Andrushia, A.D., Neebha, T.M., Umadevi, S., Anand, N., Cashell, K.A. Shrinkage Crack Detection in Expansive Soil using Deep Convolution Neural Network and Transfer Learning. KSCE Journal of Civil Engineering, 2022

Shrinkage Crack Detection in Expansive Soil using Deep Convolution Neural Network and Transfer Learning

A.Diana Andrushia[a], T.Mary Neebha[b], S.Umadevi[c], N.Anand[d*], Katherine A. Cashell[e]

[a]Associate Professor, Department of ECE, Karunya Institute of Technology and Sciences,

Coimbatore - 641114, India, andrushia@gmail.com

[b]Assistant Professor, Department of ECE, Karunya Institute of Technology and Sciences,

Coimbatore - 641114, India, Neebha08@gmail.com

[c]Associate Professor, Centre for Nano Electronics and VLSI Design, SENSE, VIT University,

Chennai - 632014, India, umadevi.s@vit.ac.in

[d]Associate Professor, Department of Civil Engineering, Karunya Institute of Technology and Sciences,

Coimbatore - 641114, India, davids1612@gmail.com

[e]Senior Lecturer, Department of Civil and Environmental Engineering,

Brunel University London, UK, Katherine.cashell@brunel.ac.uk

*Corresponding author

Abstract:

The formation of shrinkage cracks is a natural phenomenon in expansive soils. The development of these cracks affects both the physical and mechanical properties of the soil. This paper proposes new procedures for predicting and detecting the formation of crack patterns in expansive soils, based on customized Convolution Neural Network (CNN) and transfer learning. A total of four different deep learning models are developed to detect the soil crack pattern by changing the convolution layers and hyper-parameters in the analysis. The novelty of the proposed detection methods lies in the use of customized CNN models in shrinkage crack detection for expansive soils. The customized CNN models are constructed by varying the number of convolution layers and the hyperparameters. The results show that the proposed CNN models provide very accurate results and are capable of detecting the presence of cracks in the soil with great accuracy. The best results are from one of the customized CNN models namely the Customized CNN Model 2 which consists of five convolution layers, three activation layers, one pooling layer, two fully connected layers, and a softmax layer. The results from this model are compared with other well-known approaches from the literature and are shown to provide improved results. Overall, the proposed deep learning methods developed in this paper produce excellent results in terms of the accurate detection of shrinkage soil cracks and can also be applied to other types of soil cracks.

Keywords: Shrinkage crack, crack detection, expansive soil, deep learning, deep convolution neural network

1. Introduction:

This paper is concerned with the detection of shrinkage cracking in expansive soils using a deep convolution neural network together with so-called transfer learning. Shrinkage cracks are a natural phenomenon and generally occur due to the evaporation of moisture and a reduction in water content in the soil. On the other hand, swelling occurs in expansive soils when the water content is reasonably high. In this context, when expansive soils such as clays are exposed to the atmosphere, their behaviour is influenced by changes in the local weather conditions. Naturally, expansive soil tends to seep low into the material, and the soil layer may produce shrinkage cracks during evaporation. Clay soils are more likely to experience greater shrinkage and lower seepage characteristics compared with non-cohesive soils, and these properties are the main causes of crack development in these materials. The upper surface of the soil layer tends to shrink as the water evaporates, while the bottom soil layer experiences

no significant change, and the resultant shrinkage is, therefore, inhomogeneous through the overall soil depth [Shi et al. (2014) and Xuhe Gao et al. (2020 a,b)].

Shrinkage cracking can be a significant issue and is, for example, the principal cause of anti-seepage failure of compacted clay liner in landfill covers [Wan et al. (2018)]. Though it is a natural phenomenon, dehydration cracking and the effects of climate change are aggravating these issues. The formation of cracks may alter the physical properties of soil which, in turn, cause geological, geotechnical, and environmental problems such as landslides, barrier dysfunction, ground subsidence, and embankment [Xu et al. (2021)]. The soil cracks induce several physical, biochemical and chemical changes in the soil [Armstrong et al. (2000)].

Researchers have generally focused on the development of soil cracks and their morphology, as well as quantitative measurements to understand the crack patterns [Kumar et al. (2015)]. The soil morphology illustrates the drying process in the material and also provides information about large-scale crack patterns. In addition, cracks in clay soils, in particular, may weaken the physical properties of the material thereby affecting the stability of nearby structures. The formation of surface cracks due to shrinkage affects the stress capacity of the soil. An uneven distribution of water content in the soil, particularly when coupled with high relative humidity, can cause significant variations in the crack density for clay soils [Lu and Likos (2004)]. It is very important to identify soil crack patterns to evaluate the permeability of the soil, which is in turn important when considering soil improvement strategies. The geometric parameters of the soil cracks are one of the important metrics for predicting soil engineering properties and therefore it is important to obtain and understand the soil crack network in a reliable manner. There are several traditional methods available to evaluate the geometric properties of soil cracks in the field such as taking manual measurements or non-destructive testing. However, these are not necessarily reliable or accurate owing to their inherent dependency on human readings. In more recent years, due to vast and rapid computational advancements, it is possible to detect the soil cracks efficiently without human intervention, using digital images. The digital image processing methodology involves many fundamental steps including pre-processing, noise removal, and crack detection. Image processing techniques can also be used to analyze the characteristics of cracks in clay subjected to drying shrinkage. The factors influencing the development of shrinkage cracks are the water content, the wetting-drying cycle, and the composition of soil [Lu and Likos (2004)]. Many researchers have used different methods to predict, analyze and quantify the crack formation for cohesive or expansive soils [e.g. Tan and Kong (2006) and Gong et al. (2009)]. Accurate detection and assessment of

shrinkage cracks in clay soils is very important as it strongly affects slope stability. Nevertheless, the traditional methods of detecting cracks, particularly for expansive soils, have significant shortcomings, as described, and therefore newer methods employing modern technologies are being developed.

In this context, a very current and topical approach to processing the data from digital imagery is to adopt a machine learning strategy. To process large quantities of input data, machine learning-based methods such as support vector machine (SVM) learning [Li et al. (2017)] and extreme learning machine (ELM) [Dai et al. (2019)] have been adopted to classify the cracks in concrete from the input images. Andrushia et al. (2019, 2020) investigated crack detection methods for concrete structures subjected to fire. In this work, the images of the fire exposed concrete were analyzed in the transform domain to identify and locate the cracks.

In these different machine learning methods, a feature extraction process is carried out as a primary step to learn the input images. If a greater number of features are to be extracted, then feature reduction techniques can be applied to find the optimal features. Principle component analysis (PCA) is then employed to perform feature selection and thereby yield an informative feature set. The optimal features are given to classifiers to detect the cracks accurately. In addition, there are many different types of classifiers available to detect the cracks from the input images, including the so-called K nearest neighbour method (KNN), artificial neural networks (ANN), fuzzy logic with SVM, and genetic algorithms with SVM [Choudhary and Dey (2012)].

The traditional steps in pattern recognition involve pre-processing, feature extraction, segmentation, classification, and recognition and many algorithms can be introduced at each stage. The feature extraction step is the most challenging of these steps because the high-level features that are related to the damage location are provided by the user, which can reduce the overall accuracy of the detection process. This is overcome by adopting a deep learning technique that can perform automatic feature extraction steps and can deal with a very large number of inputs.

