Pakistan Journal of Engineering and Technology, PakJET
Multidisciplinary & Peer Reviewed
Volume: 4, Number: 4, Pages: 50- 56, Year: 2021

Structural Crack Detection and Classification using Deep Convolutional Neural Network

Madiha Zeeshan*, Syed M. Adnan, Wakeel Ahmad, and Farrukh Zeeshan Khan

Computer Science Department, University of Engineering and Technology, Taxila, Pakistan.
Corresponding author: *Madiha Zeeshan (madiha.zeeshan@students.uettaxila.edu.pk).

Abstract- **Cracks are indicators that affect the stability and integrity of infrastructures. Fast, reliable, and cost-effective automatic crack detection methods are required to overcome the shortcomings of traditional approaches. The traditional approaches are subjective, time taking, and need a lot of human resources. Because of these factors, more focus is on deep convolutional neural network models that automatically detect cracks. In this research, we focus on the transfer learning approach based on the deep convolutional neural network model VGG19 (Visual Geometry Group 19). Our proposed method fine-tuned the last 2 convolutional blocks and optimized fully connected layers of the VGG19 model. Then the classification layer of the model is substituted with a 2-label Softmax classifier, which classifies between concrete cracked and un-cracked images. This experiment is performed on 15,000 randomly selected images of walls, pavements, and bridges from publicly available SDNET2018 annotated images dataset. It showed that the improved model provides 91.76%testing accuracy. The precision is 91.97% in detecting images without cracks. The paper concluded that the improved VGG19 model accomplishes superior results in the crack detection process as compared to other proposed methodologies.**

Index Terms -- **Crack Detection, Deep learning, DCNN, transfer learning, VGG19.**

I. INTRODUCTION

Cracks are one of the most initial signs of deprivation of the structure. They are the direct parameter to measure the structural health of infrastructures, i.e., roads, buildings, bridges, pavements, tunnels, etc. These cracks need early maintenance and rectification, avoiding any severe damage to infrastructure and its surroundings. Manual inspection of cracks is a common practice that mostly targeted the manual sketching of crack patterns pointing out the conditions of irregularities [1]. Traditionally, the manual crack detection approaches completely depend on the knowledge and experience of specialists, which is a tedious, time-consuming, costly, subjective method, and mainly unreliable.

To overcome the above shortcoming, it developed a significant interest in image processing-based crack detection methods. A long period of Image-based automatic detection well facilitates the maintenance and health of any structure[2]. But this autonomous system relies on images of the structure. Many factors, i.e. low contrast images of concrete surface and cracks, noise, random shapes, diverse sizes of cracks, and many textures, affect these images. For all such conditions, one method is impossible to work. Hence solely relying on image-based systems limits knowledge reusability, which affects the consistency of results and capability of feature extraction and detection[3]. Image-based methods are challenging to re-implement and a lot of struggles are required in feature engineering.

Recently, researchers took more interest in machine learning-based models using deep learning. DCNN (Deep Convolutional Neural Network) shows superior performance in crack detection, segmentation, and classification[4]. The deep learning methods efficiently perform if we provide them with a huge dataset. But sometimes the unavailability of such a dataset inspires us to use transfer learning; knowledge learned in one domain transferred to other. This is the recent development of the machine learning field that helps us to make the pre-trained model learn features in very little time with a limited amount of data.

In real-life applications, DCNN pre-trained models are intelligent systems whose convolutional layers automatically and precisely learn pixel-level cracked image features during the training process[5]. These models have the capability of generalization.

This paper worked on the DCNN pre-trained VGG-19 model.

- For extracting more precise and generalized features for the crack detection problem, the VGG19 model fine-tuned convolutional blocks 4 and 5.
- Later redesigned and replaced the head layer of the VGG-19 model. These modifications help detect cracks on bridges, walls, and pavements on annotated images taken from dataset SDNET2018.

We divide this paper into sections: Section I is based on Introduction, Section II is the literature review, Section III is Proposed Methodology, Section IV is experiments and results and Section V is conclusions after research.

II. LITERATURE REVIEW

In this section, we summarize various crack detection methods based on image processing and CNN.

