



# Deep learning-based visual defect-inspection system for reinforced concrete bridge substructure: a case of Thailand's department of highways

Pravee Kruachottikul<sup>1</sup> · Nagul Cooharajanone<sup>2</sup> · Gridsada Phanomchoeng<sup>3</sup> · Thira Chavarnakul<sup>4</sup> · Kittikul Kovitangoon<sup>2</sup> · Donnaphat Trakulwanont<sup>2</sup>

Received: 26 October 2020 / Revised: 4 May 2021 / Accepted: 9 May 2021 / Published online: 27 May 2021  
© Springer-Verlag GmbH Germany, part of Springer Nature 2021

## Abstract

Reinforced concrete bridge substructures are one of the most important road components, requiring routine maintenance for road safety. These structures initially require visual inspection to identify and prioritize maintenance processes based on the damage severity deteriorating the structural strength of the bridge. However, owing to the limitations concerning field inspector resources and budget management efficiency, automation technology is introduced to support field inspectors for developing faster defect detection and inspection processes with high accuracy on a large scale. It has advantages in utilizing the limited number of field inspectors more efficiently and preventing the development of defect severity, which requires more time and budget resources for maintenance. This paper describes the complete development of a deep learning-based visual defect-inspection system for reinforced concrete bridge substructures. The system consists of four main components. The first part involves the image acquisition. The second part aims to detect images with defects using a modified ResNet-50 CNN, improved from our previous research. The third part is defect classification, where different types of defects, such as cracking, erosion, honeycomb, scaling, and spalling, are classified using a previous modified ResNet-50. The last part considers the severity prediction using ANN. As a result, the inspection accuracy rate for defect detection, classification accuracy, and severity prediction are 90.4%, 81%, and 78%, respectively. The promising results of this study were acceptable by Thailand's Department of Highways for practical use. Finally, this study could support Thailand's Department of Highways and surrounding construction industries as a standard strategy toward digital transformation.

**Keywords** Bridge visual defect inspection · Crack detection · Defect classification · Defect severity prediction · Convolutional neural network (CNN)

---

✉ Nagul Cooharajanone  
nagul.c@chula.ac.th  
Pravee Kruachottikul  
pravee.k@gmail.com  
Gridsada Phanomchoeng  
gridsada.p@chula.ac.th  
Thira Chavarnakul  
thira@cbs.chula.ac.th  
Kittikul Kovitangoon  
kovitang.cern@gmail.com  
Donnaphat Trakulwanont  
donnaphat.tra@gmail.com

<sup>1</sup> Technopreneurship and Innovation Management Program, Graduate School, Chulalongkorn University, Bangkok, Thailand  
<sup>2</sup> Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand  
<sup>3</sup> Wireless Communication Research Unit, Department of Mechanical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand  
<sup>4</sup> Department of Commerce, Chulalongkorn Business School, Chulalongkorn University, Bangkok, Thailand

## 1 Introduction

Digital transformation changed the ways several industries operate nowadays to maintain their competitive advantages, productivities, and efficiencies. The construction industry is among the latest fields adopting the digital transformation trend [1]. Thus, the field should migrate rapidly to digital transformation to overcome major industry challenges such as low productivity, project overruns, skilled labor shortage, and poor project management. Big data, a result of the digital transformation, should be used in the construction industry to improve project management, safety, energy management, decision-making design, and resource management [2], as well as the overall cost control and project schedule.

This research aims to assist Thailand's Department of Highways in integrating digital transformation strategies into the visual defect-inspection process for bridge substructures. One of the biggest challenges in this field is human inspection, which is unreliable as the number of skilled human resource is limited. Furthermore, it is a routine task involving considerable preparation for both inspection planning and expertized identification [3].

This study focuses on the bridge substructure because it is the most important component of the bridge requiring visual inspection as an initial step for defect detection. Visual inspection is used to verify apparent damages affecting the bridge strength, such as cracking, erosion, honeycomb, and concrete delamination (scaling and spalling). The current inspection process is shown in Fig. 1.

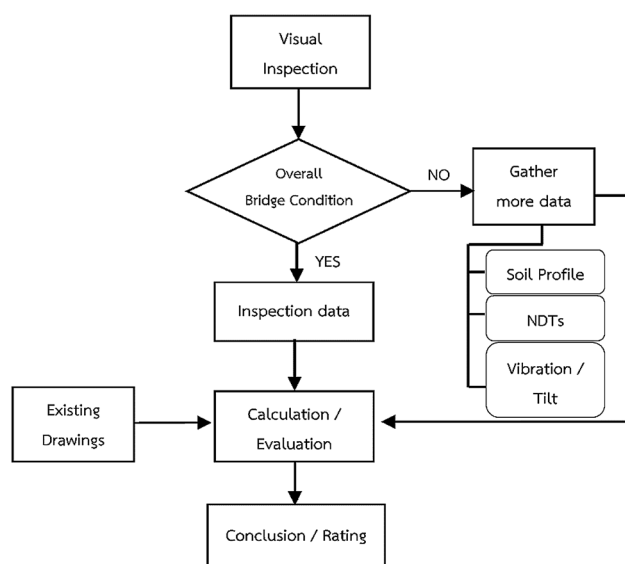


Fig. 1 Current process for bridge substructure inspection [1]

There are three types of inspection stages: basic, standard, and special. A basic inspection is the most common routine inspection in Thailand and is handled exclusively by field technicians; although a standard inspection is similar to the basic stage, it is performed under the supervision of an engineer. The limitation of performing a standard inspection is mainly due to inadequate engineering judgments. Finally, a special examination is performed when damage that might affect the bridge structure is found either in the basic or standard inspection stages. Therefore, engineers with special equipment are required to examine the bridge structure in detail, such as non-destructive testing for analyzing concrete property and bridge foundation characteristic and structure, position testing by ground penetration radar technique for determining defect size and orientation in soil or cavity in the foundation, etc.

In this study, the research approach was first used to identify the crucial factors leading to demand-driven research. That is, literature and patents were reviewed to identify essential factors. Subsequently, an in-depth interview with ten field engineers was conducted using the key findings from previous steps to identify the overall factors affecting the user adoption of this study. Finally, a quantitative analysis involving 400 questionnaires was used to confirm the earlier found factors and gain insightful information. Consequently, two key insights were obtained. First, the field inspectors require assistance tools to determine defects and provide ratings to standardize ranking judgment among the team members: technicians, engineers, and managers. Second, the tool can be used as a communication platform among field and office engineers in real time.

The advantages and innovations from this research can support Thailand in using the bridge maintenance budget efficiently each year, owing to faster visual defect identification, classification, and severity estimation. Thus, the proper maintenance and budget allocation plan can be derived to prioritize the bridge substructure that must be repaired. Moreover, it can be beneficial to Thailand in various ways: (1) to stop bridge damage from becoming severe, (2) to plan and prioritize the maintenance schedule and budget more efficiently, (3) to reduce the chances of road accidents, as Thailand is ranked number two in deaths caused by traffic accidents worldwide based on the World Health Organization record from 2018. Hence, well-maintenance road and bridge infrastructure, which is strongly linked to fatal and serious injury causation in road collisions, leads to improving overall safe road environment [4] and 4) to help Thailand's Department of Highways reduce the workload of a limited number of expert maintenance engineers.

## 2 Related works

The degradation of the material condition of reinforced concrete structure, usually resulting from weathering actions or chemical attacks, endangers the structural safety and increases maintenance and repairing costs [5]. Several studies have been conducted to investigate nondestructive evaluation techniques, including application of digital technology in bridge defect inspection in the construction industry for more precise evaluation of the reinforced concrete condition than that by the human eye [6]. Visual inspection is the first and important process of bridge defect inspection as it is a stepping-stone supporting engineers in determining whether additional specialized testing, inspections, or both are required. Currently, a limited number of engineers at the headquarter of Thailand's Department of Highways performs field inspection. Even though the limited number of engineers compared with the high number of bridges in Thailand, human decision is sometimes subjective and incomplete because it is depended on personal background, experience, and judgement. Nevertheless, its traditional human visual inspection is unavoidable. Therefore, introducing photographic evidence to assist engineers for better maintenance inspection is essential; however, developing a systematic process requires ensuring that the data collected are as consistent, accurate, and objective as those collected by field experts.