Deep learning is an emerging artificial intelligence tool that has created significant advancements in the area of pattern recognition in recent years and has been applied successfully in a wide variety of fields. The performances of these techniques are excellent in comparison with machine learning techniques [Liu et al. (2019)]. Cha et al. (2017) used deep convolution neural networks to detect damage to concrete surfaces. Also, Andrushia et al.

(2021) used U-net with residual connection architecture for damage detection in concrete structures. For pavement crack detection, Fan et al. (2020 a, b) used ensemble networks and encoder-decoder architecture; the efficiency of this approach was compared with different deep learning methods and achieved an accuracy of 92.1%. Drouyer (2020) used deep learning models to detect the cracks in concrete structures. Ali et al. (2021) compared multiple pre-trained models to detect the cracks in concrete structures. Auvray et al. (2014) discussed the image-based crack quantification for soft soils, in which the cracks were found using image processing methods.

From the above discussion, it is clear that although there are a variety of different deep learning techniques available which can detect defects in concrete structures, tunnels, dams, and pavements, there is very little available information in the literature on using these techniques for the detection of cracks in the soil. In this context, the current paper aims to explore the use of different deep learning techniques include a convolution neural network and transfer learning to detect the development of cracks in expansive soils. The farmlands under the foothills of the Western Ghat mountain range in India are identified as the local environment for this study, as the soil in this area is characterized as an expansive clay material.

Four deep neural models are employed to detect soil shrinkage cracks. The first and second approaches are Customized CNN models whilst the third and fourth models are pre-trained. The novelty of the proposed method is in the usage of Customized CNN models in shrinkage crack detection for expansive soils, which has not previously been studied. These customized CNN models are developed by varying the number of convolution layers and the hyper-parameters. The detection of soil cracks using deep-learning strategies is very rarely studied in the literature. The pre-trained architectures of VGG-19 and ResNet-50 have been used to detect the soil shrinkage cracks. A set of real time images are captured from the local environment for the analysis. The detailed steps of each model are described in the following sections. The performance of each model is analyzed and the accuracy of the results is compared with those obtained using other deep learning as well as traditional methods.

2. Methodology

The area of interest for this investigation is in farmlands near the Western Ghats mountain range in India with geographical coordinates of 10.9412° N, 76.7264° E, as shown in Fig. 1. The study area is located in the western part of Tamil Nadu. This image highlights the location under investigation, in which the soil images are captured. The average annual rainfall in this

area falls between 550 and 710 mm, of which the northeast and the southwest monsoons contribute 43.5% and 26.3% respectively to the total rainfall. The mean minimum and maximum temperatures of the location during summer and winter vary between 19 and 35°C.

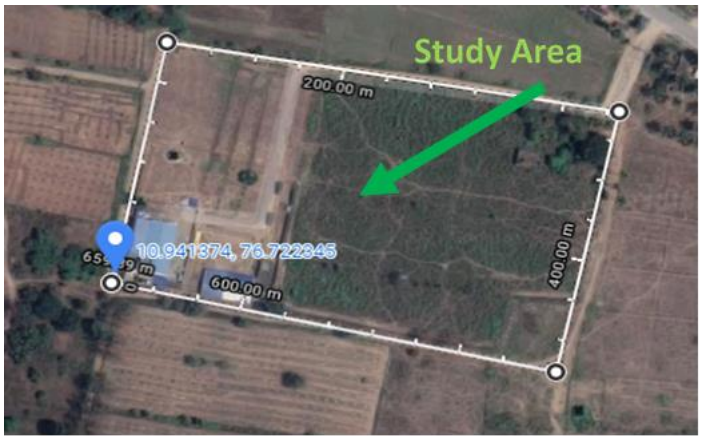


Fig. 1: Field Location of Study Area

This paper is concerned with soft clay soil, which is an example of an expansive soil. Expansive soils are significant for engineering applications because they tend to swell when they absorb water and then shrink during the drying process, causing stability issues for buildings and structures. The swelling occurs owing to the presence of expansive clay minerals in the material. In general, clay soils with a plasticity index value of greater than 25 and the liquid limit value of more than 40 are likely to experience swelling [Thomas et al. (2016)]. Expansive soils are found in many regions under the Western Ghats mountain range in India, and these soils pose settlement problems leading to foundation issues in structures. In the present study, the clay material is assumed to be from a landscape with a fine-grained soil datum, and a relatively high void ratio. The soil is in a plastic state and has a range of water content values. The details of the soil considered in the current investigation are shown in Table 1. The presented values were determined in the geotechnical laboratory at the Karunya Institute of Technology and Sciences.

Table 1: Physical Characteristics of the Soil

Soil Property	Value
Specific weight (dry state)	17.75 kN/m^3
Specific gravity	2.55 – 2.62
Size	< 0.002 mm
Particle size distribution	91.5%
Optimum moisture content	14.2 – 15.2%
Liquid limit	58-72

Plastic limit	36-42
Shrinkage index	24-32
Free swell index	66-75%

It is seen from the data presented in Table 1 that the soil samples have a high degree of expansion coupled with relatively high compressibility. Fig. 2 shows the soil samples and instruments that were considered for the laboratory testing, including (a) an image of a soil sample, (b) the device for determination of the moisture content (Atterburg limits), and (c) the tool for measuring the swelling pressure.

(a) (b) (c)

Fig. 2: Images from the experimental investigation of the soil properties including (a) the Soil sample, (b) the moisture content measuring apparatus including a grooving tool and moisture can, and (c) the device for measuring the swelling pressure.

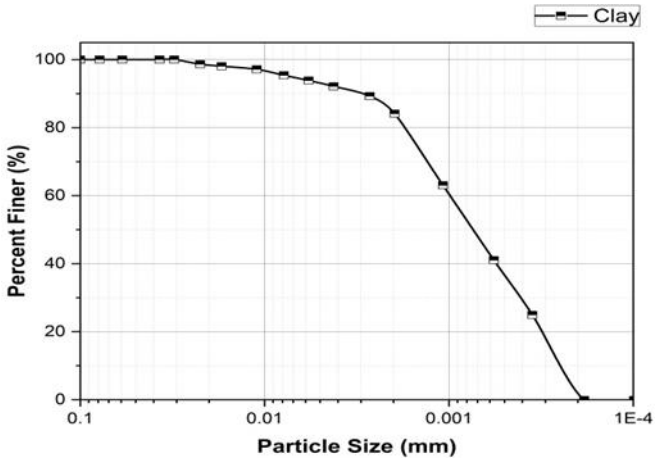
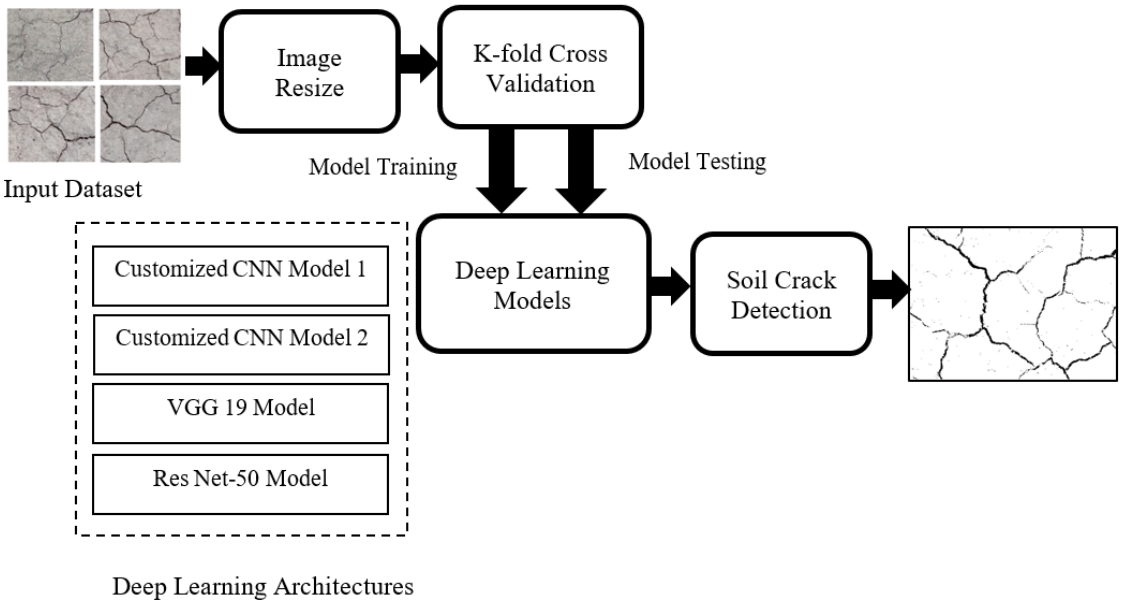


