www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

AUTOMATIC CRACK CLASSIFICATION ON ASPHALT PAVEMENT SURFACES USING CONVOLUTIONAL NEURAL NETWORKS AND TRANSFER LEARNING

SUBMITTED: September 2024
REVISED: December 2024
PUBLISHED: December 2024
GUEST EDITORS: Vito Getuli, Farzad Rahimian, Nashwan Dawood, Pietro Capone, Alessandro Bruttini
DOI: 10.36680/j.itcon.2024.055

Sandra Matarneh, Associate Professor,
Department of Civil Engineering, Al-Ahliyya Amman University, Amman, Jordan;
s.matarneh@ammanu.edu.jo

Faris Elghaish, Lecturer,
School of Natural and Built Environment, Queen's University Belfast; UK
F.Elghaish@qub.ac.uk

David John Edwards, Professor,
Department of the Built Environment, Birmingham City University, UK;
Faculty of Engineering and the Built Environment, University of Johannesburg, South Africa;
david.edwards@bcu.ac.uk

Farzad Pour Rahimian, Professor,
School of Computing, Engineering and Digital Technologies, Teesside University; UK
F.Rahimian@tees.ac.uk

Essam Abdellatef, Assistant Professor,
Department of Electrical Engineering, Faculty of Engineering, Sinai University, Arish;
essam.abdellatef@su.edu.eg

Obuks Ejohwomu, Associate Professor,
Civil Engineering and Management, The University of Manchester; UK
obuks.ejohwomu@manchester.ac.uk

SUMMARY: *Asphalt pavement cracks constitute a prevalent and severe distress of surfacing materials and before selecting the appropriate repair strategy, the type of deterioration must be classified to identify root causes. Efficient detection and classification minimize concomitant costs and simultaneously increase pavement service life. This study adopts convolutional neural networks (CNN) for asphalt pavement crack detection using secondary data available via the CRACK500 dataset and other datasets provided by GitHub. This dataset had four types of cracks viz.: horizontal, vertical, diagonal and alligator. Five pre-trained CNN models trained by ImageNet were also trained and evaluated for transfer learning. Emergent results demonstrate that the EfficientNet B3 is the most reliable model and achieved results of 94% F1_Score and 94% accuracy. This model was trained on the same dataset by performing transfer learning on pre-trained weights of ImageNet and fine-tuning the CNN. Results revealed that the modified model shows better classification performance with 96% F1_Score and 96% accuracy. This high classification accuracy was achieved by a combination of effective transfer-learning of ImageNet weight and fine-tuning of the top layers of EfficientNet B3 architecture to satisfy classification requirements. Finally, confusion matrices demonstrated that some classes of cracks performed better than others in terms of generalization. Further additional advancement with fine-tuned pre-trained models is therefore required. This study showed that the high classification results resulted from using a successful transfer learning of ImageNet weights, and fine-tuning.*

KEYWORDS: *Convolutional Neural Networks; CNN; Deep Learning; Transfer Learning, Multiclass Classification; Asphalt Pavement.*

REFERENCE: *S. Matarneh, F. Elghaish, D. J. Edwards, F. P. Rahimian, E. Abdellatef & O. Ejohwomu (2024). Automatic crack classification on asphalt pavement surfaces using convolutional neural networks and transfer learning. Journal of Information Technology in Construction (ITcon), Special issue: 'Managing the digital transformation of construction industry (CONVR 2023)', Vol. 29, pg. 1239-1256, DOI: 10.36680/j.itcon.2024.055*

1. INTRODUCTION

Asphalt pavement cracking is a serious issue that significantly impairs surface performance. A pavement crack is defined as an unintended discontinuity or a damage compromising the integrity of the pavement surface. These cracks are frequently attributed to external environmental factors, such as moisture variations, subgrade expansion, chemical shrinkage, frost, precipitation, flooding, ultraviolet (UV) radiation, and the repeated passage of heavy loads (Baduge et al., 2023). Surface cracks facilitate water infiltration into the pavement structure, further worsening the damage and ultimately compromising the pavement's base. Cracks can be classified into various types, including longitudinal, transverse, diagonal, block, alligator, and irregular patterns (Canestrari and Ingrassia, 2020).

While cracks represent a key indicator of pavement condition assessment (Guo et al., 2022), traditional crack classification mostly relies on time-consuming, costly and risky manual inspection (Yang et al., 2020). Moreover, manual inspection is inefficient and unreliable due to human subjectivity and tacit knowledge of inspectors. Indeed, extant literature reports variations in routine inspection practices and significant differences in crack classification outcomes (Dais et al., 2021). To augment inspection performance, advancements in AI offer innovative deep learning (DL) solutions to enhance crack detection accuracy and efficiency while minimizing human error (Tran et al., 2021). As an established DL technology, convolutional neural networks (CNN) have achieved advanced human-competitive performance in computer vision activities like object identification, image classification, and semantic segmentation (LeCun et al., 2015, Krizhevsky et al., 2017). Unlike traditional image analysis methods that depends on manually defined rules, CNNs automatically extract multi-level feature representations (Liu et al., 2019). Several studies have highlighted the growing efficacy of CNNs as a robust tool for crack detection (Gopalakrishnan et al., 2017, Mandal et al., 2018, Huyan et al., 2020).

However, training a CNN requires enormous volumes of properly collected data to ensure that adequate training data is available to engender accurate outcomes and avoid underfitting (i.e., CNN performs poorly on training data and new data from the problem domain) (Kim et al., 2018). Transfer learning techniques, which leverage knowledge from previously trained CNN models, are utilized to mitigate the risk of underfitting (Yang et al., 2020). These CNNs are employed in transfer learning by utilizing weights pre-trained on the ImageNet dataset. ImageNet provides a publicly accessible database of images designed to support the training of large-scale object recognition models (Feng et al., 2019). Various studies have applied transfer learning to establish classifiers for crack detection viz.: MobileNet (Hou et al., 2021, Hernanda et al., 2022); VGG (Gopalakrishnan et al., 2017, Dung et al., 2019, Rubio et al., 2019, Guzmán-Torres et al., 2022, Brien et al., 2023); GoogleNet (Jang et al., 2019, Yang et al., 2021, Elghaish et al., 2022); U-Net (Liu et al., 2019, Yang et al., 2021, Matarneh et al., 2024); Res-Net (Augustauskas and Lipnickas, 2020, Wang and Guo, 2021, Yoon et al., 2022); Inception (Feng et al., 2019, Ali et al., 2022, Wu et al., 2021a); AlexNet; YOLO (Liu et al., 2022c, Teng et al., 2022); and EfficientNet (Teng et al., 2022, Liu et al., 2022b).

Given the aforementioned context of rapidly evolving technology development, this present study introduces a modified pre-trained CNN model based on EfficientNetB3 to detect different types of asphalt pavement surface cracks. Therefore, this study investigates the classification performance of different ImageNet pre-trained models on the Crack500 dataset and employs transfer learning and fine-tuning to enhance the classification accuracy of the pre-trained models.

1.1 Convolutional neural networks for crack detection

Images can be classified using CNN in three categories namely: 1) semantic segmentation - which offers information about the specific length, width or location of any crack due to its capability of assigning a class label to each pixel (Li et al., 2018, Liu et al., 2019, Choi and Cha, 2020, Yang et al., 2021); 2) patch classification – where each patch is assigned a class label after dividing images into patches; and 3) boundary box regression – where the box bounds the identified crack and shows its location and boundaries (Zhang et al., 2019). However, the latter two categories have been extensively used for identifying cracks and have proven positive outcomes (Cha et al., 2017, Feng et al., 2019). Unlike these two categories which are implemented at block level, semantic segmentation is implemented at pixel level and has gathered application momentum in recent studies (Tang et al., 2022).

To identify and categorize cracks in asphalt pavements, Tran et al. (2021) utilized an improved faster R-CNN

(RetinaNet) to train and label asphalt images. The results of this study revealed that the RetinaNet's classification accuracy was 84.9% considering both the type and severity level of cracks. Similarly, Li et al. (2022) established a strategy for pavement surface condition index (PCI) using a genetic algorithm (GA) and CNN algorithm. The authors classified pavement crack type in five thousand pavement distress images with an accuracy reaching 98%, and image processing time of 0.047 seconds. To understand the crack features in the intricate fine-grain asphalt pavement background, Wu et al. (2021a) employed a full CNN to combine features acquired from various scales of convolutional kernels, the DenseNet and the deconvolution network to accomplish pixel-level recognition. Emergent results indicated that the adopted method reported significant segmentation results for twelve types of cracks. Teng et al. (2022) developed DeepLab_v3+, a pixel-level segmentation CNN, to segment cracks by calculating their length and width to a reported accuracy and F-score values of 80, 97.5 and 78% respectively. Fan et al. (2022) reported upon the results of a deep residual convolutional neural network (Parallel ResNet) to develop a pavement crack detection and measurement system with high-performance i.e., precision (94.27%), recall (92.52%) and F1 (93.08%). Based on a spatial channel hierarchical network, a more accurate and efficient crack detection process was proposed by Pan et al. (2020) using the Visual Geometry Group 19 (VGG19). To examine inner cracks in the turbine blade thermal barrier coating, Shi et al. (2022) used VGG19 in addition to a Multi-Scale Enhanced-Faster R-CNN (MSE-Faster R-CNN) to create prospective crack locations. Results showed that the suggested approach can precisely locate cracks on different scales (89.8%) and detect them (80.6%).

