

An Entropy-based Automated Method for Detection and Assessment of Spalling in Reinforced Concrete Bridges

Eslam Mohammed Abdelkader^{*1,2}, Osama Moselhi³, Mohamed Marzouk⁴, and Tarek Zayed⁵

ABSTRACT

Existing bridges are aging and deteriorating rapidly; elevating concerns for public safety and preservation of these valuable assets. Large numbers of bridges exist in transportation networks and the current budget limitations worsen the situation. This necessitates the development of automated condition assessment and rating methods. Spalling is a common problem that majorly influences the health, safety and structural integrity of bridges. The present study introduces a self-adaptive three-tier method for the automated detection and assessment of spalling using computer vision technologies. The first model introduces a newly-developed segmentation model that adopts multi-objective invasive weed optimization and information theory-based formalism of images for spalled concrete detection. In the second model, an integration of singular value decomposition and discrete wavelet transform are integrated for the efficient feature extraction of information in images. Additionally, Elman neural network is coupled with invasive weed optimization algorithm to enhance the accuracy of evaluation of spalling severities by amplifying exploration-exploitation trade-off mechanism of Elman neural network. The third model is developed for the purpose of structuring a rating system of spalling severity based on its area and depth. A computerized platform is developed using C#.net language to facilitate the implementation of the developed method by the users. Results demonstrated that the developed multi-objective spalling segmentation model is capable of improving detection accuracy of spalling by 12.29% with respect to the Region growing algorithm. It was inferred also the developed quantification model outperformed other prediction models such that it achieved mean absolute percentage error, root mean-squared error and root mean squared percentage error of 4.07%, 76.061 and 0.065, respectively based on the original dataset. In this regard, it is expected that the developed computer vision-based method can aid in establishing cost-effective bridge condition assessment models by transportation agencies.

Keywords: Bridges; Spalling; Computer vision; multi-objective invasive weed optimization; information theory; Elman neural network; C#.net

¹ Ph.D. candidate, Department of Building, Civil, and Environmental Engineering, Concordia University, Montreal, QC, Canada. Corresponding author, E-mail: eslam_ahmed1990@hotmail.com.

² Assistant lecturer, Structural Engineering Department, Faculty of Engineering, Cairo University, Egypt.

³ Professor and Director of the Centre for Innovation in Construction and Infrastructure Engineering and Management (CICIEM), Department of Building, Civil, and Environmental Engineering, Concordia University, Montreal, QC, Canada.

⁴ Professor of Construction Engineering and Management, Structural Engineering Department, Faculty of Engineering, Cairo University, Egypt.

⁵ Professor, Department of Building and Real Estate, the Hong Kong Polytechnic University, Hung Hom, Hong Kong.

1. INTRODUCTION

Bridges are vital links in transportation networks. However, large portion of bridges are deteriorating rapidly as a result of excessive loads, cycles of freeze and thaw, deferred maintenance and at times poor construction practices. These deterioration agents induce different modes of failure such as corrosion, delamination, cracking and spalling. The Canadian infrastructure report card indicated that 26% of the bridges are either “Fair”, “Poor” or “Very Poor” (Felio et al., 2016). Moreover, One-third of Canada’s bridges were reported to have structural or functional deficiencies with short remaining service life (National Research Council Canada, 2013). Furthermore, the backlog of bridge maintenance, rehabilitation and replacement is estimated to be equal to \$10 billion. The constant increase in the backlog results in a significant deterioration in the condition of the bridge elements (Sennah et al., 2011). In view of the above, it is very important to develop methods that support condition driven intervention actions in order to preserve the structural integrity and functionality of the assets within allocated budget constraints.

Recently, the use of computer vision technologies gained attention in evaluating the severities of the surface defects in reinforced concrete structures. These technologies can help transportation agencies to overcome the drawbacks of the visual inspection-based methods. Currently, inspections are conducted at equal time intervals by engineers to evaluate the severities of the bridge defects. However, these inspections are subjective, labour-intensive, time-consuming and hazardous in some cases. Spalling is a surface defect, which is induced in the form of a fragment of concrete detached from a larger concrete. It can cause serious structural damage and sometimes it can significantly contribute to the structural collapse of the concrete structure. It is reported that surface defects play an important role in visual inspection

manuals because they are able to establish an accurate reflection of condition ratings of structural members (Koch et al., 2015). As such, the present study introduces a novel self-adaptive optimization-based method for the automated detection and evaluation of spalling in reinforced concrete bridges.

2. LITERATURE REVIEW

Several studies were conducted to evaluate the severities of surface defects. Adhikari et al. (2014) developed an artificial neural network-based model to predict the depth of the crack given a certain crack width based on an input dataset of 101 images. They developed a method the cracking depth computation, which helps in providing a more accurate condition rating of concrete elements. They presented an approach based on spectral analysis to detect the change in crack patterns over time by converting digital images to the frequency domain using Fast Fourier Transform (FFT). Lee et al. (2017) developed a bridge inspection system using unmanned aerial vehicle (UAV). Otsu method was applied to segment the images to objects of interest and background. Otsu algorithm was proposed by Otsu in 1979 and it is a widely-used un-supervised approach that is utilized to segment an image by maximizing the variances between the segmented classes. Then, the crack properties in the HSV space were used to distinguish between cracks and other surface irregularities. They highlighted that their model was capable of detecting cracks measured in micrometers. HSV is a color space that is identified using hue, separation and value of the color. Hue stands for pure color resemblance. Saturation stands for how white a color is and the value of the color indicates its lightness.

Su (2013) proposed a computer vision method based on Charge Coupled Devices (CCD) cameras to automatically detect cracks during the bridge inspection. Weighted mean filter was utilized to remove environmental noises from images. They highlighted that 5×5 weighted

median filter provides better accuracy when compared to 3×3 weighted median filter. Otsu thresholding is adopted for the binary transformation of the image. They stated that their model was capable of achieving accuracies of 90% and 84% for the training dataset and testing dataset, respectively. Xue-jun and Xiao-ning (2013) designed a computer program using C++ programming language to detect and quantify cracks in reinforced concrete bridges. They compared the median filter, mean filter and combination of both filters by adding noise to the original image and they concluded that the median filter provided better noise removal efficiency as per the mean-squared error. Sobel algorithm was adopted to automatically identify the edges of the crack. Sobel operator algorithm is implemented by computing the gradient of image intensity at each pixel within the image. It determines the direction of largest increase from light to dark and the rate of increase in this direction. The developed method was capable of achieving a relative percentage error of 6% for cracks of width less than 0.3 mm, and error more than 30% for cracks of width less than 0.1 mm.

Zhang et al. (2018) proposed an improved watershed algorithm to enhance the detection process of bridge cracks. They utilized a combination of H-minima method, morphological forced minimum operation and watershed algorithm to enhance the segmentation capacity and avoid the over segmentation. H-minima is used to capture the local minima values of the target region which are highly correlated with bridge cracks from low frequency components in the image. They demonstrated that the developed method introduced higher accuracy and robustness when compared to the conventional watershed segmentation method. Xu et al. (2019) introduced a method to automatically detect cracks in images using convolutional neural network. Atrous Spatial Pyramid Pooling was applied to detect the input feature map with multiple sampling rates. In it, Atrous convolutions of size 3×3 and with rates of 2, 4 and 8 are applied in parallel

whereas the extracted feature maps are further processed by the depthwise separable convolution in a separate branch. Atrous convolution was utilized to obtain larger receptive field and more context information without reducing the resolution and increasing number of parameters. The multi-scale feature maps of Atrous convolutions are integrate with the globally pooled input feature map to generate the final feature map. The depthwise separable convolution in the Atrous spatial pyramid pooling module was used to minimize the computational complexity and enhance the computational efficiency. They highlighted that the proposed method achieved an accuracy of 96.73% without pre-training.

Kim et al. (2018) proposed a crack identification method using UAV equipped with a high-resolution vision sensor. A 3D point-cloud based background model was generated for the visualization of cracks on the inspection map. Then, region-based convolutional neural networks (R-CNN) were applied for the detection of the length and width of the cracks. Tao et al. (2014) presented a method to detect bridge cracks in underwater conditions. Spatial median filtering is applied to enhance the smoothing of the images. Then, the images are segmented based on an iterative method that requires the computation of maximum and minimum gray-level intensities. They illustrated that the developed method achieved promising accuracies for the computation of length and width of cracks.

Lee et al. (2008) designed a machine vision robotic system to automate the inspection process of bridges. The developed system enables the user to evaluate the cracks in real time, whereas it was evaluated using 100 noisy images. Median smoothing filter was applied to remove noises and to ensure uniform brightness through the image. Then, dilation and thinning morphological operations were utilized to maintain the connections between the crack segments. They demonstrated that the developed method yielded higher detection accuracies when

compared against Sobel, Canny and Fujita methods. Yang et al. (2019) proposed a transfer learning-based model for crack detection in infrastructures. Visual Geometry Group of 16 layers (VGG16) was utilized for the sake of improving detection performance and reducing the training time. The training and validation of the model was carried out using a set of well-known crack datasets, namely SDNET dataset, CCIC dataset and BCD dataset. It was found that the developed model could achieve an improvement in the crack detection accuracies by 2.33% and 5.06% for the SDNET and BCD dataset, respectively when compared against other deep convolutional neural networks.

Dorafshan et al. (2018) studied the implementation of two modes of deep convolutional neural network in concrete crack detection. In the first mode, the AlexNet architecture was fully-trained from scratch capitalizing on the dataset captured using small unmanned aerial systems. In the second mode, a transfer learning-based network of same topology was pre-trained using ImageNet dataset. The performances of the deep neural networks were assessed using three datasets. It was reported that the transfer learning-based network had higher training accuracy than the fully trained network. Furthermore, it accomplished higher validation accuracies for the three datasets by values ranged from 5.3% to 10%.

Based on the aforementioned studies, current practices of visual inspection-based methods are characterized by their inherent subjective nature, which lead to inefficient condition assessment models. In addition to that, most of the previous studies focused on the investigation of severities of cracking meanwhile there is lack of exploration of other surface defects such as spalling and scaling, which eventually leads to incomprehensive and unreliable condition assessment models. Furthermore, classical segmentation models including Otsu and watershed methods diverge and provide poor segmentation results in the case of complex and low contrast

images alongside irregular lighting conditions. It can be also observed that there is absence of a rating scale which enables decision makers to interpret the level of severity of bridge defects based on some descriptors such as length, width, area and depth.

It is also noticed that some previous models capitalize on feeding the whole input image directly to serve as an input to train the machine learning model for bridge defects detection. This implicates high computational cost designated to explore the multi-parameter search space. In this context, the absence of efficient feature extraction algorithm undermines the learning capacity of the prediction model resulting from its failure to map the important and significant features present in the image. Additionally, some of the prediction models reported in the literature either utilize support vector machines and different types of artificial neural network in order to automatically quantify the surface defects severities. These models suffer from local minima entrapment and premature convergence in the case of large search minima. Moreover, the existence of wide range of parameters which are usually manually tuned, can lead in some cases to inferior prediction accuracy.

In the recent years, deep learning has been widely adopted for the purpose of defects detection in infrastructures. Nevertheless, they suffer from the followings shortcomings. Deep learning requires a huge training dataset to map the important features for the sake of building the relationships between the set of independent variables and the dependent variables, which sometimes can be difficult and tedious to create. Another disadvantage of deep learning is that it is a high resource demanding paradigm that necessitates the presence of highly computational infrastructure demonstrated in the form of high graphics processing units in addition to the presence of efficient data storage. The high computational cost is elicited from the need of large amount of data, time and extensive training cycles to learn the deep learning model. It is worth

mentioning that the deep neural networks sometime induce detection latency as a result of the high processing demands. Deep learning models may cause over-fitting triggered by the presence of small dataset, which may drastically influence its recognition capacity (Demertzis and Iliadis, 2020; Al Najada et al., 2018; Dundar et al., 2015).

Another shortcoming of the deep learning models is the presences of wide range of hyper parameters that substantially influence the performance of the deep learning models. This encompasses the number of filters, stride size, padding size and kernel sizes to create the feature map, number of convolutional layers, type of transfer function, type of pooling operation, number of fully connected layers, number of neurons, and weights of the connections between neurons. There are infinite numbers of possible combinations that need exhaustive search to be carried out, which causes the manual tuning of the hyper parameters to be a very challenging task. In this context, the absence of automated systematic method to define the optimum parameters of the convolutional neural network can lead to its entrapment in local minima which yields inferior solutions triggered by the long computational time and slow convergence (Alhamdoosh and Wang, 2014; AL-Allaf, 2011). Transfer learning has recently attracted significant research attention to be utilized in the detection of surface defects. However, transfer learning-based deep convolutional neural networks are vulnerable to negative transfer, which allude to the situations where the transfer of information from the source domain has a detrimental implication on the prediction of the target domain. The absence of sufficient degree of similarity between the features of the source domain and target domain undermines the learning performance of the target errand (Dhillon and Verma, 2019; Stamate et al., 2015). As such the present study proposes a novel self-adaptive optimization-based method for the

automated detection and assessment of spalled concrete in bridges in an attempt to circumvent the limitations referred to above.

3. PROPOSED METHOD

Reinforced concrete bridges are subjected to several types of surface defects. As such, bridge defects detection and recognition models need to be developed for the sake of identifying the existence of bridge defects and classifying the type of bridge defects whether cracking, spalling and scaling. More details about the automated bridge defects detection and classification models can be found elsewhere (Mohammed Abdelkader et al., 2020). In this regard, image-based model is essential to evaluate the severities of each surface defect based on a set of descriptors. The primary objective of the present research study is to develop a computerized platform for the automated detection and assessment of spalling in reinforced concrete bridges. The framework of the proposed method is depicted in Figure 1. The proposed method consists of three main models namely: spalling detection, spalling assessment and spalling severity scale. Each will be described and evaluated separately. The first model is a multi-objective invasive weed optimization-based method for the detection of spalling in reinforced concrete images. The second model utilizes a hybridization of singular value decomposition, discrete wavelet transform, Elman recurrent neural network and invasive weed optimization algorithm for the automated evaluation of severities of spalling. It is self-adaptive, where the architecture of the neural network and its parameters are automatically tuned for its optimum performance. The third model aims at constructing a severity scale of spalling based on its area and depth.

In the first model, close-range photographs are captured with proper focus on the object of interest or the defect. The natural targets are used for calibration of dimensions such as sharp corners of walls, beams and piers. In case of their inconvenience, colored cards and painted rods

are used as artificial targets for calibration purposes (Jáuregui et al., 2006). The captured true-color image RGB is first converted to grayscale image, whereas the intensity values of the grayscale image vary from 0 to 255. For the RGB image, R stands for red, G stands for green, and B stands for blue. The gray-scale images can speed up the image processing while preserving important features of the image. Then, the converted images are standardized to a size of 200×200 pixels to ensure same size images in the training and testing processes of the neural network. The digital images are prone to degradation by noise during the image acquisition phase or during unfavorable conditions during image transmission. These conditions can substantially affect the subsequent image processing phases such as image segmentation and feature extraction. Therefore, Wiener filter is then utilized as a frequency domain filter which adopts discrete Fourier transform (DFT) to remove the noise present in the image while preserving its significant features (Hasan and El-Sakka, 2015).