Fig. 3. Particle Size Distribution of the Clay Soil

The tests to determine the Atterburg limits were conducted following the Indian standards [IS 2720 Part 5. (2006) and IS 2720 Part 6. (2001)]. Sieve analysis was conducted for the soil samples in accordance with the guidance given in IS 2720 Part 4. (2015). Fig. 3 presents the particle distribution of the clay soil samples that were considered in this analysis.

3. Crack detection procedure

The proposed detection method comprises three distinct stages, as indicated in Fig. 4. Firstly, the images in the dataset are arranged. Then, the customized CNN architecture and deep learning algorithms are generated before finally, in the third stage, the model is evaluated and comparisons are made. These steps are described in more detail in the following sub-sections. As indicated in the figure, four different models are studied in this work, namely (i) the Customized CNN Model 1, (ii) the Customized CNN Model 2, (iii) the VGG-19 Model and the (iv) ResNet-50 Model.



Fig, 4: Flowchart for the proposed soil crack detection

Once a study area has been identified and selected, the dataset can be generated by taking a series of images. The images are captured using a camera with a resolution of 3200×4800, after which they are resized to 64×64. Fig. 5 shows some sample images of soil shrinkage cracks.

Fig. 5: Sample Images of Soil Shrinkage Cracks

3.1. Convolution Neural Networks (CNN) Architecture

The CNN consists of four major segments, which are the convolution layer, pooling layer, flattening layer, and fully connected layer. The CNN was designed as a feed-forward neural network. The convolution and pooling layers perform feature extraction steps.

3.2 Convolution Layer

The primary purpose of the convolution layer is to extract the desired features from the input data. This is done by finding local contacts of the data sample from the input layer. The convolution operation is performed between the matrix of input and feature detector to obtain the feature map. In this later, the dimensions of the input data are reduced thus simplifying the processing requirements. Many feature vectors are applied to the input, to get multiple feature maps. The derived feature maps are then varied with different convolution kernels. The final feature vector (FV) is applied to the activation layer, given as:

$$FV = \sum(I_{n\times n} + W_{n\times n}) + A \quad (1)$$

where $I_{n\times n}$ is the input to which the convolution operation is being applied, $W_{n\times n}$ is the filter weights, *n* is the kernel size and *A* is the filter bias of the convolution layer.

3.2.1 Activation layer

The activation function defines how the weighted sum of the input is transmitted into an output from the nodes in a layer of the network. Two different activation functions are employed in this, namely a rectified linear unit (ReLU) and Softmax. ReLU is a linear function that is applied to the output of the convolution layer (i.e. the feature maps). For a positive input, ReLU produces the output as unity whereas, for a negative input, the output is zero. ReLU is specifically employed because it accelerates the computational processes compared with other activation functions such as tanh and sigmoid. The prime advantage of the ReLU function is that it does not trigger all the neurons at the same instant. The ReLU function is defined as:

$$\alpha(I) = \max(0, I) \quad (2)$$

where 'I' is the input matrix elements and '$\alpha'$is the mathematical function of the activation layer. The second activation function Softmax normalizes the input and produces a probability function that ranges from zero to unity. The cross-entropy cost function is generally used with the Softmax activation function.

3.2.2 Pooling Layer

The pooling layer is used to reduce the dimensions of the input further [Simonyan et al., 2015]. It separates the feature map into non-overlapping pooling kernels. Specifically, it down-

samples the input and also reduces the number of model parameters. This has the positive effect of simplifying the computational complexity and improving the potential for generalization of the model. The sub-sampling process uses max-pooling and average pooling techniques. For max-pooling, the maximum value of every smaller region is considered whilst the remaining values are removed. On the other hand, the average pooling process considers the average of all values in each smaller region. In general, max-pooling is used to find the prime features of an input image, whereas, average pooling helps the model to assess the full features of the input image. It is noteworthy that average pooling is a higher workload with a greater amount of data analyzed compared with maximum pooling.

3.2.3 Fully Connected Layer

In this final layer, the feature maps that are generated in the previous layers are aligned into columns and then feed into the neural network. The fully connected layer is considered to be a traditional neural network that gives logical inference. It converts a three-dimensional matrix into a single-dimensional vector by using a full convolution operation. The mathematical equation that defines this layer is given as:

$$C_{V_0 \times 1} = wt_{V_0 \times V_i} \, I_{V_i \times 1} \times A_{V_0 \times 1} \quad (3)$$

where V_0 and V_i are the input and output matrix size, respectively C denotes the outcome of the fully connected layer and wt and A represent the weight and bias matrices, respectively. The network generates a prediction and upon which the cost function is calculated which then reveals the performance of the network. The cost function calculation is followed by backpropagation, weights tuning, and feature maps optimization. The Adam optimization algorithm is employed which effectively updates the weights of the network depending upon the training data.

4. Model Architecture

As stated before, four different modelling approaches are used for the training and testing of the soil cracks dataset. A detailed description of each of these models (two customized models and two pre-trained models) is provided hereafter.

4.1 Customized CNN Model 1

The first CNN model is constructed by tuning different hyperparameters including the number of filters, pooling locations, number of convolution layers, stride, sizes, and number of fully connected layers. The hyper-parameter selections are done on a trial-and-error basis and there

is no standard mathematical formulation for setting the parameters for the specific dataset. The first, second, and third convolution layers have 32, 64, and 128 feature maps, respectively, employing the ReLU activation function. It is noteworthy that convolutional layers in a convolutional neural network summarize the presence of features in an input image.

After each convolution layer, a 2×2 dimensions max pooling layer is used. Pooling layers provide an approach to down sampling feature maps by summarizing the presence of features in patches of the feature map. Two common pooling methods are average pooling and max pooling and the latter is employed in the work presented herein. There are two dense layers used. The initial dense layer has 256 output perceptrons with ReLU and the second dense layer has output perceptron with a softmax function. The learning rate is given as 0.0001. As discussed before, the Adam optimizer with the cross-entropy cost function is used. Fig 6 shows the detailed architecture of the Customized CNN Model 1.

