Journal of Engineering Science and Technology
Vol. 16, No. 4 (2021) 3057 – 3071

TILE DAMAGE DETECTION IN TEMPLE FACADE VIA CONVOLUTIONAL NEURAL NETWORKS

KRISADA CHAIYASARN*, APICHAT BUATIK

Department of Civil Engineering, Thammasat School of Engineering, Thammasat University, 99 Moo 18 Khlong Nueng, Khlong Luang, Pathum Thani 12121, Thailand
*Corresponding Author: ckrisada@engr.tu.ac.th

Abstract

Many tiles on facade of historic temples are damaged and deteriorated due to aging and environmental factors. Regular inspection is required for proper maintenance and automatic inspection offers a significant advantage over manual inspection as it is efficient and accurate. Many previous studies are focus on detecting damages in factory tiles for the quality control purposes, although there has not been much study on tile damages for historical temples. This paper proposed an image-based system to detect damages in tiles using Convolutional Neural Networks (CNN) on a temple facade. The dataset was created by an Unmanned Aerial Vehicle (UAV) and a digital camera from a historical temple in Bangkok, Thailand. In the proposed work, CNN was trained on various image patches sizes. The detection accuracy of the system was found to be 95% in the validation data and 91% on the testing data. The results of the proposed system were compared with a system using handcrafted features, including 2D wavelet transforms with the Artificial Neural Network (ANN). The proposed system shows that the CNN approach is more accurate than the traditional handcrafted method.

Keywords: Computer vision, Convolutional neural network, Historical temple, Tile's damage detection, Unmanned aerial vehicle.

1. Introduction

Condition assessment plays a vital role in maintaining structural health and reliability for historical structures. Discoloration and defects are signs of degradation in tiles, which are commonly used on a temple facade. The detection of defects in structures can help to prevent further damages. Visual inspection is a common procedure for identifying the first sign of problems in structure, and it is normally carried out by a team of experts. However, this procedure is time-consuming, expensive, requires an expert knowledge, and is normally prone to human error. Furthermore, some parts of structures are not easily accessible by human inspectors, therefore; inspection cannot be carried out in these places. Image-based inspection methods aided by Unmanned Aerial Vehicles can help carrying out the visual inspection task in this age of disruptive technologies. Unmanned Aerial Vehicles (UAVs) can be used to collect images from the areas which are not easily accessible by human, and then damages in the collected images can then be detected by various crack detection algorithms using handcrafted features and classifiers, or by deep learning techniques.

Generally, damage detection systems consist of two main steps, feature extraction and classification. In the first step, features are extracted from input images using various techniques such as Gaussian filters [1], edge detection [1], and morphological operation [2]. Then, in the second step, classifiers such as Support Vector Machine (SVM) [3], Random Forest (RF) [4] and Neural Networks [5] are used to classify the extracted features. However, these techniques are often affected by noise, distortion and lighting variation. These conditions are limiting factors when working with computer vision systems based on these methods. Feature extraction from images containing complex structures as shown in Fig. 1 is usually non-trivial and therefore, the crack detection problem remains a challenging problem, which requires robust and industry-ready solutions. Therefore, in this research work, Convolutional Neural Network (CNN) is applied to detect damages in ceramic tiles from a historical site. CNN is an attractive method due to its ability to learn features automatically from input images, and in many applications, it has proven to be somewhat suitable techniques for dealing with classification problems.

For the tile damage detection problem, many early works are focus on detecting tile damages in factories for the quality control purposes. However, these algorithms are not suitable for the tiles in historical temples as the images or videos acquired from the temples are from the uncontrollable environment, unlike the factory tiles, which are usually performed in the controlled laboratory environment. Therefore, this paper is focus on developing a CNN-based damage detection system to attempt to detect damages in tiles on a temple façade offering a new paradigm shift for the assessment of such valuable structures, like these historic temples. The data used in this work is obtained by an Unmanned Aerial Vehicle (UAV) and a digital camera. The example images of the temple are shown in Fig. 1. The rest of the paper is organized as follows, Section 2 describes related works in tile damage detection, Sections 3 and 4 explains the proposed methodology and experiments, respectively. Finally, discussion and conclusion are drawn in Sections 5 and 6.

