Multi-layers deep learning model with feature selection for automated detection and classification of highway pavement cracks

Faris Elghaish
Faculty of Engineering and Physical Sciences, Queen’s University Belfast, Belfast, UK

Sandra Matarneh
Department of Civil Engineering, Al-Ahliyya Amman University, Amman, Jordan

Essam Abdellatef
Delta Higher Institute for Engineering and Technology, Mansoura, Egypt

Farzad Rahimian
Teesside University, Middlesbrough, UK and
Department of Engineering, Teesside University, Middlesbrough, UK

M. Reza Hosseini
Department of Architecture and Built Environment, Faculty of Science Engineering and Built Environment, Deakin University, Geelong, Australia, and

Ahmed Farouk Kineber
Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia

Abstract

Purpose – Cracks are prevalent signs of pavement distress found on highways globally. The use of artificial intelligence (AI) and deep learning (DL) for crack detection is increasingly considered as an optimal solution. Consequently, this paper introduces a novel, fully connected, optimised convolutional neural network (CNN) model using feature selection algorithms for the purpose of detecting cracks in highway pavements.

Design/methodology/approach – To enhance the accuracy of the CNN model for crack detection, the authors employed a fully connected deep learning layers CNN model along with several optimisation techniques. Specifically, three optimisation algorithms, namely adaptive moment estimation (ADAM), stochastic gradient descent with momentum (SGDM), and RMSProp, were utilised to fine-tune the CNN model and enhance its overall performance. Subsequently, the authors implemented eight feature selection algorithms to further improve the accuracy of the optimised CNN model. These feature selection techniques were thoughtfully selected and systematically applied to identify the most relevant features contributing to crack detection in the given dataset. Finally, the authors subjected the proposed model to testing against seven pre-trained models.

Findings – The study’s results show that the accuracy of the three optimisers (ADAM, SGDM, and RMSProp) with the five deep learning layers model is 97.4%, 98.2%, and 96.09%, respectively. Following this, eight feature selection algorithms were applied to the five deep learning layers to enhance accuracy, with particle swarm optimisation (PSO) achieving the highest F-score at 98.72. The model was then compared with other pre-trained models and exhibited the highest performance.

Practical implications – With an achieved precision of 98.19% and F-score of 98.72% using PSO, the developed model is highly accurate and effective in detecting and evaluating the condition of cracks in pavements. As a result, the model has the potential to significantly reduce the effort required for crack detection and evaluation.

Originality/value – The proposed method for enhancing CNN model accuracy in crack detection stands out for its unique combination of optimisation algorithms (ADAM, SGDM, and RMSProp) with systematic
application of multiple feature selection techniques to identify relevant crack detection features and comparing results with existing pre-trained models.

**Keywords** CNN, Highway surface cracks, Feature selection, Optimisation algorithms, Particle swarm optimisation (PSO)

**Paper type** Research paper

1. Introduction

Pavement distress, notably in the form of cracks, represents a prevalent phenomenon. Due to the influence of vehicle load, environmental conditions and other factors, cracks occur to varying degrees in the pavement’s surface and base layers. Failure to promptly detect and address these cracks can compromise the comfort and safety of road users. Moreover, it can exert adverse effects on the overall performance and anticipated service life of the road infrastructure (Beckman et al., 2019). Thus, the accurate detection of pavement cracks provides useful information for the maintenance strategy to extend the pavement’s service life. Traditional crack detection methods generally rely on manual inspection, which is tedious, expensive, time consuming, inconsistent, and potentially dangerous (Wang et al., 2017; Yang et al., 2020). Due to human subjectivity, efficiency and reliability are two key concerns about the use of manual inspection. Moreover, these two key concerns are affected by the substantial variability found in routine inspections (Dais et al., 2021).

With the rapid advancement of artificial intelligence (AI), deep learning (DL)-based crack detection techniques have gained increased attention as a solution to the problems experienced with manual inspection. A common deep learning technique for computer vision applications is the convolutional neural network (CNN). This has attained state-of-the-art and human-competitive performance in image classification, semantic segmentation, object detection and other areas (LeCun et al., 2015; Dais et al., 2021).

In contrast to machine learning techniques, CNN is capable of automatically extracting multi-level feature representations without the need for manual rule creation (Liu et al., 2019). Various studies have found that the CNN is becoming a more widely used and effective technique for detecting cracks (Gopalakrishnan et al., 2017; Mandal et al., 2018; Liu et al., 2019; Huyan et al., 2020; Majidifard et al., 2020; Liu and Wang, 2022).

However, studies using CNNs generally need substantial amounts of well-collected data to train a network. If the amount of training data is not enough, underfitting will probably occur, leading to inaccurate results (Hongjo et al., 2018). To avoid any possible underfitting, transfer learning methods focus on transferring knowledge from earlier trained CNNs (Yang et al., 2020). AlexNet is a well-trained CNN, based on more than a million images from the ImageNet database (Krizhevsky et al., 2017). It can classify images into 1,000 object categories based on eight layers, the first five being convolutional layers and the last three being fully connected layers.