Recent studies have employed a variety of CNNs, including Tan et al. (2021) who utilized the YOLOv3 to automate the recognition of sewage pipe deficiencies. The study concentrated on enhancing the model structure, bounding box prediction, data expansion and loss function. The proposed model achieved a mean average precision (mAP) value of 92%, which is greater than the level of accuracy achieved in previous related studies. Lu et al. (2022) utilized YOLOv5 and the multiple sliding windows method to detect defects in ceramic tiles surfaces.

Similarly, Yao et al. (2022) developed twelve distinct attention models for pavement crack detection using YOLOv5. Their study demonstrated that the model could process images at a rate of 13.15 ms per image while achieving a precision of 94.4%. Li et al. (2020) employed U-Net with alternately updated clique (CliqueNet), called U-CliqueNet to separate cracks from the background of tunnel images. The developed network was trained on a large dataset containing 50,000 images and tested on 10,000 images. The model attained positive results with 92.25% mean pixel accuracy (MPA), 86.96% mean intersection over union (MIoU), 86.32% precision and 83.40% F1-score. Liu et al. (2019) employed U-Net for concrete crack detection in raw images, achieving precision values of 0.9 across various complex scenarios. Huyan et al. (2020) constructed a U-shaped model structure by utilizing convolution, pooling, transpose convolution, and concatenation operations. Their results indicated that the model achieved 99.01% accuracy, 98.56% precision, 97.98% recall, and 98.42% F-measure with a learning rate of 10^2.

Recently, several studies adopted the generative network along with the DL models. For instance, Mazzini et al. (2020) proposed a CNN to augment data of highly textured images within the framework of semantic segmentation. In their study a Generative Adversarial Network (GAN) was used to develop a semantic layout, and a texture synthesizer (based on a CNN), to produce a new image. This approach was evaluated using the dataset of German Pavement Distress and the results of evaluation revealed a substantial improvement in prediction performance. Pei et al. (2021) developed a method using advanced deep convolutional generative adversarial networks (DCGANs) to solve the problem of crack identification in asphalt pavement small size images. The outcomes presented that the average precision of the proposed model is 90.32%. Dong et al. (2022) used StyleGAN and a feature fusion model to solve the issue of realizing precise and effective pixel-level segmentation of pavement damage. Results showed that the suggested model segmentation could achieve above 0.918 mAP (mean Average Precision) value for segmentation, and it has evident benefits for the map and cross cracks segmentation.

1.2 Pre-trained CNN for crack detection

Existing CNN models have a tendency for being quite complex and require a huge dataset to prevent overfitting and therefore, to solve these problems transfer learning is increasingly utilized (Dawson et al., 2023). Transfer learning is a machine learning technique that uses pre-trained DL models to resolve a new problem that is related to the original problem the model was initially trained to solve (Pan and Yang, 2010). Recently, transfer learning in DL has been widely applied across various approaches. For example, to assess the effectiveness of DCNNs in the pavement cracks classification, Ranjbar et al. (2021) used different pre-trained networks, namely: DenseNet-201, GoogleNet, AlexNet, SqueezNet, ResNet-18, ResNet-50, ResNet-101 and Inception-v3. In this study, a more efficient technique for crack segmentation was created using a wavelet transform module with more regularized metrics. According to the study's findings, retrained classifier models produced accurate results with a confusion matrix-based performance range of 94% to 99%. Moreover, the constructed wavelet module was capable of clearly segmenting crack pixels [ibid]. Wu et al. (2021b) used DenseNet, multi-scale CNN and SVM classifier to segment 12 different crack types and produce superior pavement crack segmentation comparing to the most sophisticated alternative methods Hernanda et al. (2022) employed a DL CNN with a pre-trained SSD MobileNetV2 network after altering the hyperparameter. Results showed a higher mean average precision (mAP) of 0.0869 and a lower total loss training of 0.6028. Ha et al. (2022) classified five types of cracks (alligator crack, longitudinal crack, transverse crack, pothole and patching) by developing a system utilizing SqueezeNet, U-Net and Mobilenet-SSD models together. Emergent results reported that the system accuracy was 91.2%. Fan et al. (2022) adopted the Parallel ResNet to develop a deep residual CNN to establish a high-performance system to detect and measure pavement crack. The highest scores for precision (94.27%), recall (92.52%) and F1 (93.08%) were achieved [ibid]. Another pretrained deep CNN model for crack identification is presented using hybrid images and GoogLeNet (Jang et al., 2019).

The proposed model was tested and validated utilizing various sizes of concrete sampling and showed that macro- and micro-cracks were effectively identified utilizing hybrid images. Qu et al. (2022) also modelled crack segmentation using a CNN network and transformer. To increase the performance of the feature representation, the authors employed UNet++ and polarized self-attention. In addition, they replaced the last layer of feature extraction by the transformer. The study results exhibited that the developed model showed its efficacy with F-score values of 0.856, 0.700 and 0.637 on three different datasets. Xu and Liu (2022) proposed a detection method under small samples by combining a generative adversarial network (GAN) and a CNN architecture. After training and analyzing the dataset using the transfer learning technique, results showed that the detection accuracy improved considerably from 80.75% to 91.61% thus, proving the effectiveness of the extended data.

Recently, several studies compared various CNNs and reported different results (Cha et al., 2017, Feng et al., 2019, Liu et al., 2022b, Matarneh et al., 2024). Yang et al. (2021) evaluated the performance of three CNNs, namely: AlexNet, ResNet18 and VGGNet13 and concluded that ResNet18 (accuracy 98.8%) outperforms the other two models. Another study assessed four CNNs (i.e., Inception-V3, VGG-16, VGG-19 and Resnet-50) and compared them with the proposed customized CNN model (Ali et al., 2022). The authors developed their CNN model based on spatial or sequential features and Adam optimizer for crack localization and detection in concrete structures using eight datasets. The research concluded that all models performed well on a small set of various training data; though, as the quantity and diversity of the training data increased, generalization performance decreased, and overfitting resulted. Additionally, the tailored CNN and VGG-16 models showed superior cracking localization and identification for concrete structures. (Wang and Guo 2021) developed transfer learning-based methods for fatigue crack initiation sites identification using three CNNs (i.e., feature pyramid network (FPN), ResNet-101 and VGG-16), to extract features along with the backbone model - a faster R-CNN.

The ResNet model outperformed the other two in terms of accuracy and training expense. Pozzer et al. (2021) used thermographic and regular photos taken from a variety of distances and angles to test the effectiveness of multiple deep neural network models in concrete cracks identification. The MobileNetV2 performed admirably in detecting multiclass damages in thermal pictures while in contrast, the VGG 16 model showed improved precision by reducing the rate of erroneous detections [ibid]. Elsewhere, Elghaish et al. (2022) evaluated four pre-trained CNN models (i.e., AlexNet, VGG16, VGG19 and GoogleNet) in crack detection for highways and observed that the accuracies of all pre-trained models are higher (97.72%) than averages and the calculated accuracies for AlexNet and GoogleNet models by more than 5%.

The application of AlexNet, SE-Net, and ResNet with a variety of configurations was explored by Liu et al. (2022a) who applied a two-step data pre-processing to decrease the bias of the CNN model. The study used data augmentation in the beginning to increase the dataset. The original image was then transformed into a binary black-and-white image using a crack extraction technique; the ResNet with 50 layers exhibited the highest test accuracy. Loverdos and Sarhosis (2022) evaluated different DL networks (U-Net, DeepLabV3+, U-Net (SM), LinkNet (SM) and FPN (SM)) to advance automation in brick segmentation and crack detection of masonry walls. Emergent findings revealed that DL offers superior outcomes than conventional image-processing techniques for brick segmentation.

Table 1: Comparison between existing studies used transfer learning and optimisation methods for crack detection and classification. Part One.

Reference	Aim	Method	Results	Limitations
Dawson et al., (2023)	To evaluate the performance of different CNN architectures for carbonate rock classification using transfer learning.	Used datasets with varying sizes (7k, 42k, 104k images), trained nine CNN architectures, and applied transfer learning.	Inception-v3 achieved 92% accuracy on the largest dataset; dataset size strongly affects performance.	Models showed overfitting on smaller datasets; findings highlight challenges for small geological datasets.
Ranjbar et al. (2021)	To develop a reliable system for pavement crack detection and classification using pre-trained deep CNNs and wavelet transform.	Retrained pre-trained DCNNs (e.g., AlexNet, ResNet) using transfer learning, with wavelet transform for segmentation.	Achieved reliable crack classification with accuracies ranging from 94% to 99%; wavelet transform improved segmentation.	Some models showed slower performance; requires large labeled datasets for training despite using transfer learning.
Hernanda et al. (2022)	Optimize road pothole and crack detection using CNN with SSD MobileNet V2 and hyperparameter tuning.	Adjust hyperparameters in pre-trained CNN and compare mAP and loss values with prior systems.	Improved detection with optimal mAP and reduced loss values. Validated effectiveness of the approach in road inspections.	Dataset and pre-trained model dependency may limit adaptability to new conditions or datasets.
Ha et al. (2022)	Automate detection, classification, and severity assessment of road cracks using deep learning models.	Use SqueezeNet, U-Net, and Mobilenet-SSD models for segmentation and severity evaluation.	Achieved 91.2% accuracy for crack type and severity detection. Demonstrated enhanced performance for pavement management systems.	Limited crack types and specific dataset constraints reduce generalizability.
Fan et al. (2022)	Develop a high-performance pavement crack detection method addressing noise and topology issues.	Introduced Parallel ResNet to minimize noise and accurately identify cracks in public datasets (CrackTree200, CFD).	Achieved high precision and recall scores (e.g., F1 score ~95%). Enhanced ability to segment crack features amidst noise.	Dependency on specific datasets and computational complexity for large-scale real-world deploymen
Qu et al. (2022)	Address challenges in long dependencies and global context loss in crack segmentation using CNN and transformer techniques.	Developed CrackT-Net using enriched features (RF UNet++) with transformers for better feature representation and segmentation.	Improved F1 scores across datasets (e.g., DeepCrack: 0.859). Enhanced global context understanding and segmentation capabilities.	Computational cost of transformer models and reliance on specific datasets.
Xu and Liu (2022)	Develop crack detection methods using small sample datasets with GANs for data augmentation and CNN for detection	Applied GANs for data augmentation and CNN models with transfer learning to improve accuracy.	Increased accuracy from 80.75% to 91.61% after GAN-based augmentation. Significant improvement in crack detection for limited datasets.	Limited effectiveness for highly diverse datasets. Dependency on GAN-generated data quality.