Processing images of reinforced concrete bridges is complex due to the enormous amount of information present in them. Moreover, they experience low contrast and non-uniform illumination conditions. As such, min-max gray level discrimination method is applied for contrast enhancement, and amplification of differences in gray level intensities of the spalling and non-spalling regions. This method increases the gray level intensities of the spalling pixels so that they become darker while it reduces the gray level intensities of the non-spalling pixels so they become lighter (Hoang et al., 2018). The enhanced image can be computed as follows.

$$I_a(x,y) = \begin{cases} \min(M, T) & \text{if } I_o(x,y) > I_o \min + \tau \times (S) \\ \max(N, F) & \text{if } I_o(x,y) \leq I_o \min + \tau \times (S) \end{cases} \quad (1)$$

Such that;

$$T = I_o(x,y) \times R_a, F = I_o(x,y) \times R_a^{-1}, S = I_o \max - I_o \min, M = I_o \max, \text{ and } N = I_o \min$$

Where;

$I_a(x, y)$ and $I_o(x, y)$ represent the adjusted image and original image, respectively. $I_o\max$ and $I_o\min$ represent the maximum and minimum gray level intensities in the original image. R_a and τ refer to the adjusted ratio and margin parameter, respectively. R_a and τ are assumed 1.1 and 0.5, respectively. It is worth mentioning that the grey level discrimination function is applied to each image separately.

Image segmentation is the process of partitioning the digital image into multiple segments based on some attributes such as colour, intensity and texture. The present study adopts bi-level thresholding (binarization) in order to divide the image into two segments which are: foreground (spalling) and background. Threshold values differ from one image to the other depending on its gray level intensities. In this regard, it is a challenging task to find the optimum thresholds in image segmentation problems. In the present study, the image segmentation process is handled as an optimization problem, whereas the optimum threshold is computed based on a bi-objective optimization problem. The two objective functions are the minimization of Kapur entropy and minimization of Renyi's entropy. These two objective functions are deemed a more adequate search functions because there is no single objective function can optimally handle different types of optimization problems. Moreover, these two functions proved their superior capabilities in solving complex and exhaustive optimization problems such as spalled concrete images. The proposed method utilizes invasive weed optimization algorithm to search for the optimum threshold in images. Invasive weed optimization algorithm is exhaustive search engine that demonstrated its capabilities in exploring complex and multi-local search spaces. Moreover, it manifested its superiority over some of the best-performing optimization algorithms such as genetic algorithm, particle swarm optimization algorithm, artificial immune system, artificial bee

colony and harmony search algorithm (Mohammed Abdelkader et al., 2019; Uyar and Ülker, 2017; Razmjoooy and Khalilpour, 2015).

Some imperfections may be present in the image after the segmentation process. These imperfections include some noise, holes, and non-uniform edges. These imperfections should be removed before the spalling quantification process using some morphological operations meanwhile preserving the shape of spalling. The noise and protrusions in images are reduced by removing the isolated unconnected pixels of area less than 50. Additionally, it utilizes operations like filling holes, closing operation with disk structuring element of size 6 and bridging unconnected pixels for the purpose of filling cavities within boundaries of objects and smoothing their borders. In this regard, the present study utilizes the standard 3×3 neighborhood in the bridging and filling holes operations. The bridging operator sets zero-value pixels (background pixels) to one (foreground pixel) if they have two non-zero neighbors that are not connected. The filling operation is used to fill isolated interior pixels in the spalling, which are the individual zero pixels that are surrounded by one-value pixels. It is worth mentioning that closing operation is dilation followed by erosion using a predefined structuring element. Suppose an image A of size $M \times N$ and structuring element SE. The closing operation can be described as follows (Pal and Chatterjee, 2017; Lv et al., 2014).

$$A \oplus SE = (A \oplus SE) \odot SE \quad (2)$$

Such that,

$$A \oplus SE = \max \{A(i, j) + SE(i, j)\} \quad (3)$$

$$A \odot SE = \min \{A(i, j) - SE(i, j)\} \quad (4)$$

Where;

The operators \ominus , \oplus and \odot refer to closing, dilation and erosion operations, respectively. The size and shape of structuring element are essential for performing the morphological operations. For example, the size of the structuring element has to be considerably smaller than the image to be processed and at the same time it has to be of a closer size to the object of interest in the image. The structuring element is expressed in the form of a binary image, which takes the value of zero and one. The structuring element can take the form of different shapes such as line, disk, square, diamond. In the present study, the size and type of structuring element alongside the minimum area of objects to be removed; are manually tuned for some images based on their implication on the quality of the segmented image. This setting of parameters is then appended to be applied to other images.

The proposed spalling detection method is validated through comparison against Otsu method, K-means clustering, region growing, fuzzy C-means clustering, expectation maximization such that the comparison is conducted as per a set of image quality indicators and classification evaluation metrics. The image quality metrics encompass mean-squared error, peak signal to noise ratio and mean absolute error. The classification evaluation metrics include overall accuracy, F-measure, balanced accuracy and Matthews's correlation coefficient. Finally, the statistical significance of the performances of the segmentation methods is investigated using two-tailed paired student's t test. Student's paired t-test is a parametric statistical test utilized for the purpose of analyzing the statistical significance of the difference between two population means in a research study encompassing paired samples. In the present study, the student's t test is used to investigate whether or not the image quality indicators and classification evaluation metrics of two spalling detection models are different from each other based on a set of images.

Student's t test is selected because it is a commonly utilized approach that proved its efficiency in analyzing statistical data and examining dissimilarities between different clusters in diverse applications including such as comparing core strength results (Reddy and Wanjari, 2018), verifying the effect of software on youth's learning process (Di Biasi et al., 2020) and analyzing international roughness index from multiple sources (Samsuri et al., 2019).

The second model is the automated assessment of spalling area. Feature extraction is a fundamental pre-processing stage in machine learning and pattern recognition problems because it enables to extract the features required as an input for the regression model. Two types of feature descriptors are utilized for the automated spalling assessment including: spatial and frequency domain descriptors. The spatial features encompass: color, texture and shape while the frequency features include the coefficients of discrete wavelet transform. The fusion of both types of features provides an informative and comprehensive description of the information present in images, which can eventually enhance the recognition capacity of the proposed assessment model.

The proposed method adopts the energies of all discrete wavelet transform (DWT) sub-bands are combined with non-negative singular values to create the feature vector set. The DWT and singular value decomposition (SVD) are utilized to extract the frequency domain and spatial domain features, respectively which enable to speed up the computational process by eliminating the insignificant features and reducing the computational complexity of the input dataset of gray-level images. In the present study, singular value decomposition is utilized to convert the input image of size 200×200 pixels into a spatial domain feature vector composed of singular values of size 1×200 . In it, the singular values are ranked in a descending order, whereas the first entries of the singular value matrix contain the most substantial information while the last entries at the

vector contain the least significant information (Nagarajan and Devendran, 2012; Jha et al., 2014). With respect to the frequency domain features, DWT-based feature extraction exhibits three levels of decomposition, whereas N levels of decomposition result in $(3 \times N) + 1$ sub-bands. Thus, the frequency domain feature vector is composed of ten sub-bands. The increase in the levels of DWT increases the length of extracted features, which creates more computational burden. As such, the designated number of decompositions provides a trade-off between the computational complexity and computational accuracy. In this regard, the frequency domain feature vector is of size 1×10 encompassing the energies of ten sub-bands. As such, the spatial domain feature vector is cascaded with a frequency domain feature vector to establish a feature vector set of size 1×210 to serve as an input to build the automated spalling assessment model. The energy of each sub-band can be computed using the following equation (Shanavaz and Mythili, 2016).

$$E_k = \frac{1}{M \times N} \sum_{w=1}^M \sum_{h=1}^N |X_k(x, y)| \quad (5)$$

Where;

E_k represents the energy of k – th sub-band. $X_k(x, y)$ indicates the pixel value of k – th sub-band. M and N represent the width and height of sub-band, respectively.

After designing the feature vector set, it is used to feed the automated assessment model. The proposed prediction model involves hybrid parameter-structural learning, whereas it utilizes invasive weed optimization to enhance the training process of the Elman neural network (ENN) by addressing the exploration-exploitation trade-off dilemma. The ENN is trained by designing a variable-length single-objective optimization problem which encompasses a fitness function for

minimizing the mean absolute error. The steps of the invasive weed optimization algorithm are repeated until reaching the desired number of iterations. The optimum transfer functions, number of hidden and context layers, number of hidden and context neurons, and weights and bias terms establish the optimized Elman neural network. These optimum parameters are appended for later use in performance evaluation.

The performance of the proposed automated assessment model is compared against a set of widely-used machine learning models and high-performing deep learning models reported for their higher accuracies. The machine learning models comprises back-propagation artificial neural network, Elman neural network, generalized regression neural network and radial basis neural network. The deep learning models encompass a deep convolutional neural network trained from scratch (CONVNET) alongside a group of pre-trained deep neural network architectures, namely AlexNet, VGG16, VGG19 and CaffeNet. The performances of the prediction models are assessed based on mean absolute error percentage error (MAPE), root mean-squared error (RMSE) and root mean squared percentage error (RMSPE). The comparative analysis is carried out using the original and augmented datasets for the purpose of testing the robustness of the proposed method in dealing with different sizes of datasets. The original dataset is augmented for the sake of its enlargement by creating new training and testing instances (Dung et al., 2019). The data augmentation techniques can be applied alone or combined in computer vision. In the present study, different forms of data augmentation techniques are utilized including rotation, flipping and cropping. In this context, rotation and flipping are introduced to establish rotational invariant models, whereas the images are rotated by 90°, 180° and 270°.

The third model is designated for establishing a severity rating systems of spalling based on its area and depth. The spalling area is interpreted from the previous model while the scaling depth is adopted from the third-order polynomial regression function introduced by Dawood et al. (2017). The severity levels of spalling area are expressed in the form of percentage to reduce the effect of camera angle on crack detection. In this regard, spalling percentage is equal to number of pixels occupied by spalling divided by the total number of pixels. In order to establish efficient and robust rating system, sufficient amount of records should be present. As such, the spalling area and depth are assumed as random variables that follow certain probability distributions based on the available dataset. The best-fit distribution is identified based on Anderson Darling statistic. Then, Latin hypercube sampling is employed to generate large number of scenarios to be used to structure the spalling severity rating system. Latin hypercube is stratified sampling scheme that enables better coverage and exploration of the domain of the variations of the input variables. It is preferred over conventional Monte Carlo sampling because of its time-efficiency and faster convergence resulting from its capability to construct efficient probability distributions using less number of iterations and less sampling error (García-Alfonso and Córdova-Esparza, 2018; Gupta and Kumar, 2016). Consequently, fuzzy C-means clustering is utilized to generate the thresholds of the severity levels of spalling area and depth. Fuzzy C-means is one of the soft clustering algorithms in which the data point is assigned to more than one cluster with different degrees of membership. In this context, Fuzzy C-means clustering is selected over other clustering algorithms due to its capability in dealing with vagueness and uncertainties encountered during the bridge inspection process. Furthermore, it outperformed K-means hard clustering in terms of establishing more compact homogenous clusters as well as well-separated thresholds (Bhattacharjee et al., 2017; Hooda et al., 2014). It is worth mentioning

that the severity of spalling in each image is evaluated based on its area and depth, and the worst case scenario is selected to establish a more conservative assessment of spalling severities. Eventually, a computerized platform is developed using C#.net and Matlab programming languages to facilitate the decision-making process by transportation agencies.

INSERT FIGURE 1

4. AUTOMATED DETECTION MODEL

4.1 Kapur Entropy

The automated detection model is a Kapur entropy-based segmentation approach that utilizes invasive weed optimization algorithm to find the optimum thresholds. As mentioned earlier each image has a different threshold depending on its intensity values. In this regard, the values of optimum thresholds range from 0 to 255. Kapur entropy approach is based on selecting the optimum threshold by maximizing the entropy of the segmented classes. Kapur entropy deals with images as a gray-level histogram and the objective is to divide the image into two classes C1 and C2. Thus, two probability distributions are developed: one for the object and one for the background, whereas the purpose of Kapur method is to find the optimum threshold by maximizing the entropy of the partitioned segments. Assume an image I that contains L gray-levels $\{0, 1, 2, 3, \dots, L - 1\}$. The probability of occurrence of each gray level i is calculated using Equation (3). The probability of occurrence of a certain gray level is equal to occurrence frequency of gray level i divided by the total frequencies of all gray levels (Khairuzzaman and Chaudhury, 2017; Bhandari et al., 2015).

$$P_i = \frac{H_i}{\sum_{i=0}^{L-1} H_i} \quad (6)$$

Where;

P_i denotes the probability of occurrence of gray-level i . H_i represents the number of occurrences of gray-level i .

Thus, the objective is to find the optimum threshold T by maximizing the entropy of the segmented classes as per Equation (7). Then, the entropy of the segmented classes can be computed using Equations (8) and (9).

$$F(T) = \max (H_1 + H_2) \quad (7)$$

$$H_1 = - \sum_{i=0}^{T-1} \frac{P_i}{W_1} \times \ln \left(\frac{P_i}{W_1} \right), \quad W_1 = \sum_{i=0}^{T-1} P_i \quad (8)$$

$$H_2 = - \sum_{i=T}^{L-1} \frac{P_i}{W_2} \times \ln \left(\frac{P_i}{W_2} \right), \quad W_2 = \sum_{i=T}^{L-1} P_i \quad (9)$$

Where;

$F(T)$ represents the summation of entropy of all the classes. H_1 and H_2 represent the entropies of $C1$ and $C2$, respectively. W_1 and W_2 represent the class probabilities for $C1$ and $C2$, respectively. More details about the Renyi's entropy method can be found in Mishra and Panda (2018), and Li et al. (2006). The computation steps of the Kapur entropy and Renyi's entropy functions are coded in Matlab environment.

4.2 Invasive Weed Optimization

Invasive weed optimization (IWO) is a meta-heuristic bio-inspired optimization algorithm that was introduced by Mehrabian and Lucas in 2006. IWO algorithm is based on simulating the invasive and robust behaviour of weed in colonizing and occupying territories in an attempt to find the most suitable place for growth and reproduction. In the present study, invasive weed

optimization algorithm is employed for the purpose of spalling segmentation through determining the optimum thresholds that maximizes Kapur entropy function. The computational procedures of the invasive weed optimization algorithm are divided into four stages (Mehrabian and Lucas, 2006).

The first stage is to create an initial population of weeds that are spread in the i -dimensional search space. The fitness of each weed within the population is calculated based on a predefined objective function. The second stage is the production of seeds such that the production process associated with each weed is calculated based on a linear function, where the number of seeds varies between the minimum and maximum number of seeds. Each weed in the population produces seeds based on its own comparative fitness value, maximum and minimum fitness values within the population, and the maximum and minimum number of seeds. The reproduction of seeds is shown in Equation (10) where the higher the fitness of the weed, the more seeds it produces.

$$\text{Seed}_i = \frac{f_i - f_{\min}}{f_{\max} - f_{\min}} \times (s_{\max} - s_{\min}) + s_{\min} \quad (10)$$

Where Seed_i represents number of seeds associated with the i – th weed. f_i represents the current fitness of the weed. f_{\max} , and f_{\min} represent the maximum and minimum fitness of the current population, respectively. s_{\max} , and s_{\min} denote the maximum and minimum number of seeds, respectively.