Dorafshan et al. (2018) showcased the superiority of utilising the AlexNet network within deep CNNs across various modes, including fully trained, transfer learning, and classifier modes. The AlexNet CNN architecture achieved an accurate detection rate of 86% for cracked images and was capable of detecting cracks coarser than 0.04 mm. Likewise, Kim and Cho (2018) devised an automated crack detection method based on a CNN, specifically utilising AlexNet in conjunction with a probability map. The training set was divided into the following five classes: cracks, intact surfaces, two types of similar patterns of cracks, and plants. The proposed detection method presented a prediction accuracy of 90%. Li and Zhao (2019) presented crack detection based on the CNN architecture of binary-class outputs by modifying AlexNet, with their method able to be used by a smartphone application. The method was trained in 205 images with resolution of $3120 \times 4160$ pixels.
In a study conducted by Li et al. (2021), pavement cracks were classified using a CNN architecture in conjunction with the K-means clustering algorithm, aiming to achieve unsupervised feature learning. The base CNN model employed for constructing the architecture was AlexNet. The results indicated the superiority of the proposed method, with average accuracies of 0.806 for transverse cracks, 0.792 for longitudinal cracks, and 0.913 for alligator cracks. In another study, Santos et al. (2022) introduced a technique for the automated detection of exposed steel rebars. This approach relied on region-CNNs (R-CNNs) built upon the CNN model, AlexNet. The training process resulted in an impressive accuracy of 99.1%.

To automatically detect cracks in hot-mix asphalt (HMA) and Portland cement concrete (PCC)-surfaced pavement images, researchers have explored various approaches. Gopalakrishnan et al. (2017) developed a pre-trained VGG-16 CNN with a single-layer neural network classifier based on deep transfer learning vectors. Rubio et al. (2019) devised a fully convolutional network (FCN) using a dense CNN algorithm without fully connected layers, using a VGG-16 network as the backbone, for the detection of delamination and exposed steel rebars. Their approach achieved mean accuracies of 89.7 and 78.4%, respectively. Dung et al. (2019) adopted a transfer learning approach with the pre-trained VGG-16 model, achieving 94% accuracy. A slight fine-tuning of a well-trained fully convolutional layer using the VGG-16 network resulted in an impressive 98% accuracy (Deng et al., 2009; Dung et al., 2019; Zhu and Song, 2020).

Many studies have focused on utilising a single algorithm to detect cracks in pavements and assess its performance. Consequently, there arises a need for a study that evaluates the feasibility of employing various deep learning optimisers and explores the potential of existing tools, such as feature selection algorithms, to enhance performance. Additionally, the evaluation of performance in comparison to existing solutions or other CNN algorithms can provide valuable insights for future researchers when selecting algorithms for different crack datasets.

In this context, the current study introduces an optimised CNN model-based feature selection approach for the detection of various types of highway surface cracks. The CNN model was developed in sequential stages with the aim of maximising its accuracy. To optimise the CNN model, three optimisation algorithms were applied, namely, adaptive moment estimation (ADAM), stochastic gradient descent with momentum (SGDM), and RMSProp. Subsequently, the study determined the highest accuracy scores for detecting three different types of cracks: horizontal, vertical, and diagonal. The final stage involved the application of eight distinct feature selection methods to further enhance the performance of the created CNN predictive model.

Results showed that particle swarm optimisation (PSO) could enhance the performance of the CNN detection model to achieve accuracy (98.19%) and precision (99.02%) and that this could be considered as highly reliable deep learning-based detection model. The proposed PSO-based Optimised CNN for Highway Cracks (OCNNH) model was tested against other pre-trained CNN models, and the F-score variation showed an approximate 19.84% improvement over AlexNet. This demonstrates that the proposed model offers a reliable tool for utilising deep learning in crack classification and detection. Moreover, the outcome of this paper contributes to the body of knowledge by presenting a comparative analysis of the accuracy achieved using eight pre-trained deep learning models and the level of accuracy achieved by combining feature selection optimisation with deep learning for crack detection in highway pavements.

Given that the proposed optimised PSO-based OCNNH model achieves an accuracy of 98.19%, it can assist industry practitioners in automatically detecting and evaluating cracks in long-distance highways without the need for manual detection, which consumes significant time and resources.
The remainder of the paper is structured as follows: Section 2 presents the related works, Section 3 discusses the methodology, Section 4 provides the results analysis and evaluation, Section 5 discusses the findings, and Section 6 concludes the paper.