Table 1 provides a comprehensive summary of various studies utilizing transfer learning (TL) techniques for crack detection and classification. Although TL methods have achieved significant success in this domain, a critical review of the literature reveals several limitations. Notably, only a limited number of studies have undertaken a systematic evaluation of multiple pre-trained CNN models, highlighting the need for a more thorough investigation into their comparative performance in crack classification. Furthermore, prior research has addressed the optimization of pre-trained CNN architectures, with most efforts confined to applying existing optimization algorithms to enhance these models. A detailed comparative analysis of pre-trained CNN models, however, can yield valuable insights into the potential of fine-tuning their architectural layers to improve performance. This approach not only refines the models but also contributes to the development of robust, automated TL-based systems for crack classification.

To address these gaps, the present study goes beyond the standard application of transfer learning by emphasizing comprehensive model evaluation, layer optimization, and generalizability. Unlike prior works that primarily test pre-existing architectures, this study explores the structural adaptability of pre-trained CNNs, demonstrating how fine-tuning specific layers can significantly improve model performance. Additionally, the study introduces a robust methodology for selecting optimal architectures tailored to crack detection, which fills a critical gap in the literature.

Table 1: Comparison between existing studies used transfer learning and optimisation methods for crack detection and classification. Part Two.

Reference	Aim	Method	Results	Limitations
Matarneh et al., (2024)	Optimize pre-trained CNNs for detecting and classifying pavement cracks with a focus on DenseNet201 and other architectures.	Compared ten pre-trained models, optimized DenseNet201 with feature selection and noise-resistance methods.	DenseNet201 and GWO optimizer achieved highest accuracy and noise robustness. Highlighted effectiveness of transfer learning in reducing training time and errors.	Inconsistent performance of other CNN models (e.g., VGG16) and dependency on pre-trained models for scalability.
Lu et al. (2022)	Develop an intelligent YOLOv5-based system for defect detection in ceramic tiles.	Optimized YOLOv5 with Shufflenetv2 backbone and multiple sliding windows for feature extraction and classification.	Improved mAP to 96.73%, with reduced parameters and computational cost compared to baseline YOLOv5.	Limited to ceramic tiles with specific defect characteristics; effectiveness may vary for other surface types.
Liu et al. (2022a)	Classify asphalt pavement crack severity levels using CNNs and thermography.	Compared thirteen CNN models using datasets of visible, infrared, and fused images; applied transfer learning.	EfficientNet-B3 achieved the best accuracy for all image types. Infrared imaging improved detection in low-light conditions.	Misclassification occurred primarily at medium and high severity levels. Dataset biases and varying imaging conditions may affect reliability.
Wang and Guo (2021)	Automate identification of fatigue crack initiation sites in engineering structures.	Employed Faster R-CNN with VGG-16, ResNet-101, and FPN as feature extractors for transfer learning.	ResNet-101 achieved the highest accuracy (95.9%), balancing precision and training costs.	Limited by small dataset size and image characteristics like low contrast and resolution.
Yang et al. (2021)	Use deep learning to detect and recognize structural cracks.	Compared AlexNet, VGGNet13, ResNet18, and YOLOv3 for crack detection and recognition.	ResNet18 provided superior image recognition, while YOLOv3 excelled in real-time crack area detection with high precision.	Dependency on manually collected and augmented datasets; performance varies under adverse environmental conditions like poor lighting.
Pozzer et al. (2021)	Assess deep learning methods for detecting defects in damaged concrete using thermal and regular imaging.	Trained CNN models (MobileNetV2, VGG16) on thermal and regular images for multiclass defect detection (cracks, spalling).	MobileNetV2 achieved 79.7% accuracy in thermal imaging for multiclass damage detection; VGG16 reduced false positives effectively.	Limited generalizability to diverse conditions; model performance varies significantly with viewpoint and distance in real-world settings.
Louati et al., (2022)	Optimize CNN architecture and compression jointly using bi-level optimization.	Employed co-evolutionary migration-based algorithm (CEMBA) for architectural design and filter pruning.	Achieved lightweight, optimized CNN architectures with reduced parameters while maintaining high accuracy on CIFAR and ImageNet datasets.	High computational cost for the optimization process; limited evaluation on small-scale datasets and specific architectures.

In summary, this study addresses these gaps by:

- Conducting a systematic comparison of five pre-trained CNN models, evaluating their transfer learning potential for crack detection to identify the most effective architecture.
- Introducing a novel approach to optimize CNN layers, focusing on refining internal structures to enhance both accuracy and computational efficiency.
- Evaluating performance using diverse datasets, including varying crack types and environmental settings, to ensure the model's robustness and generalizability.
- Highlighting fine-tuned transfer learning as a strategy to mitigate overfitting and enhance the adaptability of pre-trained CNNs for crack detection tasks.

2. METHODOLOGY

This empirical research adopted a positivist philosophy and deductive reasoning (Edwards et al., 2020) to analyze secondary image data obtained from open-source databases to accurately model the phenomenon under investigation. Such an approach has been extensively used previously to evaluate risk factors impacting upon public-private partnerships (Kukah et al., 2022); assess the risk associate with sustainable housing (Adabre et al., 2022); and model construction machinery stability (Edwards et al., 2019). Therefore, this well-established approach is deemed suitable for the present study.

2.1 Transfer learning

The primary objective of transfer learning is to improve the learning of the target prediction function by leveraging knowledge from the source domain and associated learning tasks (Pan and Yang, 2010). Transfer learning typically employs four key techniques viz.: transferring instances knowledge; transferring feature representations knowledge; transferring parameter knowledge; and transferring relation knowledge (Pan and Yang, 2010). There are four basic types of transfer learning used in CNN networks viz.: network transfer learning; adversarial transfer learning; instance-based transfer learning; and mapping transfer learning (Tan et al., 2018). The use of transfer learning has proven effective in pavement crack detection (Gopalakrishnan et al., 2017, Joshi et al., 2022, Liu et al., 2022b, Ranjbar et al., 2021, Xu and Liu, 2022).

When using CNN to categorize the asphalt pavement cracks severity, there are two primary processes (Figure 1). First, very large picture datasets for the source domain (frequently ImageNet which has 1.2 million images with 1000 categories) are used to train the CNN model (Deng et al., 2009). The pre-trained CNN network (otherwise known as the CNN trained network), is available via web resources (Paszke et al., 2019). Secondly, the architecture for the target domain is the pre-trained CNN architecture. Unlike existing studies which focused on utilizing the existing pre-trained CNN architectures, this study retrained the pre-trained convolutional layers parameters, then modified the fully connected layers (FC layers) to satisfy the requirements of the stated output labels (four labels in this case viz.: alligator; diagonal; horizontal; and longitudinal). The modified pre-trained CNN model is then improved (trained and assessed) using the study dataset.

