous pavement crack identification. A deep convolutional neural network (DCNN) architecture used in this paper is suggested for automatic crack detection and pavement crack categorization. We assess the effectiveness of our suggested technique and contrast it with the outcomes produced by two alternative filter sizes. What follows in the essay is structured as follows: The research approach for part 2 of our study provides a summary. The experimental findings and discussion are presented in Section 3, and the paper's conclusion and recommendations for further research are presented in Section 4.



**Fig. 1** (a) Lateral crack (b) Prolonged crack (c) Crocodile crack

Pavement detection of cracks using image processing is now both affordable and effective. In practise, methods for hand-crafted feature extraction from sub-windows based on intensity thresholding, edge detection, and other image processing techniques are frequently used. The possibility of crack identification using camera-captured images has

increased due to the quick advancement of deep learning techniques when it comes to computer vision. In challenging scenarios, techniques for detecting objects, such as regions-based convolutional neural networks (R-CNN) [10], Faster R-CNN [11], and you only look once (YOLO) [12] have outperformed humans. Although these systems have proven they can reliably detect pavement surfaces, there is always potential for improvement, especially when it comes to classifying various pavement fracture kinds. The inability to extract fracture features from the complex pavement background and shadows, as well as the requirement to extract fractures using low-level picture signals, are some of the limitations of non-computer vision and computer vision technologies. The goal of Malaysian researchers is to develop mass-produced, completely automated, affordable gadgets for monitoring road conditions. Furthermore, no system has ever been developed to categorize known cracks into different types, such as lateral(transverse), prolonged(longitudinal), and crocodile(alligator) cracks. Surveyors can easily benefit from it to use these study findings in real-world contexts right soon. As compared to other computer vision approaches, Deep Convolutional Neural Network (DCNN), developed in 1980, provides the most potent technology and exceptional performance for image classification, picture segmentation, and object recognition. Moreover, DCNN has demonstrated a special aptitude at handling visual input, such as pictures and movies.

Comparisons were made between the proposed DCNN and traditional edge detection techniques using Canny and Sobel in terms of performance. Despite the severity of the background's complexity and inhomogeneity, their investigation performed exceptionally well at categorising pavement cracks. A few studies have also recommended utilising DCNN to detect degradation to the road's surface. This led the study to provide a comprehensive pavement's identification and categorization of cracks strategy based on the deep convolutional fissures in the pavement caused by a neural network and validate network utilising the recently produced dataset of pavement cracks with respect to the process duration.

## 2. METHODOLOGY

The proposed technique will be covered in detail in this section. The three (3) parts of the technique include establishing a pavement crack picture capture system, creating a dataset of pavement cracks for DCNN training, and utilizing DCNN.

### 2.1 Image acquisition technique for pavement cracks

The first stage in producing raw images for the network to use as static images is image acquisition. Using a 16 Megapixel Nikon using a digital camera the optical axis parallel to the surface of the earth, images of pavement cracks were acquired. The crack photos were photographed in broad daylight at a distance of between 70 and 100 cm from the ground. The photos include background variations, noises like

shadows, water and oil stains, and other noises. A total of 4000 RGB photos were randomly chosen from the dataset of image data obtained from different road crossings in the districts of Kedah and Penang, with 1000 images each for no crack, lateral(transverse), prolonged(longitudinal), and crocodile(alligator). To conserve memory and quicken processing, the original photographs' resolution—roughly 3500 x 4500 pixels—was reduced to 1024 x 768 pixels.

#### 2.1 Dataset of Pavement Cracks for DCNN Training

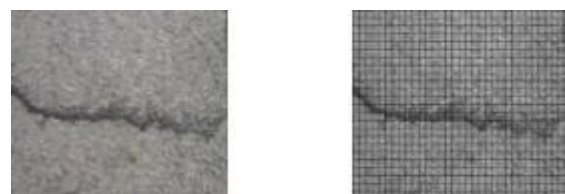
A binary dataset made up of no crack, lateral(transverse), prolonged(longitudinal), and crocodile(alligator), as well as a dataset with cracks and those without cracks, were used as the training and testing sets for the suggested technique

#### 2.2.1 Dataset for Training Crack and Non-Crack

The whole collection of collected photos need not be trained to create a training datasets for crack and non-crack. For this project, a patch grid with a size of 32x32 pixels was used to divide images with a 1024x768 pixel resolution into 768 patches for each picture. As a result, 400 RGB photos (10% of the entire image capture) were used to create a patches with and without cracks. Out of the 307,200 patches received from the grid scale, there were 282,624 patches without cracks and 24,576 crack patches. Datasets for both training and testing were made up within these patches since the input picture was segmented. A total of 9000 patches, comprising 4500 crack patches and 4500 non-crack patches, were selected for the training dataset.