McRobbie et al. [3] proposed a prototype system to collect systematically a full image record of bridges that can be aligned and displayed for analysis. However, the prototype does not automatically perform defect detection or identification. Furthermore, a smartphone-based bridge inspection and management system to support inspectors collecting data, including images and defect descriptions, was proposed [7]. Subsequently, the use of sensors and deep learning to monitor the status of bridges was also proposed [8].

Driven by the digital transformation and big data trends, wide-ranging interdisciplinary studies considering construction industry applications using big data techniques are receiving increasing attention. Especially in the area of image-based classification [9]. However, existing studies addressed only cracks because they relied on sensor data and did not include the structure geometry [10]. Moreover, the initial studies of bridge crack inspection were commonly focused on edge detection algorithms [11]. For example, Mohan et al. [12] compared different image processing techniques such as fast Haar transform, Canny, Sobel, and fast Fourier transform. However, the limitation for image processing techniques is that they depend on rule-based algorithms; thus, artificial intelligence technology was proposed recently to overcome the limitations. Similarly, research by Moon and Kim [13] applied both image processing for image enhancement and

artificial neural network (ANN) for classifying crack and non-crack images. Bu et al. [14] proposed a machine learning algorithm based on support vector machines to classify crack and non-crack images. However, both studies present limitations in terms of the controlled environmental conditions and binary classification output. A study conducted by Kim et al. [15] proposed a machine learning algorithm to identify cracks in the presence of non-crack objects sharing similar image characteristics in terms of shape and color. The author used image binarization to identify the crack candidate region, applying a convolutional neural network (CNN) to extract the crack from non-crack objects.

Furthermore, Wang et al. [16] proposed a crack detection framework based on a CNN using a new sampling and training method inspired by active learning, including a sliding window technique, annotation, and data augmentation methods to gain the advantage of data integrity and save a large amount of annotation work. Similarly, Zhang et al. [17] presented a new model based on CNN with new loss functions for automated pixel-level crack detection. Subsequently, the authors tested the functions with open datasets. Moreover, some studies applied feature extraction from image processing and ANN to predict the depth of defects, and then, used a backpropagation neural network with previous depth information to predict a condition state, including a rating of three levels scaling defects [18]. After exploring related works, there are increasing numbers of studies that have been involved with the integration of deep learning with practical engineering application using actual experiment in bridge condition assessment for non-destructive approach domain. Even though recently there have been more number of studies that applied CNN, most of the previous works were limited to the area of crack detection with binary classification (crack or not crack). Thus, there is a research opportunity especially in the area of defect classification with more than 2 types such as crack, erosion, and concrete delamination (scaling and spalling). Furthermore, combining between defect classification and defect severity scaling can be another major contribution in this study. Additionally, this study focuses on applied AI in the field of bridge substructure. There is a real need for an implementation from Thailand Department of Highways as there is no application research exists in this area before. This application study can be a novel for Thailand and its industry.

Deep learning is a standard tool based on CNN that enables object classification. The object is classified by searching for low-level features, such as edges and curves; thus, more abstract concepts of these features are developed automatically during the training. More details regarding the basics of a CNN architecture can be found elsewhere [19]. Despite the default CNN setup, some techniques are adopted to improve the accuracy of the CNN.

### 3 Deep residual neural network

In image recognition, computer programming can solve various tasks, such as rule-based algorithms and deep learning; nevertheless, they do not provide the best performance. It is common knowledge that most results obtained using a very deep layer model tends to provide high performance and high score in some tasks, but not for other tasks. As the layers become deeper, more past information is lost. This problem is called “gradient vanishing,” which occurs when the model converges; thus, the training error degrades. The model cannot obtain the learning information to finish its task at the highest performance. Deep residual learning is the method used to solve the gradient vanishing problem [20]. By applying a shortcut connection, as layers become  $k$  steps deeper, the model recalls the input value from  $k$  previous steps to use them in the calculation as well. Deep residual learning includes the past input value as it calculates the current weight layer, ensuring that the past information is not lost. When the CNN calculates the weight in a current layer, the identify function creates a shortcut connection to obtain the input value  $x$  from previous layers. This input value  $x$  will also be added into the current layer calculation to solve the gradient vanishing problem. The new equation of the actual layer can be written as follows:

$$y = F(x) + x. \quad (1)$$

A concept of Residual block is described by Eq. (1); where  $x$  is an input of the module,  $y$  is an output, and  $F(x)$  is a function that processes the input  $x$ .

### 4 CNN models

Neural networks with more than one convolution operations are called CNN. In recent years, this approach has obtained achievements in the computer vision field. Although several scholars might consider CNN a black box feature extraction model, its advantages in image feature representation and classification are still incomparable [21]. CNN uses a feature detection layer to learn from training data. Moreover, because of the weight sharing mechanism, neurons on the same feature mapping surface have the same weight to avoid explicit feature extraction. For image classification application, the input of the CNN consists of images with numerical values in each pixel organized spatially along the width, height, and depth, which typically contains three RGB layers. The total architecture aim is to obtain a probability score of an image belonging to a certain class by learning from these numerical values arranged spatially by performing

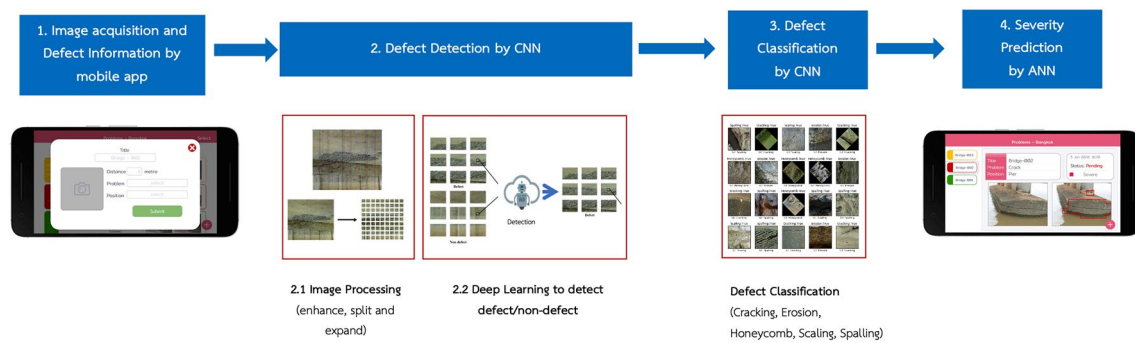
operations such as pooling and convolutions. There are many CNN methods developed in the AI communities. Each one has their strengths and weaknesses. The five famous CNN methods were chosen in this study to find the best accuracy. More details of five CNN methods and some techniques used in this study are explained in the following.

**ResNet** was proposed by He et al. [20] in 2016. It is a model that can solve the gradient vanishing problem by including the residual technique to recover or repeat the current data for a model. ResNet architecture has an identity function to receive input data from an earlier convolution layer to be combined with input data in the current convolution layer. There are several types of ResNet model based on the number of layers used, such as ResNet-18, ResNet-34, and ResNet-50.

**AlexNet** was proposed by Krizhevsky et al. [22] in 2012. The authors contributed the CNN using five convolution layers with each max-pooling layer and a fully connected layer to classify data of 1000 classes. This model was trained on ImageNet dataset with 1.2 million images and 1000 different categories. Moreover, they proposed their regularization method called “drop out,” i.e., dropping neurons in the last layer to avoid the overfitting problem.

**GoogLeNet** (or InceptionV1) was proposed by Szegedy et al. [23] in 2015. The authors proposed a method to reduce the size of the model parameters using a  $1 \times 1$  convolution layer with reducing dimensions of inputs. Thus, this technique can reduce the computation cost and help the model to increase its depth and width. Moreover, the authors proposed an inception module focusing on the spatial information of images using three types of convolution layers:  $1 \times 1$ ,  $3 \times 3$ , and  $5 \times 5$ . The reduced dimension, using  $1 \times 1$  convolution and inception module, improves resource usage and increases the depth and width of the model to achieve high performance. GoogLeNet has 22 convolution layers with max-pooling, average-pooling, dropout, and fully connected layers.

**InceptionV4 and Inception-ResNet** were proposed by Szegedy et al. [24] in 2017. Although the authors still maintain the basic concepts of the inception (or GoogLeNet) module, they use different convolution layers to extract features from different sizes of an object in an image, constructed to be uniform, reducing computation resources. Inception-ResNet includes a module called “residual layer,” which is similar to the ResNet model. It includes this computation into the inception module to reduce the gradient vanishing problem and improve the performance of the model as it becomes deeper.