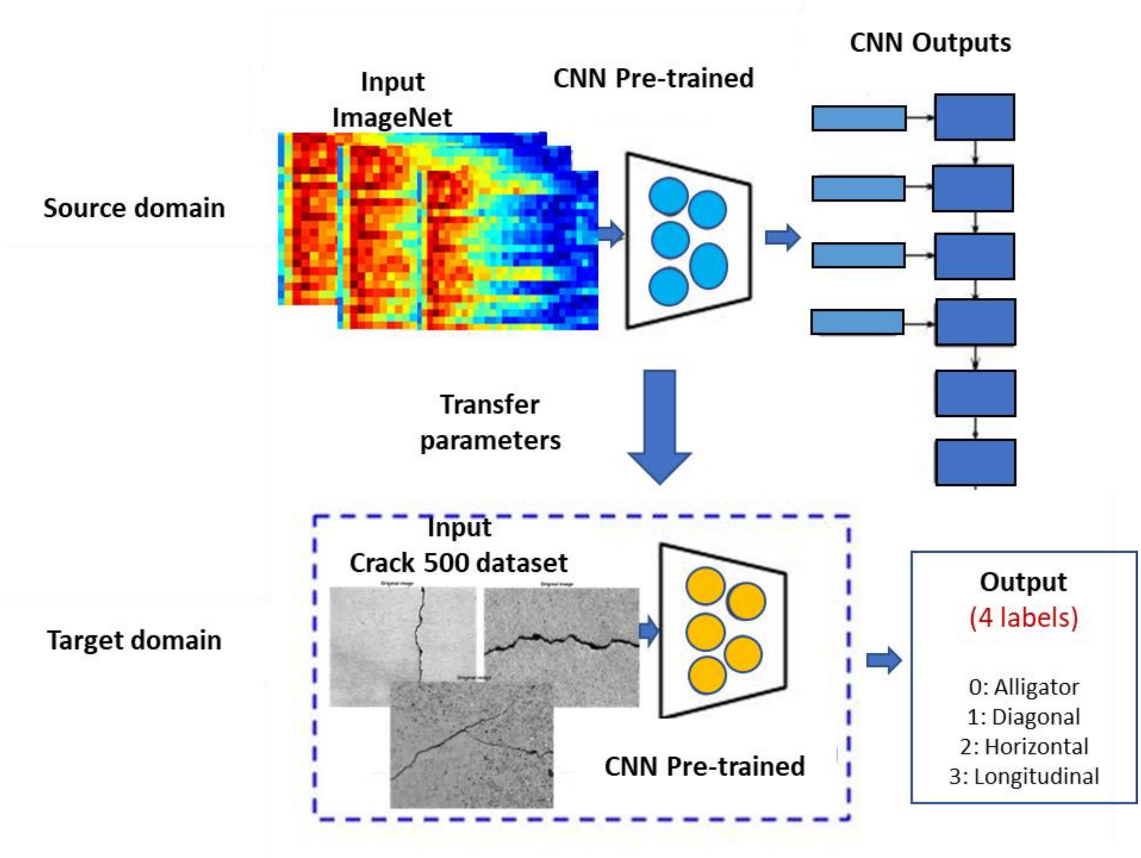


Figure 1: A transfer learning strategy using a trained CNN for identifying asphalt pavement cracks.

2.2 Performance evaluation metrics

The most typical accuracy measures for the crack classification task involve precision, accuracy, recall, F1 score and confusion matrix (Yacouby & Axman, 2020). Classification accuracy is defined as the total number of correct predictions divided by the total number of predictions made for a given dataset (Ilse et al., 2020) and is calculated as follows:

$$Accuracy = \frac{True\ Positive\ + True\ Negatives}{True\ Positive + True\ Negatives + False\ Positive + False\ Negative} \quad (1)$$

Precision as shown in equation (2) is another metric which calculates the total of correct positive predictions made. Thus, precision is a metric used to calculate the accuracy for the marginal class (Ilse et al., 2020).

$$Precision = \frac{True\ Positives}{(True\ Positives + False\ 1his1ives)} \quad (2)$$

Recall as shown in equation (3) is a metric that measures the proportion of correct positive predictions out of all possible positive instances. In contrast to precision, which reflects the accuracy of positive predictions, recall provides insight into the number of missed positive predictions. Therefore, recall gives an indication of the model's coverage of the positive class (Ilse et al., 2020).

$$Recall = \frac{True\ Positives}{(True\ Positives + False\ Negatives)} \quad (3)$$

The F1-score combines both precision and recall into a single metric, effectively capturing both aspects of model performance. After calculating precision and recall for either a binary or multiclass classification task, these two metrics are used to compute the F1-score (Ilse et al., 2020) as shown in equation (4).

$$F1 = \frac{(2 * Precision * Recall)}{(Precision + Recall)} \quad (4)$$

The classification problem is frequently related to the multi-class classification, which classifies instances into three or more classes (Ilse et al., 2020). Precision, recall and F1 score can be classified into three categories for multi-class classification viz.: 'macro', 'micro' and 'weighted'. In macro averaging, the multiclass predictions are reduced into various sets of binary predictions, the corresponding metric is determined for each of the binary cases, and the results are averaged for k classes (Luo & Uzuner, 2014) as shown in equations (5), (6) and (7).

$$Macro - averaged\ Precision = \frac{\sum_1^k Precision(P1, P2, Pk)}{Total\ number\ of\ samples} \quad (5)$$

$$Macro - averaged\ Recall = \frac{\sum_1^k Recall(R1, R2, Rk)}{Total\ number\ of\ samples} \quad (6)$$

$$Macro - averaged\ F1 = \frac{\sum_1^k F1(Fs1, Fs2, Fsk)}{Total\ number\ of\ samples} \quad (7)$$

Micro averaging treats the entire set of data as an aggregate result and calculates 1 metric rather than k metrics that get averaged together. Like macro, weighted determines the weighted mean by taking label imbalance into account (Luo & Uzuner, 2014) as shown in equations (8), (9), and (10).

$$Weighted - averaged\ Precision = \frac{\sum_1^k Precision(P1 * NP1 * NP2, Pk * NPk)}{Total\ number\ of\ samples} \quad (8)$$

$$Weighted - averaged\ Recall = \frac{\sum_1^k Recall(R1 * NR1, R2 * NR2, Rk * NRk)}{Total\ number\ of\ samples} \quad (9)$$

$$Weighted - averaged\ F1 = \frac{\sum_1^k F1(Fs1 * NFs1, Fs2 * NFs2, Fsk * NFsK)}{Total\ number\ of\ samples} \quad (10)$$

2.3 Pre-trained models hyperparameters

The CNN model was developed based on web resources (Paszke et al., 2019). This online repository contains an op-for-op PyTorch reimplementation of various pre-trained models. Five pre-trained models are used in this study; details about the models are provided in Table 2.

Table 2: The pre trained CNN model details.

No.	Model name	Total params
1	DenseNet121	7M
2	EfficientNetB0	4M
3	EfficientNetB3	10.8M
4	MobileNet	3.2M
5	MobileNetV2	2.2M

In these pretrained models the raw images are pre-processed by resizing and normalising all images to become 224 × 224. The learning rate (0.001), training epoch (10) and batch size of 64.

2.4 Dataset

The study utilised public datasets that were obtained from the GitHub website. The main dataset used is CRACK500, consisting of 500 images, each with a size of around 2,000 × 1,500 pixels. The images were taken on the main campus of Temple University using mobile phones. Yang et al. (2019) augmented the images number by splitting each image into 16 separate regions without overlap. Only the regions that included more than 1,000 pixels of crack were retained. This method increased the size of the training dataset by incorporating 1,896 additional photos. To evaluate the pre-trained CNN models, additional datasets from GitHub are combined with the primary dataset to create 2,380 pavement crack images. The dataset's images are divided into four groups (to represent the four types of cracks: longitudinal, horizontal, diagonal and alligator) as shown in Table 3. A random selection of 20% of the collected images were pooled to create the test set (see Table 3).

Table 3: Dataset groups summary.

Dataset	Longitudinal	Horizontal	Diagonal	Alligator	Total
Train	547	546	543	209	1845
Test	160	160	160	55	535
Total	707	706	703	264	2380

3. RESULTS

This section presents and discusses the results of the study, including results of the loss, and accuracy metrics.

3.1 Loss

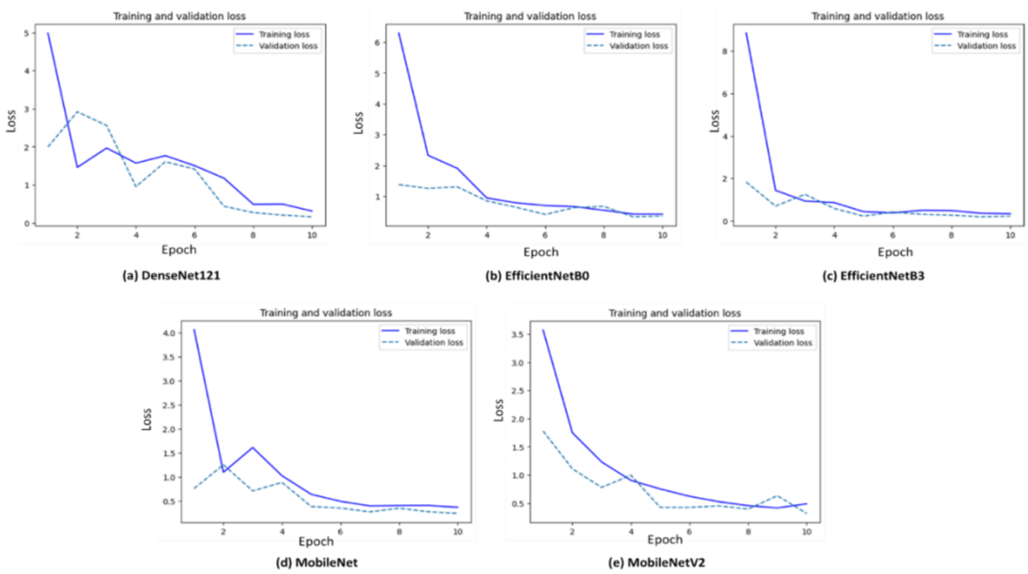


Figure 2: Loss curves of the training and validation for (a) DenseNet121, (b) EfficientNetB0, (c) EfficientNetB3, (d) MobileNet, and (e) MobileNetV2.

The loss value varied throughout the first few epochs but as the epoch increased, it decreased and proceeded to converge (refer to Figure 2). Among all models, DeneseNet121 was the most volatile, whereas the variations of the other four models were smoother as shown in Figure 2 (a,b,c,d and e). However, the EfficientNet-B3 model has the smoothest variation among all other models and tended to converge faster than the other four models.