In [2] automatic crack detection method is proposed based on UAV with image processing. UAV is used to measure the crack image connected distance. Then image processing is applied to an image by subtracting with median filter, image binarization via Sauvola's method after that image revision by the connection of pixel finally crack decomposition and width calculation. The crack width calculation is like measurements by a crack gauge. Then [6] bridge inspection and crack detection were carried out via image processing techniques. Initially, Fourier transforms were applied to images for crack quantification and change detection. Then the neural network is to map between crack depths with width. For visual routine inspection, a 3D visualization model was presented. Then [7] image-based crack detection technique in cement extracts foreground and background image pixels. Initially, they apply the Sobel filter and then the area of region filter for noise removal from the image. A more refined and clearer detection of major cracks from surrounding objects applies the Otsu thresholding method. This method detects cracks more accurately with less noise in comparison with the Kittler method.

As more research in automatic crack detection, it strengthened deep CNNs based models, which better detect cracks and classify them. In [8] first CNN model Conv-Net was proposed for road crack detection. This model has 4 convolutional layers and 2 fully connected layers. It trained on 500 pavement pictures of size 3264 × 2448 image patches and then further classification was performed whether the image patch was cracked or non-cracked. It marked the cracked image patches as positive and vice versa. We need optimization of this model to introduce a low-cost real-time crack detection method.

In [9] pre-trained CNN state of art model was trained on 300 images with fewer epochs. After the experiment, they concluded that the pre-trained model having 16-22 convolutional layers (VGG16, VGG19, and Google Net) gives 90% of training accuracy. Hence, analyzed that pre-trained networks require fewer training samples. It also analyzed those features learned during the training process are transferable to other types of materials as well.

In [10] damage detection was performed via pertained VGG16 model on images taken by hexacopter UAV. They performed pavement crack detection via pre-trained VGG16 with transfer learning to exclude a fully connected layer. Instead of an FC classifier, it employs machine learning classifiers such as NN, SVM, and RF, achieving 89-90% accuracy in crack detection.

Further [11] is based on the deep learning framework YOLO v2 for road crack detection. They train it on 1813 images. YOLO v2 framework is beneficial because it reframes the detection of the object. This approach achieves a precision of 88.5% and an F1 score of 87.8%. They can improve its accuracies by increasing the number of images in the training set.

In [12] transfer learning was applied to the SDNET2018 dataset using DCNN AlexNet.SDNET2018 dataset comprises images of walls, bridges, and pavements. It gives almost 90-95% accuracy on Alex Net in all classes. Later [13] faster-RCNN applied on 5000 cropped images of bridges that contain crack, non-crack, including handwriting. They divide the dataset into training and testing, where 20% testing and 80% training set. Later faster-RCNN compared with YOLO v2, which shows it has better detection accuracy in detecting cracks and handwriting scripts.

In [14] another deep convolutional neural network model has been proposed for automatic pavement crack detection and classification. DCNN has 3 convolutional, 3 pooling, and 2 fully connected layers. Images are taken via camera and then these images are resized according to requirement. 9k of training dataset for classification of crack and non-crack images and 5.7K training set based on 4 classes. This model provides 99% accuracy in the crack and no crack image classification. Further longitudinal and traverse crack classification has lower accuracy of 94% as compared to other classes' classification.

In [15] Deep Crack neural network is proposed for training the whole images and generating predictions on the pixel level. It has 13 convolutional layers of VGG16. The batch normalization (BN) layer is inserted in between each convolutional and RELU layer for improvement in model generalization. FC layer and pool5 layer discarded. They introduced a side output layer and refinement module based on a guided filter to better segment crack images for detection. The dataset of 537 images provided 88% accuracy and AUC > 0.98. For improvement, more false crack regions need to be added.

In [16] U-Net model based on the FCN framework was proposed. The model is based on an encoder and decoder architecture. This model provides 90% accuracy on the total of 84 images (57 training and 27 testing's) on epoch = 80. This model reduces the personnel work, despite that there is a need to work on a high input size to feed in the model. Further, a lot of hyper-parameters have artificially adjusted that need to check the effect via experimenting.