2. Related works
Transfer learning in deep learning has been extensively implemented in different approaches. To detect and classify various types of cracks in asphalt pavements, Tran et al. (2021) labelled and trained asphalt images using an updated version of a Faster R-CNN called RetinaNet. Results indicated that the detection and classification accuracy of the trained model was 84.9% when considering both the crack type and severity level. Wu et al. (2021) employed DenseNet and a deconvolution network framework to achieve pixel-level detection and fuse features. These features were learned from different scales of convolutional kernels through a full convolutional network to learn the crack characteristics in the complex fine-grained background of asphalt pavement. The research findings demonstrated that the suggested method achieved outstanding crack segmentation across 12 different types, showcasing an enhancement in asphalt pavement crack segmentation when compared to the most cutting-edge techniques. Another showcase study conducted by Teng and Chen (2022) presented a pixel-level segmentation CNN called DeepLab_v3+ to segment cracks and calculate physical properties (the length and width) of cracks based on segmentation results. Results clarified that DeepLab_v3+ achieved accuracy and F-score values of 0.80%, 97.5 and 0.78%, respectively. A study adopting the Parallel ResNet was conducted by Fan et al. (2022) who developed a deep residual CNN (Parallel ResNet) to create a high-performance pavement crack detection and measurement system. The Parallel ResNet reached maximum scores in precision (94.27%), recall (92.52%) and F-score value (93.08%). A study conducted by Li et al. (2022) developed a vision-based pavement crack detection strategy to provide a Pavement Surface Condition Index (PCI) for asphalt pavement. The authors employed a combination of a Convolutional Neural Network (CNN) algorithm and a Genetic Algorithm (GA) to enhance the accuracy of the network in classifying pavement crack types in a dataset containing 5,000 pavement distress images. Accuracy in classifying different types of cracks reached 98%, with processing time for an image of 0.047 s.

A more accurate and efficient crack detection process based on a spatial channel hierarchical network, with a base VGG-19 network, was proposed by Pan et al. (2020). An FCN model with a VGG-19 network was employed by Yang et al. (2018) to simultaneously detect and quantify various cracks at the pixel level, for which performance was examined in 800 images with the number of pixels varying from 1 to 100. A VGG-19 network was used to generate candidate crack regions based on multi-scale-enhanced-Faster R-CNN (MSE-Faster R-CNN) to evaluate internal cracks in turbine blade thermal barrier coating (Shi et al., 2022). The presented results indicated that the proposed approach could accurately locate (0.898) and detect (0.806) of cracks at different scales.

The above-mentioned CNN models were mainly built based on AlexNet and VGGNet which involve a substantial number of parameters and amount of memory, as well as a large data set to realise good performance. To address this issue, new models are introduced to improve CNNs’ performance and efficiency such as.

2.1 GoogLeNet model
A pre-trained deep CNN model, namely, GoogLeNet, was employed for crack detection using hybrid images (Jang et al., 2019). The proposed approach was validated using a laboratory-scale concrete specimen with cracks of various sizes. Results revealed that macrocracks and microcracks were successfully detected using hybrid images. Wu et al. (2021) developed an
approach based on a CNN using the pre-trained GoogLeNet Inception V3 for crack detection. Experimental results demonstrated that the rate of accuracy of the trained model on the test data set reached 98.5%.

2.2 YOLO model
Other studies detected cracks using the “you only look once” (YOLO) algorithm: these studies included the work of Tan et al. (2021) who developed an improved CNN-based YOLOv3 method for automatic detection of sewage pipe defects. The main improvements focused on loss function, data augmentation, bounding box prediction and network structure. Results indicated that the developed model outperformed Faster R-CNN and YOLOv3, reaching a mean average precision (mAP) value of 92%, which was higher than the findings of existing research on the automatic detection of sewage pipe defects. Lu et al. (2022) constructed an intelligent detection method for surface defects of ceramic tiles based on an improved YOLOv5 algorithm and the multiple sliding windows detection method. Similarly, Yao et al. (2022) constructed a pavement crack detection model based on the YOLOv5, the spatial and channel squeeze and excitation (SCSE) module and the convolutional block attention module (CBAM) to develop 12 different attention models to detect cracking. The study’s results revealed the model could process images at 13.15 ms/pic while maintaining 94.4% precision.

2.3 U-net model
To detect cracks in pavement surfaces, U-Net was adopted by several scholars (Jenkins et al., 2018; König et al., 2019). For example, Liu et al. (2019) employed U-Net to detect cracks in concrete. The trained U-Net could detect cracks in raw images under different conditions (e.g. brightness, disturbed background) with high efficacy. The precision of the model, which was trained by 57 images, could reach 0.9 for different complex situations. A deep learning algorithm-based U-Net and a CNN with alternately updated clique (CliqueNet), called U-CliqueNet, was proposed to separate cracks from the background in tunnel images (Li et al., 2020). The proposed method could separate cracks from images with noise-like cracks, such as patchwork joints, and wires. The proposed model was trained on an extensive data set consisting of 50,000 and 10,000 images for training and testing, respectively, and achieved promising results. The mean pixel accuracy (MPA), mean intersection over union (MIoU), precision and F-score value were 92.23%, 86.96%, 86.32% and 83.40%, respectively. Huyan et al. (2020) developed CrackU-net which used convolution, pooling, transpose convolution and concatenation operations, forming the “U”-shaped model architecture. Results showed that CrackU-net had the following values: performance of loss = 0.025; accuracy = 0.9901; precision = 0.9856; recall = 0.9798; and F-score value = 0.9842, with a learning rate of $10^{-2}$. Qu et al. (2022) conducted another study in which a CNN and a transformer named CrackT-net were proposed for crack segmentation. To enhance feature representation capabilities, the authors used richer features (RF) UNet++ and polarised self-attention. In addition, they replaced the last feature extraction layer by the transformer. Results showed that the proposed method proved its effectiveness with F-score values of 0.856, 0.700 and 0.637 on three data sets.