The next stage is the spatial dispersion, where the seeds are randomly scattered in the search space based on a normal distribution of a mean equal to zero and an adaptive varying standard deviation. This stage ensures that the seeds are accumulated around the weed plant, which leads to a local search around each parent weed. The standard deviation of the seed dispersion is

reduced from an initial predetermined maximum value to a final predetermined smaller value based on a non-linear function as shown in Equation (11). The probability of finding a seed far from the weed plant is high at the beginning of the optimization process and it decreases within a predefined number of iterations.

$$\sigma_i = \sigma_{\min} + \left(\frac{\text{iter}_{\max} - \text{iter}}{\text{iter}_{\max} - \text{iter}_{\min}} \right)^p \times (\sigma_{\max} - \sigma_{\min}) \quad (11)$$

Where;

σ_i indicates the standard deviation of the current iteration. σ_{\max} , and σ_{\min} indicate the initial and final standard deviation of the optimization process, respectively. iter_{\max} represents the maximum number of iterations. p represents non-linear modulation index, and usually, it is a number between two and three.

The final stage is the competitive exclusion which is performed because the number of weeds and seeds reaches the maximum population size due to the fast reproduction (exponential increase in the number of plants). The parent weeds alongside with the seeds are ranked based on the fitness value in order to eliminate the solutions with the least fitness values to keep the number of the weed plants and seeds within the maximum allowable population size. The seeds and their parent weeds with higher fitness survive, and become reproductive. The process continues until the convergence criteria is satisfied, i.e., reaching the desired number of iterations. The previous computation steps are modeled in Matlab environment to enhance the practicality and flexibility of the proposed spalling detection model.

5. Automated Assessment Model

5.1 Feature Extraction Model

As mentioned earlier, the proposed method utilizes singular value decomposition and discrete wavelet transform to map the spatial domain and frequency domain features. These two

feature vectors are fused together to generate an overall feature vector set of size 1×210 that is employed as an input to the automated model for modeling spalling area. The basic theory of discrete wavelet transform is described in the following lines.

Wavelets refer to mathematical functions that allow the data analysis of signals and images into different frequency components by incorporating a set of band-limited filters. For example, DWT is used to analyze and synthesize the input signal into multiple frequency sub-bands. Two-dimensional discrete wavelet transform is utilized to divide the original image into four frequency sub-bands (sub-images) namely LL_n , LH_n , HL_n and HH_n , whereas the information of the image stored is decomposed in the form of frequency sub-bands. N-level wavelet decomposition for a digital image of size $2^n \times 2^n$ is depicted in Figure 2 (Kumar et al., 2015).

At the primary level, there are four frequency sub-bands namely LL_1 , LH_1 , HL_1 and HH_1 . Then, the frequency sub-band LL_1 is used as an input for the next level of decomposition, whereas the frequency sub-band is divided into frequency sub-bands namely LL_2 , LH_2 , HL_2 and HH_2 . Then, the process continues until reaching the desired number of decomposition levels. As mentioned before, the frequency sub-bands are obtained by applying analysis filters during the different levels of decomposition. The four frequency sub-bands (sub-images): LL_n , LH_n , LH_n and HH_n can be computed using Equations (12), (13), (14) and (15), respectively. LL_n is obtained by applying low-pass filter for both rows and columns. LH_n is obtained by applying low-pass filter for rows and high-pass filter for columns. HL_n is computed by applying low-pass filter for columns and high-pass filter for rows. HH_n is computed by applying high-pass filter for both rows and columns. A low-pass filter is a filter that passes all the frequencies lower than a cut-off frequency. However, a high-pass filter is a filter that passes all the frequencies higher

than a cut-off frequency. High-pass filter helps in providing detailed image pixels while low-pass filter provides coarse approximation for the image pixels (Malonia and Agarwal, 2016).

$$\Phi_{LL}(x, y) = \Phi(x)\Phi(y) \quad (12)$$

$$\psi_{LH}(x, y) = \Phi(x)\psi(y) \quad (13)$$

$$\psi_{HL}(x, y) = \psi(x)\Phi(y) \quad (14)$$

$$\psi_{HH}(x, y) = \psi(x)\psi(y) \quad (15)$$

Where;

$\Phi(.)$ refers to the scaling function while $\psi(.)$ refers to the wavelet function. The scaling function denotes the low-pass filter while the wavelet function denotes the high-pass filter. $\Phi_{LL}(x, y)$ is a frequency sub-band that represents an approximation for the digital image, which can be further decomposed into multiple levels. $\psi_{LH}(x, y)$ is a frequency sub-band that represents the horizontal details of the image. $\psi_{HL}(x, y)$ is a frequency sub-band that represents the vertical details of the image. ψ_{HH} is a frequency sub-band that represents the diagonal details of the image. With respect to singular value decomposition, more information about it can be adopted from Chang et al. (2016) and Jha et al. (2014).

INSERT FIGURE 2

5.2 Modeling of Spalling Area

Elman neural network is a recurrent neural network that was introduced by Jeffrey Locke Elman in 1990. The main difference between the conventional feedforward neural networks and recurrent neural networks (RNN) is that in the case of RNNs, the output at each time step depends on the previous inputs and previous computations by memorizing previous events. ENN consists of: input layer, hidden layer, context layer, and output layer, whereas number of neurons

in the context layer is the same as that of the hidden layers. ENN is considered as a recurrent neural network because it has a feedback loop, which has a substantial impact on improving the learning capability of the network. The feedback loop utilizes unit-delay element (Z^{-1}), which provides a non-linear dynamic behaviour to the neural network (Kurach and Pawlowski, 2016; Bianchi et al., 2017).

The output of the hidden layer is fed into two layers: context layer and output layer. The output from the hidden layer is sent into the context layer, stored, and fed through the weights into the hidden layer in order to rely on this information in the next iterations, so the neural network is constantly remembering the output from the hidden layer and re-feeding this output from the previous iteration into the hidden layer. This behaviour enables the neural network to maintain short term memory, which amplifies the network performance (Köker 2013; Wang et al., 2014; Yu and Xu, 2014).

In the proposed spalling area interpretation model (ENN – IWO), invasive weed optimization algorithm is applied to train the ENN based on the computed area from the automated detection model. The input of this model is the feature vector of the gray-level images of size 1×210 . In this regard, IWO algorithm is utilized for two main reasons which are: inferior accuracy of the gradient descent algorithm and manual tuning of hyper parameters. The proposed training method aids in overcoming premature convergence, local minima and over fitting issues elicited from traditional training mechanisms. Additionally, there are wide ranges of hyper parameters of Elman neural network which significantly affect the prediction performance of Elman neural network. In this regard, there is no exact method reported in the literature to compute the number of context layers and context neurons. Most of these methods are case dependent and cannot be generalized to be applied in other case studies in addition to their time

exhaustive nature. For instance, if the numbers of hidden layers and neurons are less than the optimal number this will result in a substantial decrease in the prediction accuracy. Additionally, if the numbers layers and neurons are more than their optimum number this creates a lengthy training computational time. In this context, the blindness in determination of such hyper parameters and absence of systematic efficient method for their computation will have a considerable negative impact on the computational efficiency and accuracy of the prediction model. Thus, a self-adaptive method is formulated for the sake of autonomous and dynamic tuning of the Elman neural network parameters and hyper parameters based on the present available dataset.

Invasive weed optimization algorithm is adopted for the sake of optimizing the parameters of Elman neural network and the configuration of its architecture. The development made in the use of Invasive weed optimization as a training mechanism encompasses selection of most suitable transfer functions between the network layers, number of hidden layers and hidden neurons, number of context layers and context neurons, and values of weights and bias terms. In this study eight types of transfer activation functions are investigated: hyperbolic tangent sigmoid transfer function, log-sigmoid transfer function, Elliot symmetric sigmoid transfer function, linear transfer function, positive linear transfer function, triangular basis transfer function, radial basis transfer function and normalized radial basis transfer function.

Optimality theory is fundamentally focused on fixed-length assumption, such that most of the optimization models involve a fixed vector length of decision variables in order to simulate a particular set of possible solutions in the design space. Nonetheless, few cases reported in the literature comprised variable length optimization models. In them, the length of vector of solutions changes iteratively over the course of training epochs. It should be highlighted herein

that the variable-length optimization problems are of more complex nature and they require more computational time and resources during the training process when compared against the fixed-length optimization problems. There is no clear definition for the gradient vector of the variable-length problem in the variable-length optimization problems. Hence, gradient-based methods are inefficient in dealing with the dynamic vector of solutions. One of the approaches to deal with the variable-length optimization models is to presume a fixed length for the decision variables and to tune iteratively the decision variable that causes variability in length. Nevertheless, this approach often leads to suboptimal solutions. Additionally, it is time inefficient and impractical method especially in the presence of wide ranges of decision variables. This necessitates the formulation of a new approach which enables the estimation of the varying length of vector of possible solutions over the course of iterations (Ryerkerk et al., 2016).

Due to the parameter and structural learning nature of the proposed training method, the current optimization model is a variable-length optimization problem, whereas its length varies iteratively as per the number of hidden layers, hidden neurons, context layers and context neurons. As such an estimator is designed to handle the dynamism of the configuration of the ENN by computing the number of weights and bias terms in each training epoch. The estimator can be defined using Equation (16).

$$\text{Num} = ((I + 1) \times N) + ((N \times C \times P + ((N + 1) \times N \times (P - 1)) + ((N + 1) \times O)) \quad (16)$$

Where;

Num represents the total number of weights and bias terms. I represents the number of input neurons. N indicates the number of hidden neurons. C represents the number of neurons in the context layer. P represents number of hidden and context layers. O depicts the number of output

neurons. For simplification purposes, the number of context layers is assumed to be equal to the number of hidden layers.

The training process is performed based on a single objective function which minimizes the mean absolute percentage error. Thus, the developed optimization model establishes a dynamically changing optimum configuration and characteristics of Elman neural network triggered by the number of images and their spatial and frequency features. The mean absolute percentage error is selected because it is a well-known good performing performance indicator, unitless, and un-biased performance metric. Moreover, it is usually more practical to deal with cost functions. MAPE of the spalling area can be computed using Equation (17).

$$MAPE = \frac{100}{k} \times \sum_{i=1}^K \frac{|P_i - O_i|}{O_i} \quad (17)$$

Where;

O_i indicates the actual spalling area. P_i indicates the predicted spalling area. K represents the number of images. The variables O_i is computed from the automated detection model.

6. TRANSFER LEARNING-BASED DEEP NEURAL NETWORKS

This section describes a set of widely-utilized pre-trained deep neural networks that are utilized to validate the developed prediction model.

6.1 AlexNet

AlexNet architecture was developed by krizhevsky et al. (2012) which won the ImageNet Large Scale Visual Recognition Challenge in 2012. It encompasses five convolutional layers and three fully connected layers, whereas the input images are of size $224 \times 224 \times 3$. In the first

convolutional layer, 96 kernel filters of size $11 \times 11 \times 3$ and a stride of 4 pixels are applied to generate a stack of 96 feature maps. In the second convolutional layer, 256 kernels of size $5 \times 5 \times 48$ are employed. The third convolutional layer utilizes 384 kernels of size $3 \times 3 \times 256$ while the fourth uses 384 kernel filters of size $3 \times 3 \times 192$. In the fifth layer, 256 kernels of size $3 \times 3 \times 192$ are adopted. It is worth mentioning that the third, fourth and fifth layers are connected without intervening normalization or pooling layers. Additionally, each of the three fully connected layers encompasses 4096 neurons.

6.2 VGG16

VGG16 is a typical pre-trained deep neural network that was introduced by Simonyan and Zisserman in 2014 from Visual Geometry Group (VGG) at university of Oxford. They studied the implication of increasing the depth of the network while maintaining small sizes of kernel filters. It was found that the learning capacity of deep neural network can be significantly increased by establishing 16 to 19 weighted layers. A typical VGG16 architecture is composed of 16 weighted layers that involve a stack of 13 convolutional layers and 3 fully connected layers. In this context, some of the convolutional layers are followed by max pooling layers for the purpose of reducing the dimension of the input variable. The default size of input image is 224×224 RGB image. The convolutional layers utilize smaller kernel filters of size 3×3 and with stride and padding of 1 pixel when compared against AlexNet. The max pooling layer is of size 2×2 and stride of 2 pixels (Simonyan and Zisserman, 2015).

6.3 VGG19

VGG19 is considered as a deeper neural network than AlexNet and VGG16. It is composed of 19 trainable layers that involve 16 convolutional layers and 3 fully connected layers. The input images are of size 224×224 and three channels. The number of filter kernels

starts with 64 in the first layer and it is doubled after each max pooling layer until reaching 512 kernels. In this context, the convolutional layers use small kernels of size 3×3 with 1 pixel for stride and padding. The convolution operations are followed by rectified linear unit to add the dimension of non-linearity to the network structure. The max pooling layers utilizes kernel filters of size 2×2 and stride of 2 pixels for downsizing the feature map. The stack of convolutional layers is followed by three fully-connected layers, whereas the first one involves 4096 channels while the last two fully-connected layers involve 1000 channels. The last fully-connected layer is followed by softmax layer which generates probabilities for 1000 classes present in the Image Net dataset. It is important to highlight that VGG19 is a larger network than VGG16 and more computationally exhaustive to train (Simonyan and Zisserman, 2015).

6.4 CaffeNet

CaffeNet is a reproduction of AlexNet with two main important differences. First, it does not use data augmentation. Second, the pooling layer is applied before the local normalization layer. It is composed of five convolutional layers, three fully connected layers and three max pooling layers, whereas the input size of the image is $227 \times 227 \times 3$. The max pooling layers are applied after the first, second and fifth convolutional layers. The activation function in all the layers is the rectified linear unit. For instance, the first layer comprises 96 filters of size 11×11 and stride of 4 pixels. The fifth convolutional layer utilizes 256 kernel filters of size 3×3 with stride and padding of 1 pixel (Jia et al., 2014).

7. PERFORMANCE EVALUATION METRICS

The proposed method employs two different types of performance evaluation metrics to study quantitatively and automatically the efficiency of the different spalling segmentation models instead of their evaluation through human eye perception. In the first type, a set of image

quality indicators are applied, namely mean-squared error (MSE), peak signal to noise ratio (PSNR) and mean absolute error (MAE). An efficient image segmentation approach is the one, which sustains higher peak signal to noise ratio and lower mean-squared error and mean absolute error. In this context, the mean-squared error, peak signal to noise ratio and mean absolute error can be mathematically expressed using the following Equations (Mishra and Panda, 2018; Chakraborty et al., 2017).

$$\text{MSE} = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n [X(i, j) - Y(i, j)]^2 \quad (18)$$

$$\text{PSNR} = 10 \times \log_{10} \left(\frac{R^2}{\text{MSE}} \right) \quad (\text{in decibels}) \quad (19)$$

$$\text{MAE} = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n |X(i, j) - Y(i, j)| \quad (20)$$

Where;

$X(i, j)$ and $Y(i, j)$ denote the original and segmented images, respectively. R represents the maximum intensity value in the original image.