3.2 Accuracy metrics

The purpose of this research was to identify and classify cracks utilising a dataset of pavement images and five pre-trained CNN models: DenseNet-121, EfficientNet-B0, EfficientNet-B3, MobileNet, and MobileNetV2. These models were trained and tested, and subsequently validated. Each pre-trained CNN model performance was measured using several key metrics, including, F1-score, recall, accuracy, and precision, which are commonly utilised to assess CNN model performance.

Accuracy, for example, serves as an indicator of the model's ability to correctly identify and categorize cracks relative to the total number of detection iterations performed. Figure 3 presents the accuracy metric values of MobileNet started out very high with an accuracy of 0.6475 during the whole training procedure. Contrarily, DenseNet-121 started with a moderate accuracy of 0.4580, whereas MobileNetV2, EfficientNetB0 and EfficientNet-B3 started with comparatively low accuracy of 0.3263, 0.3581 and 0.3870, respectively. Conversely, the EfficientNet-B3 accuracy metrics increased quickly as the epoch augmented and suddenly exceeded that of DenseNet-121 around Epoch 9.

The recall metric reflects the ability of the model to correctly detect positive instances. A strong relationship exists between the number of positive instances detected and recall; with the greater recall values indicating a greater number of correctly identified positive instances.

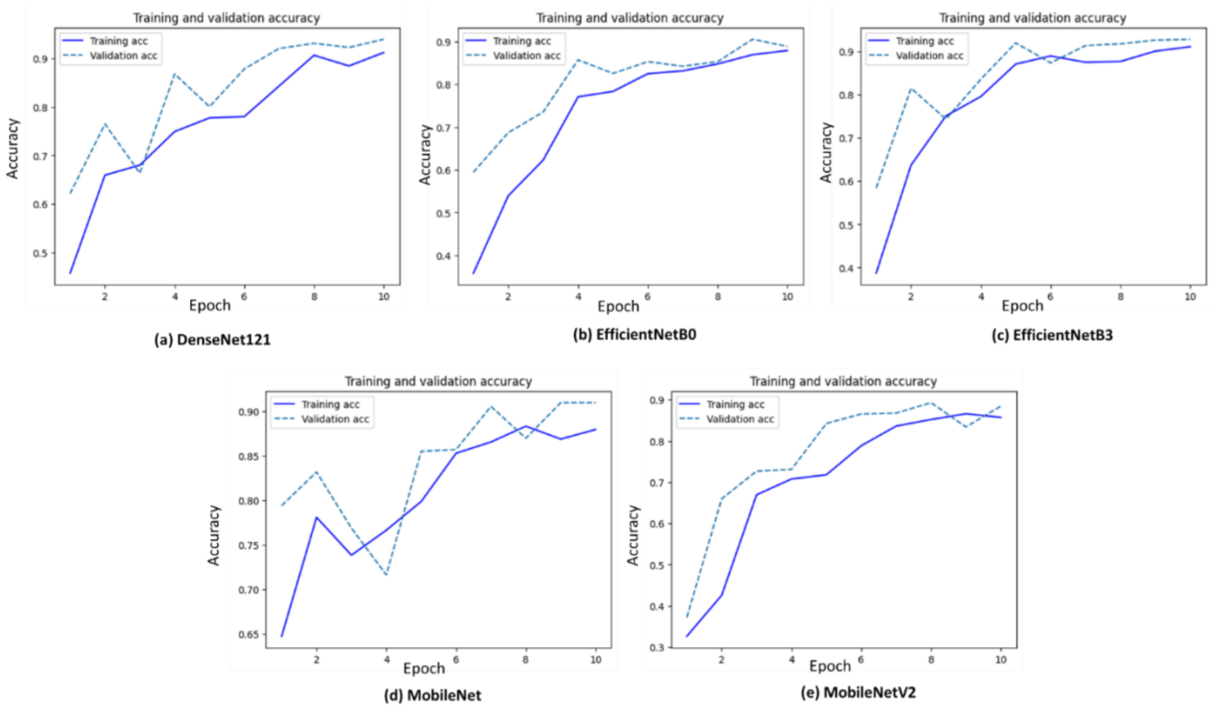


Figure 3: Accuracy metrics of the training and validation for (a) DenseNet121, (b) EfficientNetB0, (c) EfficientNetB3, (d) MobileNet, and (e) MobileNetV2.

The results indicate that EfficientNet-B3 realized the highest recall rate at 94%, followed by DenseNet-121 at 92%, EfficientNet-B0 at 89%, MobileNet at 86% and MobileNet V2 at 84%. F1-score is a performance measure that integrates both precision and recall into a single value, presenting a balanced assessment of a model's accuracy in categorization tasks. As shown in Table 4, EfficientNet-B3 achieved the highest F1-score at 94%, followed by DenseNet-121 at 92%, EfficientNet-B0 at 89%, MobileNet at 86%, and MobileNetV2 at 84%.

The accuracy performance parameters for the pre-trained CNNs are showed and compared in Table 4. Notably, MobileNetV2 and MobileNet exhibited the lowest performance among all accuracy measures when compared to the other models. In contrast, EfficientNet-B3 outperformed all other models, attaining the highest scores across all accuracy metrics, with values of 94% for accuracy, precision, recall, and F1-score.

Table 4: Accuracy metrics summary.

	Accuracy	**Precision**		**Recall**		**F1 score**	
		Macro	Weighted	Macro	Weighted	Macro	Weighted
DenseNet-121	0.92	0.91	0.92	0.91	0.92	0.91	0.92
EfficientNet-B0	0.84	0.91	0.90	0.88	0.89	0.89	0.89
EfficientNet-B3	**0.94**	**0.92**	**0.94**	**0.94**	**0.94**	**0.93**	**0.94**
MobileNet	0.87	0.84	0.87	0.87	0.86	0.85	0.86
MobileNetV2	0.84	0.86	0.86	0.81	0.84	0.83	0.84

Learning rate = 0.01, batch size = 64, and Max. Epochs = 10

4. DISCUSSION

The classification classes of the evaluated dataset were classified with varying degrees of accuracy using the same pre-trained CNN models. The confusion matrix illustrates how well-trained models perform while categorizing various classes. Figure 4 (a,b,c,d,e) shows the normalised confusion matrix, out of these confusion matrices, the pre-trained models performed very effectively on the diagonal, longitudinal and alligator crack images. However, misclassifications appeared on the horizontal crack for EfficientNetB0.

Moreover, the confusion matrices show that the alligator crack has less accuracy in the DenseNet121 and MobileNetV2 models. This could be because the alligator data set is smaller than other datasets.

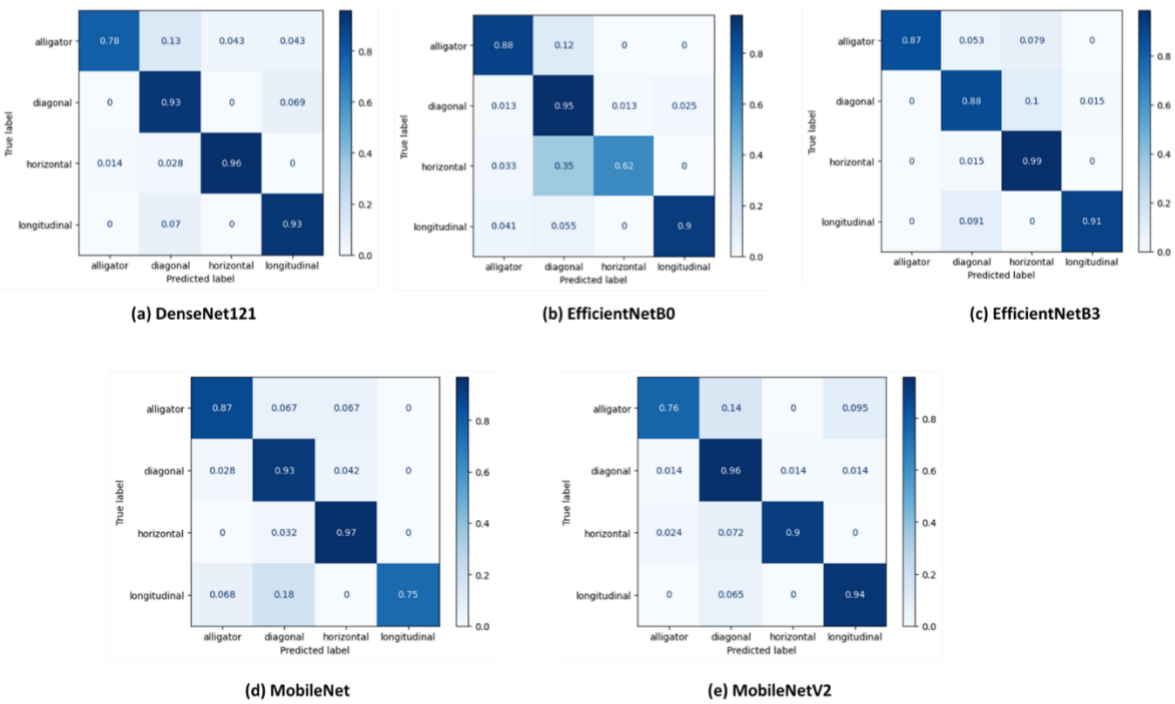


Figure 4: Normalised confusion matrix for (a) DenseNet121, (b) EfficientNetB0, (c) EfficientNetB3, (d) MobileNet, and (e) MobileNetV2.