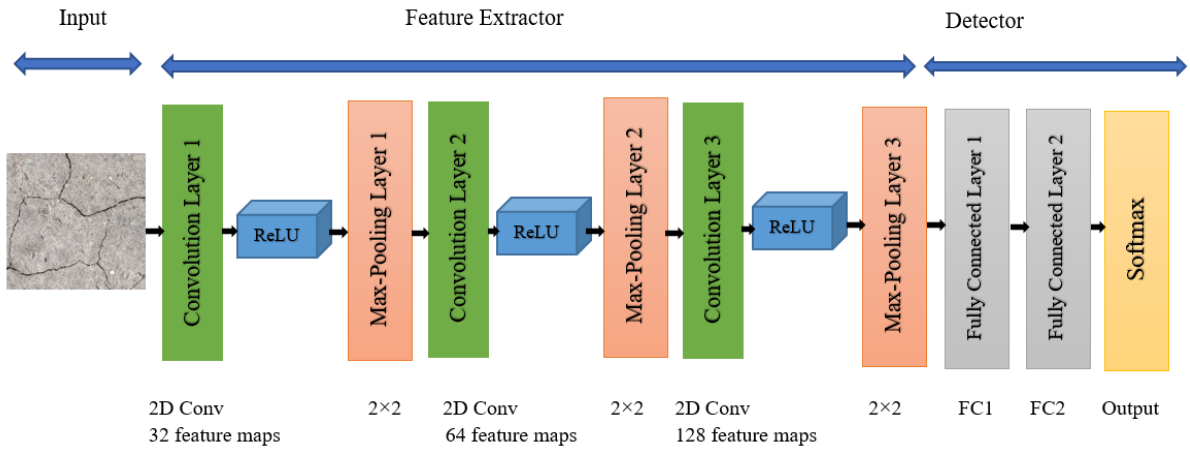


Fig. 6: Flow chart of the Customized CNN Model 1

4.2 Customized CNN Model 2

The customized architecture of the second CNN Model consists of five convolution layers, three activations, one pooling layer, 2 fully connected layers, and a softmax layer (Fig. 7). Each of these layer arrangements are used to increase the performance of the model and to extract the features and reduce the dimensionality requirements of the inputs. The convolution layers are used to improve the spatial invariance property and to recognize the key features from the input image. The CNN architecture relies on the sequential features of the input data. If the input is highly sparse then the training ability of the neural network should be downgraded [Luo et al. (2020)]. In this context, the Adam optimizer is again used in this model to handle the sparse input data. It handles the sparse data well in comparison with Adadelta and RMSprop optimizer by adding bias correction and momentum to RMSprop. The learning rate of the

second model is reduced to 0.00001. The detailed steps involved in the implementation of both of the customized CNN models (in Section 4.1 and 4.2) are outlined as:

Step 1: The input image with size of 64×64 is applied through the convolution layers which produce 32 feature maps (ReLU activation function). The output of the convolution layer is passed into the maximum pooling layer with 2×2 dimensions:

$$\alpha(I) = \max(0, I) \quad (4)$$

Step 2: The input image with a size of 64×64 is applied through the convolution layer which produces 64 feature maps (the ReLU activation function). The output of the convolution layer is passed into the maximum pooling layer with 2×2 dimensions.

Step 3 Part (i): For the Customized CNN Model 1, the input image with a size of 64×64 is applied through the convolution layers which produce 128 feature maps (ReLU activation function). The output of the convolution layer is passed into the maximum pooling layer with 2×2 dimensions and then flattened.

Step 3 Part (ii): For the Customized CNN Model 2, the input image with a size of 64×64 is applied through the convolution layers which produce 512 feature maps (ReLU activation function). The output of the convolution layer is passed into the maximum pooling layer with 2×2 dimensions and then flattened.

Step 4: The outcome of the final layer is passed through a fully connected layer with 256 perceptrons (for model 1) and 1024 perceptrons (for model 2) as well as the ReLU function.

Step 5: The model is compiled using the Adam optimizer and a learning rate of 0.0001.

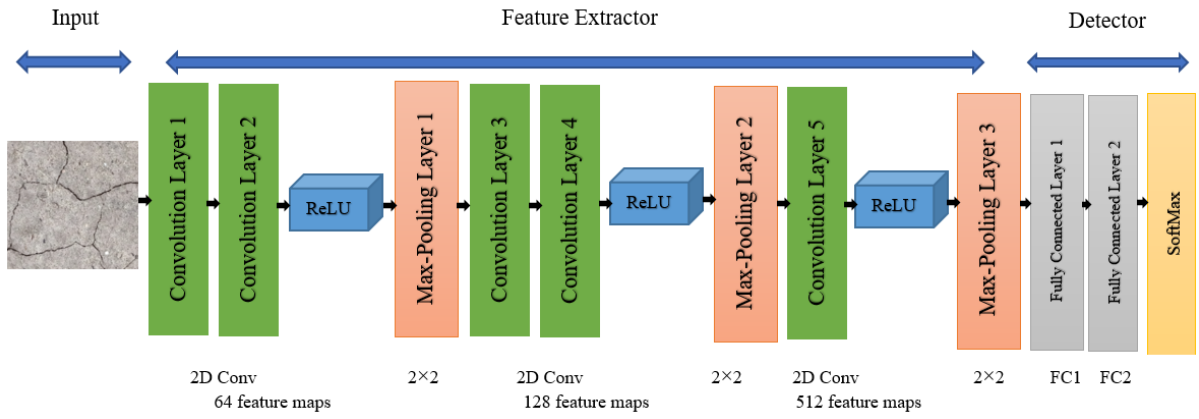


Fig. 7: Flow chart of Customized CNN Model 2

4.3 VGG-19 Model

The VGG-19 transfer learning model was first proposed by Simonyan et al (2015) and consists of 16 convolution layers and 3 fully connected layers with a 3×3 filter size. The smaller kernel size is used to reduce the number of parameters and to access the entire input image. The use of multiple kernels increases the depth of the neural model. The model performs 2×2 maximum pooling with stride 2. This means that the pooling layer reduces the size of each feature map by a factor of 2, e.g. each dimension is halved, reducing the number of pixels or values in each feature map to one quarter the size. As before, the Adam optimizer is employed with the cross-entropy cost function to reduce the learning rate to 0.00001. The VGG-19 Model emphasizes the concept that CNN model need to have a deep layer network which can interpret the visual data hierarchically.

4.4 ResNet-50 Model

The fourth approach is the ResNet-50 Model which employs pre-trained neural networks. He et al. (2016) originally proposed the residual network 'ResNet' which is an artificial neural network (ANN) primarily used for image classification. It introduces residual connections between the layers which help the model to reduce losses and improve the performance during the training phase. The ResNet convolution layer has 3×3 filters and the downsampling is performed by convolution layers with a stride of 2. The final fully-connected layer employs ReLU and Softmax activation functions. The learning rate of the model is 0.000001 and, as before, the Adam optimizer is used. The residual connection in the ResNet network is used to nullify the issues of diminishing gradient and degrading accuracy which occur in deep networks. Hence, the training error is reduced and the convergence rate is higher in comparison with other deep networks. ResNet networks have been developed in different architectures (e.g ResNet 34, ResNet Xt, ResNet V2 and ResNet-50) and ResNet-50 is a residual network that consists of 50 layers. A graphical representation of the ResNet architecture is presented in Fig. 8.