2.4 GAN model
Recent studies implemented the deep learning technique along with a generative network. Mazzini et al. (2020) presented a method based on CNNs for data augmentation in the context of semantic segmentation of highly textured images. In this research, the authors employed a Generative Adversarial Network (GAN) to create a semantic layout and a texture synthesiser that relied on a Convolutional Neural Network (CNN) to generate a new image following the
The study involved an assessment of their proposed method using the German Pavement Distress dataset, revealing substantial performance enhancements, especially in scenarios involving low cardinality classes, when training CNNs with augmented datasets as opposed to the original datasets. A virtual image set generation method for asphalt pavement cracks was proposed by Pei et al. (2021), based on improved deep convolutional generative adversarial networks (DCGANs) to address the issue of the small size of crack images during intelligent road detection. The results showed that the augmented data set of the proposed method with this detection model had an average precision of 90.32%, with this being higher than that of the conventional method when evaluated using the same test data set.

To address the problem in existing segmentation methods of the loss of edge information for the damaged area, Dong et al. (2022) proposed a feature fusion model combined with a StyleGAN, called Road-Seg-CapsNet, for the segmentation of multiple complex forms of damage to asphalt pavements. A StyleGAN was first used to amplify the data set and then padding convolution was used in the convolution layer of Road-Seg-CapsNet to retain more image edge information. Results reported that the segmentation of the proposed model could reach 0.942, with the minimum accuracy for the damaged areas found to be 0.903. Another detection method using small samples, as proposed by Xu and Liu (2022), combined a generative adversarial network (GAN) and a CNN model. After the data set was trained and tested by the transfer learning method to separate verify the effectiveness of the expanded data, detection accuracy was found to have been improved from 80.75% to 91.61%, which considered a substantial improvement.

2.5 Comparison between various CNN models

Different studies have compared various CNNs and have reported different results. Sharma et al. (2020) provided a comparison between the AlexNet network with GoogLeNet and ResNet18 in transfer learning mode. Their results showed that GoogLeNet and ResNet18 introduced significant improvement in the case of images of bridge deck and walls but achieved little improvement in the case of pavement images. Augustauskas and Lipnickas (2020) employed a CNN with different architectures (U-Net; ResU-Net; ResU-Net + ASPP; ResU-Net + ASPP_Waterfall (WF); ResU-Net + ASPP + AG; and ResU-Net + ASPP_WF + AG). Their study concluded that the utilised neural network configurations showed a segmentation performance improvement over U-Net, but no significant improvement using the different architecture of other neural networks.

In another study, three kinds of deep neural networks, AlexNet, ResNet18 and VGGNet13, were compared, with findings showing that ResNet18 (accuracy 98.8%) performed well compared to the other two models (Yang et al., 2021). Five deep learning models (i.e. VGG-16, VGG-19, Resnet-50, Inception-V3 and a proposed customised CNN model) were evaluated for the detection and localisation of cracks in concrete structures using eight data sets (Ali et al., 2021). The findings indicated that initially, all models exhibited strong performance when tested with a small and varied training dataset. Nevertheless, as the training dataset’s size grew and its diversity decreased, the ability to generalise and prevent overfitting decreased. The results highlighted that the suggested customised Convolutional Neural Network (CNN) - tailored to either spatial or sequential characteristics and employing an ADAM optimiser - along with the VGG-16 models, excelled in detecting and pinpointing cracks in concrete structures (Ali et al., 2021). In their 2021 study, S. Y. Wang and Guo employed a transfer learning approach to train models. They utilised three cutting-edge Convolutional Neural Networks (CNNs): VGG-16, ResNet-101, and a Feature Pyramid Network (FPN) to extract features, while adopting a Faster R-CNN as the backbone for developing models focused on full convolutional instance segmentation (FCIS). Among these three models, ResNet
showcased the best accuracy and the most efficient training cost. Furthermore, the performance of the FPN closely trailed ResNet’s results, excelling particularly in terms of recall (Wang and Guo, 2021). Pozzer et al. (2021) investigated the performance of different deep neural network models to detect the main anomalies in concrete, including delamination, cracks, spalling and patches in thermographic and regular images captured from various distances and viewpoints. The results showed that the MobileNetV2 had promising performance in identification of multiclass damages in the thermal images, identifying 79.7% of the total delamination, cracks, spalling and patches on test images of highly damaged concrete areas. The VGG-16 model showed better precision by reducing the number of false detections. Several pre-trained models (AlexNet, GoogLeNet, SqueezNet, ResNet-18, ResNet-50, ResNet-101, DenseNet-201 and Inception-v3) have been employed in retraining, based on pavement images, using transfer learning. Five performance metrics were used to assess and compare the efficiency of crack detection models, including accuracy, sensitivity, specificity, precision, and F-score values. According to performance metrics, SqueezNet and GoogLeNet generally have better performance than the other models (Ranjbar et al., 2021). Hou et al. (2021) developed a CNN structure model, called MobileCrack, for fast object classification in cracked pavements, including pavement matrix, pavement marking, cracks and sealed cracks. Compared with the three common CNN models (AlexNet, VGG-16 and MobileNet), MobileCrack could achieve higher accuracy (0.9375 after dropout).