In [17] Improved I-UNET proposed that follows the same encoder-decoder architecture which has the ability for road crack images classification at the pixel level. To overcome drawbacks of U-Net replacement of the convolutional layer in U-Net with dilated convolutional layer and RELU with ELU. The dataset comprises 173 images from different road sections and I-UNET provides 91.8% accuracy, which is better than U-Net for crack segmentation. I-UNET is a crack segmentation model that accurately separates the crack area from the pavement background and remains unaffected by environmental conditions.

In [18] VGG-16 was applied on images from SDNET2018, CCIC, and BCD datasets. They propose a method of transfer learning that relied on DCNN for crack detection. This method transfers knowledge based on sample knowledge, parameter knowledge, and model knowledge. There proposed model achieves 90% on crack detection images from the SDNET dataset. However, it improved by quantitatively representing knowledge transferred and performance evaluation as compared to other CNN structures.

In [19] quantum transfer learning model was developed, which is the concatenation of pre-trained feature extraction with the quantum circuit as a classifier. VGG19 performs best on the SDNET2018 dataset to detect cracks on concrete images with 91.2% accuracy. This model lacks multi-level classification.

In [20] focus is on NDT (non-destructive testing) to check the bond performance and pull-out strength via Schmidt hammer and ultrasonic pulse velocity (UPV) of the concrete anchor bolt. Then ANN (artificial neural network) was developed to analyze various parameters to predict the load-carrying capacity and pullout strength of anchor bolts. To evaluate the NDT of anchor bolt ANN adds new aspects in the research field.

In [21] experimental research was conducted to estimate the bond/crack condition of concrete surrounded by steel via UPV test. A multi-layer ANN was developed to predict the width of the crack and analyze various influential parameters that lead to bond deterioration. Further, research can be enhanced by having more experiments to know the maximum width of crack on which the UPV test reliably evaluates the condition of the bond.

III. METHODOLOGY

Our proposed model is based on the DCNN state-of-the-art model VGG19 for crack detection and classification of images from the SDNET2018 dataset.

A. DATASET

SDNET2018 publicly available annotated image dataset that compromises 56,000 images of crack and non-crack walls, pavement, and bridges. The pixel size of images is 256 x 256 RGB. Because of its diversity, we used it for training, validation, and benchmarking for AI-based crack detection algorithms. We randomly collect almost 15k images. We train the model on 10,000 images, which include 5000 cracked and 5000 un-cracked images of walls, pavements, and bridges from SDNET2018. We divide these sets of images into 70% training images, 20% validation set, and 30% testing set of cracked and non-crack surfaces. More details about the dataset are shown in Table I.

TABLE I
NUMBER OF IMAGES FROM SDNET2018 DATASET

	Number of Images
Train	10,000
Validation	2,588
Test	3,000

B. IMAGE PREPROCESSING

The dataset comprises 256 x 256 RGB images. Before inputting an image into the VGG19 model, images are resized into 224 x 224 to match the size of the input layer of the CNN model.

C. TRANSFER LEARNING APPROACH

Transfer learning is the knowledge adaptation of pre-trained deep learning models. It is an optimization that improved the performance and progress of the targeted task. These pre-trained models are trained enough on large ImageNet datasets, so they work at that place effectively where we have limited training time and data. The DCNN architecture is based on several layers that focus on feature extraction, segmentation, classification, etc.

CNN architecture has many pre-trained state-of-the-art models. In this paper, we focus on the transfer learning technique of fine-tuning in 2 ways for better detection of cracked and non-cracked surfaces.

- We freeze the weights of the base model and retrain the head layers of architecture composed of fully connected layers.
- We fine-tune convolution block 4 and block 5 layers to efficiently and more accurately learn the features specifically for crack detection.

The basic architecture of VGG-19 is in Fig. 1.