In the second type, the spalling segmentation is modelled as a binary classification problem. The classification evaluation metrics involve overall accuracy, F-measure, balanced accuracy, Matthews's correlation coefficient. Overall accuracy and F-measure are of well-known measures reported for their efficiency in symbolizing the performance of segmentation models. Additionally, each image can sometimes be dealt as imbalanced dataset because the number of foreground pixels is different from the number of background pixels. As such, the overall accuracy and F-measure can be overoptimistic and lead in return to misinterpretation of the performances of spalling segmentation models. In this regard, the proposed method utilizes

balanced accuracy, Matthews's correlation coefficient to handle the data imbalance. The overall accuracy (Ov_ACC), F-measure, balanced accuracy (BACC) and Matthews's correlation coefficient (MCC) can be mathematically computed using Equations (21), (22), (23) and (24), respectively (Jain et al., 2018; Shang et al., 2017; Jiao et al., 2016).

$$\text{Ov_ACC} = \frac{\sum_{i=1}^R X_{ii}}{\sum_{i=1}^R \sum_{i=1}^R X_{ii}} \quad (21)$$

$$\text{F-measure} = \frac{2 \times \text{sensitivity} \times \text{precision}}{\text{sensitivity} + \text{precision}} \quad (22)$$

$$\text{BACC} = \frac{\text{Sensitivity} + \text{Specificity}}{2} \quad (23)$$

$$\text{MCC} = \frac{(\text{TP} \times \text{TN}) - (\text{FP} \times \text{FN})}{\sqrt{(\text{TP} + \text{FP}) \times (\text{TP} + \text{FN}) \times (\text{TN} + \text{FP}) \times (\text{TN} + \text{FN})}} \quad (24)$$

Such that;

$$\text{Precision} = \frac{\sum_{i=1}^C \text{TP}}{\sum_{i=1}^C \text{TP} + \sum_{i=1}^C \text{FP}} \quad (25)$$

$$\text{Sensitivity (recall)} = \frac{\sum_{i=1}^C \text{TP}}{\sum_{i=1}^C \text{TP} + \sum_{i=1}^C \text{FN}} \quad (26)$$

$$\text{Specificity} = \frac{\sum_{i=1}^C \text{TN}}{\sum_{i=1}^C \text{TN} + \sum_{i=1}^C \text{FP}} \quad (27)$$

Where;

C indicates number of classes. TP, TN, FP and FN denote the number of true positive pixels, true negative pixels, false positive pixels and false negative pixels, respectively. The values of TP, TN, FP and FN are obtained from the confusion matrix, which consists of two columns and two rows since it is used to report a binary classification model. R represents number of rows or columns (classes). X_{ii} represents the element of i – th row and i – th column in the confusion

matrix (diagonal elements). MCC returns values between -1 and +1, whereas +1 implies perfect agreement, 0 implies similarity to random prediction and -1 indicates total disagreement between the ground truth image and segmented image. True positives are pixels labeled as spalling in the ground truth and they correctly classified by the segmentation model as spalling pixel. True negatives are background pixels in the ground truth that are correctly interpreted by the segmentation model as background pixels. False positives refer to pixels that are annotated as non-spalling pixels in the ground truth and they are incorrectly detected as spalling pixels. False negatives are pixels labeled as spalling in the ground truth and they are falsely detected as background pixels. In the present study, the pixels are manually labeled to generate the ground truth image. Furthermore, a pixel to pixel comparison is carried out to compute the evaluation metrics. It is worth mentioning that higher overall accuracy, F-measure, balanced accuracy, Matthews's correlation coefficient symbolize a more efficient spalling segmentation model.

METHOD IMPLEMENTATION

The proposed method utilizes a dataset composed of 60 real-world images, whereas 50 and 10 images are used for training and testing purposes, respectively. These images were captured from three bridge decks in Montreal and Laval, Canada using Sony DSC-H300 digital camera of 20.1 megapixel resolution. All the computations and optimization algorithms took place on a laptop with an Intel Core i7 CPU, 2.2 GHz and 16 GB of memory. Sample of the spalling images used are shown in Figures 3 and 4. As can be seen, different lighting conditions are considered for the purpose of establishing a robust model that is invariant to lighting conditions. The computerized platform gives the user the flexibility to select between the single restoration process and hybrid restoration process in the image restoration. In some cases, the images may be corrupted by a combination of noises, whereas a single standalone filter fails to remove the

noises from the images. The proposed method enables the user to select among a combination of spatial domain and frequency domain filters in order to provide more in-depth evaluation and better restored images in addition to defining the size of the filter. The investigated filters include: median filter, mean filter, mode filter, Wiener filter, Gaussian filter, Lee filter and Frost filter. In the present case study, Wiener filter of size 3×3 is applied to restore the images. Then, by clicking the “Import” button, the restored image is displayed. As shown in Figure 5, Wiener filter is capable of removing the noise present in the image and build the reconstructed image without losing its important features.

INSERT FIGURE 3

INSERT FIGURE 4

INSERT FIGURE 5

The next module is the contrast enhancement, whereas the user is asked to specify the adjusting ratio and margin parameter. As shown in Figure 6, the min-max gray level discrimination method is capable of discriminating the gray intensities in the potential defected and non-defected regions. The following module is the spalling detection, whereas the user is asked to specify the entropy index of the Renyi’s entropy segmentation model as well as the parameters of the invasive weed optimization algorithm. The outcome of this module is the optimum threshold, Kapur and Renyi’s entropy fitness function values, set of performance indicators in addition to the segmented image (See Figure 6). The number of iterations and the initial population size are assumed 30 and 10, respectively. The maximum and minimum numbers of seeds are 5 and 0, respectively. The initial and final standard deviations are assumed 0.5 and 0.001, respectively. The non-linear modulation index is assumed 2. The proposed

detection model computed an optimum threshold of 146, which achieved mean-squared error, mean absolute error and peak signal to noise ratio of 0.408, 0.626 and 3.364, respectively. Finally, Figure 6 demonstrates the final outcome of the proposed method after the application of the multi-objective invasive weed optimization-based segmentation method and the morphological operations. It is worth mentioning that the proposed detection model has successfully recognized the defected region present in the image.

INSERT FIGURE 6

A comparative analysis of the different segmentation models for the 60 images is recorded in Table 1. The comparison is carried out based on the average values of MSE, MAE, PSNR, Ov_ACC, F – measure, BACC and MCC. It is worth mentioning the differences in the values of the image quality indicators between the different spalling segmentation models due to the fact that the developed detection method deals with bi-level thresholding problems. As shown in Table 1, the proposed model achieved very promising results. It can be observed that the proposed model outperformed other spalling detection models achieving MSE, PSNR, MAE, Ov_ACC, F – measure, BACC and MCC of 0.399, 0.621, 3.56, 90.448%, 90.981%, 91.343% and 0.818, respectively. It can be also inferred that K-means clustering, fuzzy C-means clustering and expectation maximization algorithms provided similar efficiency in detecting spalling pixels in images. Region growing failed to detect efficiently the spalling such that it attained MSE, PSNR, MAE, Ov_ACC, F – measure, BACC and MCC of 0.411 0.633, 3.433, 76.538%, 81.033%, 83.319% and 0.595, respectively.

INSERT TABLE 1

Two-tailed Student's t-tests were performed to evaluate the significance levels of the segmentation performances for the different spalling detection models. The tests were conducted using the image quality indicators and classification performance metrics. The performed student's t-tests examine the null hypothesis (H_0), which is that there is no significant difference between the segmentation performances obtained from each pair of spalling detection models. On the contrary, the alternative hypothesis (H_1) assumes that there is a significant difference between the segmentation performances produced by each pair of spalling detection models. Thus, if the P – value is less than the significance level, then the null hypothesis is rejected in favor of the alternative hypothesis. However, if the P – value is more than the significance level, then the null hypothesis is accepted. Statistical comparisons between the spalling detection models capitalizing on peak signal to noise ratio and balanced accuracy are reported in Tables 2 and 3, respectively.

INSERT TABLE 2

INSERT TABLE 3

As shown in Tables 2 and 3, the P – values of the pairs (proposed model, Otsu), (proposed model, K-means clustering), (proposed model, fuzzy C-means clustering), (proposed model, region growing) and (proposed model, expectation maximization) are less than 0.05, which implies that the null hypothesis (H_0) is rejected, which manifests that the proposed model significantly outperformed other spalling segmentation models stepping on peak signal to noise ratio and balanced accuracy. In Table 2, P – values of the pairs (K-means clustering, fuzzy C-means clustering), (K-means clustering, region growing) and (fuzzy C-means clustering, region growing) are less than 0.05. Thus, there are no statistical significant differences between these

segmentation models based on peak signal to noise ratio. It can be also concluded from Table 3 that the P – values of the pairs (K-means clustering, fuzzy C-means clustering), (K-means clustering, expectation maximization) and (fuzzy C-means clustering, expectation maximization) are less than 0.05, which implies that the null hypothesis (H_0) is accepted. Thus, there are no statistical significant differences between these spalling segmentation models based on the balanced accuracy.

Clearer visual comparisons between the different spalling detection models are conducted, whereas any pixel that has a value equal to one is appended as spalling. Otherwise, if the pixel is equal to zero it is considered as non-spalling. Figures 7, 8 and 9 depict the binarized images using the different spalling segmentation models for images 3.a, 3.b and 4.d, respectively. By examining the differences between the segmented images, it can be inferred that the proposed spalling detection method generates a better well-separated spalling pixels from the background when compared against other segmentation models.

INSERT FIGURE 7

INSERT FIGURE 8

INSERT FIGURE 9

In order to provide an in-depth evaluation of the performances of the spalling detection models, Table 4 presents the thresholds obtained using the developed multi-objective optimization model, Otsu, K-means clustering, fuzzy C-means clustering and expectation maximization algorithms. It should be mentioned that no threshold is reported for region growing algorithm because it capitalizes on mapping the image into a set of connected pixels called “Regions” according to a pre-defined homogeneity indicator. Then, these groups of pixels are

grouped into larger regions based on this attribute. As shown in Table 4, the optimum thresholds generated from the developed multi-objective optimization model is different from other segmentation models, and in some cases it is very distant such as images of ID 3, 4 and 6. This not only creates noise and distortion in the segmented image by other spalling detection models but also this necessitates the presence of more computational procedures of morphological operations in an attempt to improve the quality of segmented image.

INSERT TABLE 4

The second model is the quantification of spalling area. The developed ENN – IWO model is compared against a set of baseline models. In the back-propagation artificial neural network, the number of hidden layers, number of hidden neurons and momentum coefficient are assumed 4, 2 and 0.8, respectively. The Elman neural network model is composed of 4 hidden and context layers in addition to 2 hidden and context neurons. In the generalized regression neural network, the spread of the Gaussian activation function is assumed 1. For the radial basis neural network, the maximum number of neurons in the hidden layer is assumed 10 while the spread of the Gaussian activation function is assumed 1. With respect to the deep learning models, the architecture of CONVNET is composed of two convolutional layers whereas the first one utilizes 16 filters of size 3×3 and padding of 1 pixel. The second one uses 32 filters of size 3×3 and padding of 1 pixel. The first convolutional layer is followed by a max pooling layer of window size 2×2 and stride of 2 pixels. Each convolutional layer is followed by a rectified linear unit activation function to establish more effective training through mapping the negative values to zero and maintaining the positive values. Finally, one fully connected layer of one neuron followed by a regression layer. The deep learning models are trained using the stochastic

gradient descent algorithm. The number epochs, initial learning rate and MiniBatchSize are assumed 500, 0.01 and 5, respectively. Training performance of AlexNet for interpretation of spalling area based on original dataset is presented in Table 5. It reports the elapsed time of training process and the generated prediction accuracy.

INSERT TABLE 5

The interface of the developed ENN – IWO model for feature extraction is depicted in Figure 10. The automated platform gives the user the flexibility to select between Haar, Daubechies 1, Symlet 1 and Coiflet 5 wavelet transforms to build the frequency domain feature vector set. In the present study, Haar wavelet transform is selected. By clicking “View” button, the singular values of the spatial domain features alongside the energies of wavelet sub-bands of frequency domain features retrieved from the gray-level images are computed. The performance of the Elman neural network is substantially governed by its parameters and hyper parameters. Thus, the proposed method adopts IWO algorithm to determine the optimum topology of the ENN and its optimum characteristics. The ENN is composed of one output neuron for the prediction of mean absolute percentage error. In the automated assessment model, the user can select the lower and upper bounds for the different structural and parametric learning parameters in addition to the parameters of the invasive weed optimization algorithm. The maximum numbers of hidden layers and hidden neurons are equal to 10. Thus, the maximum length of the optimization problem is 4114, which is considered as a large search space that substantiates the employment of exhaustive training mechanism. The parameters of the IWO algorithm are as follows: the number of iterations and the initial population size are assumed 400 and 200, respectively. The maximum and minimum numbers of seeds are 5 and 0, respectively. The initial and final standard deviations are assumed 0.5 and 0.001, respectively. As shown in Figure 11,

the optimum numbers of hidden and context layers are two while the optimum numbers of hidden and context neurons are three. Normalized radial basis is the optimum transfer function. The convergence curve of the developed ENN – IWO model for predicting spalling area based on original dataset is shown in Figure 12. As shown in Figure 12, the least mean absolute percentage error achieved by the invasive weed optimization algorithm is 3.97%.

INSERT FIGURE 10

INSERT FIGURE 11

INSERT FIGURE 12

An illustration of the actual and predicted values of the spalling area is demonstrated in Figure 13. As shown in Figure 13, the proposed prediction model succeeded in simulating the actual values of spalling area, whereas it provided very close predicted values to the actual values. A performance comparison between the different prediction methods based on the original dataset is described in Table 6. It is worth mentioning that the RMSE is measured in cm^2 . The proposed ENN – IWO model attained the highest prediction accuracy with respect to other prediction models reported in the literature, whereas MAPE, RMSE and RMSPE are equal to 4.07%, 76.061 and 0.065, respectively. Pre-trained deep learning networks outperformed the trained from scratch deep neural network, whereas VGG16 achieved MAPE, RMSE and RMSPE of 6.774%, 93.176 and 0.184, respectively. On the level of conventional machine learning models, Elman neural network provided the highest prediction accuracies followed by radial basis neural network and then generalized regression neural network. In this regard, back back-propagation artificial neural network yielded the lowest prediction accuracies attaining MAPE, RMSE and RMSPE of 26.203%, 276.936 and 0.519, respectively.

INSERT FIGURE 13

INSERT TABLE 6

In order to further evaluate the prediction models, the original dataset is augmented for the sake of establishing more robust and generalized spalling interpretation models. In this regard, the augmented dataset is composed of 565 images, whereas 450 images (80%) are utilized for training the prediction models while the remaining 115 images (20%) are used for their testing. The present comparison adopts the same setting of hyper parameters of machine learning and deep learning models except the initial learning rate and MiniBatchSize which are assumed 0.0001 and 50, respectively. A sample of training performance based on VGG16 for modeling of spalling area based on augmented dataset is reported in Table 7.