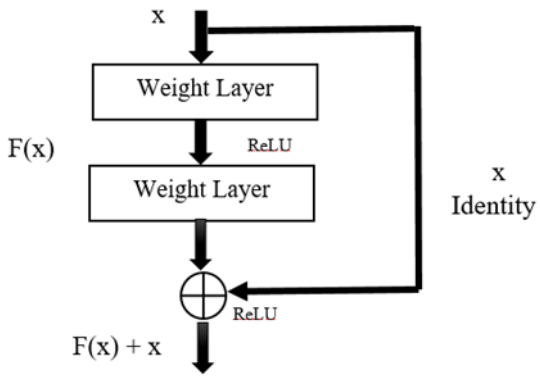


Fig. 8: Residual Learning block of ResNet-50 Architecture

5. Experimental Results

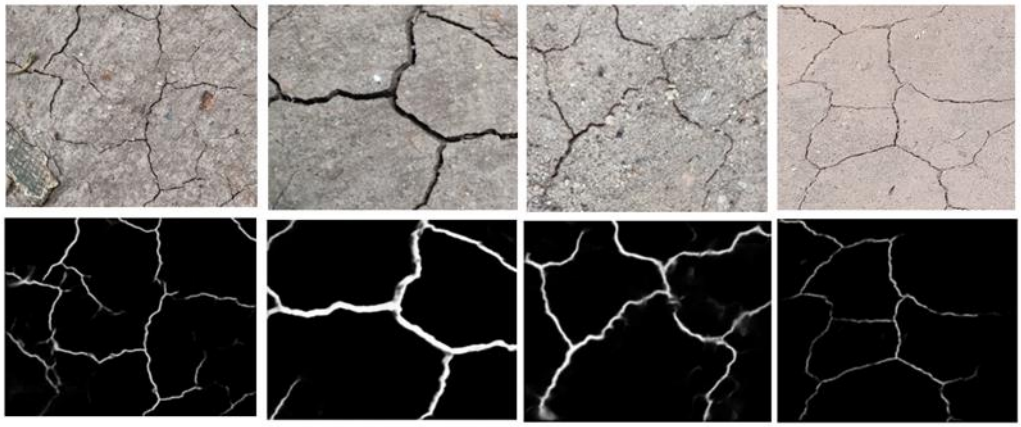
The performances of each of the four deep learning models which were developed for soil crack detection are evaluated hereafter based on the input size of the dataset, model parameters, and the convergence rate. The experiments were implemented using MATLAB 2018, a 2.90 GHz processor with a 12 GB GPU card NVIDIA GTX1050. A total of 5200 images were obtained during the analysis. The approach adopted herein employs k-fold cross-validation as the resampling procedure to evaluate the machine learning models for the sample set. Instead of training fixed data only as 'train and test', different combinations of data have been used. Initially, different k-fold-scenarios selected the training and testing dataset. After a number of trials, a fivefold cross-validation technique was adopted to evaluate the proposed methods. In the fivefold cross-validation, the entire dataset is segregated into 5 mutual folds. Each fold is used once to access the model performance, generated from the collective data of the remaining 4 subsets. It results in 5 independent performance estimations. This method assigns all soil crack images to a validation subset once, thus can be utilized for cross-validation evaluation. The cross-validation procedure is repeated after every 5 epochs. 80% of input images are used for training, 10% of input images are used for testing, and 10% are employed for validation assists in avoiding the overfitting problem.

5.1 Analysis of the Customized and Pre-trained Models

The number of convolution layers is lower in the customized model in comparison with the pre-trained models. The VGG-19 models are three to four times deeper than the customized model. The architecture of the VGG-19 and ResNet models are quite similar but the ResNet model is deeper than the VGG-19 Model. In terms of the number of parameters, the ResNet

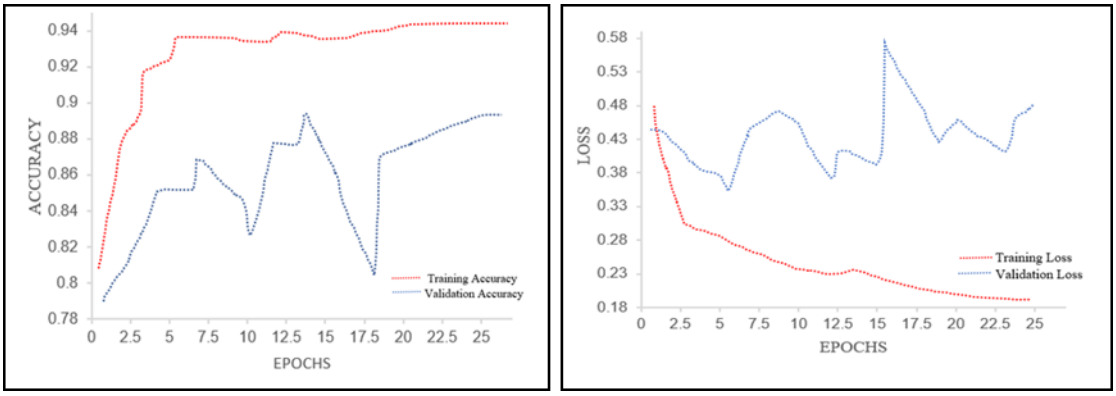
model is less complex. The number of epochs in the proposed work is tuned as 25 for all four models. The losses of all models are minimized at the 25^{th} epoch. Following this, there was found to be no further improvement in the accuracy of the models. The relatively low number of epochs helps to reduce the computational complexity of the architecture. Also, it has been shown that a relatively larger number of epochs may tend to over-fit the models [Yan et al. (2019)].

Fig. 9 presents a comparison between the soil samples (in top images) and the results predicted by each of the four modelling approaches (lower images). It is observed that all of the major and minor cracks are detected with different lighting conditions. Figs. 10-13 presents the model accuracy and loss results for each of the four modelling approaches, respectively; the data is also presented in Table 2. With reference first to Fig. 10, it is observed that Customized CNN Model 1 has a training accuracy of 93.82% and a validation accuracy of 88.83%. The training and validation loss for Customized CNN Model 1 is 20.01% and 39.14%, respectively. The corresponding training accuracy figures for Customized CNN Model 2 are 96.91% and the testing accuracy is 94.15%. On the other hand, the training and validation loss figures are 10.15% and 26.18%, respectively. By comparing the two customized models, it is shown that although both CNN approaches provide an excellent prediction, the second model is the most accurate and provides more consistent and efficient scores.