Elghaish et al. (2022) developed a CNN model for detection and classification of highway cracks by testing four pre-trained CNN models (i.e. AlexNet, VGG-16, VGG-19 and GoogLeNet). Results indicated that the accuracies of all pre-trained models were higher (97.72%) by more than 5% than the averages and calculated accuracies for AlexNet and GoogLeNet models. A two-step data pre-processing method was designed by Hou et al. (2022b) to decrease the bias of the CNN model. The study first applied the data augmentation method to enlarge the data set. A crack extraction method was then applied to convert the original image into a binary black-and-white image. The study explored the application of AlexNet, SE-Net and ResNet with a variety of configurations: results indicated that ResNet with 50 layers had the highest test accuracy. Tang et al. (2022) developed the iteratively optimised patch label inference network (IOPLIN) framework and compared the proposed framework with CNN models such as GoogLeNet and EfficientNet. Results provided evidence of IOPLIN’s superiority over the state-of-the-art image classification approaches in automatic pavement distress detection. To advance automation in brick segmentation and crack detection in masonry walls, Loverdos and Sarhosis (2022) assessed several deep learning networks (U-Net, DeepLabV3+, U-Net, LinkNet and FPN). The study’s results indicated that, for brick segmentation, the use of deep learning provided better outcomes than typical image-processing applications. Hou et al. (2022a, b) developed a deep learning method named FS-Net by integrating the FReLU structure and the strip pooling method for detecting pavement distress. The results revealed that the average accuracy of the proposed method was 4.96% higher than the Faster R-CNN and 3.67% higher than the YOLOv3 networks.

Yang et al. (2020) used a VGG-16 network as a base model for transfer learning to detect cracks. The detection accuracy for the three data sets was 99.83%, 99.72 and 97.07%. The VGG-16 network was used as the backbone network, with the dilated convolution added to the backbone network to track slab cracks. Experimental results showed that the proposed method achieved 81.84%, 67.68 and 84.55% in precision, Intersection over Union (IoU) and F-score value, respectively, in the self-made data set (Li et al., 2021). Malini et al. (2021) presented a Faster R-CNN with a modified VGG-16 network for automatic assessment of road cracks. The detections were made with 62.3% mean average precision @ Intersection over Union (IoU) ≥ 0.5 for the generation of 300 region proposals which is a good score for object detections. Xu et al. (2022) developed a network based on an encoder–decoder structure and employed a VGG-16 network as the basic feature extraction network for concrete crack classification of highway pavement cracks.
segmentation. Feature activation visualisation proved that the proposed method had solid linear topology structure capture capability. Guzmán-Torres et al. (2022) tested several CNN models to find the one that best performed concrete crack detection. Based on their investigation, the authors proposed a new transfer learning method to improve the accuracy of a micro-crack and macro-crack deep learning classifier. The implemented method’s architecture was based on an improved VGG-16 model and reached an accuracy of 99.5%, and an F-score value of 100%.

Recent research has introduced the utilisation of pre-trained CNNs in the realm of crack detection and categorisation. It’s noteworthy that only a limited number of studies have investigated the performance of different pre-trained CNN models and very limited studies have explored the integration of various optimisers to enhance the pre-trained CNNs performance in crack detection and classification (Samma et al., 2021; Elghaish et al., 2022; Islam et al., 2022; Liu et al., 2022; Li et al., 2023; Qayyum et al., 2023). In Table 1, an overview of various studies employing transfer learning techniques for crack detection and classification is provided. While TL methods have demonstrated considerable success in crack classification, a critical examination of existing studies has unearthed certain limitations. Notably, only a few studies have assessed multiple pre-trained CNNs, and a comprehensive assessment of distinct pre-trained CNN models is essential in gaining insights into their performance in crack classification.

Moreover, a noteworthy point is that previous studies have rarely delved into the concept of optimising pre-trained CNN models. Most existing research has limited itself to employing one optimiser to enhance a single CNN model. However, conducting a thorough comparison of various feature selection optimisers can offer a comprehensive perspective on how a finely tuned model using a feature selection optimiser can contribute to the development of an automated transfer learning-based crack classification model. Consequently, our study explores the performance of various pre-trained CNNs using different optimisation techniques, presenting a well-refined pre-trained CNN model tailored for asphalt crack classification.

3. Methodology

Chen et al. (2019) state that the development of a machine learning/deep learning model consists of specific stages, including problem selection, data collection, model development, validation, assessment of impact, and deployment. Therefore, the process of developing the optimised deep learning model for highway pavement crack detection and classification has four steps as follows.