**Fig. 2** Research methodology of the deep learning-based visual defect-inspection system for reinforced concrete bridge substructure

**Transfer Learning** was proposed by Yosinski et al. [25] in 2014. Sometimes teaching the model to achieve its objective requires high computing resources. If the problem requires more computation resources than the system limit, the model could not reach its maximum potential. Instead of training the neural networks from the initial stage or for every new random parameter, Yosinski et al. transfer the ability of a model that already achieves its task with high performance to a new model using the following methods:

**Fixed feature extractor** this method consists of using the model already trained on some datasets such as ImageNet and MS-COCO. Subsequently, the last fully connected layer is removed, and all parameters for the remaining part of the model are maintained constant. Next, a new fully connected layer is created that will complete a new task, and the new model is trained only on the new part of the new dataset.

**Fine tuning** this strategy is similar to the fixed feature extractor strategy; however, instead of training only the new part of the modified model, it trains the entire model to fine-tune each parameter in each layer closing to the best point.

## 5 Proposed methods and experimental setup

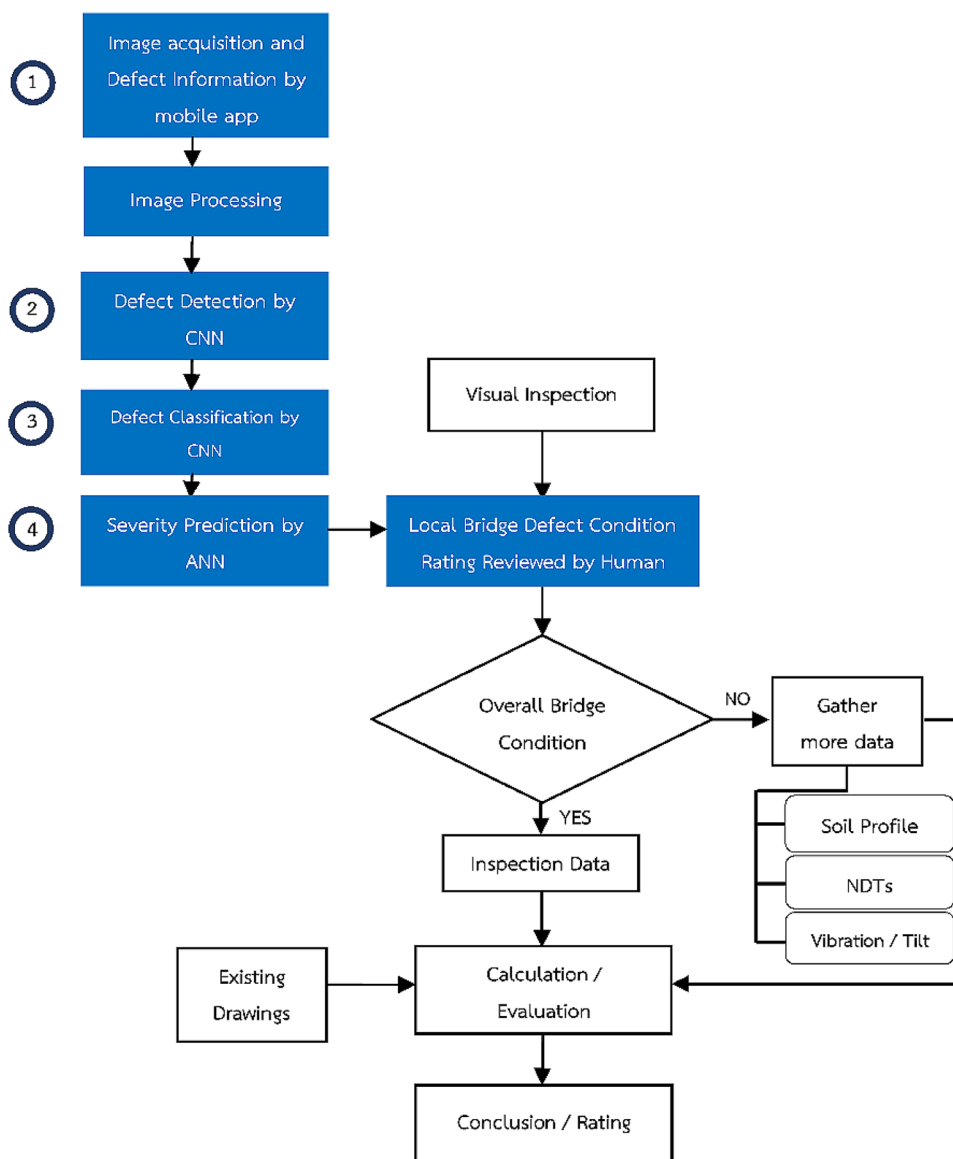
A demand-driven application is proposed in this study; specifically, a deep learning-based visual defect-inspection system for reinforced concrete bridge substructure is proposed. The system can assist the field inspectors and engineers of Thailand's Department of Highways. The proposed research methodology is shown in Fig. 2, and the modified version of the process for bridge substructure inspection of Thailand's Department of Highways is shown in Fig. 3, highlighted in blue.

The proposed system integrates four main components: image acquisition, defect detection, defect classification, and severity prediction. The first system involves the image acquisition of the bridge substructure and other data such as defect location using a mobile application. The second component involves defect detection and analysis using our modified ResNet-50 CNN model. Subsequently, image processing is performed to obtain defect information such as the coordinate of grid and pixel, defect size, the number of defects, and the number of small regions of defects. The third component aims to classify defects and determine the type of defects such as cracking, erosion, honeycomb, scaling, and spalling using our modified ResNet-50 CNN model. Finally, severity prediction using ANN is performed to predict severe and non-severe results.

## 6 Image acquisition

The first part is used for acquiring images and the required information. The image acquisition system consists of a mobile phone application developed on Android OS, as shown in Fig. 4. The application is used for the site inspectors to capture the images of the bridge substructure and input other necessary information, such as defect location and bridge location. After capturing the images, the data are uploaded to the cloud server. However, regarding the image acquisition protocol, it is recommended to take a photo within 3–8 m away from the approximated center of the defect using a camera with at least 5-MP resolution. Although several high-performance mobile phones are available in the market, a high-resolution camera is still expensive for every technician on the field. Therefore, in practice, the application is designed to be used on low cost and aged mobile devices. Moreover, when training the model for completing a task, the particular problem that scholars would

**Fig. 3** Modified inspection process of a reinforced concrete bridge substructure using a deep learning-based visual defect-inspection system

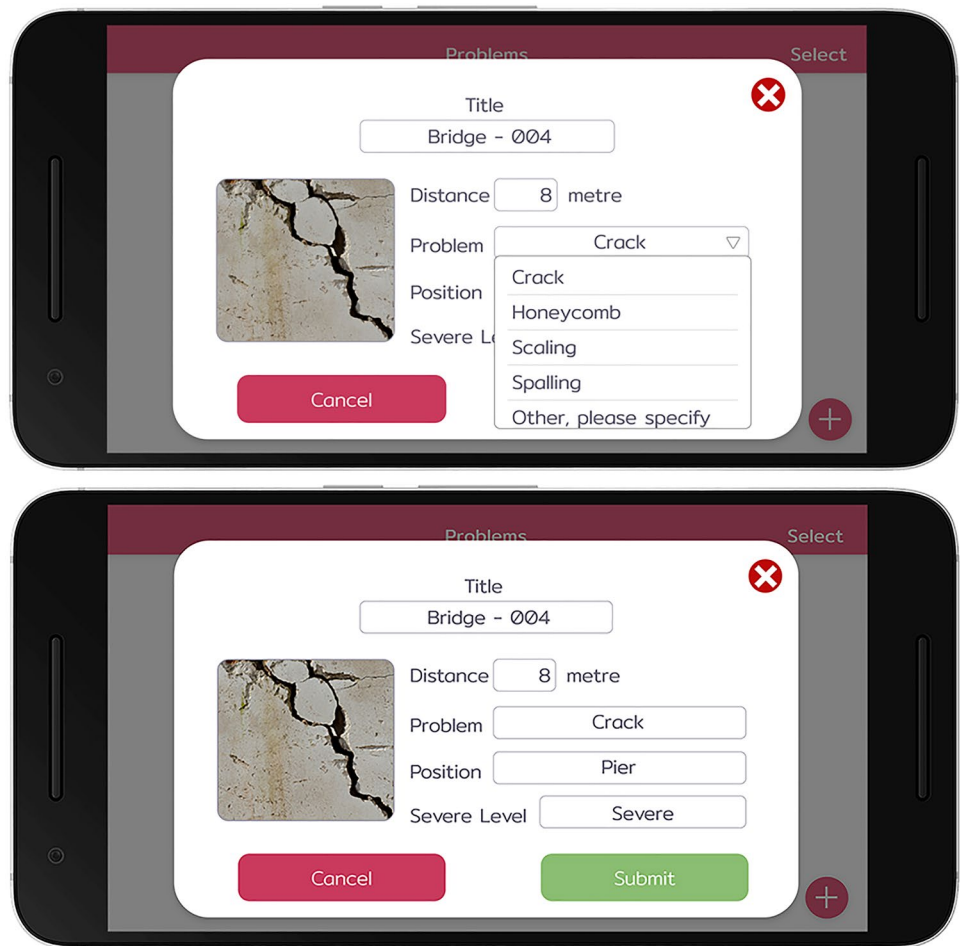


face is an overfitting problem. Overfitting is when a model functions only on a small amount and less characteristic training dataset.