The pre-trained CNNs accuracy is illustrated in Figure 5. The EfficientNetB3 realized the superior accuracy among the five pre-trained CNNs followed by DenseNet121 and MobileNet.

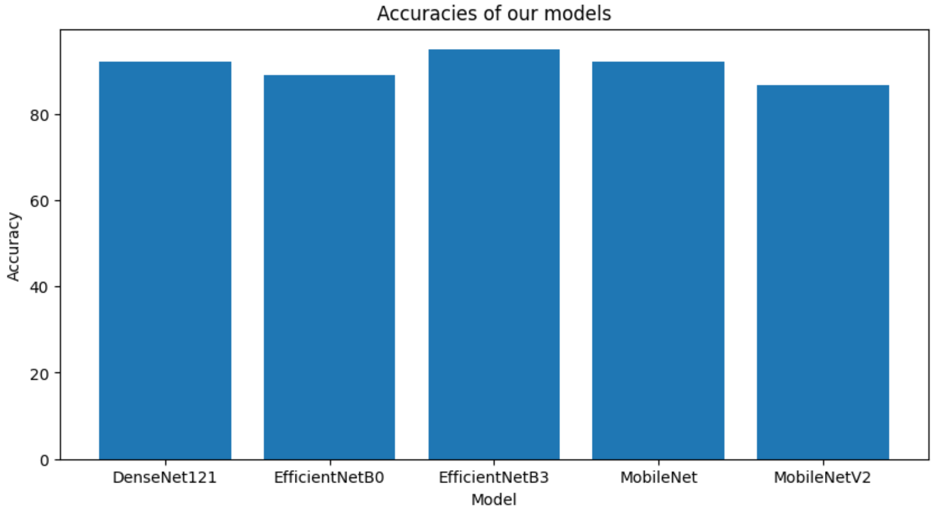
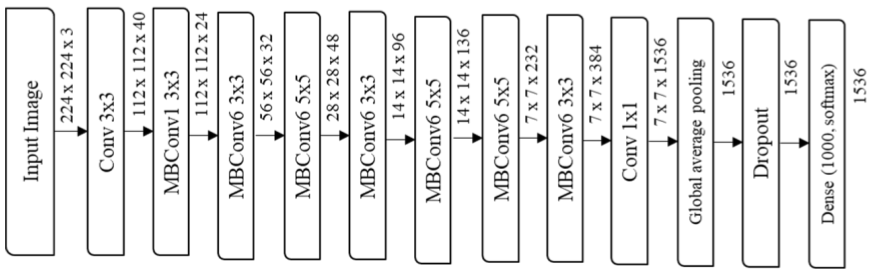
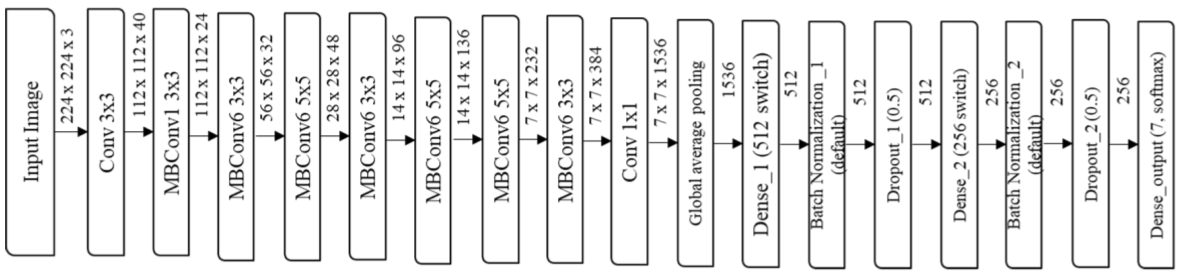


Figure 5: Accuracy of five pre-train CNN models.

4.1 Proposed approach: modifications in network architecture



(a) Original EfficientNetB3 architecture diagram



(a) Modified EfficientNetB3 architecture diagram

Figure 6: The original and modified architecture diagrams of EfficientNet B3.

In this study, some modifications in EfficientNetB3 architecture are proposed to enhance the model performance. Specifically, the top layers of the model were replaced with dense, batch normalization, and dropout layers, as suggested by Ali et al. (2022). These modifications were applied to the B3 base architecture, utilizing the Swish activation function for the dense (fully connected) layers, as recommended by Ramachandran et al. (2018) and employed by Tan and Le (2019) instead of the conventional ReLU activation function. Figure 6 illustrates the original EfficientNet-B3 baseline architecture alongside the proposed improvements. The original model was preserved, while the top layers of EfficientNet B3 architecture were restructured. In the original model, the top layers—comprising global average pooling 2D, dropout, and a dense layer—were prone to overfitting. In the modified architecture, these layers were replaced with dense, batch normalization, and dropout layers to improve performance and prevent overfitting.

4.2 Implementation and performance evaluation metric

In this experiment, a learning rate annealer was employed to reduce the learning rate after a specified number of epochs if the error rate remained unchanged. The validation accuracy was closely monitored to detect any potential plateau over a span of three epochs. If a plateau was observed, the learning rate was subsequently decreased by 0.01. After, the EfficientNet B3 convolutional layers received fine-tuning using a Stochastic Gradient Descent (SGD) optimiser. The modified EfficientNets B3 was tested an evaluated using same dataset (Crack500). The improved EfficientNet B3 reports have better results than the original EfficientNet B3 architecture. The improved EfficientNet B3 achieved an F1 score of 96% and accuracy, precision, recall and recall of 96%.

4.3 Contributions to theory and practice

The theoretical contributions of the study lie in its innovative use of convolutional neural networks (CNNs) and transfer learning to address the critical task of asphalt pavement crack classification. This research addresses gaps in current methodologies by systematically evaluating five pre-trained CNN models: EfficientNet-B0, EfficientNet-B3, DenseNet-121, MobileNet, and MobileNetV2, all fine-tuned for crack detection tasks using the Crack500 dataset. The study's results highlight EfficientNet-B3 as the most effective model, achieving a 96% F1 score and 96% accuracy after applying advanced transfer learning techniques.

Key Theoretical Contributions:

1. Transfer Learning Optimization: By leveraging ImageNet pre-trained weights and fine-tuning the CNN layers, the study demonstrates the effectiveness of transfer learning for domain-specific tasks. This approach minimizes the need for extensive domain-specific data, a common challenge in pavement inspection.
2. Evaluation of Multiple Models: The comparison of five CNN architectures provides a deeper understanding of their strengths and weaknesses in multiclass classification. The study highlights EfficientNet-B3's architecture as optimal due to its ability to balance performance and computational efficiency.
3. Scalability and Generalization: The study establishes a foundation for future research by demonstrating that transfer learning techniques can generalize well to different types of cracks (e.g., longitudinal, diagonal). It sets a roadmap for expanding datasets and testing additional pre-trained models to enhance robustness further.
4. Integration into Practical Applications: Beyond theoretical advancements, the study offers insights into integrating AI-driven crack detection into real-world workflows, such as automated road inspections, reducing manual effort and improving accuracy.

By combining theoretical rigor with practical implications, this research provides a scalable framework that can be built upon for broader pavement distress classifications and applied in various infrastructure management scenarios.

5. CONCLUSION

Convolutional neural networks (CNNs) have quickly emerged as a prominent and effective method for crack detection. While numerous studies have explored crack detection using CNNs, limited research has focused on leveraging pre-trained models for this purpose. This study applied transfer learning to classify types of asphalt pavement cracks. The main contributions of this research are: 1) the evaluation of five pre-trained CNN models, originally trained on the ImageNet dataset, on the collected asphalt pavement crack dataset; 2) the assessment of the impact of transfer learning on multiclass crack classification using metrics such as Precision, Recall, Accuracy, F1 Score, and Confusion Matrices; and 3) the selection of the most reliable model for further fine-tuning by incorporating additional layers to improve its performance.

The results show that the most reliable model is the EfficientNet B3, achieved scores of 94% F1_ Score and 94% accuracy. Thus, this model was selected and trained on the same dataset by applying transfer learning with pre-trained weights from ImageNet and fine-tuning the CNN. The modified model showed better classification performance with 96% F1_Score and 96% accuracy. The pre-trained models performed very effectively on the diagonal, longitudinal, and alligator crack images. However, misclassifications appeared on the horizontal crack for EfficientNetB0. Moreover, the confusion matrices show that the alligator crack has less accuracy on two models namely, DenseNet121 and MobileNetV2. The reason for this could be the that the alligator data set is smaller than other datasets. EfficientNet B3 had the highest accuracy for both TL and transfer learning with fine tuning.

In future studies, additional crack classes, for example, fatigue cracks, can be examined to further assess the effectiveness of TL and pre-trained models for the classification of asphalt pavement cracks. A larger dataset can be collected, encompassing more crack classes and distresses, to test and assess various pre-trained models. Furthermore, the optimised EfficientNetB0 will undergo additional training to identify a wider range of highway pavement distress, including joint reflection cracking, slippage cracking, corrugation, shoving, block cracking etc. The model's performance metrics will be compared across these distress types.

Additionally, a user-friendly interface will be developed to enable novice users without a programming background to effectively utilise the model. The integration of automated crack detection and classification systems can potentially replace human intervention due to their superior accuracy, cost-effectiveness, and automatic reporting functionality. Future studies may explore the use of AI to develop a tool for suggesting corrective actions for detected and classified cracks.