1. A collection of 4,663 images was obtained from various sources on GitHub and subsequently categorised manually into distinct crack types: horizontal, vertical, and diagonal. According to Gholamy et al. (2018), empirical studies indicate that the best results are achieved when 20–30% of the data is reserved for testing, with the remaining 70–80% used for training. Therefore, for this research, the 4,663 crack images were divided into 70% for training (3,264) and 30% for testing (1,399).

2. After categorising the data, a five-layer CNN model was developed, and three optimisation algorithms—ADAM, SGDM, and RMSProp—were applied to enhance the accuracy of the CNN model. The selection of these algorithms is based on their demonstrated detection accuracy with similar datasets (Jiang et al., 2020; Reyad et al., 2023).

3. Subsequently, the most accurate algorithm was identified and subjected to evaluation using eight feature selection algorithms. This step aims to determine the most
an effective method for maximising the accuracy of the entire feature selection-based CNN crack detection model.

(4) A testing sample comprising 1,399 data points was utilised to assess the accuracy of the optimised feature selection-based CNN model in comparison to other eight pre-trained models. The evaluation was conducted using the F-score, which measures a
model’s accuracy on a dataset as the harmonic mean of the model’s precision and recall percentages. Consequently, we will compare the F-score of the optimal CNN model, achieved through the best feature selection algorithms, with the F-scores of eight other selected feature selection optimisers.

Figure 1 illustrates the developmental and testing phases of the proposed optimised optimised feature selection-based CNN crack detection model.

4. Results and evaluation
This section presents and discusses the results of the proposed model for detecting and classifying pavement cracks. The development of the fully connected deep learning model involved three steps, each accompanied by experiments.

In the first step, a five-layer deep learning CNN was constructed using three optimisation algorithms: ADAM, SGDM, and RMSProp. This initial model was designed without feature selection and aimed to evaluate its accuracy in classifying and detecting three different types of cracks (horizontal, vertical, and diagonal).

For the second step, an optimisation layer was incorporated along with various feature selection algorithms. The modified architecture employed several performance measures, including true positives, true negatives, false positives, false negatives, accuracy, specificity, precision, recall, and F-score values. This step involved the implementation of eight different feature selection optimisers.
In the third step, the fully connected model was tested against seven pre-trained models, and the variation in F-scores was assessed. These models were ranked based on their F-score values.

4.1 Developing five deep layers connected CNN model

Figure 2 shows that a deep learning model including five layers were developed in order to detect cracks from images, five layers were developed to enhance accuracy starting from 32 \((3 \times 3)\) to 512 \((3 \times 3)\), then using three optimisers such as ADAM, SGDM and RMSProp to adjust training parameters and enhance the detection accuracy and precision.

An empirical comparison of optimisers was conducted in the first step, involving an investigation of the best performance of three optimisers across a variety of pavement image classification tasks.

In the current study’s first step, the proposed OCNNH model was tested using three optimisation algorithms at different learning rates but without feature extraction. Firstly, the OCNNH architecture was tested with the ADAM optimiser upon various configurations of parameters with a maximum classification accuracy of 93.68% and an F-score value of 97.4%, as shown in Table 2.

<table>
<thead>
<tr>
<th>Optimisation algorithm</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>TP</th>
<th>Accuracy</th>
<th>Specificity</th>
<th>Precision</th>
<th>Recall</th>
<th>F-score value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAM</td>
<td>298</td>
<td>4</td>
<td>13</td>
<td>318</td>
<td>93.65</td>
<td>98.68</td>
<td>98.76</td>
<td>96.07</td>
<td>97.4</td>
</tr>
<tr>
<td>SGDM</td>
<td>325</td>
<td>5</td>
<td>7</td>
<td>328</td>
<td>96.33</td>
<td>98.48</td>
<td>98.5</td>
<td>97.91</td>
<td>98.2</td>
</tr>
<tr>
<td>RMSProp</td>
<td>283</td>
<td>23</td>
<td>4</td>
<td>332</td>
<td>91.67</td>
<td>92.48</td>
<td>93.52</td>
<td>98.81</td>
<td>96.09</td>
</tr>
</tbody>
</table>

Note(s): Learning rate = 0.01; mini-batch size = 128; max. epochs = 10
Source(s): Table created by author

Table 2. OCNNH performance evaluation with three optimisation algorithms

Source(s): Figure created by author
The experiments were repeated using the OCNNH model with ADAM, SGDM and RMSProp optimisers and with the same input size, number of layers and convolution filters to evaluate their performance. The maximum classification accuracy for these three optimisers was 97.4%, 98.2 and 96.09% respectively, as shown in Table 2.

4.1.1 Performance evaluation of the adjusted CNN model. Figures 3–5 display the confusion matrix, where the values represent counts derived from predicted and actual values. The term “TN” stands for True Negative, indicating the number of accurately classified negative examples. Similarly, “TP” stands for True Positive, representing the number of accurately classified positive examples. The abbreviation “FP” corresponds to the False Positive value, signifying the number of actual negative examples classified as positive. Meanwhile, “FN” represents the False Negative value, denoting the number of actual positive examples classified as negative (Kulkarni et al., 2020).