#### 2.2.2 Dataset for training for No Crack, Lateral(or transverse), Longitudinal, and Crocodile(or alligator)

Manually gathered RGB pictures were used to train the DCNN for the categories of no crack, transverse, longitudinal, and alligator pictures, resulting in 5700 binary pictures in the training and 460 in the testing. Using a starting learning rate, the default optimizer was employed, of 0.001, and 10 batches were created for each iteration. Layers of convolution with filter sizes of 3x3 and 5x5 were incorporated in the network architecture. The network was trained for 500 epochs, and the accuracy was measured using the confusion matrix.



**Fig. 2** (a) Original photo (b) 32x32 grid size for the image



**Fig. 3** (a) A crack patch example (b) Illustration of non-patches

No of image	No of patch	Crack	Non-crack
1	768	66	702
400	307,200	24,576	282,624

**Table 1.** Number of patches for crack and non-crack

	Crack	Non-crack
Training set	4500	4500
Testing set	500	500

**Table 2.** Number of training and testing dataset

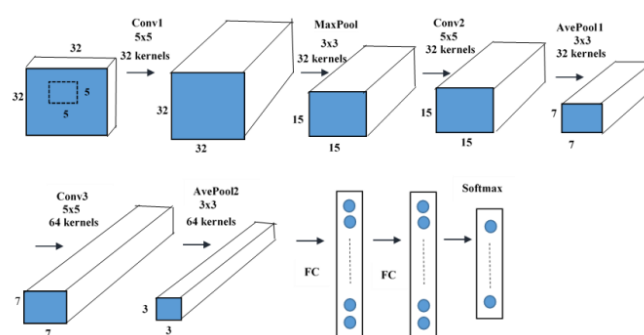
improve the quality of the data. These steps included resizing the images to a uniform size of 224 x 224 pixels and normalizing the pixel values between 0 and 1. This ensures that the input images are consistent and comparable, and also helps to reduce the impact of lighting and color variations in the images.

### 2.3.1 The Network Architecture

The ResNet-50 architecture was chosen by the authors for crack detection and classification due to its capability of effectively representing complex features using a deep neural network with 50 layers. For the task, the DCNN was trained using transfer learning, starting with a pre-trained ResNet-50 model that was fine-tuned for the specific purpose. The model's last layer, which was initially designed for ImageNet classification, was substituted with a new fully connected layer to classify cracks.

### 2.3.2 Training

The DCNN was trained over 50 epochs using batches of 570 pre-processed pictures, Adam's optimizer, and a starting learning rate of 0.001. Ten batches of the training dataset were created, and Figure 4 shows the suggested DCNN's design.



**Fig.4** Proposed DCNN architecture

Figure 4 depicts a DCNN consisting of three levels of convolution, three layers of pooling, and two layers that are totally connected. The network receives input data from both patches with and without cracks using a three-channel configuration. The 3x3 or 5x5 filters used in the convolution layers Conv 1, Conv 2, and Conv 3 have steps. Filters larger than 7x7 are avoided to minimise training issues. Max pooling (MaxPool) and average pooling (AvePool) are employed over a 3x3 window with a 2 stride in the proposed network's convolutional layers. Additionally, a pooling layer is a critical element of the DCNN. The last layer outputs were flattened and fed into several fully connected (FC) layers to anticipate detecting pavement in two categories: crack and non-crack, as well as pavement crack classification into crocodile, lateral and prolonged cracks.

The rectifier linear unit follows the convolution layer (ReLU) is a nonlinear activation function used to handle nonlinearity for multi-classification. As recall, precision, and accuracy were calculated for performance indicator to evaluate how well the DCNN classified the pavement patches. The percentage of the actual fracture grid that the DCNN accurately categorised is known as recall. The proportion of correctly recognised crack pixels in relation to the total number of cracks in the dataset may be used to measure precision. The percentage of accurately recognised cracked and uncracked instances in relation to the overall number of cracked and uncracked instances in the dataset is known as accuracy.