Thus, data augmentation is a simple solution used to solve such a problem. [22] The method generates more data based on the existing data, using simple image processing algorithms (e.g., image rotations, adding noises, and flipping images, as shown in Fig. 5). The generated images from data augmentation are included in the training set. The model will obtain the ability to avoid the overfitting

problem. Zhang et al. [26] proposed a method to train the CNN, called “auxiliary training.” The method uses augmented images to train the auxiliary model; thus, the parameters from the auxiliary model are transferred to the primary model, enhancing robustness. Gontijo et al. [27] proposed another augmentation method, “patch Gaussian augmentation,” by adding Gaussian noise into a small square area of the original images for training the model to achieve robustness. The robustness of a CNN model refers to the ability of a model to predict the result from non-existing characteristic input.

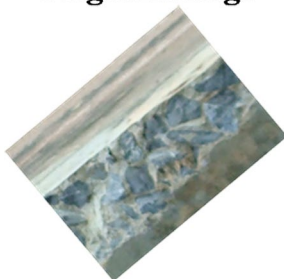
**Fig. 4** Mobile application for site inspectors to acquire image data and input the required defect information



**Original image**



**Add noise**



**Rotation**



**Flip in horizontal and vertical**

## 7 Defect detection by CNN

This process is used to identify defect and non-defect images. A total of 180 initial pictures of 5-MP (2592 × 1944 pixels) of bridge substructure images with defects, provided by the Department of Highways, were used in the process. After image enhancement and pre-processing, each 5-MP image was enhanced and filtered to increase contrast and remove noise. Each image was divided into 9 rows × 9 columns (288 × 216 pixels in each section), as shown in Fig. 6. Moreover, each section was expanded by 8% to cover details at the image edge and corner. Consequently, each 5-MP image became 81 small images with 311 × 233 pixels. Thus, the total number of images was 14,580. Finally, environmental images such as water, soil, and trees were removed; therefore, the final data consisted of 3618 bridge images with and without defects with a resolution of 311 × 233 pixels.

**Fig. 5** Examples of data augmentation



**Fig. 6** Examples of image partition for defect detection

The previous research was developed based on HALCON software, which used CNN's modified ResNet model [19]. However, improving the accuracy of the model was challenging owing to a lack of information regarding the type of ResNet and parameters used in the software. Therefore, in this study, the system overview was designed as shown in Fig. 7. Moreover, our modified ResNet-50 model, as shown in Fig. 8, was selected and developed in Pytorch to replace the HALCON software.

After the defects were detected, an image processing algorithm was used to obtain further defection information by combining neighborhood defected sub-images and assigning a group number to each combined defect images. Subsequently, defect information, such as the coordinate in grid and pixel and the number of small regions of defects, was calculated.



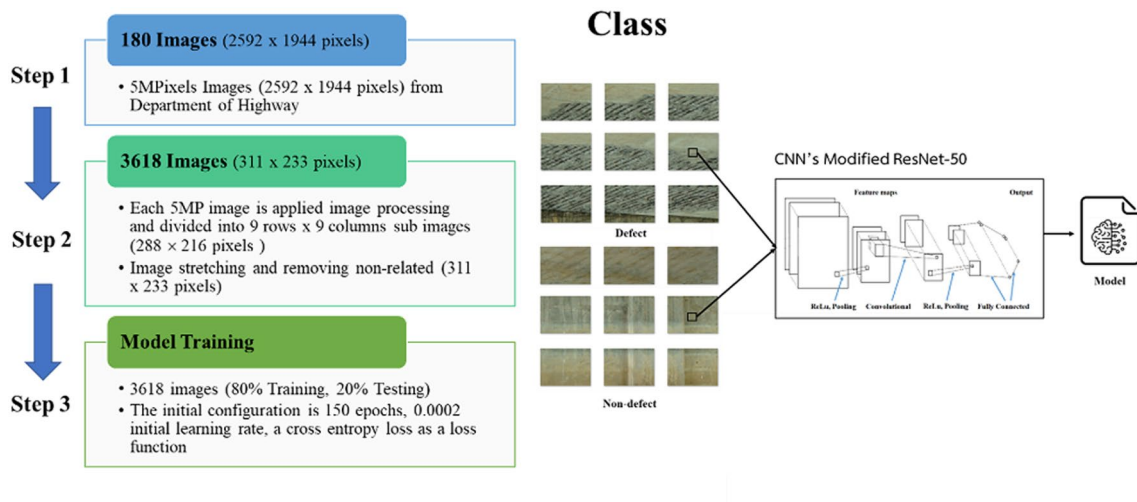


Fig. 7 Defect detection system overview

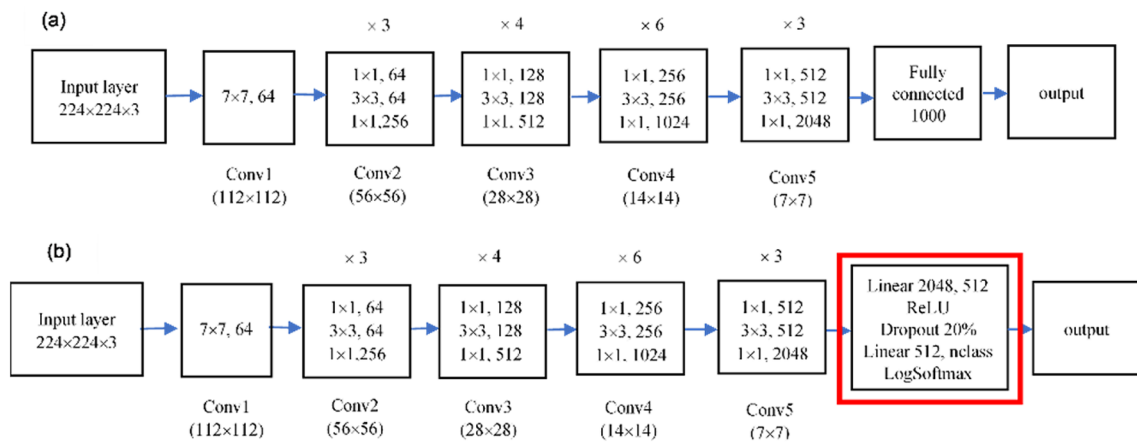


Fig. 8 a Example of the default ResNet-50 CNN architecture, b Our Modified ResNet-50 CNN that are replaced by the last fully connected layer with linear layer, ReLU layer, 20% dropout layer, another linear layer, and LogSoftmax layer

### 8 Defect classification by CNN

This module was used to automate the classification of five types of defects. At the beginning of this study, there was uncertainty in which CNN models would be the most suitable for analyzing the dataset. Thus, five models, namely, ResNet-50, AlexNet, GoogLeNet, InceptionV4, and InceptionResNetV2, were selected for an evaluation owing to their reputation based on their accuracy and operational performance [28]. As a result, ResNet-50 was selected as a base classifier due to its highest accuracy among five selected models. After that