Figure 3 shows the confusion matrix for proposed OCNNH model with the ADAM optimiser: accuracy values of the true class (target class) for the OCNNH model are 97.6% for vertical cracks, 94.6% for horizontal cracks and 88.7% for diagonal cracks. The average accuracy of the model in evaluating and detecting the three cracks, based on the ADAM optimiser, is 93.65%. Moreover, higher percentages of precision are achieved in detecting the three different types of pavement cracks, with an average precision value of 98.76%, thus proving the validity of using the OCNNH model with the ADAM optimiser.

![Confusion Matrix](image)

**Source(s):** Figure created by author
Similarly, the OCNNH model was tested using the SGDM algorithm, with Figure 4 depicting the confusion matrix for the three types of cracks. Horizontal cracks were the most accurately detected type of distress at 97.6%, then diagonal cracks at 96.1%, followed by vertical cracks at 95.2%. The average accuracy of the OCNNH model, based on the SGDM optimiser, was 96.3%, which was 2.6% higher than the ADAM algorithm. The SGDM algorithm was more accurate than the ADAM algorithm for two types of cracks, namely, diagonal and horizontal. At the same time, the ADAM algorithm was 2.4% more accurate than the SGDM algorithm in detecting vertical cracks.

Finally, Figure 5 shows that the average accuracy of the OCNNH model using the RMSProp algorithm was 91.7%, which was significantly lower than the OCNNH model based on using the ADAM and SGDM optimisers. However, the accuracy of detecting horizontal cracks using the RMSProp optimiser was higher than when using the ADAM and SGDM optimisers by 4.2 and 1.2%, respectively.

4.1.2 Evaluating three optimisers at learning curve 0.01. Figure 6 shows the learning curve of the OCNNH model, based on the ADAM optimisation algorithm, at LR 0.01, indicating that increasing the number of iterations increases the accuracy percentages and decreases the loss percentage reach by 6.3%.
Figure 7 shows the learning curve for the OCNNH model based on the SGDM optimisation algorithm. The maximum accuracy was 96.3% with a minimum loss of 3.7%, which was lower than that for the ADAM algorithm by 2.6%.

In addition, Figure 8 illustrates the learning curve for the OCNNH model based on the RMSProp optimisation algorithm, including accuracy and loss percentages against each epoch. The loss using the RMSProp optimiser is the highest value among the three optimisers at about 8.3%, which is relatively high compared to the other two optimisation algorithms.

To summarise the first step, and as shown in Figure 9, the experimental work revealed that RMSProp was the best optimiser for detecting horizontal cracks with accuracy at more than 98.2%. At the same time, ADAM was the best optimiser for detecting vertical cracks with an overall accuracy of 97.6%. Finally, SGDM was the best optimiser for detecting diagonal cracks with an overall accuracy of around 96.1%. The user can therefore adopt the appropriate optimisation algorithm according to the types of images: moreover, multiple optimisation algorithms can be employed to ensure that a wide range of cracks can be detected with minimal loss of data.
Classification of highway pavement cracks

Figure 6. Learning curve for OCNNH-based Adam algorithm

Source(s): Figure created by author

Figure 7. Learning curve for OCNNH-based Adam algorithm

Source(s): Figure created by author
4.2 Using feature selection optimisers to enhance accuracy

In the first step, to maximise the accuracy level, the OCNNH model was developed without integrating feature selection algorithms. In the second step, feature selection algorithms were integrated into the OCNNH model to optimise the input variables, as well as to enhance the developed OCNNH model's overall performance. Figure 10 depicts the process of integrating feature selection extraction into the OCNNH model.

Feature selection is one of the most critical components of deep learning, with most real-world data sets having many features. However, not all features are necessary for a specific deep learning model. Utilising many redundant features may lead to several problems, the most significant of which is the computation cost. Due to the unreasonably huge data set, it...
will take an unnecessarily long time to run the model. At the same time, it may cause an unexpected overfitting problem. Several feature selection methods are available. The current study’s experimental work demonstrates eight popular feature selection methods comprising: principal component analysis (PCA); independent component analysis (ICA); sparse filtering (SF); features transformation (FT); discrete cosine transform (DCT); discrete wavelet transform (DWT); particle swarm optimisation (PSO); and grey wolf optimiser (GWO).

Table 3 presents a comparison between the eight feature selection optimisers to evaluate the most accurate one and to provide a comprehensive comparison of the accuracy of each feature selection algorithm.

As shown in Table 3, the precision (99.02%), specificity (98.45%) and accuracy (98.19%) for PSO achieve better performance than those of the other optimisers.

As shown in the results presented in Figure 11, PSO outperforms all other optimisers in terms of values for F_score, recall, precision, specificity, and accuracy, followed by GWO, DWT, FT, ICA, SF, DCT and PCA, in descending order.