Fig. 1. Investigated shapes of projectiles (Geometry and dimensions).
Pictures of a temple and the tiles façade:
(a) Top left and right: The overview image of the temple.
(b) Bottom left: A close-up image of the temple façade.
(c) Bottom right: a close-up image of the tile façade on the temple.

2. Literature Review

Tile decoration on temple facades not only provides aesthetic appearance, but also protects the concrete surface beneath [6]. Tile colours, shapes and patterns must remain intact and undamaged to fulfil their intended use and regular maintenance are generally required. Ar et al. [7] applied a Gabor filter and connected component analysis to extract patterns on marble tiles, which were then used for classifying different types of tile textures. Najafabadi and Pourghassem [8] used an image-based technique to obtain corners in ceramic tiles and the tiles, whose corner angles are between 92 degree and 89 degrees were classified as non-defective, otherwise they were considered as defective. Ghazvini et al [9] applied a rotation invariant wavelet transform, statistical features and an Artificial Neural Network (ANN) classifier for tiles defect detection. The statistical features included the median of maximum points, minimum points, mean, and standard deviation, and the average accuracy was 90%, and the proposed method was faster than other previous methods.

Hanzaei et al. [10] proposed a tile defect detection system with the accuracy of 93% using Rotation Invariant of Local Variance (RIMLV), Median filtering, morphological operations and Support Vector Machine (SVM). Thomas et al. [11] proposed a photogrammetry technique, which is in the form of image-based rendering (IBR), to create a 3D models for post-processing. The result showed that the rendering method can create a new way of inspection by the image recognition technique. Samarawickrama et al. [12] proposed an image-based defect detection technique in ceramic tiles using colour, edges, and morphological operation to

separate cracked tiles, corner-damaged tiles, and discoloured tiles from good tiles. The proposed method claimed to have the detection accuracy up to 96% with the detection rate of 2 seconds per tile. However, the proposed techniques were still largely affected by colour variation in images, scale changes, and viewpoints.

Convolutional Neural Network (CNN) for crack detection can be an alternative method for detecting tile defects. CNN can classify many classes and therefore, it became popular for the image classification problems [13]. Zhang et al. [14] used a deep CNN for automatic crack detection in road images. In this work, the authors used 500 images collected using a low-cost smartphone for classification. The results by the proposed CNN method offered a superior performance than the handcrafted feature methods. Cha et al. [15] applied a deep convolutional neural network for concrete crack detection. The authors used 40 thousand images of size 256x256 for training and testing the system, which gave 98% accuracy. The proposed CNN was then compared with the traditional Canny and Sobel edge detection methods. The proposed CNN showed a much better performance than the other methods. Recently, Faster region-based convolutional neural network (Faster R-CNN) and Mask R-CNN were implemented for the detection of damage in tile. Mask R-CNN has higher capability to detect damages of the tiles with complex geometric features [16]. The MCuePush U-Net model was used to detect magnetic tile surface defects. The proposed model was able to detect multiple surface defects in low-contrast images [17]. Faster R-CNN was proposed to detect the damage in the historic masonry buildings. The deep learning model was able to detect damages with small proportion of errors, which was particularly sensitive to light [18]. A deep fully convolution neural network was implemented to detect pixel-level cracks in concrete structure [19]. Another studied was carried to detect concrete crack using the deep convolutional network with VGG 16, Inception V3, and ResNet as classifiers. The result from the VGG 16 outperform the remaining classifiers to detect cracks [20]. The Mask-RCNN with Inception-ResNet-V2 performs best for the crack detection and segmentation problems [21].

3. Methods

The outline of the proposed system consists of the modules as shown in Fig. 2(a). The images are acquired by an Unmanned Aerial Vehicle (UAV) or by a DSLR camera. After the data acquisition step, a CNN algorithm is used for detecting damages in ceramic tiles on a temple facade. The architecture of the proposed damage detection system is shown in Fig. 2(b). The detail of each module in the proposed system is explained below.

3.1. Image acquisition

A UAV has been used for surveying as they are simple to use, fast and cheap to acquire image or video data. The main advantage is that it can obtain images from high altitude, which is not accessible easily by human. In this paper, a UAV is used to acquire images from the façade of a historic temple using a DJI Phantom 4 drone. The drone was manually flown along the temple's facade to collect images. Sample images taken from the drone are shown in Fig. 3. Some images are also taken by a DSLR camera.