REFERENCES

Adabre, M. A., Chan, A. P. C., Edwards, D. J., & Osei-Kyei, R. (2022). To build or not to build, that is the uncertainty: Fuzzy synthetic evaluation of risks for sustainable housing in developing economies. Cities, 125, 103644. https://doi.org/https://doi.org/10.1016/j.cities.2022.103644

Ali, K., Shaikh, Z. A., Khan, A. A., & Laghari, A. A. (2022). Multiclass skin cancer classification using EfficientNets – a first step towards preventing skin cancer. Neuroscience Informatics, 2(4), 100034. https://doi.org/https://doi.org/10.1016/j.neuri.2021.100034

Augustauskas, R., & Lipnickas, A. (2020). Improved Pixel-Level Pavement-Defect Segmentation Using a Deep Autoencoder. Sensors, 20(9), 2557. https://www.mdpi.com/1424-8220/20/9/2557

Baduge, S. K., Thilakarathna, S., Perera, J. S., Ruwanpathirana, G. P., Doyle, L., Duckett, M., Lee, J., Saenda, J., & Mendis, P. (2023). Assessment of crack severity of asphalt pavements using deep learning algorithms and geospatial system. Construction and Building Materials, 401, 132684. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2023.132684

Brien, D. O., Andrew Osborne, J., Perez-Duenas, E., Cunningham, R., & Li, Z. (2023). Automated crack classification for the CERN underground tunnel infrastructure using deep learning. Tunnelling and Underground Space Technology, 131, 104668. https://doi.org/https://doi.org/10.1016/j.tust.2022.104668

Canestrari, F., & Ingrassia, L. P. (2020). A review of top-down cracking in asphalt pavements: Causes, models, experimental tools and future challenges. Journal of Traffic and Transportation Engineering (English Edition), 7(5), 541-572. https://doi.org/https://doi.org/10.1016/j.jtte.2020.08.002

Cha, Y.-J., Choi, W., & Büyüköztürk, O. (2017). Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. Computer-Aided Civil and Infrastructure Engineering, 32(5), 361-378. https://doi.org/https://doi.org/10.1111/mice.12263

Choi, W., & Cha, Y. J. (2020). SDDNet: Real-Time Crack Segmentation. IEEE Transactions on Industrial Electronics, 67(9), 8016-8025. https://doi.org/10.1109/TIE.2019.2945265

Dais, D., Bal, İ. E., Smyrou, E., & Sarhosis, V. (2021). Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning. Automation in Construction, 125, 103606. https://doi.org/https://doi.org/10.1016/j.autcon.2021.103606

Dawson, H. L., Dubrule, O., & John, C. M. (2023). Impact of dataset size and convolutional neural network architecture on transfer learning for carbonate rock classification. Computers & Geosciences, 171, 105284. https://doi.org/https://doi.org/10.1016/j.cageo.2022.105284

Deng, J., Dong, W., Socher, R., Li, L. J., Kai, L., & Li, F.-F. (2009, 20-25 June 2009). ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition,

Dong, J., Wang, N., Fang, H., Hu, Q., Zhang, C., Ma, B., Ma, D., & Hu, H. (2022). Innovative method for pavement multiple damages segmentation and measurement by the Road-Seg-CapsNet of feature fusion. Construction and Building Materials, 324, 126719. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2022.126719

Dung, C. V., Sekiya, H., Hirano, S., Okatani, T., & Miki, C. (2019). A vision-based method for crack detection in gusset plate welded joints of steel bridges using deep convolutional neural networks. Automation in Construction, 102, 217-229. https://doi.org/https://doi.org/10.1016/j.autcon.2019.02.013

Edwards, D., Parn, E. A., Sing, M. C. P., & Thwala, W. D. (2019). Risk of excavators overturning. Engineering, Construction and Architectural Management, 26(3), 479-498. https://doi.org/10.1108/ECAM-03-2018-0125

Edwards, D. J., Rillie, I., Chileshe, N., Lai, J., Hosseini, M. R., & Thwala, W. D. (2020). A field survey of hand–arm vibration exposure in the UK utilities sector. Engineering, Construction and Architectural Management, 27(9), 2179-2198. https://doi.org/10.1108/ECAM-09-2019-0518

Elghaish, F., Talebi, S., Abdellatef, E., Matarneh, S. T., Hosseini, M. R., Wu, S., Mayouf, M., Hajirasouli, A., & Nguyen, T.-Q. (2022). Developing a new deep learning CNN model to detect and classify highway cracks. Journal of Engineering, Design and Technology, 20(4), 993-1014. https://doi.org/10.1108/JEDT-04-2021-0192

Fan, Z., Lin, H., Li, C., Su, J., Bruno, S., & Loprencipe, G. (2022). Use of Parallel ResNet for High-Performance Pavement Crack Detection and Measurement. Sustainability, 14(3), 1825. https://www.mdpi.com/2071-1050/14/3/1825

Feng, C., Zhang, H., Wang, S., Li, Y., Wang, H., & Yan, F. (2019). Structural Damage Detection using Deep Convolutional Neural Network and Transfer Learning. KSCE Journal of Civil Engineering, 23(10), 4493-4502. https://doi.org/10.1007/s12205-019-0437-z

Gopalakrishnan, K., Khaitan, S. K., Choudhary, A., & Agrawal, A. (2017). Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection. Construction and Building Materials, 157, 322-330. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2017.09.110

Guo, W., Zhang, J., Cao, D., & Yao, H. (2022). Cost-effective assessment of in-service asphalt pavement condition based on Random Forests and regression analysis. Construction and Building Materials, 330, 127219. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2022.127219

Guzmán-Torres, J. A., Naser, M. Z., & Domínguez-Mota, F. J. (2022). Effective medium crack classification on laboratory concrete specimens via competitive machine learning. Structures, 37, 858-870. https://doi.org/https://doi.org/10.1016/j.istruc.2022.01.061

Ha, J., Kim, D., & Kim, M. (2022). Assessing severity of road cracks using deep learning-based segmentation and detection. The Journal of Supercomputing, 78(16), 17721-17735. https://doi.org/10.1007/s11227-022-04560-x

Hernanda, Z. S., Mahmudah, H., & Sudibyo, R. W. (2022, 6-9 June 2022). CNN-Based Hyperparameter Optimization Approach for Road Pothole and Crack Detection Systems. 2022 IEEE World AI IoT Congress (AIIoT), Hou, Y., Li, Q., Han, Q., Peng, B., Wang, L., Gu, X., & Wang, D. (2021). MobileCrack: Object Classification in Asphalt Pavements Using an Adaptive Lightweight Deep Learning. Journal of Transportation Engineering, Part B: Pavements, 147(1), 04020092. https://doi.org/doi:10.1061/JPEODX.0000245

Huyan, J., Li, W., Tighe, S., Xu, Z., & Zhai, J. (2020). CrackU-net: A novel deep convolutional neural network for pixelwise pavement crack detection. Structural Control and Health Monitoring, 27(8), e2551. https://doi.org/https://doi.org/10.1002/stc.2551

Ilse, M., Tomczak, J. M., & Welling, M. (2020). Chapter 22 - Deep multiple instance learning for digital histopathology. In S. K. Zhou, D. Rueckert, & G. Fichtinger (Eds.), Handbook of Medical Image Computing and Computer Assisted Intervention (pp. 521-546). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-816176-0.00027-2

Jang, K., Kim, N., & An, Y.-K. (2019). Deep learning–based autonomous concrete crack evaluation through hybrid image scanning. Structural Health Monitoring, 18(5-6), 1722-1737. https://doi.org/10.1177/1475921718821719

Joshi, D., Singh, T. P., & Sharma, G. (2022). Automatic surface crack detection using segmentation-based deep-learning approach. Engineering Fracture Mechanics, 268, 108467. https://doi.org/https://doi.org/10.1016/j.engfracmech.2022.108467

Kim, H., Kim, H., Hong, Y. W., & Byun, H. (2018). Detecting Construction Equipment Using a Region-Based Fully Convolutional Network and Transfer Learning. Journal of Computing in Civil Engineering, 32(2), 04017082. https://doi.org/doi:10.1061/(ASCE)CP.1943-5487.0000731

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Commun. ACM, 60(6), 84–90. https://doi.org/10.1145/3065386

Kukah, A. S., Owusu-Manu, D.-G., Badu, E., & Edwards, D. (2022). Evaluation of risk factors in Ghanaian public-private- partnership (PPP) power projects using fuzzy synthetic evaluation methodology (FSEM). Benchmarking An International Journal, 30. https://doi.org/10.1108/BIJ-09-2021-0533

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521, 436-444. https://doi.org/10.1038/nature14539

Li, G., Ma, B., He, S., Ren, X., & Liu, Q. (2020). Automatic Tunnel Crack Detection Based on U-Net and a Convolutional Neural Network with Alternately Updated Clique. Sensors, 20(3), 717. https://www.mdpi.com/1424-8220/20/3/717

Li, J., Liu, T., Wang, X., & Yu, J. (2022). Automated asphalt pavement damage rate detection based on optimized GA-CNN. Automation in Construction, 136, 104180. https://doi.org/https://doi.org/10.1016/j.autcon.2022.104180

Li, Y., Li, H., & Wang, H. (2018). Pixel-Wise Crack Detection Using Deep Local Pattern Predictor for Robot Application. Sensors, 18(9), 3042. https://www.mdpi.com/1424-8220/18/9/3042