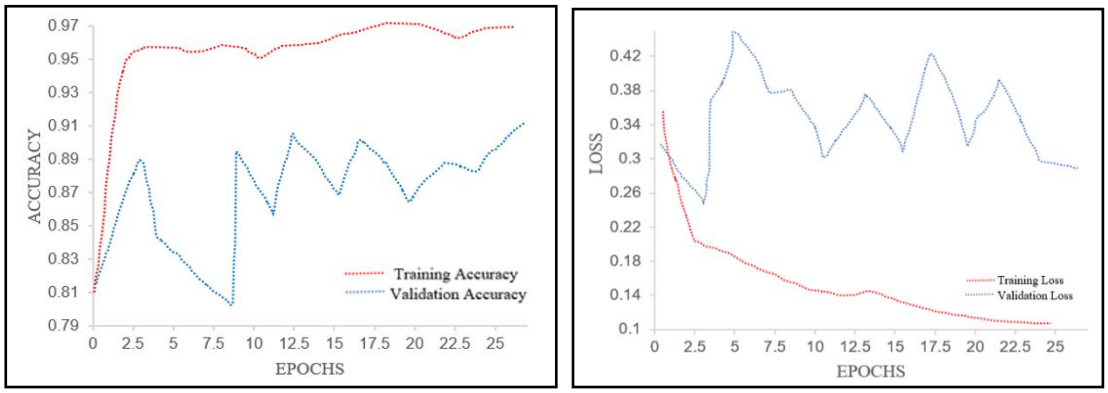
(a) (b) (c) (d)

Fig. 9: Output of soil shrinkage crack detection from each of the four approaches including (a) the Customized CNN Model 1, (b) the Customized CNN Model 2, (c) the VGG-19 pre-trained approach and (d) the Res-Net-50 CNN Model 1



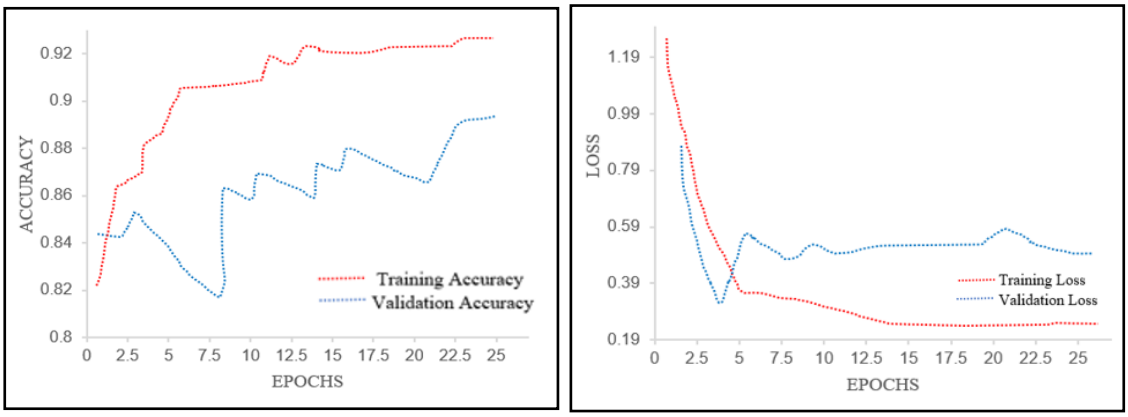
(a) (b)

Fig. 10. Results for Customized CNN Model 1 including (a) training accuracy versus validation accuracy and (b) training loss versus validation loss



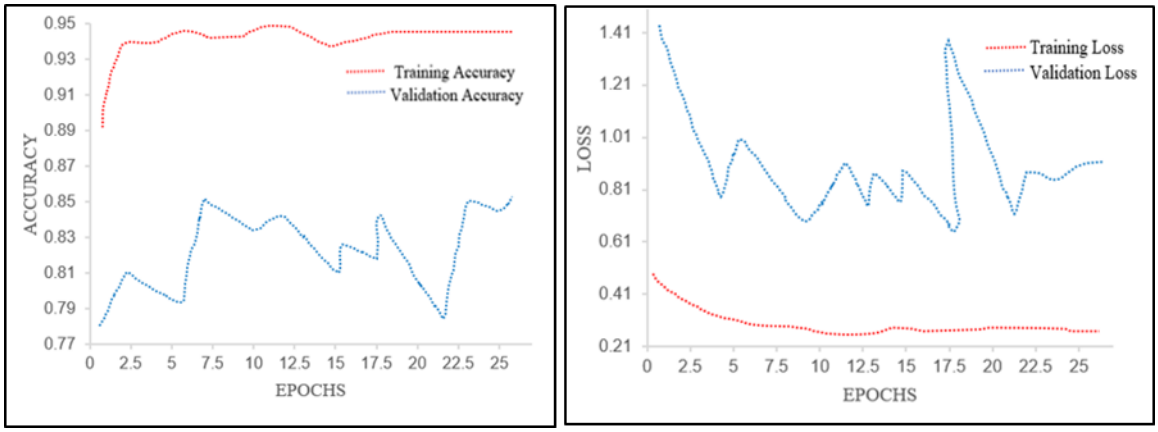
(a) (b)

Fig. 11. Results for Customized CNN Model 2 including (a) training accuracy versus validation accuracy and (b) training loss versus validation loss



(a) (b)

Fig. 12. Results for the VGG-19 pre-trained model including (a) training accuracy versus validation accuracy and (b) training loss versus validation loss



(a) (b)

Fig. 13. Results for the ResNet-50 pre-trained model including (a) training accuracy versus validation accuracy and (b) training loss versus validation loss

With reference to the pre-trained models, the model accuracy and loss graphs are shown in Figs. 12 and 13 for the VGG-19 and ResNet-50 Models, respectively. The validation accuracy of the ResNet-50 Model is relatively low and the validation loss is also quite high compared with the VGG-19 Model. The results highlight the variation in accuracy and loss values for increasing epochs. The VGG-19 Model has less overfitting and greater testing accuracies compared with the ResNet-50 Model. Overall, the pre-trained models have a higher number of layers and their testing accuracy is shown to be less than for the customized CNN models. Smaller datasets are used for both training and testing so that the pre-trained models produce lesser accuracy. When a larger dataset is employed, the transfer learning models generally give improved outputs.

Table 2: Accuracy and Loss Values for the Four Models

Name of the Model	Training Accuracy (%)	Training loss (%)	Validation Accuracy (%)	Validation Loss (%)
Customized CNN Model 1	93.82	20.01	88.83	39.14
Customized CNN Model 2	96.91	10.15	94.15	26.18
VGG-19 Model	92.92	18.95	87.63	36.41
ResNet-50 Model	94.71	15.81	84.11	70.11

Table 3 highlights the accuracies of the Customized CNN models as well as the pre-trained models at the first and twenty-fifth epochs. The accuracy of the VGG-19 Model differs

significantly at the first and twenty-fifth epochs. The ResNet-50 Model had lower accuracy due to the identity mapping concepts which resulted in its architecture being quite complex.

Table 3. Accuracy of the Models at the First and 25th Epoch

Name of the Model	1st Epoch	25th Epoch
Customized CNN Model 1	0.873	0.891
Customized CNN Model 2	0.934	0.927
VGG-19 Model	0.839	0.884
ResNet-50 Model	0.832	0.850