The images collected using the UAV and DSLR are converted into patches of various sizes. The total of 3600 images are used in this work, in which 2880 are used for training and validation, while the remaining 720 image patches are used

to test the proposed system. The damage and non-damage tiles are manually labelled as 0 and 1, respectively. The samples of damage and non-damage image patches are shown in Fig. 4.

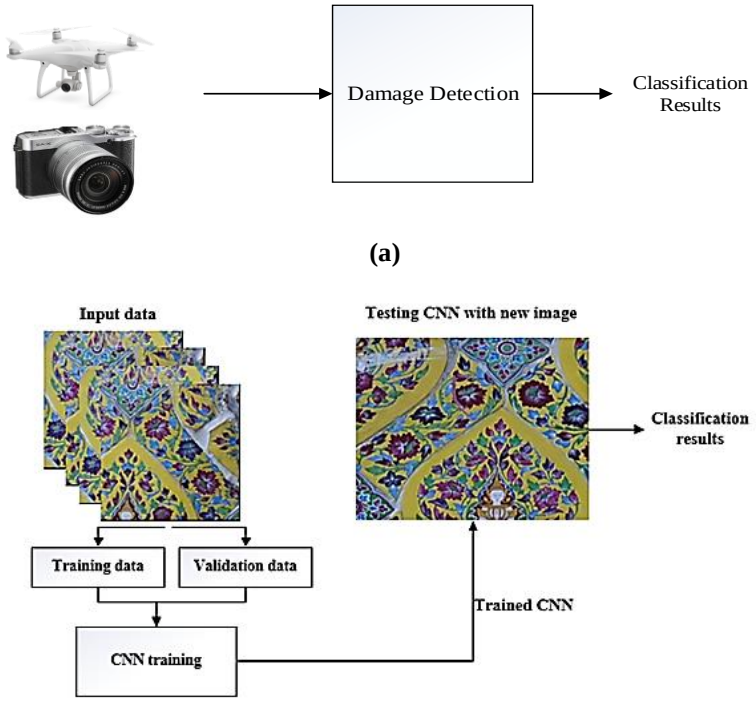

Fig. 2. (a) System outline and (b) CNN damage detection system.

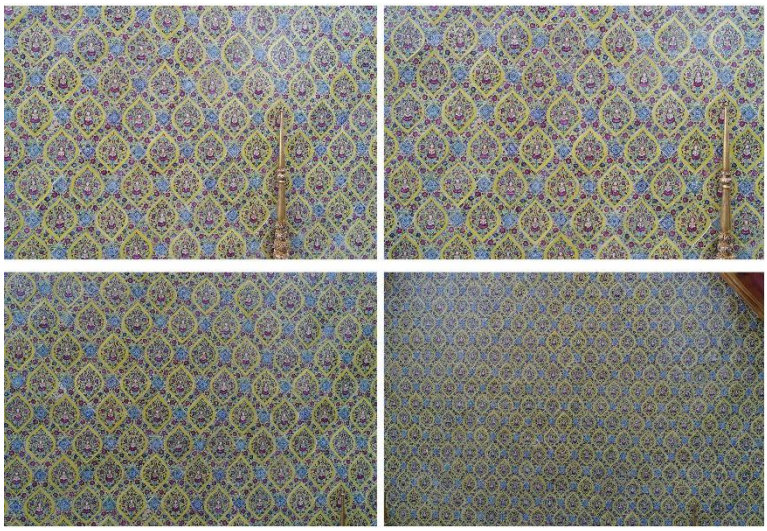

Fig. 3. Example images from the temple.