Liu, F., Liu, J., & Wang, L. (2022a). Asphalt pavement fatigue crack severity classification by infrared thermography and deep learning. Automation in Construction, 143, 104575. https://doi.org/https://doi.org/10.1016/j.autcon.2022.104575

Liu, F., Liu, J., & Wang, L. (2022b). Deep learning and infrared thermography for asphalt pavement crack severity classification. Automation in Construction, 140, 104383. https://doi.org/https://doi.org/10.1016/j.autcon.2022.104383

Liu, J., Zhao, Z., Lv, C., Ding, Y., Chang, H., & Xie, Q. (2022). An image enhancement algorithm to improve road tunnel crack transfer detection. Construction and Building Materials, 348, 128583. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2022.128583

Liu, Z., Cao, Y., Wang, Y., & Wang, W. (2019). Computer vision-based concrete crack detection using U-net fully convolutional networks. Automation in Construction, 104, 129-139. https://doi.org/https://doi.org/10.1016/j.autcon.2019.04.005

Louati, H., Bechikh, S., Louati, A., Aldaej, A., & Said, L. B. (2022). Joint design and compression of convolutional neural networks as a Bi-level optimization problem. Neural Computing and Applications, 34(17), 15007-15029. https://doi.org/10.1007/s00521-022-07331-0

Loverdos, D., & Sarhosis, V. (2022). Automatic image-based brick segmentation and crack detection of masonry walls using machine learning. Automation in Construction, 140, 104389. https://doi.org/https://doi.org/10.1016/j.autcon.2022.104389

Lu, Q., Lin, J., Luo, L., Zhang, Y., & Zhu, W. (2022). A supervised approach for automated surface defect detection in ceramic tile quality control. Advanced Engineering Informatics, 53, 101692. https://doi.org/https://doi.org/10.1016/j.aei.2022.101692

Luo, Y., & Uzuner, O. (2014). Semi-Supervised Learning to Identify UMLS Semantic Relations. AMIA Joint Summits on Translational Science proceedings AMIA Summit on Translational Science, 2014, 67-75.

Mandal, V., Uong, L., & Adu-Gyamfi, Y. (2018, 10-13 Dec. 2018). Automated Road Crack Detection Using Deep Convolutional Neural Networks. 2018 IEEE International Conference on Big Data (Big Data),

Matarneh, S., Elghaish, F., Pour Rahimian, F., Abdellatef, E., & Abrishami, S. (2024). Evaluation and optimisation of pre-trained CNN models for asphalt pavement crack detection and classification. Automation in Construction, 160, 105297. https://doi.org/https://doi.org/10.1016/j.autcon.2024.105297

Mazzini, D., Napoletano, P., Piccoli, F., & Schettini, R. (2020). A Novel Approach to Data Augmentation for Pavement Distress Segmentation. Computers in Industry, 121, 103225. https://doi.org/https://doi.org/10.1016/j.compind.2020.103225

Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345-1359. https://doi.org/10.1109/TKDE.2009.191

Pan, Y., Zhang, G., & Zhang, L. (2020). A spatial-channel hierarchical deep learning network for pixel-level automated crack detection. Automation in Construction, 119, 103357. https://doi.org/https://doi.org/10.1016/j.autcon.2020.103357

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., . . . Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. ArXiv, abs/1912.01703.

Pei, L., Sun, Z., Xiao, L., Li, W., Sun, J., & Zhang, H. (2021). Virtual generation of pavement crack images based on improved deep convolutional generative adversarial network. Engineering Applications of Artificial Intelligence, 104, 104376. https://doi.org/https://doi.org/10.1016/j.engappai.2021.104376

Pozzer, S., Azar, E. R., Rosa, F. D., & Pravia, Z. M. C. (2021). Semantic Segmentation of Defects in Infrared Thermographic Images of Highly Damaged Concrete Structures. Journal of Performance of Constructed Facilities, 35(1), 04020131. https://doi.org/doi:10.1061/(ASCE)CF.1943-5509.0001541

Qu, Z., Li, Y., & Zhou, Q. (2022). CrackT-net: a method of convolutional neural network and transformer for crack segmentation. Journal of Electronic Imaging, 31(2), 023040. https://doi.org/10.1117/1.JEI.31.2.023040

Ramachandran, P., Zoph, B., & Le, Q. V. (2018). Searching for Activation Functions. ArXiv, abs/1710.05941.

Ranjbar, S., Nejad, F. M., & Zakeri, H. (2021). An image-based system for pavement crack evaluation using transfer learning and wavelet transform. International Journal of Pavement Research and Technology, 14(4), 437-449. https://doi.org/10.1007/s42947-020-0098-9

Rubio, J. J., Kashiwa, T., Laiteerapong, T., Deng, W., Nagai, K., Escalera, S., Nakayama, K., Matsuo, Y., & Prendinger, H. (2019). Multi-class structural damage segmentation using fully convolutional networks. Computers in Industry, 112, 103121. https://doi.org/https://doi.org/10.1016/j.compind.2019.08.002

Shi, L., Long, Y., Wang, Y., Chen, X., & Zhao, Q. (2022). Evaluation of Internal Cracks in Turbine Blade Thermal Barrier Coating Using Enhanced Multi-Scale Faster R-CNN Model. Applied Sciences, 12(13), 6446. https://www.mdpi.com/2076-3417/12/13/6446

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018, 2018//). A Survey on Deep Transfer Learning. Artificial Neural Networks and Machine Learning – ICANN 2018, Cham.

Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ArXiv, abs/1905.11946.

Tan, Y., Cai, R., Li, J., Chen, P., & Wang, M. (2021). Automatic detection of sewer defects based on improved you only look once algorithm. Automation in Construction, 131, 103912. https://doi.org/https://doi.org/10.1016/j.autcon.2021.103912

Tang, W., Huang, S., Zhao, Q., Li, R., & Huangfu, L. (2022). An Iteratively Optimized Patch Label Inference Network for Automatic Pavement Distress Detection. IEEE Transactions on Intelligent Transportation Systems, 23(7), 8652-8661. https://doi.org/10.1109/TITS.2021.3084809

Teng, S., Liu, Z., & Li, X. (2022). Improved YOLOv3-Based Bridge Surface Defect Detection by Combining High- and Low-Resolution Feature Images. Buildings, 12(8), 1225. https://www.mdpi.com/2075-5309/12/8/1225

Tran, K., Kondrashova, O., Bradley, A., Williams, E., Pearson, J., & Waddell, N. (2021). Deep learning in cancer diagnosis, prognosis and treatment selection. Genome Medicine, 13. https://doi.org/10.1186/s13073-021-00968-x

Wang, S. Y., & Guo, T. (2021). Transfer Learning-Based Algorithms for the Detection of Fatigue Crack Initiation Sites: A Comparative Study [Original Research]. Frontiers in Materials, 8. https://doi.org/10.3389/fmats.2021.756798

Wu, L., Lin, X., Chen, Z., Lin, P., & Cheng, S. (2021). Surface crack detection based on image stitching and transfer learning with pretrained convolutional neural network. Structural Control and Health Monitoring, 28(8), e2766. https://doi.org/https://doi.org/10.1002/stc.2766

Wu, Y., Yang, W., Pan, J., & Chen, P. (2021). Asphalt pavement crack detection based on multi-scale full convolutional network. Journal of Intelligent & Fuzzy Systems, 40, 1495-1508. https://doi.org/10.3233/JIFS-191105

Xu, B., & Liu, C. (2022). Pavement crack detection algorithm based on generative adversarial network and convolutional neural network under small samples. Measurement, 196, 111219. https://doi.org/https://doi.org/10.1016/j.measurement.2022.111219

Yacouby, R., & Axman, D. (2020). Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models. https://doi.org/10.18653/v1/2020.eval4nlp-1.9

Yang, C., Chen, J., Li, Z., & Huang, Y. (2021). Structural Crack Detection and Recognition Based on Deep Learning. Applied Sciences, 11(6), 2868. https://www.mdpi.com/2076-3417/11/6/2868

Yang, F., Zhang, L., Yu, S., Prokhorov, D., Mei, X., & Ling, H. (2019). Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection. IEEE Transactions on Intelligent Transportation Systems, PP, 1-11. https://doi.org/10.1109/TITS.2019.2910595

Yang, Q., Shi, W., Chen, J., & Lin, W. (2020). Deep convolution neural network-based transfer learning method for civil infrastructure crack detection. Automation in Construction, 116, 103199. https://doi.org/https://doi.org/10.1016/j.autcon.2020.103199

Yao, H., Liu, Y., Li, X., You, Z., Feng, Y., & Lu, W. (2022). A Detection Method for Pavement Cracks Combining Object Detection and Attention Mechanism. IEEE Transactions on Intelligent Transportation Systems, 23(11), 22179-22189. https://doi.org/10.1109/TITS.2022.3177210

Yoon, H.-Y., Kim, J.-H., & Jeong, J.-W. (2022). Classification of the Sidewalk Condition Using Self-Supervised Transfer Learning for Wheelchair Safety Driving. Sensors, 22(1), 380. https://www.mdpi.com/1424-8220/22/1/380

Zhang, J., Lu, C., Wang, J., Wang, L., & Yue, X.-G. (2019). Concrete Cracks Detection Based on FCN with Dilated Convolution. Applied Sciences, 9(13), 2686. https://www.mdpi.com/2076-3417/9/13/2686