INSERT TABLE 7

With respect to the developed ENN – IWO model, it utilizes the same parameters as the ones used to train the original dataset except the number of iterations and initial population size are assumed 500 and 125, respectively. In this regard, the optimum numbers of hidden and context layers while the optimum numbers of hidden and context neurons are nine. Additionally, normalized radial basis is selected as the optimum transfer function. The convergence curve of the developed ENN – IWO model for predicting spalling area based on augmented dataset is depicted in Figure 14. As can be seen, the lowest training mean absolute percentage error achieved by the developed ENN – IWO model is 4.06%. A performance comparison between the prediction models based on the augmented dataset is reported in Table 8. It can be inferred that the developed model provided the highest prediction accuracies such that it attained MAPE, RMSE and RMSPE of 4.625%, 81.257 and 0.087, respectively. On the level of deep learning

models, CONVNET outperformed the pre-trained deep learning networks achieving MAPE, RMSE and RMSPE of 5.218%, 83.797 and 0.105, respectively. It is observed also that AlexNet, VGG16, VGG19 and CaffeNet generated very close prediction accuracies. With respect to the conventional machine learning models, Elman neural network and radial basis neural network accomplished the highest prediction performance while artificial neural network and generalized regression provided the highest prediction error. In this regard, it can be interpreted the developed model outperformed the conventional Elman neural network based on original and augmented datasets. This can be explained by the facts that the developed model offers two layers of improvement over the conventional Elman recurrent neural network, which are the application of singular value decomposition and discrete wavelet transform for efficient feature extraction in addition to employing invasive weed optimization algorithm to enhance the search mechanism of Elman neural network by boosting the exploration-exploitation trade-off.

INSERT FIGURE 14

INSERT TABLE 8

Another fold of comparison is carried out to validate the developed model for predicting spalling area. In this regard, it is compared against the model developed by Dawood et al. (2017) that utilized artificial neural network for quantifying spalling area in subway networks. It is understood that it is difficult to evaluate the performance metrics of two models based on two different datasets. However, the size of original dataset of the present study is less than the one used by Dawood et al. (2017) which is a substantial factor which could lead to a worse performance based on the prediction accuracies metrics. Nevertheless, it is found that the developed ENN – IWO model achieved higher prediction performance. The metrics stdv_MAPE and N_img refer to the standard deviation of the MAPE and number of images, respectively. As

described in Table 9, the model developed by Dawood et al. (2017) achieved MAPE and stdv_MAPE of 11% and 7.1%, respectively using 75 images. On the other hand, the proposed model achieved stdv_MAPE of 5.127% using 60 images. This evinces that the proposed achieved lower MAPE and stdv_MAPE using less number of images.

INSERT TABLE 9

The third model is designated for establishing a severity rating system of spalling based on its area and depth. In this regard, spalling area and depth are assumed as random variables in order to generate large number of possible scenarios. The best-fit distribution is the one which is associated with the lowest Anderson Darling statistic (A^2). The probability distributions and their corresponding rankings and Anderson Darling statistics are depicted in Table 10. As shown in Table 10, logistic and exponential achieved the lowest and highest Anderson Darling statistics, respectively. In this context, logistic is the best-fit distribution and its location and scaling factors are assumed 979.37 and 102.25, respectively. It is worth mentioning that loglogistic is the best-fit distribution of spalling depth and its location and scaling factors are assumed 1.58 and 0.28, respectively. A sample of the cluster memberships of spalling area obtained from the fuzzy C-means clustering algorithm are displayed in Table 11. In fuzzy C-means clustering, the data point is assigned to the cluster with the highest degree of membership. As shown in Table 11, the data point ١٠١٣,١٨٦ is assigned to “Cluster 2” since it is accompanied by the maximum degree of membership of 0.461. Additionally, the data point 760.253 is assigned to “Cluster 1” because it has the maximum degree of membership of 0.901. The thresholds utilized to describe the severity levels of spalling area and depth are presented in Table 12. The severity level of spalling is expressed in the form of percentage of zone area. It can be observed that if the spalling area is between 45% and 55%, this implies that the bridge deck is in a “Poor” condition. Furthermore, it

can be concluded that if spalling depth is more than 10 millimeters. Then, the bridge deck is in a “Very Poor” condition category.

INSERT TABLE 10

INSERT TABLE 11

INSERT TABLE 12

It is worth noting that the main limitation of the present study is the long computational time required to train the developed ENN – IWO model for modeling spalling area. However, almost most of the execution time is consumed in the learning process of the underlying pattern between inputs and outputs. Furthermore, the testing time is nearly the same for all the prediction models. The long computational training time of the developed model can be explained by the fact that it has capacity to optimize all of the parameters and hyper parameters of Elman neural network. The structure and parameter learning is a very exhaustive search process, which requires more processing time to explore the search space efficiently. It should be pointed out that the testing time is a more useful performance evaluation metric than training time in practical applications because the training process is only needed to be performed once. Additionally, the issue of long computational time can be resolved using parallel computing or cloud computing platforms.

8. CONCLUSION

Current bridge inspections are carried out through visual inspection. However, these inspections are time-consuming, involves safety risks and inefficient in terms of both cost and accuracy. As such, the present study introduces a computerized platform that involves a newly-developed two-tier self-adaptive method for the automated detection and assessment of spalling

in reinforced concrete bridges. In the first model, a multi-objective invasive weed optimization-based method is designed to detect spalled concrete. In this study, the optimum threshold values are computed by maximizing the Kapur entropy and Renyi's entropy objective functions. The second model is designed for the purpose of automated quantification of spalling area. In it, a hybridization of singular value decomposition and discrete wavelet are designed for the efficient use of the dimensional feature space. Invasive weed optimization algorithm is applied to identify the optimum topology and parameters of Elman neural network during the training process in an attempt to improve its prediction accuracy of spalling area. The validation was conducted through comparison against other well-performing machine learning and deep learning models reported in the literature capitalizing on the original and augmented datasets. The third model is designed for building a severity scale of spalling based on its area and depth. In this model, fuzzy C-means clustering algorithm was adopted to compute the thresholds of severity rating system of spalling.

It was derived that the developed multi-objective spalling segmentation provided significant superior segmentation accuracies when compared against conventional segmentation methods, whereas it accomplished an improvement in the detection of spalling by 9.79% and 12.29% with reference to Otsu and region growing algorithms. Additionally, experimental results exemplified that the proposed ENN – IWO model achieved higher prediction accuracies, whereas it yielded MAPE, RMSE and RMSPE of 4.625%, 81.257 and 0.087, respectively based on the augmented dataset. With respect to the third model, it was found that the thresholds of severity rating system of spalling area are 35%, 45% and 55% of zone area. It is expected that the proposed automated method can improve the evaluation accuracies of spalling severities, which eventually aids in the establishment of more accurate condition assessment models and optimum maintenance actions

by transportation agencies. The developed method can be extended in the future to study the implications of surface defects and their propagation over time on stiffness reduction of bridge deck. Additionally, more recent optimization algorithms can be investigated to train the Elman neural network in order to compare their prediction accuracies with the ones reported in the present study.

Data Availability Statement

Some or all data, models or code that support the findings of this study are available from the corresponding author upon request.

REFERENCES

1. Adhikari, R. S., Moselhi, O., and Bagchi, A. (2014). “Automation in Construction Image-based retrieval of concrete crack properties for bridge inspection”. *Automation in Construction*, 39(1), 180–194.
2. Alhamdoosh, M., and Wang, D. (2014). “Fast decorrelated neural network ensembles with random weights.” *Information Sciences*, 264, 104–117.
3. Al-Allaf, O. N. A. (2011). “Fast BackPropagation Neural Network algorithm for reducing convergence time of BPNN image compression.” *2011 International Conference on Information Technology and Multimedia*, Kuala Lumpur, Malaysia, 14-16 November, 1-6.
4. Al Najada, H., Mahgoub, I., and Mohammed, I. (2019). “Cyber Intrusion Prediction and Taxonomy System Using Deep Learning and Distributed Big Data Processing.” *Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence*, Bangalore, India, 18-21 November, 631-638.

5. Bhandari, A. K., Kumar, A., and Singh, G. K. (2015). Tsallis Entropy Based Multilevel Thresholding for Colored Satellite Image Segmentation using Evolutionary Algorithms. *Expert Systems With Application*, 42(22), 8707-8730.
6. Bhattacharjee, P. S., Fujail, A. K. M., and Begum, S. A. (2017). "A Comparison of Intrusion Detection by K-Means and Fuzzy C-Means Clustering Algorithm over the NSL-KDD Dataset." *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, Coimbatore, India, 14-16 December, 1-6.
7. Bianchi, F. M., Maiorino, E., Kampffmeyer, M. C., Rizzi, A., and Jenssen, R. (2017). "An overview and comparative analysis of Recurrent Neural Networks for Short Term Load Forecasting", *arXiv preprint*, 1-41.
8. Chakraborty, S., Chatterjee, S., Dey, N., Ashour, A. S., Ashour, A. S., and Shi, F. (2017). "Modified Cuckoo Search Algorithm in Microscopic Image Segmentation of Hippocampus." *Microscopy Research and Technique*, 1–22.
9. Chang, C. C., and Chen, S. H. (2019). "Developing a Novel Machine Learning-Based Classification Scheme for Predicting SPCs in Breast Cancer Survivors." *Frontiers in Genetics*, 10, 1–6.
10. Dawood, T., Zhu, Z., and Zayed, T. (2017). "Machine vision-based model for spalling detection and quantification in subway networks". *Automation in Construction*, 81, 149–160.
11. Demertzis, K., and Iliadis, L. (2020). "GeoAI : A Model-Agnostic Meta-Ensemble Zero-Shot Learning Method for Hyperspectral Image Analysis and Classification." *Algorithms*, 13(3), 1–24.
12. Dhillon, A., and Verma, G. K. (2019). "Convolutional neural network: a review of models, methodologies and applications to object detection." *Progress in Artificial Intelligence*, 1–28.

13. Di Biasi, M. R., Valencia, G. E., and Obregon, L. G. (2020). “A new educational thermodynamic software to promote critical thinking in youth engineering students.” *Sustainability*, 12, 1–17.
14. Dorafshan, S., Thomas, R. J., Coopmans, C., & Maguire, M. (2018). “Deep Learning Neural Networks for sUAS-Assisted Structural Inspections: Feasibility and Application.” In *2018 International Conference on Unmanned Aircraft Systems*, 1-9, Dallas, United States of America, 12-15 June.
15. Dundar, M., Kou, Q., Zhang, B., He, Y., and Rajwa, B. (2015). “Simplicity of kmeans versus deepness of deep learning: A case of unsupervised feature learning with limited data.” *Proceedings of 2015 IEEE 14th International Conference on Machine Learning and Applications*, Miami, USA, 9-11 December, 883-888.
16. Dung, C. V., Sekiya, H., Hirano, S., Okatani, T., and 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.
17. Felio, G. (2016). “Canadian Infrastructure Report Card”. Canadian Construction Association, Canadian Public Works Association, Canadian Society for Civil Engineering, and Federation of Canadian Municipalities, Canada.
<[www.canadainfrastructure.ca/downloads/Canadian Infrastructure Report 2016.pdf](http://www.canadainfrastructure.ca/downloads/Canadian_Infrastructure_Report_2016.pdf)>
(06.05.2016).
18. García-Alfonso, H., and Córdova-Esparza, D.-M. (2018). “Comparison of uncertainty analysis of the Montecarlo and Latin Hypercube algorithms in a camera calibration model.” *2018 IEEE 2nd Colombian Conference on Robotics and Automation (CCRA)*, Barranquilla, Colombia, 1-3 November, 1-5.

19. Gupta, A., and Kumar, A. (2016). “ATC calculation including wind : A probabilistic study and a comparison of MCS and LHS.” *2016 IEEE 6th International Conference on Power Systems (ICPS)*, New Delhi, India, 4-6 March, 1-6.
20. Hasan, M., and El-Sakka, M. R. (2015). “Structural Similarity Optimized Wiener Filter : A Way to Fight Image Noise.” *International Conference Image Analysis and Recognition*, 60–68, Waterloo, Canada, 27-29 August.
21. Hoang, N. (2018). “Detection of Surface Crack in Building Structures Using Image Processing Technique with an Improved Otsu Method for Image Thresholding”. *Advances in Civil Engineering*, Article ID 3924120, 10 pages.
22. Hooda, H., Verma, O. P., and Singhal, T. (2015). “Brain tumor segmentation: A performance analysis using K-Means, Fuzzy C-Means and Region growing algorithm.” *Proceedings of 2014 IEEE International Conference on Advanced Communication, Control and Computing Technologies*, Ramanathapuram, India, 8-10 May, 1621-1626.
23. Jain, S., Kotsampasakou, E., and Ecker, G. F. (2018). “Comparing the performance of meta-classifiers — a case study on selected imbalanced data sets relevant for prediction of liver toxicity”. *Journal of Computer-Aided Molecular Design*, 32(5), 583–590.
24. Jáuregui, D. V., Tian, Y., and Jiang, R. (2006). “Photogrammetry applications in routine bridge inspection and historic bridge documentation.” *Transportation Research Record*, 1958(1), 24–32.
25. Jha, R. K., and Chouhan, R. (2014). “Noise-induced contrast enhancement using stochastic resonance on singular values”. *Signal, Image and Video Processing*, 8(2), 339–347.
26. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S. and Darrell, T. (2014). “Caffe: Convolutional architecture for fast feature embedding.”

Proceedings of the 22nd ACM International Conference on Multimedia, 675–678, Florida, United States of America.