## 3. EXPERIMENTAL RESULTS

The trained DCNN was used to categorize 32x32 patches of input images as cracked or non-cracked, with filter sizes of 3x3 and 5x5 applied to 1000 testing datasets. The results of the network's recall, precision, and accuracy, displayed in Table 3, indicate that the network's performance was consistent for both filter sizes. Subsequently, the same DCNN architecture was applied to a fresh testing dataset to categorize 360 test pictures into three categories: lateral, prolonged, and crocodile cracks, following the cracked and uncracked detection. Table 4 summarizes the network's overall lateral, prolonged, and alligator performances, indicating high-accuracy classification using both 3x3 and 5x5 filter sizes. Table 5 presents the DCNN network's precision, accuracy, and pavement classification performance for pavement identification. These methods

perform well for pavement detection in terms of precision, accuracy, and pavement classification (A) concerning transverse (T), longitudinal (L), and alligator. According to Table 4, Paul et al. [13] obtained precision and accuracy results of 91.9% and 90.2%, respectively, for cracked and uncracked detection using 500 RGB images of pavement as their test dataset.

Filter size	Transverse	Longitudinal	Alligator
3x3	97.1%	97.4%	99.0%
5x5	97.0%	97.1%	99.0%

**Table 3.** Three type of crack classification with different filter size

Method	Crack detection		Crack classification		
	Pr	Acc	T	L	A
Pauly et al., 2017	91.9%	90.2%	-	-	-
Cha, Choi, & Büyükoztürk, 2017	-	97.9%	-	-	-
X. Wang & Hu, 2017	-	-	97.6%	97.2%	90.1%
Our method	99.1%	99.0%	97.0%	97.1%	99.0%

**Table 4.** Comparison result using DCNN architecture

Cha et al. [14] produced a training dataset of 40K images using 332 photos and a testing dataset of 55 images using 55 photos using the same 332 photos. Their training on CNN had a 97.9% success rate in differentiating between crack and non-crack. With only 9K training photographs and 1000 testing images, the network was able to achieve 99.1% and 99.0% accuracy and precision when compared to our recommended technique. Utilizing 310 test images, X. Wang et al. [15] classified transverse, longitudinal, and alligators accurately 97.6%, 97.2%, and 90.1% of the time, respectively. The practical findings of the proposed study for classifying into three classifiers that provide promising results of 97.0%, 97.1%, and 99.0% for lateral, prolonged, and crocodile, respectively, were accomplished utilizing just 4.5K training images and a filter size of 5x5.

In order to accurately describe the situation to the surveyors, additional categories must be provided. As a result, crack classification also applied to pavement without cracks.

Filter size	Transverse	Longitudinal	Alligator	No crack
3x3	91.2%	90.0%	99.0%	98.0%
5x5	91.0%	90.0%	99.0%	98.0%

**Table 5.** Four type pavement classification with different filter size

According to the results presented in Table 6, the accuracy of classifying lateral and prolonged cracks was depreciated compared to other crack categories. It occurred when the

DCNN failed to discriminate between longitudinal and transverse cracks in pavement without fractures. One possible reason for this lower accuracy could be the complexity of the background in the images, which includes noise, oil stains, and water spots. These elements are similar in intensity and contrast to the darker backgrounds, making it more difficult for the DCNN to distinguish between cracked and non-cracked pixels.

Experiment 3					
Method	NC	T	L	A	Overall
Li et al., 2018	93.0%	97.7%	97.3%	89.3%	94.3%
Our method	98.0%	91.0%	90.0%	99.0%	94.5%

**Table 6.** Comparison result using DCNN with different filter size

Table 6 presents a summary of classifying pavement cracks according to latest studies. Overall, Li et al. [20] accomplished excellence by utilising similar 5x5 measurement of the filter for the classification of no crack, lateral, prolonged, and crocodile types, with an overall accuracy of 94.3%. The suggested approach in this study acquired an accuracy of 94.5%, a little bit higher than Li et al. [16]. With a minimal training dataset (as shown in Table 8), the suggested approach performs with consistency for classifying cracks in pavement into four groups.