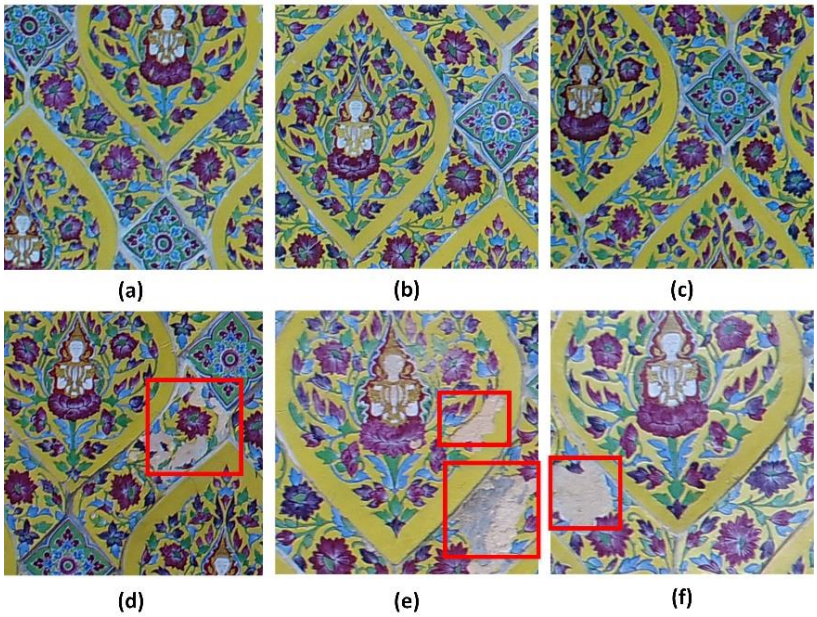

Fig. 4. Non damaged image patches (a-c), Samples of damaged tiles (d-f) (highlighted in red).

3.2. Damage detection using CNN

In this work, a Convolutional Neural Network (CNN) is used for automatic damage detection in ceramic tiles due to its capability to learn features automatically unlike the traditional handcrafted features techniques. The architecture of the CNN used is this work is shown in Fig. 5 and the detail of the architecture is shown in Table 1. The CNN consists of three main parts, input images, deep features extractor, and a classifier. The role of the feature extractors is to extract meaningful features through automatic learning from raw input images. The learned features are then given to the SoftMax classifier to classify between damaged and non-damaged tiles. Keras Sequential model [22] was used in this work, the architecture of the CNN used in this work is shown in Fig. 5. Each layer is explained below.

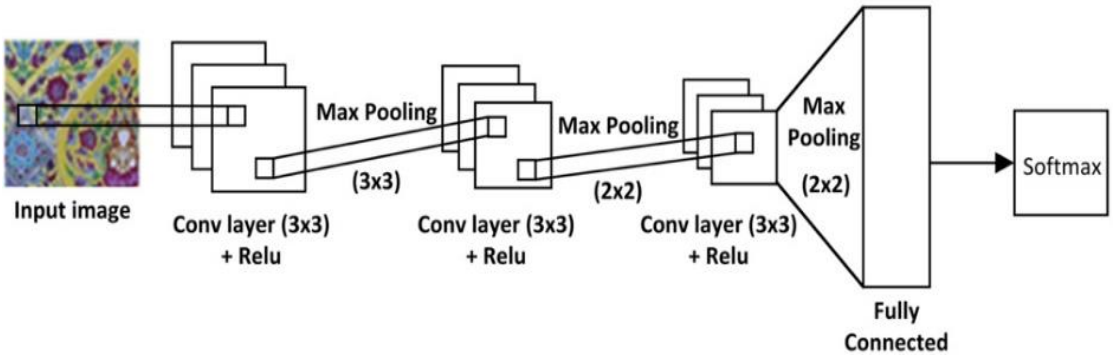

Fig. 5. CNN architecture used in the proposed system.

Table 1. Proposed CNN architecture.

Operation Layer		#Filters	Filter Size	Stride Value	Padding Value	Size of Output Image		
Input Image		-	-	-	-	100×100×3	64×64×3	32×32×3
Convolution Layer	Convolution	64	3×3×3	1×1	1×1	100×100×64	64×64×64	32×32×64
Pooling Layer	Max Pooling	1	3×3	0	0	49×49×64	31×31×64	15×15×64
Convolution Layer	Convolution	128	3×3×64	1×1	1×1	49×49×128	31×31×128	15×15×128
Pooling Layer	Max Pooling	1	2×2	0	0	24×24×128	15×15×128	7×7×128
Convolution Layer	Convolution	256	3×3×128	1×1	1×1	24×24×256	15×15×256	7×7×256
Pooling Layer	Max Pooling	1	2×2	0	0	12×12×256	7×7×256	3×3×256
Fully Connected Layer		4056	1×1	1×1	1×1	36,864	12,544	2,304
Softmax Layer		2	1×1	1×1	1×1	2	2	2

3.2.1. Convolution layer

Three steps involved in the convolution layer are shown in Fig. 6. Firstly, an element-by-element dot product multiplication is performed between the sub-array of an input image and a kernel, which is called a filter matrix or a receptive field. The initial weight of the filter matrix is randomly generated. The filter size is smaller than the input array size of the input image, but the sub-array of input images and the filter size are identical in each operation. Different filter sizes can be used in each convolution layer. In the second step, output values are multiplied and then added. And finally, an output feature map is obtained from the output values from the second step.