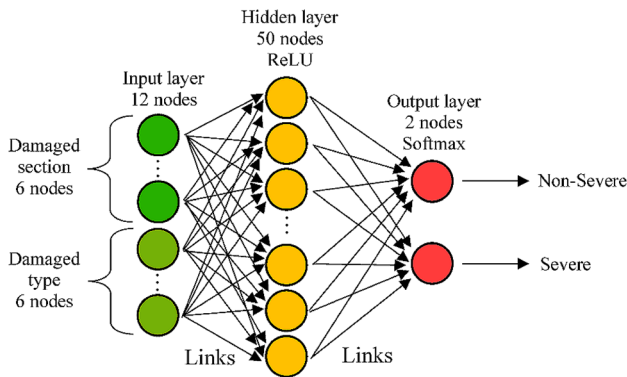
the algorithm setup for ResNet-50 was modified at the fully connected layer, for transfer learning purpose, by substituting with a linear layer, ReLU layer, 20% dropout layer, another linear layer, and LogSoftmax layer. The default setup and modified setup used in this study are shown in Fig. 8. Experts from Thailand’s Department of Highways verified and labeled the initial 3618 images as 2270 non-defect images and 1348 defect images. They also classified the defects into five categories, namely, cracking, erosion, honeycomb, scaling, and spalling with 97, 219, 253, 117, and 540 images, respectively. As the number of images was not equally distributed between

**Table 1** Ranking definitions of the bridge damage (Thailand Department of Highways Maintenance Manual, 2008)

Ranking	Definition
5	Good condition, no need to repair
4	Good condition with minor damage to the strength of the structure
3	Moderate condition, but found some damage requiring preventive maintenance
2	Moderate condition with some damage requiring immediate repair
1	Poor condition, but found some damage to the structure requiring urgent reinforcement
0	Severe condition. Bridge replacement is required

**Table 2** New ranking definitions of bridge damage for preliminary studies

Ranking	Definition
Non-severe	Does not require immediate attention
Severe	Requires immediate attention



**Fig. 9** System architecture of severity prediction based on ANN

all categories, image augmentation, as mentioned earlier, was applied to solve this problem. In this regard, images were rotated by 90°, 180°, 270°, and 135° to increase the number of images in each category, except for the spalling, to obtain approximately 500 images for each category. However, 135° rotation increases the black areas around an image. Thus, these images were cut and modified to reduce the black areas. The final dataset contained 485 cracking, 438 erosion, 506 honeycomb, 585 scaling, and 540 spalling images.

### 9 Severity prediction by ANN

An important aspect, required by Thailand’s Department of Highways, is to develop a severity prediction system. Owing to the reduced number of inspection experts and limited maintenance budget, the prediction system is used to plan and prioritize resources and efficiently allocate maintenance budget. The manual severity-ranking system currently includes five ratings, as listed in Table 1, where an inspection expert makes the decision based on each image. However, sometimes even among experts,



**Fig. 10** Complete inspection system user interface. **a** login page, **b** main menu page, **c** image acquisition and defect information page, **d** acquired bridge substructure image and defect information with

predicted severity page, **e** defect information with predicted severity waiting for review page, **f** defect information with predicted severity after review page

indecisiveness and contradicting judgments can occur because of the overwhelming number of images inspected and different experiences. In this regard, a standardized ranking system is developed based on ANN to eliminate such inconveniences. As a small number of images (only 180 defect images) can be currently used, the outputs of the severity-ranking system are initially developed with a two rating format, severe and non-severe, as shown in Table 2, to study the possibility of a standardized ranking system. The setup for this study, as shown in Fig. 9, was performed using Keras [29], which is a Python library framework for developing and evaluating AI models. It is an interface for TensorFlow library and also contains the standard building blocks used by neural network such as layers, activation functions, etc. Keras is the AI tool to make working with AI easier and simplifying the coding necessary. Next, twelve inputs identified by inspection experts were categorized into two groups as follows.

- The bridge substructure parts, i.e., parapets, girders and beams, deck bottom surfaces, cap beams, piers, and footings.
- The defect types, i.e., cracking, efflorescence, erosion, honeycomb, scaling, and spalling.

## 10 Complete inspection system

All four main components are integrated into the inspection system named deep learning-based visual defect-inspection system for reinforced concrete bridge substructure. The user experience and user interface (UX/UI) are primarily designed to maximize user adoption of the application, as shown in Fig. 10, based on the key findings obtained from in-depth interviews, as well as from the real activity observation with expert engineers and field inspectors. Using the mobile application, field inspectors can login with their accounts, and subsequently, capture a new image using the mobile phone camera or select an image recorded in the mobile phone. The field inspectors also provide basic information such as bridge name and defect location. All data are transferred to a cloud system for improving the CNN and ANN models or automating defect classification in the future. The early architecture, as shown in Fig. 11a, is drafted on Amazon Web Service (AWS) [30].

Microservice architecture is selected because it collects small services that are straightforward for managing the complication of an application and adding more services required in the future. The UI of the application is created using the React Native framework that communicates with the back-end software using AWS Amplify via Amazon API

Gateway. The developed image processing software, CNN and ANN, are executed on AWS Lambdas, where software and Amazon S3 storage communication are performed through a programming interface or API. All microservices are supervised by AWS IAM and Amazon Cloudwatch to manage accessibility and resources efficiently.

The sequence diagram, as shown in Fig. 11b, is used to explain how microservices work at the backend, starting from the left when an image and defect information are initially uploaded by the user. Then the original picture is stored in S3 Bucket whereas the defect information is recorded on Image Table. After that the program processes image processing, defect detection, defect classification, and severity prediction. The final result is stored on Image Table while the processed image is saved into S3 Bucket. However, if there is no defect found, the program will stop.

## 11 Results

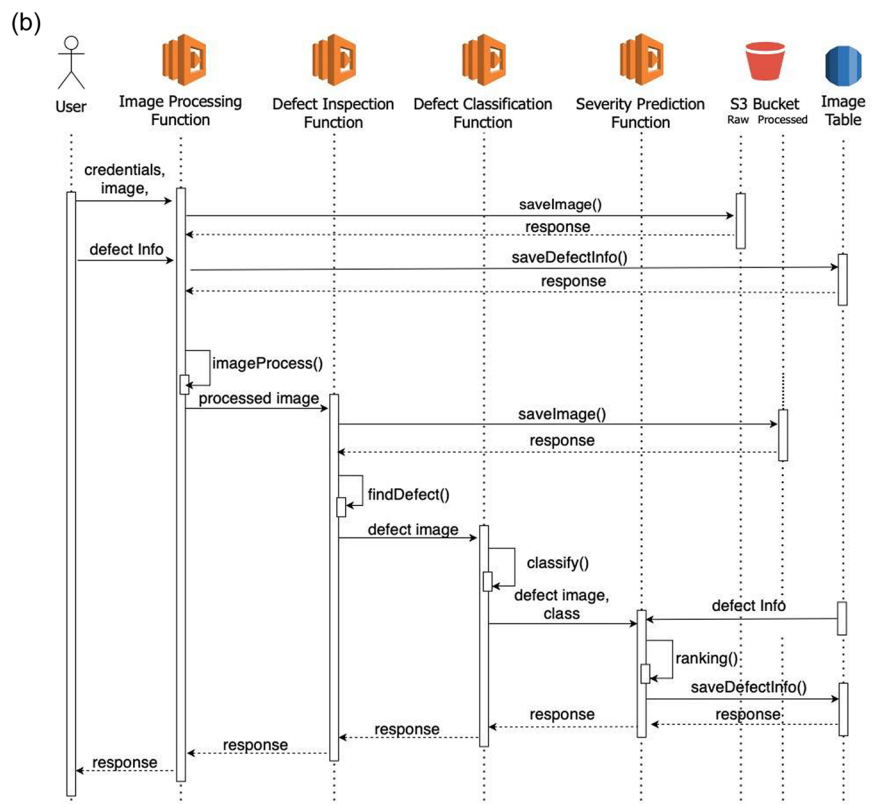
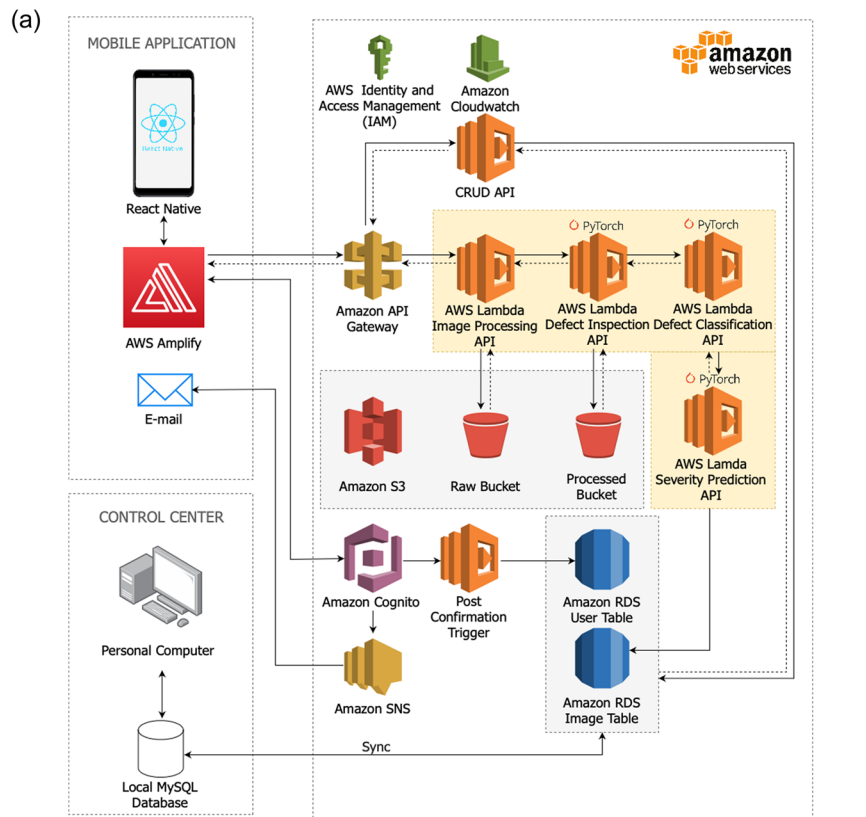
This section presents the result of each research component of the proposed deep learning-based visual defect-inspection system for reinforced concrete bridge substructure.