27. Jiao, Y., and Du, P. (2016). “Performance measures in evaluating machine learning based bioinformatics predictors for classifications”. *Quantitative Biology*, 4(4), 320–330.
28. Khairuzzaman, A. K., and Chaudhury, S. (2017). “Multilevel Thresholding using Grey Wolf Optimizer for Image Segmentation”. *Expert Systems With Applications*, 86, 64–76.
29. Kim, I.-H., Jeon, H., Baek, S.-C., Hong, W.-H., and Jung, H.-J. (2018). “Application of Crack Identification Techniques for an Unmanned Aerial Vehicle”. *Sensors*, 18, 1–14.
30. Koch, C., Georgieva, K., Kasireddy, V., Akinci, B., and Fieguth, P. (2015). “A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure.” *Advanced Engineering Informatics*, 29(2), 196–210.
31. Köker, R. (2013). “A Genetic Algorithm Approach to a Neural-network-based Inverse Kinematics Solution of Robotic Manipulators Based on Error Minimization.” *Information and Control*, 222, 528–543.
32. Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). “ImageNet Classification with Deep Convolutional Neural Networks.” *Advances in Neural Information Processing Systems*, 1097–1105.
33. Kumar, A., Rastogi, P., and Srivastava, P. (2015). “Design and FPGA Implementation of DWT , Image Text Extraction Technique.” *Procedia Computer Science*, 57, 1015–1025.
34. Kurach, K., and Pawlowski, K. (2016). “Predicting Dangerous Seismic Activity with Recurrent Neural Networks.” *Proceedings of the Federated Conference on Computer Science*, Gdansk, Poland, 11-14 September, 239–243.
35. Lee, J. H., Jin, S. S., Kim, I. H., and Jung, H. J. (2017). “Development of crack diagnosis and

- quantification algorithm based on the 2D images acquired by Unmanned Aerial Vehicle (UAV).” *The 2017 Congress on Advances in Structural Engineering and Mechanics*, Seoul, Korea, 28 August- 1 Septemeber.
36. Lee, J. H., Lee, J. M., Kim, H. J., and Moon, Y. S. (2008). “Machine Vision System for Automatic Inspection of Bridges”. *2008 Congress on Image and Signal Processing*, Sanya, China, 27-30 May, 363–366.
 37. Li, Y., Fan, X., and Li, G. (2006). “Image Segmentation based on Tsallis-entropy and Renyi-entropy and Their Comparison”. *4th IEEE International Conference on Industrial Informatics*, Singapore, 16-18 August, 943–948.
 38. Lv, Z. Y., Zhang, P., Benediktsson, J. A., and Shi, W. Z. (2014). “Morphological profiles based on differently shaped structuring elements for classification of images with very high spatial resolution.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(12), 4644–4652.
 39. Malonia, M. and Agarwal, S. (2016). “Digital Image Watermarking using Discrete Wavelet Transform and Arithmetic Progression Technique.” *2016 IEEE Students’ Conference on Electrical, Electronics and Computer Science Digital*, Bhopal, India, 5-6 March, 1-6.
 40. Mohammed Abdelkader, E., Moselhi, O., Marzouk, M. and Zayed, T. (2020). “Hybrid Elman Neural Network and Invasive Weed Optimization Method for Bridge Defects Recognition, *Submitted to: Journal of the Transportation Research Board*.
 41. Mohammed Abdelkader, E., Moselhi, O., Marzouk, M. and Zayed, T. (2019). “A Multi-objective Invasive Weed Optimization Method for Segmentation of Distress Images.” *Intelligent automation and soft computing*, 1-20.
 42. Mehrabian, A. R., and Lucas, C. (2006). “A novel numerical optimization algorithm inspired from weed colonization.” *Ecological Informatics*, 1(4), 355–366.

43. Mishra, S., and Panda, M. (2018). "Bat Algorithm for Multilevel Colour Image Segmentation Using Entropy-Based Thresholding." *Arabian Journal for Science and Engineering*, 43, 7285–7314.
44. National Research Council Canada. (2013). "Critical Concrete Infrastructure: Extending the Life of Canada's Bridge Network". <<http://www.nrc-cnrc.gc.ca/ci-ic/article/v18n1-5>> (20.12.2016).
45. Nagarajan, B., and Devendran, V. (2012). "Singular value decomposition based features for vehicle classification under cluttered background and mild occlusion." *Procedia Engineering*, 30, 529–534.
46. Pal, S., and Chatterjee, S. (2018). "Mathematical morphology aided optic disk segmentation from retinal images." *2017 3rd International Conference on Condition Assessment Techniques in Electrical Systems*, Rupnagar, India, 16-18 November, 380-385.
47. Razmjoo, N. and Khalilpour, M. (2015). "A new design for PID controller by considering the operating points changes in Hydro-Turbine Connected to the equivalent network by using Invasive Weed Optimization (IWO) Algorithm." *International Journal of Information, Security, and Systems Management*, 4, 468–475.
48. Reddy, B. S. K., and Wanjari, S. (2018). "Core strength of concrete using newly developed in situ compressive testing machine." *Magazine of Concrete Research*, 70(22), 1149–1156.
49. Ryerkerk, M. L., Averill, R. C., Deb, K., and Goodman, E. D. (2016). "Solving Metameric Variable-length Optimization Problems Using Genetic Algorithms." *Genetic Programming and Evolvable Machines*, 128(2), 247-277.
50. Samsuri, S., Surbakti, M., Tarigan, A. P., and Anas, R. (2019). "A Study on the Road Conditions Assessment Obtained from International Roughness Index (IRI): Roughometer

- Vs Hawkeye.” *Simetrikal Journal of Engineering and Technology*, 1(02), 103–113.
51. Shanavaz, K. T., and Mythili, P. (2016). “A fingerprint-based hybrid gender classification system using genetic algorithm.” *International Journal of Computational Vision and Robotics*, 6(4), 399-413.
 52. Shang, X., Li, X., Morales-esteban, A., and Chen, G. (2017). “Improving microseismic event and quarry blast classification using Artificial Neural Networks based on Principal Component Analysis”. *Soil Dynamics and Earthquake Engineering*, 99, 142–149.
 53. Simonyan, K., and Zisserman, A. (2015). “Very deep convolutional networks for large-scale image recognition.” *3rd International Conference on Learning Representations*, 1–14, San Diego, United States of America, 7-9 May.
 54. Sennah, K., Juette, B., Witt, C., and Combar, P. M. (2011). “Vehicle Crash Testing On a GFRP-Reinforced PL-3 Concrete Bridge Barrier.” *Proceedings of the 4th International Conference on Durability and Sustainability of Fibre Reinforced Polymer Composites for Construction and Rehabilitation*, Canada, 1–8.
 55. Stamate, C., Magoulas, G. D., and Thomas, M. S. C. (2015). “Transfer learning approach for financial applications.” *ArXiv:1509.02807*, 1–5.
 56. Su, T. (2013). “Application of Computer Vision to Crack Detection of Concrete Structure.” *International Journal of Engineering and Technology*, 5(4), 457–461.
 57. Tao, W., Tao, D., Xiaohong, Z., and Yiming, S. (2014). “A Research on Detecting and Recognizing Bridge Cracks in Complex Underwater Conditions.” *Frattura ed Integrità Strutturale*, 30, 537–544.
 58. Uyar, K., and Ülker, E. (2017). “B-spline curve fitting with invasive weed optimization.” *Applied Mathematical Modelling*, 52, 320–340.

59. Xu, H., Su, X., Wang, Y., Cai, H., Cui, K., and Chen, X. (2019). “Automatic Bridge Crack Detection Using a Convolutional Neural Network.” *Applied Sciences*, 9(14), 1–14.
60. Xue-jun, X.U., and Xiao-ning, Z. (2013). “Crack Detection of Reinforced Concrete Bridge using Video Image.” *Journal of Central South University*, 20, 2605–2613.
61. Yang, Q., Shi, W., Chen, J., and Lin, W. (2020). “Deep convolution neural network-based transfer learning method for civil infrastructure crack detection.” *Automation in Construction*, 116, 103199.
62. Yu, F., and Xu, X. (2014). “A Short-term Load Forecasting Model of Natural Gas Based on Optimized Genetic Algorithm and Improved BP Neural Network”. *Applied Energy*, 134, 102–113.
63. Zhang, L., Luo, W., and Xu, Y. (2018). “Bridge Crack Image Segmentation Based on Improved Watershed Algorithm”. In *2018 Chinese Control And Decision Conference (CCDC)*, Shenyang, China, 9-11 July, 3537-3541.

List of Figures

Fig. 1: Framework of the proposed spalling detection and assessment method

Fig. 2: N-level wavelet decomposition for a digital image of size $2^n \times 2^n$

Fig. 3: Sample of the spalling images

Fig. 4: Sample of the spalling images (Continued)

Fig. 5: Interface of the developed spalling preprocessing module

Fig. 6: Interface of the developed spalling detection model

Fig. 7: Segmentation outcome of image 3.a using (a) proposed method, (b) Otsu, (c) K-means clustering and (d) fuzzy C-means clustering, (e) region growing and (f) expectation maximization

Fig. 8: Segmentation outcome of image 3.b using (a) proposed method, (b) Otsu, (c) K-means clustering and (d) fuzzy C-means clustering, (e) region growing and (f) expectation maximization

Fig. 9: Segmentation outcome of image 4.d using (a) proposed method, (b) Otsu, (c) K-means clustering and (d) fuzzy C-means clustering, (e) region growing and (f) expectation maximization

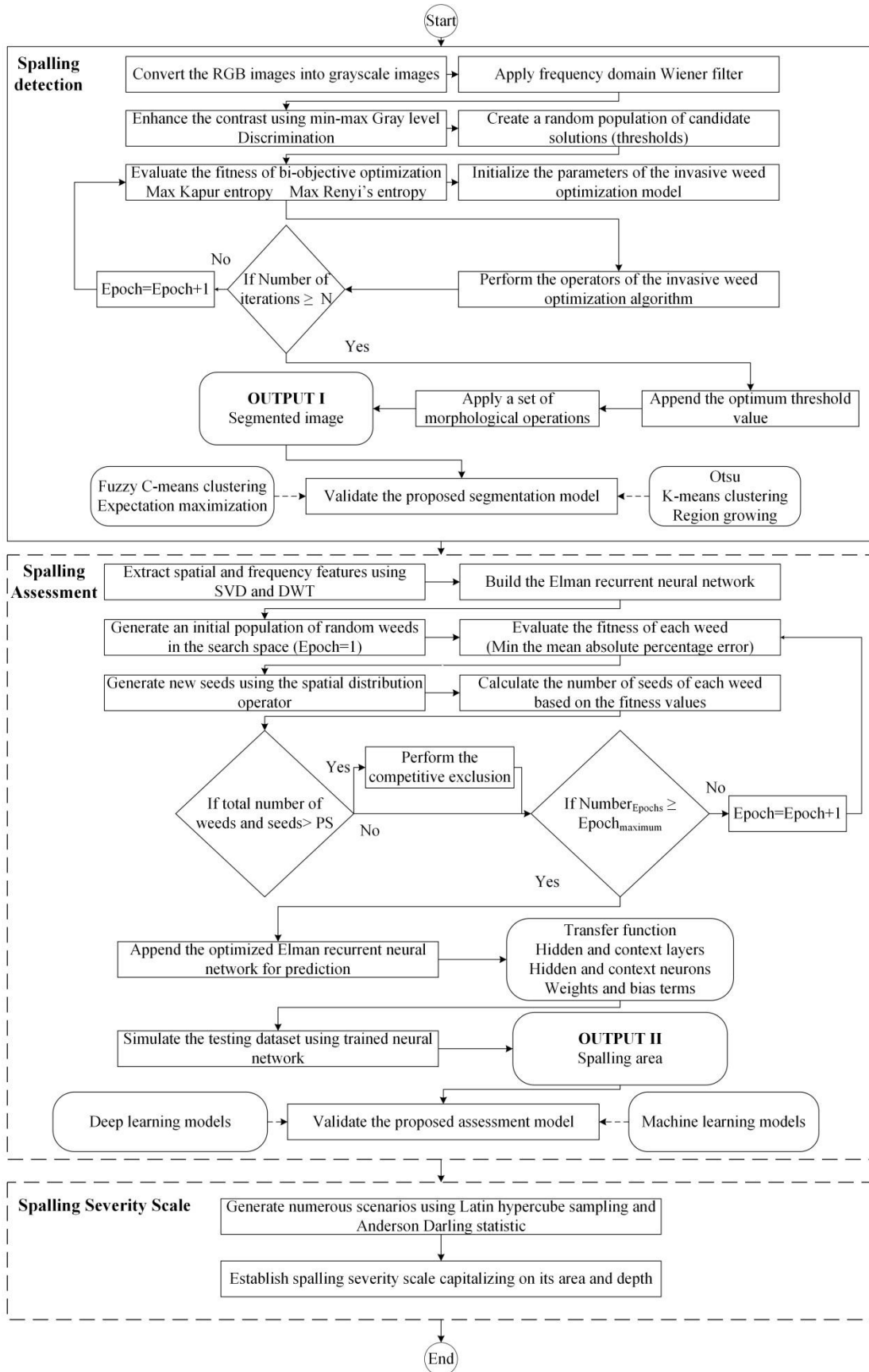
Fig. 10: Interface of the developed feature extraction model

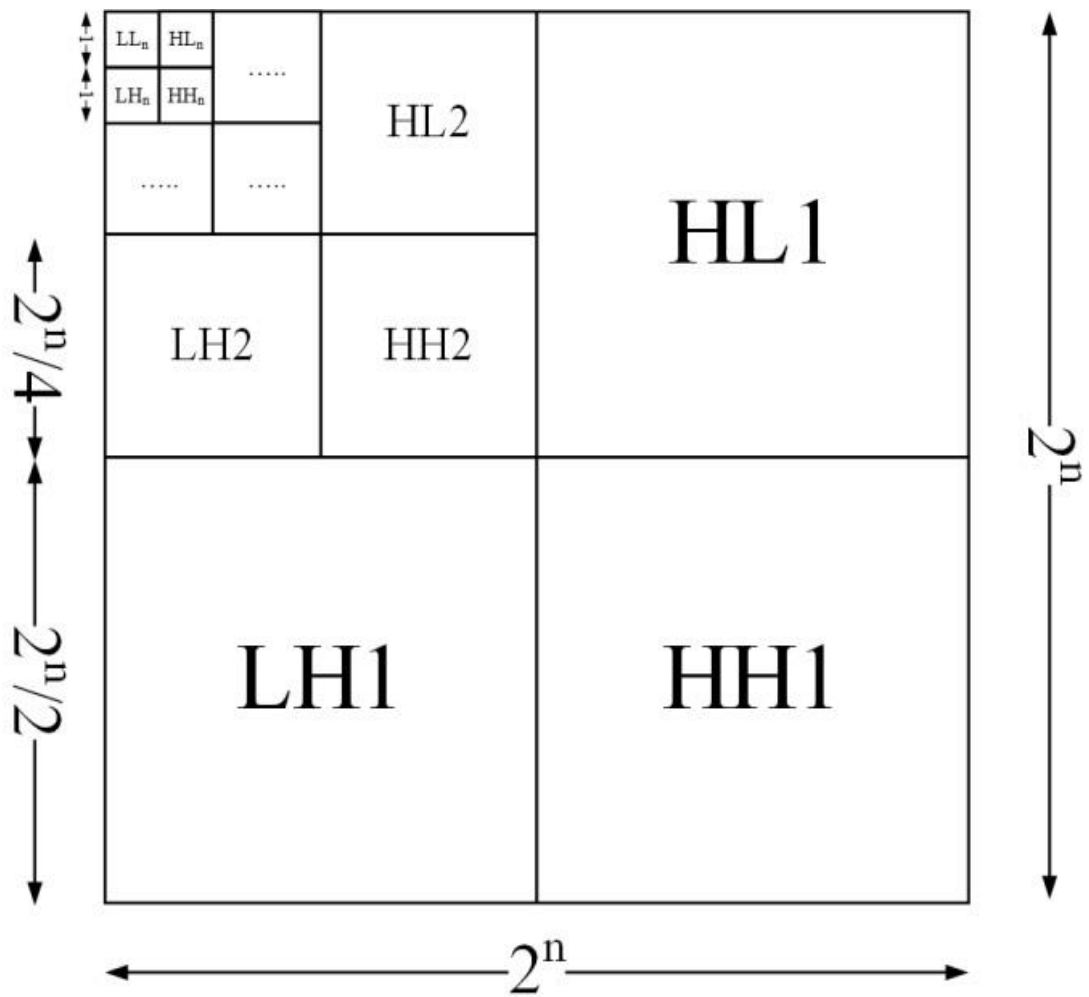
Fig. 11: Interface of the developed model for automated evaluation of spalling area