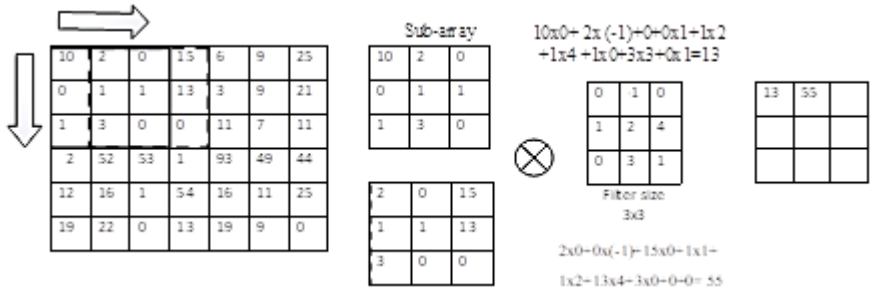

Fig. 6. An example of the computation in the convolution layer.

3.2.2. Max pooling layer

In the pooling layer of the CNN, the dimensionality of the images is reduced by reducing the number of pixels in the output obtained from the previous layer of the convolutional layers. A non-linear down sampling operation is performed, in which the spatial size of input images is reduced. In the max pooling layer, the maximum value of the sub-array of the input image is selected as shown in Fig. 7. Scherer et al. [23] showed that the max pooling layer significantly outperforms the subsampling operations, therefore; the max pooling operation is used in this study.

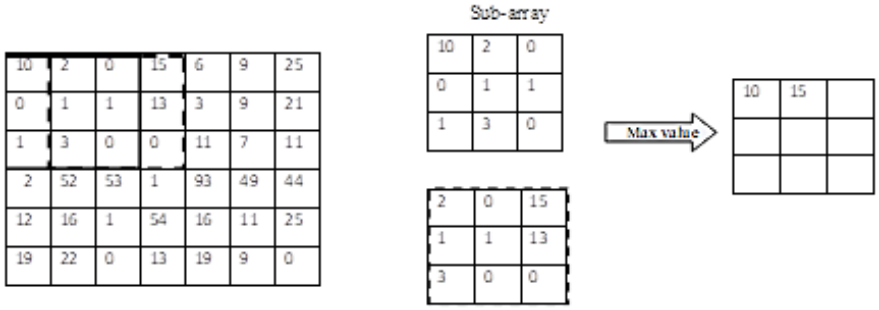

Fig. 7. The example of the computation of the max pooling layer.

3.2.3. Activation layer

Generally, in the activation layer, a non-linearity function such as, $y = tanh\ (x)$, is used, but in this work the Rectified Liner Unit (ReLu) [24] function is applied instead. Krizhevsky et al. [13] demonstrated that CNN with the ReLu function can be trained six times faster than using the standard tanh units. Equation (1) shows the mathematical formula of the Relu activation function, in which the output is either zero or positive, and has no negative values as shown in Fig. 8.

$$y(x) = max(0, x) \tag{1}$$

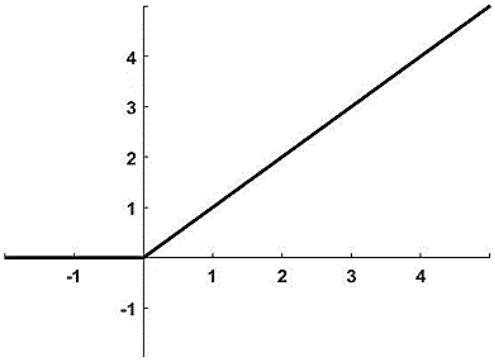

Fig. 8. Relu activation function.

3.2.4. Dropout layers

One of the main issues in machine learning is overfitting, which is when the data is learned from the training samples effectively but fail to generalize for validation and testing data. To avoid overfitting, the dropout layers are used. In the dropout layers, certain neurons are randomly disconnected with a normable dropout rate.