4.3 Comparison between the proposed optimised CNNs and traditional pre-trained models
To assess the accuracy and validity of the proposed model after selecting the most precise optimiser, which is PSO, the same datasets were tested using seven other pre-trained models, as indicated in Table 4. It can be observed that the proposed model outperforms all seven other models across all metrics, including accuracy, specificity, precision, and F-score.
The charts in Figure 12 depict the variation in F-score with the proposed OCNH model, reflecting the model’s performance. It is evident that the least accurate model in crack detection is Alexnet, with a variation of approximately 19.84%. This comparison aids prospective researchers in selecting the optimal model for feature extraction when developing deep learning models to detect various types of cracks or distress in construction elements.

5. Discussion on findings
In our study, we developed five deep layers CNN model to automatically detect and classify pavement cracks, three common optimisers were used to adjust training parameters to enhance the performance of the model. The most accurate results were recorded for the five deep learning layers with the RMSProp optimiser for horizontal cracks as it showed promising results in detecting horizontal cracks with 98.2% overall accuracy and low false positives. At the same time, the study’s results showed that the best optimisers for detecting
vertical and diagonal cracks were ADAM and SGDM, respectively, with overall accuracy of 97.6 and 96.1%.

Our study next proposed a pre-trained deep structured algorithm with feature selection and tested eight popular feature selection optimisers. As noted, the results obtained from the eight feature selection optimisers were comparable in terms of the performance measures of accuracy, recall, precision, specificity, and F-score. In comparing these performance measures, particle swarm optimisation (PSO) performed best with 98.19% accuracy, 99.02% precision, and 98.45% specificity. A good level of efficiency was provided by PSO, decreasing false positives to four and false negatives to eight, which supported feature selection, encouraged reuse of features, and reduced the number of parameters. The empirical results for classification performance, feature elimination effects, and convergence rates all indicated the superiority of the proposed PSO over other feature selection optimisers in undertaking feature selection tasks. Therefore, the proposed PSO-based OCNNH model can improve classification performance by identifying the most discerning insights, as evidenced by this study's empirical results and results from its statistical tests.

To ensure the accuracy of the proposed PSO-based OCNNH model, a comprehensive evaluation of model metrics, including key parameters such as accuracy, specificity, precision, recall, and F-score, was conducted. The results were ranked based on their F-score values. The ranking, from lowest to highest performance, is as follows: AlexNet, VGG16, VGG19, GoogleNet, Darknet19, Xception, DenseNet201, OpDark, and OCNNH. As such, this comparison provides insights for future researchers when selecting the most appropriate pre-trained CNN model to extract features and develop new models for similar purposes in different construction elements such as concrete, walls, etc.

6. Conclusion
Countries worldwide have lost vital infrastructure worth trillions of dollars via degradation (cracks) of their highways. The literature is unanimous concerning the understanding of the causes of highway degradation and its socio-economic implications on countries globally. With these factors having been extensively studied, it is evident that maintaining highways in good condition is the most cost-effective way to save highways worldwide. However, manual inspection of cracks is expensive, time consuming, and potentially dangerous. As alternatives, artificial intelligence and deep learning crack detection are being perceived as
the most optimum solutions. Consequently, this study provides a fully connected PSO-based five deep learning layers to detect highway pavement distress. Results showed that

(1) Two main steps were undertaken. In the first step, five deep learning layers with three optimisation algorithms, namely, ADAM, SGDM and RMSProp, were employed to develop an accurate deep learning model to detect cracks in highway pavement. In the first step, the experimental work revealed that RMSProp was the best optimiser for detecting horizontal cracks with more than 98.2% accuracy. Furthermore, the results revealed that the best optimisers for detecting vertical and diagonal cracks were ADAM and SGDM, respectively, with overall accuracy of 97.6 and 96.1%.

(2) In the second step, eight feature selection algorithms were employed to select the highest performance algorithm with the five deep learning layers. Results showed that particle swarm optimisation (PSO) performed best with 98.19% accuracy, 99.02% precision and 98.45% specificity and 98.72% F-score value.

(3) The five deep learning layers with adjusted parameters and PSO was then compared with seven pre-trained models and our OCNNH model, and this model showed the highest performance 98.72% F-score and the variation with the lowest model, which is Alexnet with around 19.84% F-score.

The proposed PSO-based OCNNH model provides an integrated deep learning solution for distress detection in highway pavement. In future studies, the proposed PSO-based five-layer CNN model has the potential for extension to assess distress in various other construction elements, including concrete and walls. Furthermore, the comparison presented alongside eight other pre-trained models provides valuable insights for fellow researchers, aiding them in making well-informed decisions when selecting pre-trained models for similar applications.

While the proposed PSO-based five-layer CNN model demonstrated strong capabilities, it is important to acknowledge its limitations in terms of detecting all forms of distress in highway pavement. It was primarily tested for three types of cracks — horizontal, vertical, and diagonal. The model could benefit from evaluation with additional datasets encompassing a broader range of distress types, including alligator (fatigue) cracking, bleeding and others.

References


Further reading

Corresponding author
Sandra Matarneh can be contacted at: s.matarneh@ammanu.edu.jo