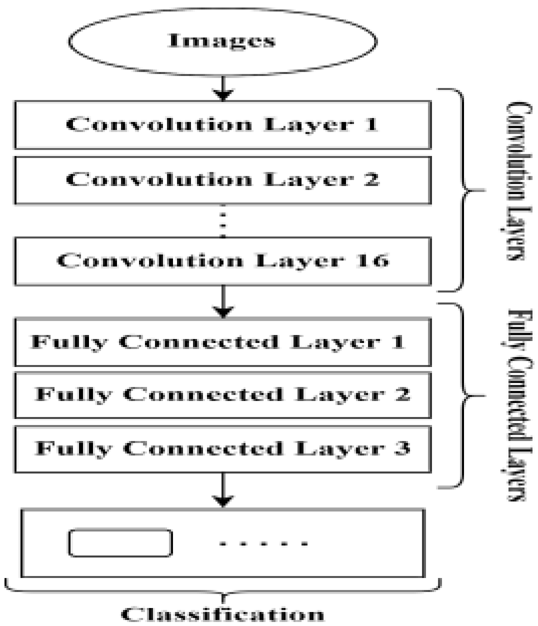

FIGURE 1. Basic Architecture of VGG-19 model

D. IMPROVED VGG19 MODEL FOR CRACK DETECTION

VGG19 trained on millions of images from the ImageNet database. The network depth of this model is 19. It has 144 million parameters processes images of input size of 224 x 224 x 3. This model has 16 convolutions, 5 max pooling, and 3 fully connected layers. The kernel/filter size of the convolutional layer is 3 x 3 with stride 1. Initially, a dot product between filter and patch of filter size input results in a single value. These values later get summed up together for extracting features maps. These features maps show the strength of detected features and their locations in input, such as cracks in cracked images.

A non-linear RELU (Rectified linear activation unit) activation function follows each convolutional layer of the network. As RELU gradients are always zeros and ones. It has no bounded output which results in faster calculation and better accuracies as compared to traditionally using sigmoid functions. We use the RELU activation function is in (1).

$$R(y) = max(0, y) \quad (1)$$

The output of feature maps is location-sensitive. To overcome this sensitivity, we required a down-sampling approach. For this, a max-pooling layer of 2 x 2 with stride 2 to keep the most activated features of an image. A pooling layer after each convolution layer leads to the parameters reduction, training time, and control over-fitting. The 3 fully connected layers determine the parameters of the model. We can classify these parameters into 1000 classes. In the improved version of this model, we replace these 3 FC (fully connected) layers with 1 Flatten layer, two fully connected layers. The last layer, called the classification layer, uses a 2-label Softmax classifier, which classifies input images into cracked and uncracked. The first fully connected layer contains 512 neurons followed by RELU, the second FC layer contains 256 neurons, and the classifier layer has 2 neurons because it is binary classification. For sparse features and to avoid over-fitting, there are several generalization techniques. In this, we include a dropout layer after FC layers. Hence, the improved crack detection VGG-19 model use fine-tuning to transfer network parameters of the pre-trained VGG19 model to a fully connected layer. Fig. 2 depicts the overall improved architecture of the VGG19 model.

The algorithm for the improved VGG-19 model illustrated in Fig. 2, is given below:

1. Enter the sample image of the cracked and un-cracked images from the training sample set.
2. Preprocessing: we resize the input image to 224 x 224 for improvement in training efficiency.
3. Build the improved model via the VGG19 model. The 3 FC (fully connected) layers were replaced by 1 Flatten layer and 2 FC layers. After each FC layer, a dropout layer for more concentrated parameters. Replace the classification layer with a 2-label Softmax classifier.
4. Fine-tune: Freeze the weights of the first 3 convolutional blocks and unfreeze convolutional blocks 4 and 5 to optimize and learn more specific parameters for the detection model by transfer learning.
5. Train the model by setting the learning rate, loss function, and optimizer. Build the model by setting epoch and then batch size for the training set.