3.2.5. Softmax layer

For classification, a softmax function is used in the last layer of the CNN. The SoftMax function is explained in Eq. (2). For a given training set $S = \{a(i), b(i)\}$ containing n image patches, let $a(i)$ be *ith* image patch and $b(i)$ be the class label, either 1 for damaged tiles, or 0 for non-damaged tiles, be the outpuft of unit j in the

last layer of $a(i)$, then the probability P of the label $b(i)$ of $a(i)$ can be calculated from Eq. (2). The scores are converted to the probabilities of the output classes.

$$P(b^i) = j\left(k_j^i\right) = \frac{e^{k_j^{(i)}}}{\sum m e^{k_m^{(i)}}} \tag{2}$$

The SoftMax loss function L corresponding to Eq. (3) is given by:

$$L = \frac{1}{n}\left[\sum_{i=1}^{n}\sum_{j=1}^{i} 1\{b^{(i)} = j\}\log\frac{e^{k_j^{(i)}}}{e^{k_m^{(i)}}}\right] \tag{3}$$

where n is the total number of patches. To optimize and reduce the loss error, the stochastic gradient descent (SGD) function is applied since the SGD function is the simplest and most efficient way to reduce the deviations [15, 25]. In this work, the weight decay and momentum parameters are set as 0.00001 and 0.9, respectively.

4. Results

To evaluate the proposed damage detection system, 2880 tile patches are used for training and validation. Out of 2880 patches, 2160 patches or 75% of patches are used to train the CNN model, while the remaining 720 tile patches are used for validation of the proposed system, and the five-fold cross validation are applied. To test the proposed system, 720 previously unseen images are used for testing.

Table 2 shows a confusion matrix that are used to evaluate the results of the proposed system. Each output value from the confusion matrix, i.e., TP, FP, FN, and TN, is obtained by comparing the classification result to the ground truth value. The ground truth is obtained by manually labelling the image patches. The ground truth is labelled for the training, validation and testing dataset so that the quantitative measures for the accuracy of the proposed CNN accuracy can be estimated.

Table 2. Confusion matrix.

Actual input label	**Predicted output label**	
	Positive (Damage)	Negative (Non-damage)
Positive (Damage)	True Positive (TP)	False Negative (FN)
Negative (Non-damage)	False Positive (FP)	True Negative (TN)

The results of validation and testing are shown in Tables 3 and 4, respectively. Different patch sizes are also tried to study the optimal patch sizes against the system accuracy. As shown in Table 3, the maximum accuracy was achieved when 32 x 32 patch size was used to train and test the proposed system. An increase in the size of a patch may make damages appear to become smaller and could fade out within the tile patterns, which resulted in a decrease in accuracy. However, if a patch size is too small, it can be difficult for CNN as an entire patch may only contain damaged pixels, which makes it indistinguishable from non-damaged areas. Therefore, in this proposed work, a 32 x 32 patch size was used since it can achieve the accuracy of up to 94%, and it offered the best performance among the other patch sizes. To evaluate the performance of the proposed system, a confusion matrix and classification report are used, and the metrics used for the calculations

are shown in Eq. (4, 5 and 6). From Tables 3 and 4, the validation accuracy of the proposed damage detection system is 95%, and the testing accuracy is 91%. The training and validation accuracy with the number of epochs are shown in Fig. 9. As can be seen, the system accuracy and loss became stable after 3 epochs.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{4}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{5}$$

$$F_1 \text{ score} = \frac{2\ x\ \text{Precision}\ x\ \text{Recall}}{\text{Precision+Recall}} \tag{6}$$

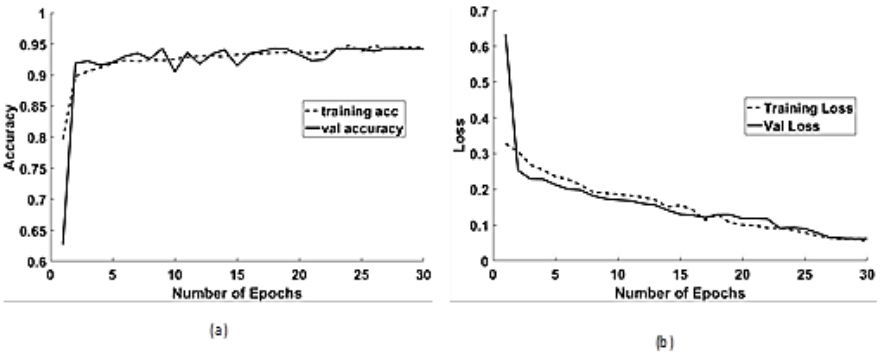

Fig. 9. (a) Training and validation accuracy, (b) Training and validation loss.