5.2 Evaluation Metrics

The performance metrics are carefully chosen to determine the best-performing model as these play a major role in finding the efficiency of the detection method. In the proposed work, the performance metrics of accuracy, precision, recall, and F1 are taken for the fine comparison of experimental results. The accuracy (*A*) is determined as:

$$Accuracy\ (A) = \frac{T_P + T_N}{T_P + F_P + T_N + F_N} \tag{5}$$

where T_P is truly positive, T_N is a true negative, F_P is false positive and F_N is a false negative. T_P and T_N represent the correctly detected soil cracks and non-cracks. F_P and F_N represents wrongly detected soil cracks and non-cracks. Accuracy is defined as the ratio of the correctly detected cracks and non-cracks to the total cracks in the input image. On the other hand, the precision (*P*) is defined as the ratio between the correctly detected cracks and the total number of cracks detected by the model. The recall (*R*) is the ratio of correct crack detections among all positive samples. The *F*1 score is the harmonic mean of precision and recall. If it is high, then the F_P and F_N are relatively low. The precision (*P*), recall (*R*) and *F*1 scores are determined in accordance with Eqs. 6-8, respectively:

$$Precision\ (P) = \frac{T_P}{T_P + F_P} \tag{6}$$

$$Recall\ (R) = \frac{T_P}{T_P + F_N} \tag{7}$$

$$F1 = 2 \times \frac{P \times R}{(P + R)} \tag{8}$$

Table 4 shows the evaluation metrics of all four models, for comparison. The customized CNN Model 2 achieves better performance metrics in comparison with the other three models. The precision, recall, and F1 scores of the Customized CNN Model 1 and 2 are greater than 9%. The Customized CNN Model 2 in particular has a recall value of 97.25% which is exceptionally high. Overall, the VGG-19 Model gives better results in comparison with ResNet-50. Hence out of four analyzed models, the Customized CNN Model 2 is shown to provide the best overall performance, following by the Customized CNN Model 1, then the VGG-19 Model and finally the ResNet-50 approach.

Table 4: Performance Metrics

Name of the Model	Precision (P)	Recall (R)	*F*1 Score
Customized CNN Model 1	0.901	0.934	0.918
Customized CNN Model 2	0.957	0.980	0.964
VGG-19 Model	0.882	0.944	0.909
ResNet-50 Model	0.805	0.894	0.850

The receiver operating characteristic (ROC) is another commonly-used performance metric for the evaluation of detection algorithms. The ROC plot of the Customized CNN Models 1 and 2 are shown in Fig. 14. The area under the ROC curve identifies that the model effectively detects crack pixels and other non-crack pixels. The highest possible area value under the ROC curve is unity. The Customized CNN Model 2 has a value 0.961 whereas the corresponding value for the Customized CNN Model 1 is 0.911. It indicates the second customized model accurately detects the cracks and non-cracks from the input.

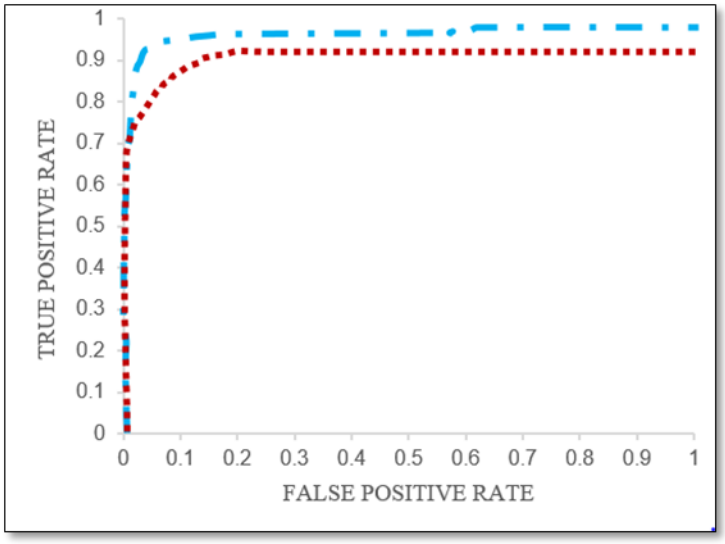


Fig. 14. ROC Curve of the Customized CNN Models

5.3 Comparative Study

The best performing of the four methods described herein, that is the Customized CNN Model 2, is compared with four state-of-the-art methods proposed by other researchers. The traditional Canny edge detector method is used to detect the cracks in the studies conducted by Zhao et al. (2010). The modified median filters are used to remove the noise in the input image. The Support Vector Machine (SVM) based machine learning method is adopted to find the cracks in the investigation proposed by Wang et al. (2016). The feature fusion technique is also used to identify the features correctly. Fan et al. (2020a) used an ensemble network for detecting cracks. The small size of input images is given to the network to identify the cracks. The encoder and decoder architecture of the U-net is used to detect the cracks in the analysis of Fan et al (2020b). The comparison of performance metrics in terms of precision, recall, and F1-score are presented in Table 5.

Table 5: Comparison of the Proposed Detection Method with State-of-the-Art Methods

Reference	Method	Precision	Recall	F1-Score
Zhao et al., (2010)	Canny	0.437	0.730	0.457
Wang et al(2016)	SVM	0.768	0.681	0.681
Fan et al (2020a)	DCNN	0.930	0.926	0.923

Fan et al (2020b)	U-Net	0.945	0.936	0.939
Proposed method	Customized CNN Model 2	0.957	0.972	0.964

The data presented in Table 5 indicates that the proposed Customized CNN Model 2 outperforms the different methods proposed by other researchers for the range of parameters examined. It has been shown that the detection results from earlier traditional models mostly depend on image quality [Zhao et al. (2010)]. Hence, the soil elements are also identified as cracks due to their color being similar to the cracks. More basic computer vision methods are also faced with this problem [Wang et al (2016)]. On the other hand, deep learning models produce better and more accurate results in terms of crack detection compared with traditional computer vision-based methods. The deep learning models proposed by Fan et al. (2020a) and the proposed Customized CNN Model 2 produces greater accuracy of all the methods examined herein in terms of accuracy, precision, recall, and the F1 score.

6. Conclusions

This paper presents a series of new, deep learning-based models for the detection of shrinkage cracks in expansive soils. This is a complex problem, and recent advances in digital methods have enabled these new, sophisticated approaches to be developed. A real-time soil crack dataset is used for illustrating the accuracy of the devised approaches. Four different models are develops including two customized CNN Models and two pre-trained models (i.e. the VGG-19 and ResNet-50 models). The proposed methods are shown to accurately detect shrinkage cracks in expansive soils. The dataset consists of images with various lighting and soil conditions as well as different field environments.

The Customized CNN Model 2 is shown to provide the best results of those presented, although all four methods are better than other approaches in the literature. The recall is a useful performance metric to minimize the number of false negatives in the automatic detection tasks. It is shown in this paper that the proposed model has a very high recall value of 98%. It can be concluded that the increasing number of convolution layers and parameters which are employed in this model result in accurate detection of soil cracks.

Even though the proposed deep learning method yields satisfactory results in shrinkage soil crack detection, there are some limitations which are noteworthy. In particular, complicated

crack geometries and their morphological conditions need to be collected and included in the datasets. In addition, testing images with different photographic angles can be encountered. Nevertheless, the proposed deep-learning method developed herein can be readily adapted and employed for other types of cracks in soils also.

References

Ali, L., Alnajjar, F., Jassmi, H. A., Gochoo, M., Khan, W., & Serhani, M. A. (2021). Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures. *Sensors*, *21*(5), 1688

Andrushia A, D., & Lubloy, E. (2021). Deep Learning-based Thermal Crack Detection on Structural Concrete Exposed to Elevated Temperature: *Advances in Structural Engineering*, Vol. 24, No. 9, pp. 1896–1909, DOI: https://doi.org/10.1177%2F1369433220986637.

Andrushia, Diana, N. Anand, and Prince Arulraj. (2019) Anisotropic diffusion based denoising on concrete images and surface crack segmentation." *International Journal of Structural Integrity*, DOI: 10.1108/IJSI-06-2019-0061

Armstrong, A., Aden, K., Amraoui, N., Jarvis, N., and Mouvet, C. (2000). Comparison of the performance of pesticide-leaching models on a cracking clay soil : results using the Brimstone Farm dataset: Vol. 44.