First, the defect detection part was experimentally evaluated. In this regard, the initial configuration was performed with 150 epochs, 0.0002 initial learning rate, a cross-entropy loss as a loss function, and dividing 3618 images into two groups, 80% for training and 20% for testing. Consequently, an accuracy of 90.4% was obtained, calculated from the confusion matrix shown in Fig. 12. The obtained result improved slightly compared to that from previous research [19], i.e., 89.3%. In this part, the research goal of 90% accuracy has been reached as required by Thailand's Department of Highways.

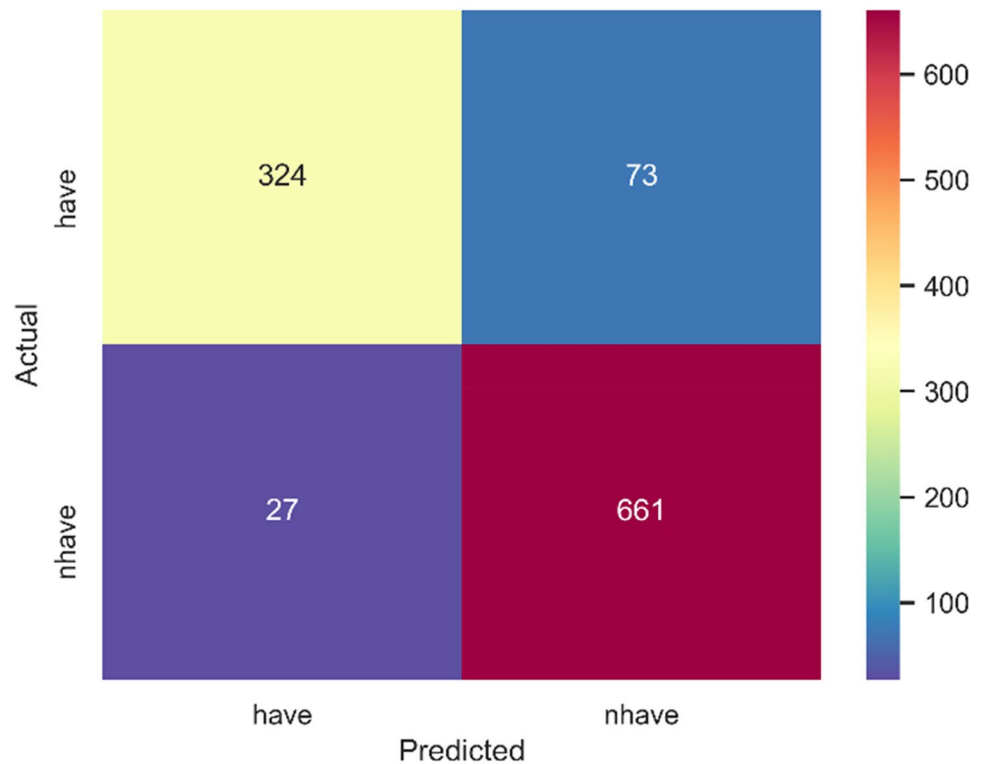
Second, the defect classification module was experimentally evaluated. After dataset preparation, 70% of the initial images, 858 images, were used for training, while 30% of initial images, 368 images, were used for validation. The increase of validation dataset portion was due to the number of available image dataset was limited, and the classification challenge required more validation data for higher model accuracy. After that, each CNN model was applied to classify the type of defect with the training configuration set to 100 epoch, 0.0002 initial learning rate, loss function of cross-entropy loss, and using Adam optimizer. Regarding the learning rate, the weight decay was set to be equal to 0.0005 with weight decay value and training with 200 epochs. The obtained outcome for each model is listed in Table 3.

Based on the results listed in the table, our modified ResNet-50 was selected as the base classifier for transfer learning owing to its highest accuracy among the entire

**Fig. 11** **a** System architecture. **b** Sequence diagram of deep learning-based visual defect-inspection system for reinforced concrete bridge substructure



**Fig. 12** Confusion matrix of defect detection using our modified ResNet-50, where “nhave” denotes no defect found in the image, whereas “have” means defect found in the image



**Table 3** Comparison results of each CNN model for defect classification

Valid Acc. (%)	Our Modified ResNet-50	ResNet-50	AlexNet	GoogLeNet	InceptionV4	Inception ResNetV2
Mean	81.1	77.2	69.8	70.2	63.5	72.3
Max	85.3	80.8	72.2	73.5	66.3	75.9

**Table 4** Accuracy of the proposed modified ResNet-50 with transfer learning for defect classification

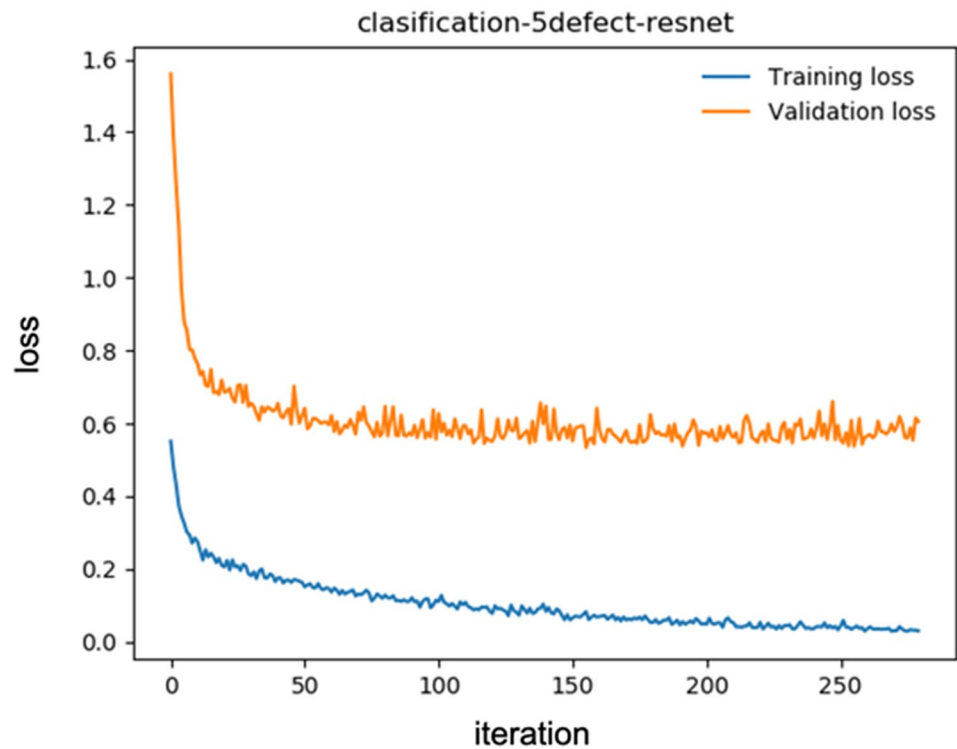
	Average accuracy	Maximum accuracy
Proposed model	81.1%	85.3%

selected CNN models. Moreover, the results, shown in Table 4, were obtained using the experimental setup.