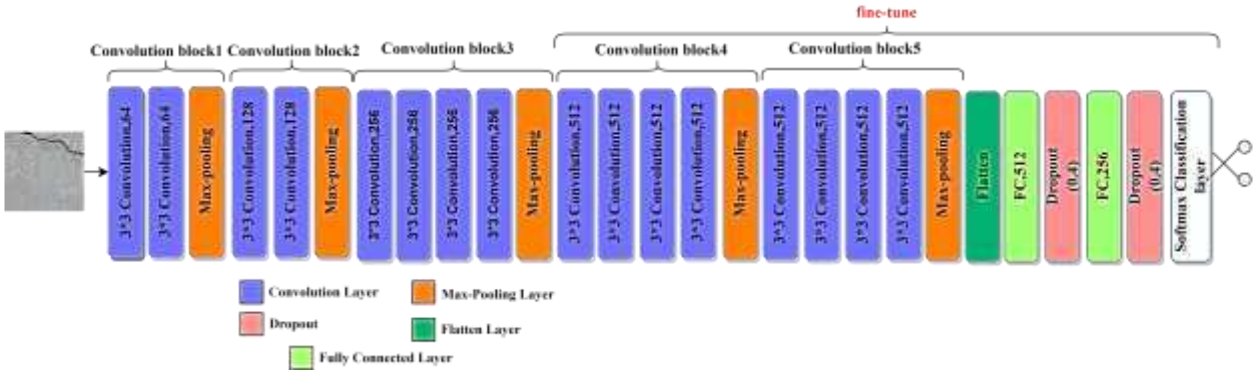

FIGURE 2. Proposed improved VGG19 model for Crack Detection and classification

6. Test the model by extracting sample images from the dataset as the testing sample set for model testing, and calculate precision, recall, and F1-score.

IV. RESULTS AND DISCUSSION

A. EVALUATION METRIC:

For quantitative evaluation of binary classification, considered performance metrics are accuracy (2), precision (3), recall (4), and F1-scores (5).

a) ACCURACY

It measures how accurately our model predicts classes. The formula for accuracy calculation is in (2).

$$Accuracy = \frac{TN+TP}{TN+FN+FP+TP} \qquad (2)$$

b) PRECISION

$$Precision = \frac{TP}{FP+TP} \qquad (3)$$

c) RECALL/SENSITIVITY

$$Recall = \frac{TP}{TP+FN} \qquad (4)$$

d) F1-SCORE

It is the harmonic mean of precision and recall. We use (5) to calculate the F1-score of the model.

$$F1 - score = \frac{2*Precision*Recall}{Precision+Recall} \qquad (5)$$

In the above metrics, TP (true positive) means a model that correctly identifies positive samples as positive. TN (true negative) is the model that correctly identified negative samples as negative. FP (false positive) means that samples are negative, but the model identifies them as positive. FN (false negative) means samples are positive, and the model identifies them as negative.

B. RESULTS

Train and test the model performance on the Google Colab environment implemented in python using Tensor Flow and Keras frameworks. For training the head of the VGG19 model, we replaced fully connected layers with 1 flatten and 2 FC layers. Each FC layer is followed by a dropout layer. Then, for binary classification of cracked and un-cracked images, we add a 2-label Softmax classifier keeping the base model as it is. After too many experiments on the model, set the threshold value of the dropout layer to be 0.4 because it better fit on crack detection problem also gets better precision and recall. For model compilation, we use an optimizer named Adam having a learning rate of 1×10^{-4}. The loss function is binary cross-entropy with a batch size of 32. Later, convolutional block 4 and block 5 are fine-tuned by partially keeping the weights of other convolution layers, which improves the crack detection accuracy.

While fine-tuning FC layers provide us with 85% testing accuracy and 86% precision in detecting cracks on 100 epochs. In another strategy of fine-tuning, unfreezing convolution block 4 and block 5 from the base model and transfer parameters to change head layers gives us 99% training accuracy. The learning rate is 1×10^{-5}, kept as low for better feature learning. It results in 91.76% testing accuracy having a precision of 91.97% on the cracked and un-cracked images. To evaluate the model, test results are in TABLE II.

TABLE II
PERFORMANCE METRIC SCORE ON TEST SET OF SDNET2018 IMAGES

Accuracy	Precision	Recall	F1-score
91.76	91.97	91.76	91.76

Fig. 3 demonstrates the confusion matrix, which shows the performance of classification of cracked and un-cracked images by the model.

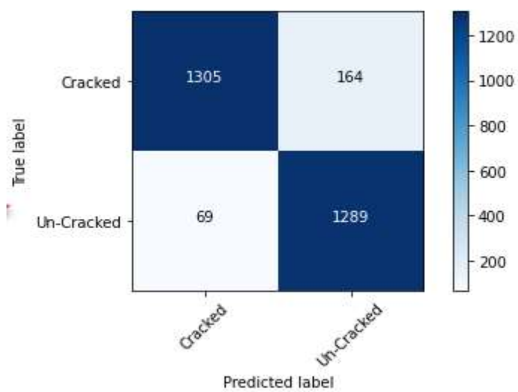

FIGURE 3. The confusion matrix of cracked and un-cracked images depicts the performance of the model.