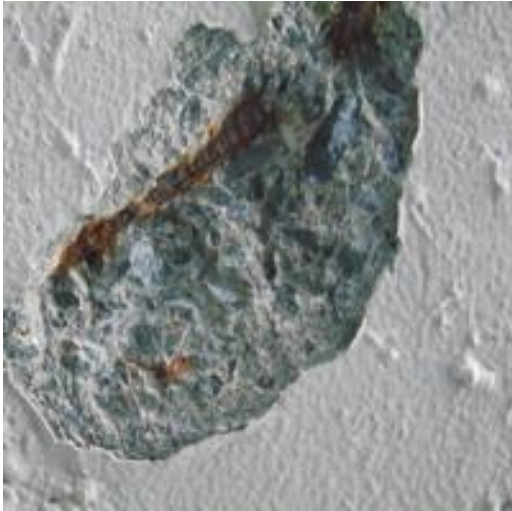
Fig. 12: Convergence Curve of the developed ENN – IWO model for predicting spalling area based on original dataset

Fig. 13: Actual and predicted spalling area for a sample of 20 images

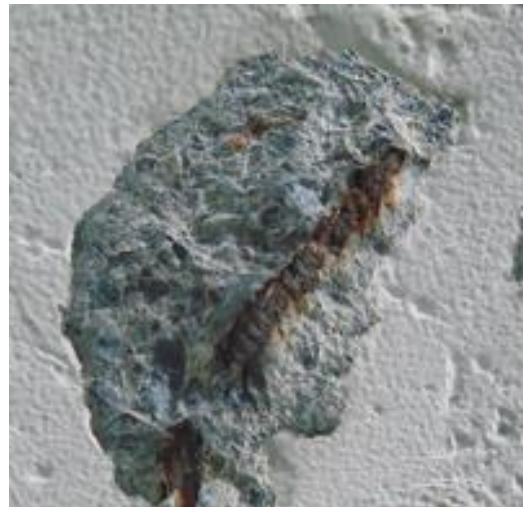
Fig. 14: Convergence Curve of the developed ENN – IWO model for predicting spalling area based on augmented dataset







(a) Spalling -Image "A"



(b) Spalling -Image "B"



(c) Spalling-Image "C"



(d) Spalling-Image "D"



(a) Spalling -Image “A”



(b) Spalling -Image “B”



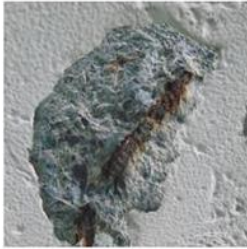
(c) Spalling-Image “C”



(d) Spalling-Image “D”

Image preprocessing

Image Retrieval



Import

Image Restoration

Single restoration

Type of filter "1"Wiener filter

Size of filter "1"3

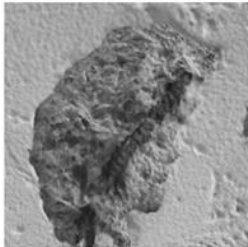
Hybrid restoration

Type of filter "1"

Size of filter "1"

Type of filter "2"

Size of filter "2"