Figure 13 depicts a graph regarding the loss for training and validation whereas Fig. 14 shows defect classification results in 5 defect categories. The results validate the model shown by the confusion matrix presented in Fig. 15 as the validation loss graph expressed the accuracy difference between the training images and validation images. The trial and error method were used to optimize the model so that the losses were stabilized.

Next, the severity prediction part was experimentally evaluated. The experiment considered a network with twelve variables in an input layer, one hidden layer with 50 nodes using rectified linear unit (ReLU), as activation function. ReLU gives the output only if the input is positive number, otherwise it will be zero. It is one for the default activation function used by many types of neural networks [31]. Moreover, two nodes of outputs using Softmax were connected as activation function. The initial configuration was set to 2000 epochs, 0.01 initial learning rate, a cross-entropy loss as a loss function, and dividing 180 images into two groups, 80% for training and 20% for testing. Although a small dataset was used, a preliminary accuracy of 78% was obtained, as shown in Table 5.

**Fig. 13** Training and validation losses of defect classification



**Table 5** Results of preliminary study of severity prediction system

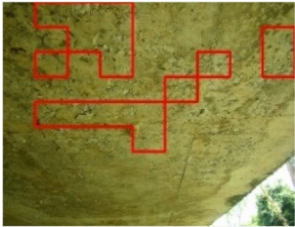
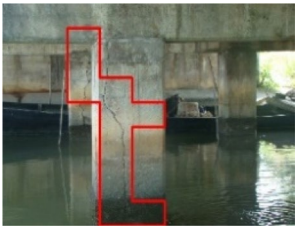
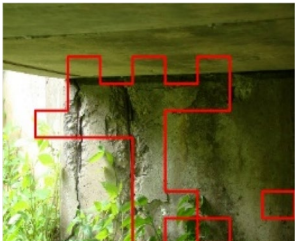
ANN results	Actual results	
	Severe	Non-severe
Severe	23	8
Non-severe	0	6

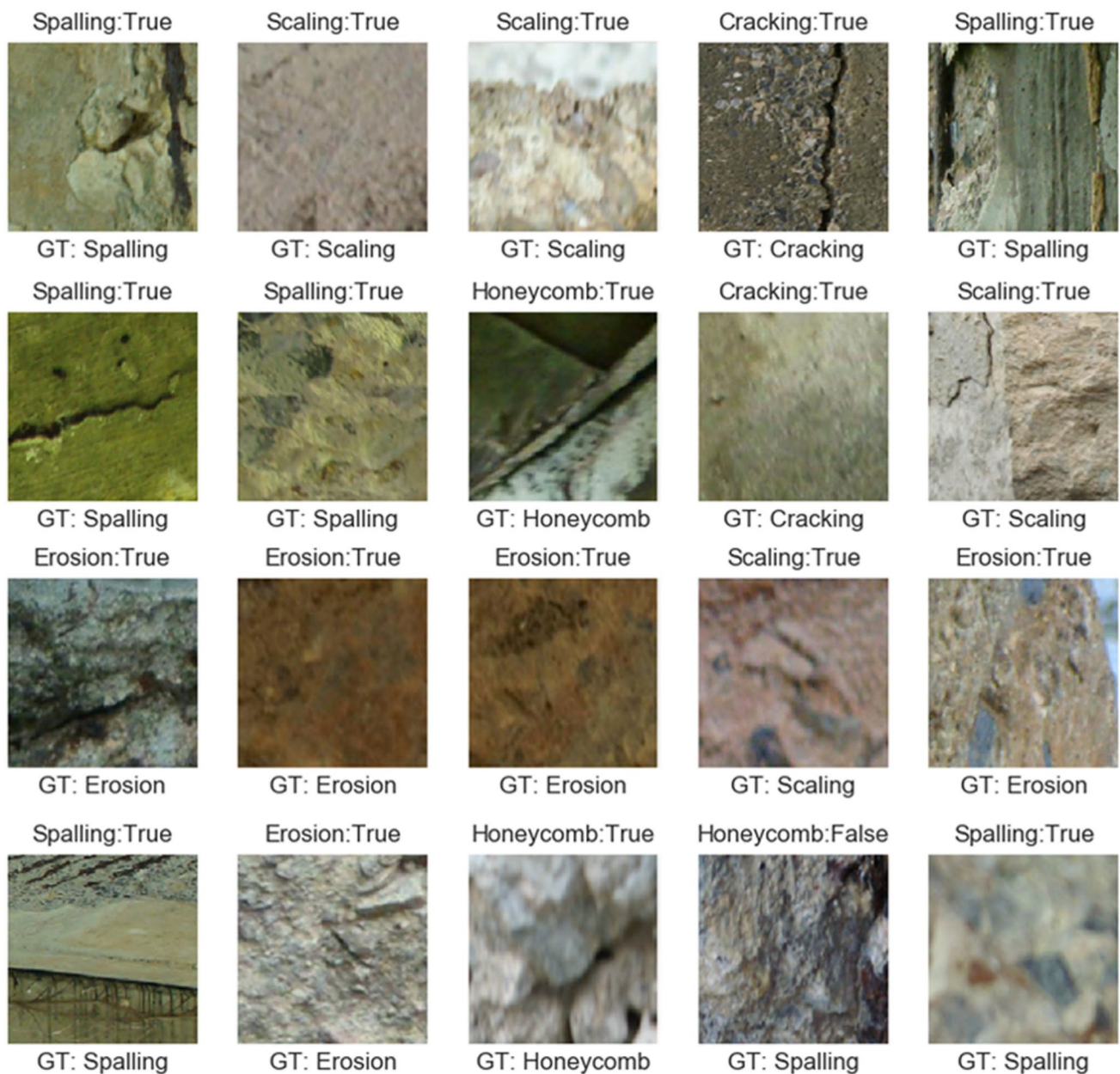
To combine total system, the results were shown in Table 6. The proposed framework and overall accuracy are acceptable by Thailand’s Department of Highways in terms of expanding on-site usage on the field, assisting on-site inspectors. Moreover, the promising results open the possibility of developing a complete automated severity prediction system when more bridge data from all over Thailand are available in the future.

## 12 Conclusion

The study aimed to assist Thailand’s Department of Highways in the digital transformation of the visual inspection processes for reinforced concrete bridge substructure by proposing a deep learning-based visual defect-inspection system. In this study, a system was developed that comprised four components: image acquisition, defect detection, defect classification, and severity prediction.

**Table 6** Results of complete system

Input	Location	Soffit	
Output	Defect type	Scaling	
Output	Severity	Non-severe	
Input	Location	Pier	
Output	Defect type	Cracking	
Output	Severity	Severe	
Input	Location	Abutment	
Output	Defect type	Spalling	
Output	Severity	Severe	



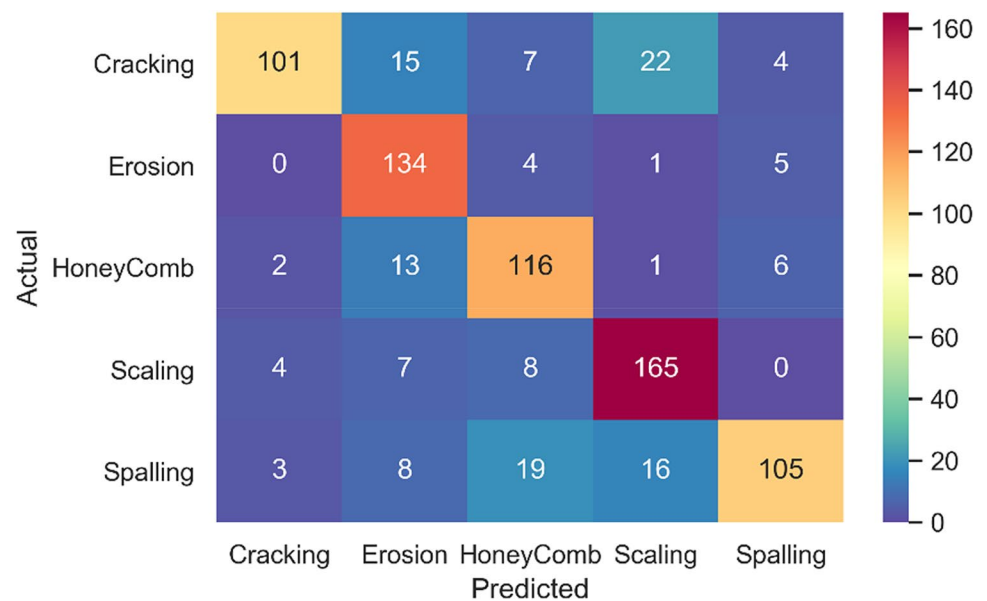
**Fig. 14** Examples of defect classification results