Fig. 4 measures the diagnostic accuracy of the model ROC (Receiver Operating Characteristic) curve of the model.

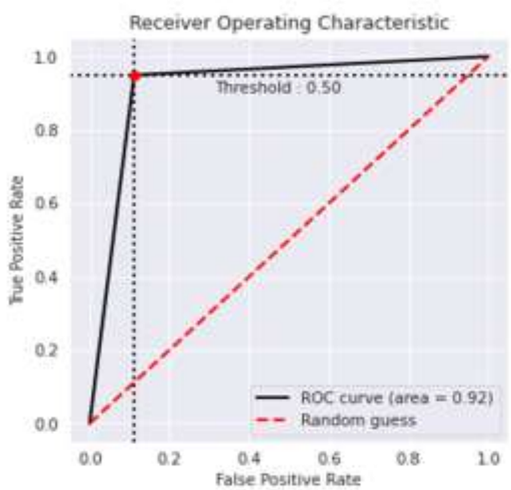

FIGURE 4. ROC Curve having True Positive Rate on the y-axis and False Positive Rate on the Y-axis to examine the performance of the model for crack detection.

The testing results and evaluation metrics depict that fine-tuning of head layer and last 2 convolutional blocks of VGG-19 outperforms other proposed methods for crack detection on images of the SDNET2018 dataset in TABLE III.

TABLE III
PERFORMANCE COMPARISON OF CRACK DETECTION ON SDNET2018

Reference	Year	Accuracy
[22]	2019	89%
[18]	2020	90.95%
[19]	2021	91.2%
Proposed method		**91.76%**

V) CONCLUSION

The study is based on fine-tuning of the VGG19 model, which replaced the head layers into 1 flatten and 2 FC layers. To avoid over-fitting, we add a dropout layer after each FC layer. For the classification of cracked and un-cracked images, utilizes a 2-label Softmax classifier. The experimental results concluded that:

- Fine-tune not only the head layer of the VGG19 model but also unfreeze the weights of convolution block 4 and block 5 for better detection. While fine-tuning to avoid overfitting dropout layer at the threshold of 0.4 incorporated after each FC layer.
- The model trained on the SDNET2018 dataset gives us 91.76% testing accuracy in detecting 3,000 cracked and uncracked images. The precision for the detection of surface cracks on concrete images is 91.97%.
- Our focus in the research is on binary classification. Later research can be extended by working on multi-level classification (i.e. categorizing cracked and un-cracked images belonging to walls, pavements, or bridges) and training the model from scratch with an increased size of the dataset.

After crack detection, the depth detection and categorization of cracks (i.e. thin cracks, wide cracks, mixed cracks, complex cracks) is the limitation of our model. Depth detection is important to check the level of severity of cracks in concrete infrastructure. To measure the depth of crack from images we need to perform cracked image segmentation or edge detection. This is another aspect for the improvement of the automatic crack detection process that needs further investigation.

REFERENCES

[1] A. Bianchini, P. Bandini, and D. W. Smith, “Interrater Reliability of Manual Pavement Distress Evaluations,” *Journal of Transportation Engineering,* vol. 136, pp. 165-172, 2010.

[2] H. Kim, S.-H. Sim, and S. Cho, “Unmanned aerial vehicle (UAV)-powered concrete crack detection based on digital image processing,” ed: EUCENTRE.

[3] A. Mohan and S. Poobal, “Crack detection using image processing: A critical review and analysis,” *Alexandria Engineering Journal,* vol. 57, pp. 787-798, 2018.

[4] S. Bhat, S. Naik, M. Gaonkar, P. Sawant, S. Aswale, and P. Shetgaonkar, “A Survey On Road Crack Detection Techniques,” *IEEE,* pp. 1-6, 2020.

[5] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial Intelligence Review,* vol. 53, pp. 5455-5516, 2020.