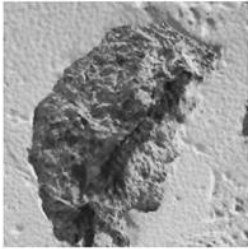
ComputeImport

Contrast Enhancement

User input

Adjusting ratio0.5

Margin parameter1.1



ExportCompute

ImportSave

Next

Spalling segmentation

User input

Entropy index

0.9

Parameters of the invasive weed algorithm

Initial population size

30

Maximum number of iterations

10

Minimum number of seeds

0

Maximum number of seeds

5

Initial standard deviation

0.5

Final standard deviation

0.001

Non-linear modulation index

2

Model output

Optimum threshold

146

Kapur entropy fitness function value

8.647

Renyi's entropy fitness function value

8.683

Mean-squared error

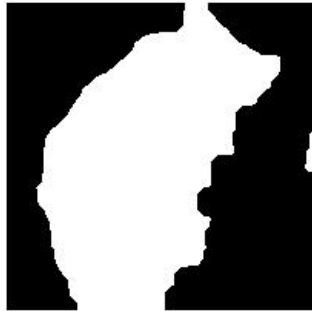
0.408

Mean absolute error

0.629

Peak signal to noise ratio

3.364



View

Compute

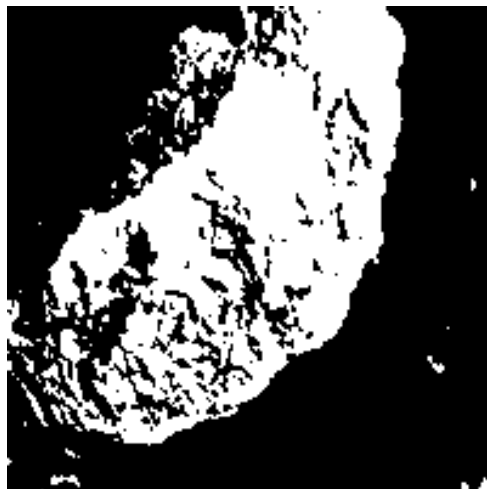
Export

Next

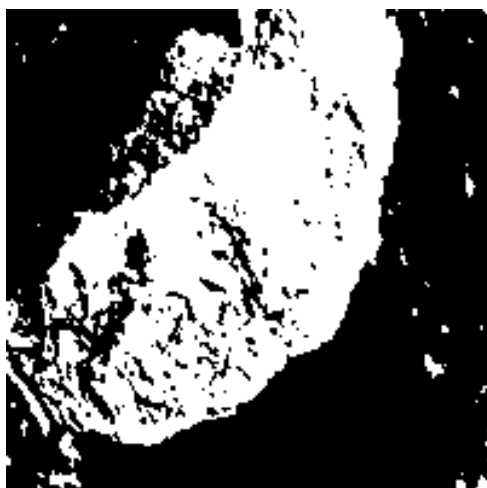
60



(b) Proposed method



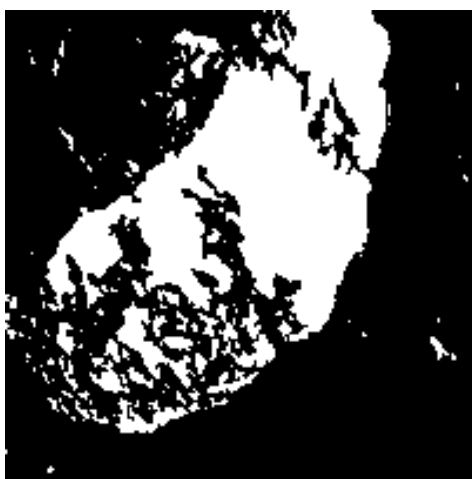
(b) Otsu



(c) K-means clustering



(d) Fuzzy C-means clustering



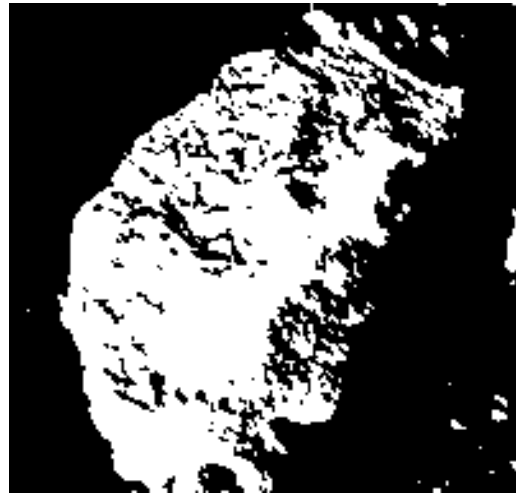
(f) Region growing



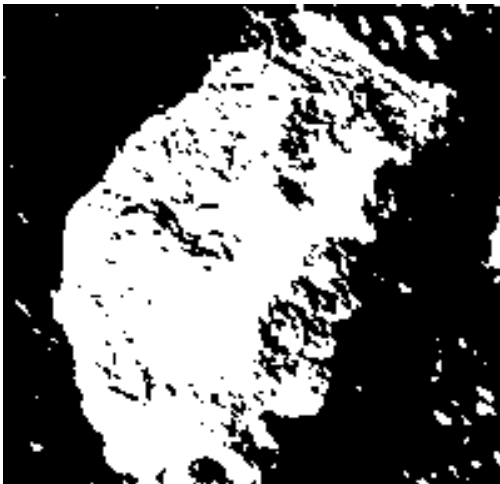
(f) Expectation maximization



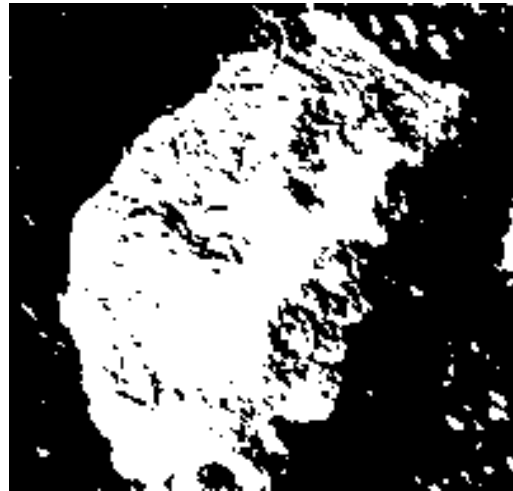
(a) Proposed method



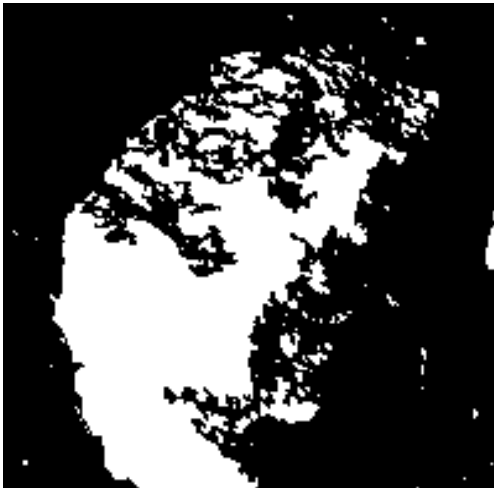
(b) Otsu



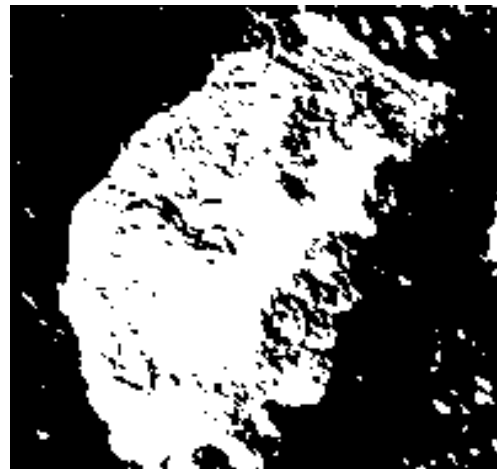
(c) K-means clustering



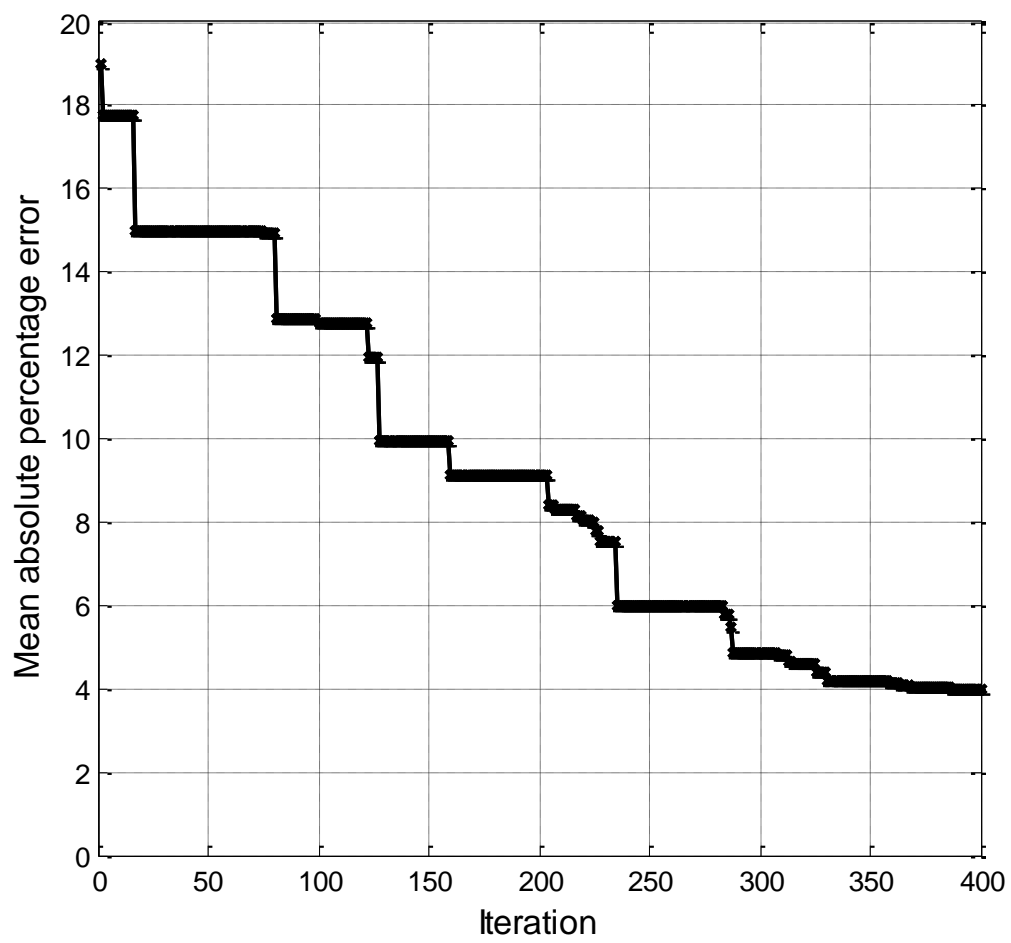
(d) Fuzzy C-means clustering

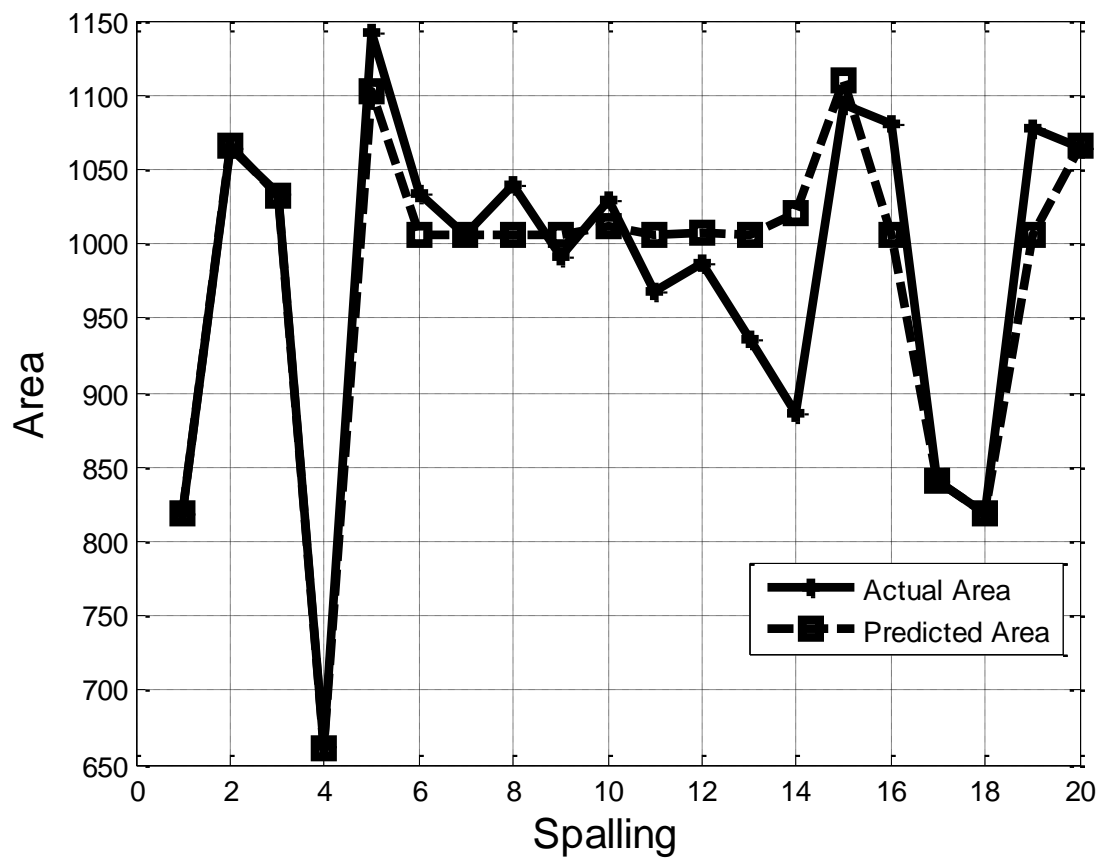


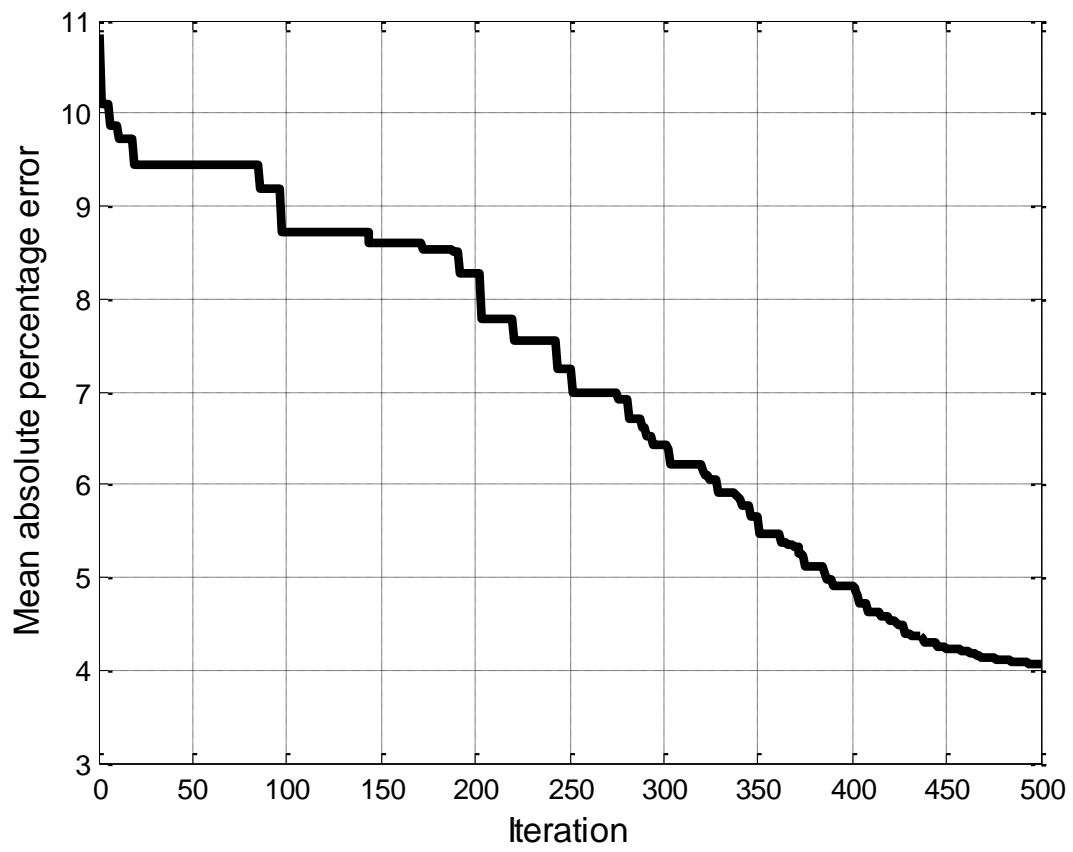
(e) Region growing



(f) Expectation maximization







List of Tables

Table 1: Performance comparison between the different spalling detection models

Table 2: Statistical comparison between the different spalling detection models using peak signal to noise ratio

Table 3: Statistical comparison between the different spalling detection models using balanced accuracy

Table 4: Sample of thresholds obtained from the spalling detection models

Table 5: Sample of training performance of AlexNet for interpretation of spalling area based on original dataset

Table 6: Performance comparison between the prediction models based on the original dataset

Table 7: Sample of training performance of VGG16 for interpretation of spalling area based on augmented dataset

Table 8: Performance comparison between the prediction models based on the augmented dataset

Table 9: Performance comparison of the developed model with previous literature

Table 10: Anderson Darling tests for probability distributions of spalling area

Table 11: Sample of the cluster memberships of spalling area obtained from fuzzy C-means clustering algorithm

Table 12: Severity rating system of spalling based on its area and depth

Table 1: Performance comparison between the different spalling detection models

Segmentation model	MSE	MAE	PSNR	Ov_ACC	F – measure	BACC	MCC
Proposed model	0.399	0.621	3.56	90.448%	90.981%	91.343%	0.818
Otsu	0.423	0.645	3.297	81.919%	83.849%	84.475%	0.659
K-means clustering	0.42	0.641	3.335	84.321%	85.225%	85.136%	0.617
Fuzzy C-means clustering	0.419	0.642	3.322	84.228%	85.208%	85.141%	0.694
Region growing	0.411	0.633	3.433	76.538%	81.033%	83.319%	0.595
Expectation maximization	0.419	0.641	3.334	84.302%	85.227%	85.141%	0.694

Table 2: Statistical comparison between the different spalling detection models using peak signal to noise ratio

Pair of segmentation models	Proposed model	Otsu	K-means clustering	Fuzzy C-means clustering	Region growing	Expectation maximization
Proposed model	H_0 (P – value=1)	H_1 (P – value = 1.33×10^{-6})	H_1 (P – value = 5.94×10^{-7})	H_1 (P – value = 1.08×10^{-6})	H_1 (P – value = 4.78×10^{-2})	H_1 (P – value = 1.29×10^{-6})
Otsu	H_1 (P – value = 1.33×10^{-6})	H_0 (P – value=1)	H_1 (P – value = 2.48×10^{-5})	H_1 (P – value = 4.3×10^{-2})	H_1 (P – value = 4.08×10^{-2})	H_1 (P – value = 4.67×10^{-5})
K-means clustering	H_0 (P – value = 5.94×10^{-7})	H_1 (P – value = 2.48×10^{-5})	H_0 (P – value=1)	H_0 (P – value = 1.96×10^{-1})	H_0 (P – value = 1.46×10^{-1})	H_1 (P – value = 4.19×10^{-2})
Fuzzy C-means clustering	H_1 (P – value = 1.08×10^{-6})	H_1 (P – value = 4.3×10^{-2})	H_0 (P – value = 1.96×10^{-1})	H_0 (P – value=1)	H_0 (P – value = 1.01×10^{-1})	H_0 (P – value = 2.34×10^{-1})
Region growing	H_1 (P – value = 4.78×10^{-2})	H_1 (P – value = 4.08×10^{-2})	H_0 (P – value = 1.46×10^{-1})	H_0 (P – value = 1.01×10^{-1})	H_0 (P – value=1)	H_0 (P – value = 1.44×10^{-1})
Expectation maximization	H_1 (P – value = 1.29×10^{-6})	H_1 (P – value = 4.67×10^{-5})	H_0 (P – value = 4.19×10^{-2})	H_0 (P – value = 2.34×10^{-1})	H_0 (P – value = 1.44×10^{-1})	H_0 (P – value=1)

Table 3: Statistical comparison between the different spalling detection models using balanced accuracy

Pair of segmentation models	Proposed model	Otsu	K-means clustering	Fuzzy C-means clustering	Region growing	Expectation maximization
Proposed model	H_0 (P – value=1)	H_1 (P – value = 5.59×10^{-7})	H_1 (P – value = 2.65×10^{-6})	H_1 (P – value = 2.31×10^{-6})	H_1 (P – value = 6.07×10^{-7})	H_1 (P – value = 2.53×10^{-6})
Otsu	H_1 (P – value = 5.59×10^{-7})	H_0 (P – value=1)	H_0 (P – value = 2.2×10^{-1})	H_0 (P – value = 2.13×10^{-1})	H_0 (P – value = 7.86×10^{-2})	H_0 (P – value = 2.12×10^{-1})
K-means clustering	H_1 (P – value = 2.65×10^{-6})	H_0 (P – value = 2.2×10^{-1})	H_0 (P – value=1)	H_0 (P – value = 8.46×10^{-1})	H_1 (P – value = 1.89×10^{-2})	H_0 (P – value = 7.89×10^{-1})
Fuzzy C-means clustering	H_1 (P – value = 2.31×10^{-6})	H_0 (P – value = 2.13×10^{-1})	H_0 (P – value = 8.46×10^{-1})	H_0 (P – value=1)	H_1 (P – value = 1.74×10^{-2})	H_0 (P – value = 9.55×10^{-1})
Region growing	H_1 (P – value = 6.07×10^{-7})	H_0 (P – value = 7.86×10^{-2})	H_1 (P – value = 1.89×10^{-2})	H_1 (P – value = 1.74×10^{-2})	H_0 (P – value=1)	H_1 (P – value = 1.71×10^{-2})
Expectation maximization	H_1 (P – value = 2.53×10^{-6})	H_0 (P – value = 2.12×10^{-1})	H_0 (P – value = 7.89×10^{-1})	H_0 (P – value = 9.55×10^{-1})	H_1 (P – value = 1.71×10^{-2})	H_0 (P – value=1)

1 **Table 4: Sample of thresholds obtained from the spalling detection models**

Image ID	Proposed model	Otsu	K-means clustering	Fuzzy C-means clustering	Expectation maximization
1	143	132	142	142	142
2	138	131	141	141	141
3	146	130	140	140	140
4	119	132	142	141	141
5	142	129	139	138	139
6	150	135	145	144	145
7	143	136	146	145	146
8	138	132	144	143	144

2
3
4
5
6
7
8
9
10
11
12
13
14

Table 5:Sample of training performance of AlexNet for interpretation of spalling area based on original dataset

Epoch	Iteration	Time Elapsed (seconds)	Mini-batch root mean- squared error	Base learning rate
1	1	0.01	17414.46	0.01
100	1000	3.74	2338.66	0.01
300	3000	10.57	2338.66	0.01
400	4000	14.04	2338.66	0.01
500	5000	17.8	2338.66	0.01

Table 6: Performance comparison between the prediction models based on the original dataset

Prediction model	Mean absolute percentage error	Root mean-squared error	Root mean squared percentage error
ENN – IWO	4.07%	76.061	0.065
Artificial neural network	26.203%	276.936	0.519
Elman neural network	18.043%	175.96	0.333
Generalized regression neural network	19.669%	198.707	0.359
Radial basis neural network	18.349%	176.041	0.335
CONVNET	9.362%	108.276	0.198
AlexNet	6.824%	93.184	0.186
VGG16	6.774%	93.176	0.184
VGG19	6.791%	93.303	0.182
CaffeNet	6.8%	93.209	0.184

Table 7: Sample of training performance of VGG16 for interpretation of spalling area based on augmented dataset

Epoch	Iteration	Time Elapsed (seconds)	Mini-batch root mean- squared error	Base learning rate
1	1	0.01	18156.69	0.0001
100	900	3.47	7926.5	0.0001
150	1350	4.93	5273.99	0.0001
300	2700	9.86	2074.68	0.0001
500	4500	16.68	1318.96	0.0001

Table 8: Performance comparison between the prediction models based on the augmented dataset

Prediction model	Mean absolute percentage error	Root mean-squared error	Root mean squared percentage error
ENN – IWO	4.625%	81.257	0.087
Artificial neural network	13.791%	225.2	0.231
Elman neural network	11.513%	216.306	0.225
Generalized regression neural network	13.775%	225.106	0.233
Radial basis neural network	12.183%	209.37	0.22
CONVNET	5.218%	83.797	0.105
AlexNet	6.394%	89.728	0.108
VGG16	6.395%	89.726	0.108
VGG19	6.393%	89.723	0.108
CaffeNet	6.393%	89.715	0.108

Table 9: Performance comparison of the developed model with previous literature

Method	MAPE	Stdv_MAPE	N_img
Proposed method	4.07%	5.127%	60
Dawood et al. (2017)	11%	7.1%	75

104 **Table 10: Anderson Darling tests for probability distributions of spalling area**

Probability distribution	Anderson Darling statistic (A^2)	Rank
Normal	2.545	3
Logistic	1.76	1
Loglogistic	2.843	4
Gamma	3.744	5
Largest extreme value	4.827	6
Exponential	18.05	7
Weibull	1.891	2

105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121

Table 11: Sample of the cluster memberships of spalling area obtained from fuzzy C-means clustering algorithm

Data point	Degree of membership				Assigned Cluster
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	
1013.186	0.048	0.461	0.445	0.044	Cluster 2
855.934	0.172	0.766	0.047	0.014	Cluster 2
1126.234	0.001	0.004	0.989	0.004	Cluster 3
1265.121	0.009	0.024	0.127	0.837	Cluster 4
760.253	0.901	0.077	0.015	0.005	Cluster 1
1140.675	0.004	0.015	0.957	0.022	Cluster 3
928.617	0.003	0.99	0.004	0.001	Cluster 2

143 **Table 12: Severity rating system of spalling based on its area and depth**

Condition category	Spalling area	Spalling depth
Good	Less than 35%	Less than 3.7 mm
Medium	Between 35% and 45%	Between 3.7 and 6.1 mms
Poor	Between 45% and 55%	Between 6.1 and 10 mms
Very Poor	More than 55%	More than 10 mm

144

145

146

147

148

149

150

151

152