Auvray, R., Rosin-Paumier, S., Abdallah, A., and Masrouri, F. (2014). Quantification of soft soil cracking during suction cycles by image processing: *European Journal of Environmental and Civil Engineering*, Vol. 18, No. 1, pp. 11–32, DOI: 10.1080/19648189.2013.840250.

Bureau of Indian Standards (1985). Indian Standard, Methods of test for soils, Part 5: Determination of liquid limit and plastic limit: *Bureau of Indian Standards, New Delhi, India.*, Vol. Reaffirmed, No. 2006, pp. 1–16.

Cha, Y.J., Choi, W., and Büyüköztürk, O. (2017). Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks: *Computer-Aided Civil and Infrastructure Engineering*, Vol. 32, No. 5, pp. 361–378, DOI: 10.1111/mice.12263.

Choudhary, G.K., and Dey, S. (2012). Crack detection in concrete surfaces using image processing, fuzzy logic, and neural networks: *2012 IEEE 5th International Conference on Advanced Computational Intelligence, ICACI 2012*, pp. 404–411, DOI:

10.1109/ICACI.2012.6463195.

Dai, B., Gu, C., Zhao, E., Zhu, K., Cao, W., and Qin, X. (2019). Improved online sequential extreme learning machine for identifying crack behavior in the concrete dam: *Advances in Structural Engineering*, Vol. 22, No. 2, pp. 402–412, DOI: 10.1177/1369433218788635.

Diana Andrushia, A., Anand, N., and Prince Arulraj, G. (2020). A novel approach for thermal crack detection and quantification in structural concrete using ripplet transform: *Structural Control and Health Monitoring*, Vol. 27, No. 11, pp. 1–20, DOI: 10.1002/stc.2621.

Drouyer, S. (2020). An All Terrain Crack Detector Obtained by Deep Learning on Available Databases. *Image Processing On Line*, *10*, 105-123.

Fan, Z., Li, C., Chen, Y., Di Mascio, P., Chen, X., Zhu, G., and Loprencipe, G. (2020a). Ensemble of deep convolutional neural networks for automatic pavement crack detection and measurement: *Coatings*, Vol. 10, No. 2, pp. 1–15, DOI: 10.3390/coatings10020152.

Fan, Z., Li, C., Chen, Y., Wei, J., Loprencipe, G., Chen, X., and Di Mascio, P. (2020b). Automatic crack detection on road pavements using encoder-decoder architecture: *Materials*, Vol. 13, No. 13, pp. 1–18, DOI: 10.3390/ma13132960.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2016-Decem, pp. 770–778, DOI: 10.1109/CVPR.2016.90.

IS:2720 (Part 4) (1985). Indian Standard, Methods of Test for Soils, Part 4: Grain Size Analysis: *Bureau of Indian Standards, New Delhi, India.*, Vol. Reaffirmed, No. 2006, pp. 1–38.

IS 2720-Part(VI) (1972). Determination of shrinkage factors: *Indian Standard Code*.

Kumar, P., Gouri, S., Bhunia, S., and Maiti, R. (2015). Soil crack morphology analysis using image processing techniques: *Modeling Earth Systems and Environment*, Vol. 1, No. 4, pp. 1–7, DOI: 10.1007/s40808-015-0036-z.

L.R. Tan and L.W. Kong (2006). *Special Geotechnical Engineer Soil Science*, Science Press, Beijing, China.

Li, G., Zhao, X., Du, K., Ru, F., and Zhang, Y. (2017). Recognition and evaluation of bridge cracks with modified active contour model and greedy search-based support vector machine: *Automation in Construction*, Vol. 78, pp. 51–61, DOI:

10.1016/j.autcon.2017.01.019.

Liu, Z., Cao, Y., Wang, Y., and Wang, W. (2019). Computer vision-based concrete crack detection using U-net fully convolutional networks: *Automation in Construction*, Vol. 104, No. January, pp. 129–139, DOI: 10.1016/j.autcon.2019.04.005.

Lu, N. and Likos, W.J. (2004). *Unsaturated Soil Mechanics*, 1st edition Ed., John Wiley & Sons, Hoboken.

Luo, X., Wu, H., Yuan, H., and Zhou, M.C. (2020). Temporal Pattern-Aware QoS Prediction via Biased Non-Negative Latent Factorization of Tensors: *IEEE Transactions on Cybernetics*, Vol. 50, No. 5, pp. 1798–1809, DOI: 10.1109/TCYB.2019.2903736.

Shi, B.X., Zhang, C.F., and Wu, J.K. (2014). Research progress on expansive soil cracks under changing environment: *Scientific World Journal*, Vol. 2014, DOI: 10.1155/2014/816759.

Simonyan, K., and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition: *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–13.

Thomas H. Brown *Highway Engineering* (T. H. B. Daniel J. Findley, Bastian J. Schroeder, Christopher M. Cunningham, Ed.).

Wan, Y., Xue, Q., Liu, L., and Wang, S.Y. (2018). Relationship between the shrinkage crack characteristics and the water content gradient of compacted clay liner in a landfill final cover: *Soils and Foundations*, Vol. 58, No. 6, pp. 1435–1445, DOI: 10.1016/j.sandf.2018.08.011.

Wang, C., Wang, X., Zhou, X., and Li, Z. (2016). The Aircraft Skin Crack Inspection Based on Different-Source Sensors and Support Vector Machines: *Journal of Nondestructive Evaluation*, Vol. 35, No. 3, DOI: 10.1007/s10921-016-0359-3.

Xu, J. J., Zhang, H., Tang, C. S., Cheng, Q., Liu, B., & Shi, B. (2021). Automatic soil desiccation cracks recognition using deep learning: *Geotechnique*, Vol. 54, No. 1, pp. 307–314, DOI: https://doi.org/10.1680/jgeot.20.P.091.

Xuhe Gao, Baocheng Cheng, Weiping Tian, Zhipei Zhang, Jiachun Li, Hongliang Qi.(2020 a) Simulation Parameter Selection and Steady Seepage Analysis of Binary Structure Slope [J]. Water 12(10), 2747. doi: 10.3390/ w12102747

Xuhe Gao, Weiping Tian, Zhipei Zhang, Jiachun Li, Hongliang Qi. (2020 b) Simulation

Parameter Test and Seepage Effect Analysis of Pile-Anchor Support for Binary Slope [J]. Advances in Civil Engineering, 8862163. doi:10.1155/ 2020/8862163

Y.K. Gong, L. Chen, and G.F.W. (2009). Electric conductance characteristics of fissures of expansive soil: *Journal of Hohai University*, Vol. 37, No. 3, pp. 323–326.

Yan, J., Downey, A., Cancelli, A., Laflamme, S., Chen, A., Li, J., and Ubertini, F. (2019). Concrete crack detection and monitoring using a capacitive dense sensor array: *Sensors (Switzerland)*, Vol. 19, No. 8, pp. 1–12, DOI: 10.3390/s19081843.

Zhao, H., Qin, G., and Wang, X. (2010). Improvement of canny algorithm based on pavement edge detection: *Proceedings - 2010 3rd International Congress on Image and Signal Processing, CISP 2010*, Vol. 2, pp. 964–967, DOI: 10.1109/CISP.2010.5646923.