[6] R. S. Adhikari, O. Moselhi, and A. Bagchi, “Image-based retrieval of concrete crack properties for bridge inspection,” *Automation in Construction,* vol. 39, pp. 180-194, 2014.

[7] A. M. A. Talab, Z. Huang, F. Xi, and L. HaiMing, “Detection crack in image using Otsu method and multiple filtering in image processing techniques,” *Optik,* vol. 127, pp. 1030-1033, 2016.

[8] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, “Road crack detection using deep convolutional neural network,” *2016 IEEE International Conference on Image Processing (ICIP),* pp. 3708-3712, 2016.

[9] A. G. Sorguç, “Performance Comparison of Pretrained Convolutional Neural Networks on Crack Detection in Buildings,” presented at the Proceedings of the 35th International Symposium on Automation and Robotics in Construction (ISARC), 2018.

[10] K. G. Gopalakrishnan, Hoda & Vidyadharan, Akash & Choudhary, Alok and A. & Agrawal, “Crack Damage Detection in Unmanned Aerial Vehicle Images of Civil Infrastructure Using Pre-Trained Deep Learning Model,” *International Journal for Traffic and Transport Engineering,* vol. 8, pp. 1-14, 2018.

[11] V. Mandal, L. Uong, and Y. Adu-Gyamfi, “Automated Road Crack Detection Using Deep Convolutional Neural Networks,” *IEEE,* pp. 5212-5215, 2018.

[12] S. Dorafshan, R. J. Thomas, and M. Maguire, “SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks,” *Data Brief,* vol. 21, pp. 1664-1668, Dec 2018.

[13] J. Deng, Y. Lu, and V. C.-S. Lee, “Concrete crack detection with handwriting script interferences using faster region-based convolutional neural network,” *Computer-Aided Civil and Infrastructure Engineering,* vol. 35, pp. 373-388, 2020.

[14] N. A. M. Yusof, A. Ibrahim, M. H. M. Noor, N. M. Tahir, N. M. Yusof, N. Z. Abidin*, et al.*, “Deep convolution neural network for crack detection on asphalt pavement,” *Journal of Physics: Conference Series,* vol. 1349, p. 012020, 2019/11 2019.

[15] Y. Liu, J. Yao, X. Lu, R. Xie, and L. Li, “DeepCrack: A deep hierarchical feature learning architecture for crack segmentation,” *Neurocomputing,* vol. 338, pp. 139-153, 2019/04/21/ 2019.

[16] Z. Liu, Y. Cao, Y. Wang, and W. Wang, “Computer vision-based concrete crack detection using U-net fully convolutional networks,” *Automation in Construction,* vol. 104, pp. 129-139, 2019/08/01/ 2019.

[17] L. Wang, X. h. MA, and Y. Ye, “Computer vision-based Road Crack Detection Using an Improved I-UNet Convolutional Networks,” in *2020 Chinese Control And Decision Conference (CCDC)*, 2020, pp. 539-543.

[18] Q. Yang, W. Shi, J. Chen, and W. Lin, “Deep convolution neural network-based transfer learning method for civil infrastructure crack detection,” *Automation in Construction,* vol. 116, p. 103199, 2020/08/01/ 2020.

[19] H. Mogalapalli, M. Abburi, B. Nithya, and S. K. V. Bandreddi, “Classical–Quantum Transfer Learning for Image Classification,” *SN Computer Science,* vol. 3, p. 20, 2021/10/27 2021.

[20] M. Saleem, “Assessing the load carrying capacity of concrete anchor bolts using non-destructive tests and artificial multilayer neural network,” *Journal of Building Engineering,* vol. 30, p. 101260, 2020/07/01/ 2020.

[21] M. Saleem and H. Gutierrez, “Using artificial neural network and non-destructive test for crack detection in concrete surrounding the embedded steel reinforcement,” *Structural Concrete,* vol. 22, pp. 2849-2867, 2021.

[22] E. Elhariri, N. El-Bendary, and S. A. Taie, “Performance Analysis of Using Feature Fusion for Crack Detection in Images of Historical Buildings,” presented at the Proceedings of the 11th International Conference on Management of Digital EcoSystems, Limassol, Cyprus, 2019.