Method	Dataset	NC	T	L	A
Li et al., 2018	Training	5552	5184	5419	5445
	Testing	300	300	300	300
Our method	Training	1200	1500	1500	1500
	Testing	100	130	130	100

**Table 7.** Comparison training and testing dataset

## 4. CONCLUSIONS

In conclusion, this study proposes an automated method for detecting and categorizing pavement cracks using DCNN. The model effectively learns four different types of cracks using training datasets of 5.7K and 9K for four and two classes, respectively. The analysis shows that filter size has an impact on training duration but not on classification performance, with a 5x5 filter size being the best. However, noisy backgrounds in pavement images remain a challenge for crack identification, requiring further research to improve the algorithm's ability to categorize transverse and longitudinal cracks and reduce noise interference in pavement images. Overall, the method shows promise for autonomous pavement recognition but requires further refinement to improve its accuracy and effectiveness. The presence of patterns on pavement surfaces, for example, shadow, oil stains, and water spots, can create detection of cracks challenging because these patterns often have a larger contrast and a similar intensity to minor cracks. This can make it difficult to differentiate between the two. As a result, further research is necessary to develop a method that can classify

pavement cracks into four groups, especially for lateral and prolonged cracks, by improving pixel algorithms. This could aid in the management of pavements with noisy patterns. Due to the presence of noisy patterns like shadows, oil stains, and water spots on concrete surfaces, finding fractures can be challenging. Most of these patterns have a similar greater intensity and contrast than small cracks, making it challenging to distinguish them [17].

## REFERENCES

1. Guan H Li J Yu Y Chapman M and Wang C 2015 Automated Road Information Extraction From Mobile Laser Scanning Data IEEE Trans. Intell. Transp. Syst. 16, 1 p. 194–205.
2. 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.
3. D. Zhang, Q. Li, Y. Chen, M. Cao, L. He, and B. Zhang, "An efficient and reliable coarse-to-fine approach for asphalt pavement crack detection," Image and Vision Computing, vol. 57, pp. 130–146, 2017.
4. H. Oliveira and P. L. Correia, "Automatic road crack segmentation using entropy and image dynamic thresholding," in Proc. Eur. Signal Process. Conf., 2009, pp. 622–626.
5. N. Tanaka and K. Uematsu, "A crack detection method in road surface images using morphology," in Proc. Workshop Mach. Vis. Appl., vol. 98, 1998, pp. 17–19.
6. H. Zhao, G. Qin, and X. Wang, "Improvement of canny algorithm based on pavement edge detection," in Proc. Int. Conf. Image, Signal process., vol. 2. IEEE, 2010, pp. 964–967.
7. Lins R G and Givigi S N 2016 Automatic Crack Detection and Measurement Based on Image Analysis IEEE Trans. Instrum. Meas. 65, 3 p. 583–590.
8. Madli R Hebbar S Pattar P, and Golla V 2015 Automatic detection and notification of potholes and humps on roads to aid drivers IEEE Sens. J. 15, 8 p. 4313–4318.
9. Adu-Gyamfi Y Kambhamettu C and Okine N A 2013 Performance Assessment of Flexible Pavements Using Active Contour Models Airf. Highw. Pavement 2013 Sustain. Effic. Pavements p. 887–902.
10. Girshick, R., et al. Rich feature hierarchies for accurate object detection and semantic segmentation. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.
11. Ren, S., et al. Faster r-cnn: Towards real-time object detection with region proposal networks. in Advances in neural information processing systems. 2015.
12. Redmon, J. and A. Farhadi, Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.
13. Pauly L Peel H Luo S Hogg D and Fuentes R 2017 Deeper Networks for Pavement Crack Detection p. 479–485.
14. Cha Y J Choi W and Büyüköztürk O 2017 Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks, Comput. Civ. Infrastruct. Eng. 32, 5 p. 361–378.
15. Wang W and Hu Z 2017 Grid-Based Pavement Crack Analysis Using Deep Learning, in 2017 4th International Conference on Transportation Information and Safety, ICTIS 2017 - Proceedings p. 917–924.
16. Li B Wang K C P Zhang A Yang E, and Wang G 2018 Automatic Classification Of Pavement Crack Using Deep Convolutional Neural Network Int. J. Pavement Eng. 0, 0 p. 1–7.
17. Chen F C and Jahanshahi M R 2018 NB-CNN: Deep Learning-Based Crack Detection Using Convolutional Neural Network and Naïve Bayes Data Fusion, IEEE Trans. Ind. Electron. 65, 5 p. 4392–4400.
18. Yusof, N. a. M., Ibrahim, A., Noor, M. H. M., Tahir, N. M., Yusof, N. M., Abidin, N. Z., & Osman, M. K. (2019). Deep convolution neural network for crack detection on asphalt pavement. Journal of Physics: Conference Series, 1349(1), 012020.