A total of 180 images with 5-MP resolution were collected in the southern part of Thailand to prepare an image dataset. After the image processing, the number of images increased to 3618 with a resolution of  $311 \times 233$  pixels. Consequently, the defect detection algorithm based on our modified ResNet-50 CNN model improved the accuracy from 89.3%, in our previous research<sup>19</sup>, to 90.4%. Moreover, the defect classification accuracy of the proposed method was 81.1% based on our modified ResNet-50 with five defect categories: cracking, erosion, honeycomb, scaling, and spalling. The last part was

severity prediction, a network with 12 variables in an input layer and a hidden layer with 50 nodes using ReLU.

This study has the potential of minimizing Thailand's bridge maintenance budget each year by prioritizing the maintenance schedule and budget and reducing the possibilities of road accidents in Thailand. Additionally, it can be beneficial to Thailand's Department of Highways as the proposed approach can be used as an assistance tool among team members for standardizing ranking judgment or as a communication platform for better project management.

**Fig. 15** Confusion matrix of defect classification



### 13 Limitation and future works

In this study, the number of bridge image datasets is limited because Thailand's Department of Highways is working on a new database system. Meanwhile, new cloud-based data acquisition is developed to collect images for an extension of this study and temporarily store data. The collaboration among various branches of Thailand's Department of Highways is established to collect nationwide bridge data, ensuring that this research will benefit Thailand's transportation system in integrating digital transformation to the organization. In addition, severity prediction in this research is still limited to a binary output; some ideas to increase the number of severity scaling to five levels, as well as its accuracy, will be considered in future works by including additional input. Furthermore, the idea of applying environmental data such as humidity and temperature for more accurate estimation of the bridge damage mechanisms will be considered in the future.

**Acknowledgements** The writers would like to thank the Bureau of Road Research and Development of Thailand, Department of Highways, Ministry of Transportation for assisting the required information, image dataset, and knowledge to complete this research. Moreover, we want to express our gratitude to Solitech Innovations and Technologies Co., Ltd., for supporting through the software and hardware used in this research.

**Funding** The authors received no financial support for the research, authorship, and/or publication of this article.

### Declarations

**Conflict of interest** The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### References

- Sullivan G (2018) Global digital transformation on the building and construction sector, Forecast to 2025. North America: Frost & Sullivan.
- Ismail SA, Bandi S, Maaz ZN (2018) An appraisal into the potential application of big data in the construction industry. *Int J Built Environ Sustain*. <https://doi.org/10.11113/ijbes.v5.n2.274>.
- McRobbie SG, Wright MA, Chan A (2015) Can technology improve routine visual bridge inspections? *Proc Inst Civil Eng*. 168:197–207. <https://doi.org/10.1680/jbren.12.00022>
- Global status report on road safety 2018: Summary (2018) World Health Organization.
- Brysse D (2010) Non-destructive evaluation of reinforced concrete structures. Woodhead Publishing, Sawston, pp 28–56
- Final inspection report of damage and load testing for slab type bridge in the area of Highway Bureau 5 (2008) Khon Kaen Province, Thailand: Thailand Department of Highway.
- Hüthwohl P, Lu R and Brilakis I (2011) Challenges of bridge maintenance inspection. In: the 16th international conference on computing in civil and building engineering, Osaka, Japan
- Ding Y, Han R, Liu H, et al. Bridge inspection and management system based on smart phone. In: Volume 1: Multifunctional Materials; Mechanics and Behavior of Active Materials; Integrated System Design and Implementation; Structural Health Monitoring. 2016.



9. Weinstein JC, Sanayei M, Brenner BR (2018) Bridge damage identification using artificial neural networks. *J Bridge Eng.* [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0001302](https://doi.org/10.1061/(ASCE)BE.1943-5592.0001302)
10. Bilal M, Oyedele LO, Qadir J et al (2016) Big data in the construction industry: a review of present status, opportunities, and future trends. *Adv Eng Inform* 30:500–521. <https://doi.org/10.1016/j.aei.2016.07.001>
11. Abdel-Qader I, Abudayyeh O, Kelly ME (2003) Analysis of edge-detection techniques for crack identification in bridges. *J Comput Civil Eng* 17:255–263. [https://doi.org/10.1061/\(ASCE\)0887-3801\(2003\)17:4\(255\)](https://doi.org/10.1061/(ASCE)0887-3801(2003)17:4(255))
12. Mohan A, Poobal S (2018) Crack detection using image processing: a critical review and analysis. *Alexandria Eng J* 57:787–798. <https://doi.org/10.1016/j.aej.2017.01.020>
13. Moon, HG, & Kim, JH. Intelligent Crack Detection Algorithm on the Concrete Crack Image Using Neural Network. In: 28th International Symposium on Automation and Robotics in Construction (ISARC) Republic of Korea, 2011.
14. Bu GP, Chanda S, Guan H et al (2015) Crack detection using a texture analysis-based technique for visual bridge inspection. *Electron J Struct Eng* 14(1):41–48
15. Kim H, Ahn E, Shin M et al (2018) Crack and noncrack classification from concrete surface images using machine learning. *Struct Health Monit* 18:725–738. <https://doi.org/10.1177/1475921718768747>
16. Wang Z, Xu G, Ding Y et al (2020) A vision-based active learning convolutional neural network model for concrete surface crack detection. *Adv Struct Eng* 23:2952–2964. <https://doi.org/10.1177/1369433220924792>
17. Zhang L, Shen J, Zhu B (2020) A research on an improved Unet-based concrete crack detection algorithm. *Struct Health Monit.* <https://doi.org/10.1177/1475921720940068>
18. Adhikari RS, Moselhi O and Bagchi A (2012) Automated prediction of condition state rating in bridge inspection. ISG\*ISARC2012 Full paper proceedings
19. Kruachottikul P, Cooharajanone N, Phanomchoeng G, et al (2019) Bridge sub structure defect inspection assistance by using deep learning. In: 10th international conference on awareness science and technology (iCAST). <https://doi.org/10.1109/ICAwST.2019.8923507>
20. He K, Zhang X, Ren S, et al (2016) Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR) Las Vegas, NV
21. Xin M, Wang Y (2019) Research on image classification model based on deep convolution neural network. *EURASIP J Image Video Proc.* <https://doi.org/10.1186/s13640-019-0417-8>
22. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. pp. 1097–1105.
23. Szegedy C et al (2015) Going deeper with convolutions. In: *The IEEE conference on computer vision and pattern recognition*
24. Szegedy C et al (2017) Inception-v4, Inception-ResNet and the impact of residual connections on learning. In: *The 31st AAAI conference on artificial intelligence*
25. Yosinski J et al. How transferable are features in deep neural networks? In: *Advances in Neural Information Processing Systems* 2014.
26. Zhang L et al (2020) Auxiliary training: towards accurate and robust models. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*
27. Gontijo LR, Yin D, Poole B, et al (2019) Improving robustness without sacrificing accuracy with patch Gaussian augmentation. [arXiv:1906.02611](https://arxiv.org/abs/1906.02611).
28. Canziani A, Paszke A, Culurciello E (2016) An analysis of deep neural network models for practical applications. [arXiv:1605.07678](https://arxiv.org/abs/1605.07678).
29. Chollet F (2020) Keras. <https://keras.io>. Accessed 5 Aug 2020
30. Amazon.com I (2020) Amazon Web Service. <https://aws.amazon.com/>. Accessed 12 Aug 2020.
31. Agarap, Abien Fred (2018) Deep learning using rectified linear units (relu). [arXiv:1803.08375](https://arxiv.org/abs/1803.08375).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.