Table 3. The results of the proposed system on the validation dataset.

Patch size	Accuracy (Validation)	Precision	Recall	F1 Score
32x32	**94.9**	**0.94**	**0.94**	**0.94**
64x64	92.4	0.92	0.91	0.91
100x100	83.3	0.83	0.83	0.83

Table 4. The results of the proposed system on the testing dataset.

Method	Accuracy	Precision	Recall	F1 Score
Wavelets with ANN	72.1	0.71	0.70	0.71
LBP, HOG	61.2	0.61	0.61	0.61
Proposed method	91.25	0.91	0.92	0.91

After training, the proposed system was tested with 720 previously unseen images. The accuracy of the testing data is 91% as shown in Table 4. From Table 3 and 4, Recall identifies the proportion of the patches that are correctly classified as cracks. In Table 3, 32×32 patch size has the highest recall of 0.94, which means that 94% of the cracked patches are identified correctly. The high recall means that most of the crack patches are detected. From Table 4, the proposed method has the highest recall against the other feature-based methods.

The proposed method was compared with other existing methods, which were based on extracting handcrafted features by Ghazvini et al. [9]. The features were wavelet transform features and other statistical features, including the median of maximum points, minimum points, mean, and standard deviation. An Artificial Neural Network (ANN) classifier was used to classify the features to detect damages on the tiles. The results were also compared with other hand-crafted features, which was based on the Local Binary Patterns (LBP) as they have been successfully applied in many texture classification problems [26, 27]. The LBP was combined with Histogram of Oriented Gradients (HOG) features [28], which were normally used in the object detection problem. The features from LBP and HOG were then classified by Support Vector Machine as shown in [27].

As can be seen from Table 4 (Testing dataset), the accuracy of the proposed method was better than other methods as it has the accuracy up 91% on the testing dataset. This suggests that automatic feature learning by CNN is better than the handcrafted method to detect damages in tiles.

To localize damages, some sample images were used to show the results. A sliding window technique is used to classify each patch as a damage or non-damage. It can be seen in Fig. 10. that most damaged regions are correctly identified, although some misclassification can still be observed. The localization of damages can be improved by increasing the number of training images.

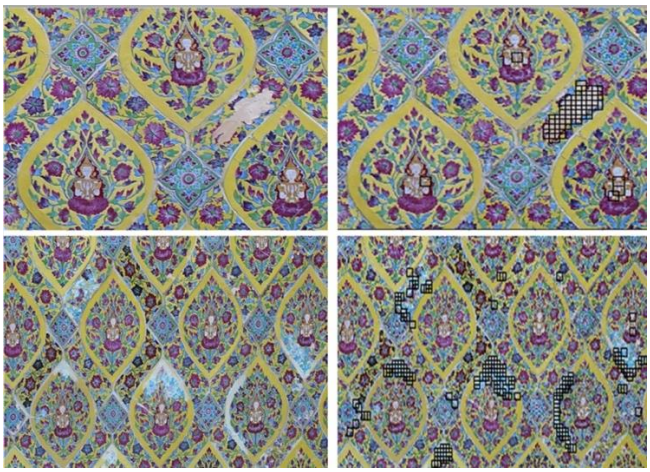

Fig. 10. Damage localization.

A Receiver Operating Characteristic (ROC) curves can be used to evaluate the classification performance. A ROC curve is a plot between True Positive Rate (TPR), also called as sensitivity, and False Positive Rate (FPR) as shown in Fig. 11. The TPR and FPR are computed for different probabilities values of the output by comparing predicted labels to ground truth values. Figure 11 shows the ROC curves between the proposed method and the other existing methods. As can be seen, the best result is achieved with the proposed CNN approach. Therefore, the proposed method by CNN is effective and can be used for detecting tile damages on a temple facade.

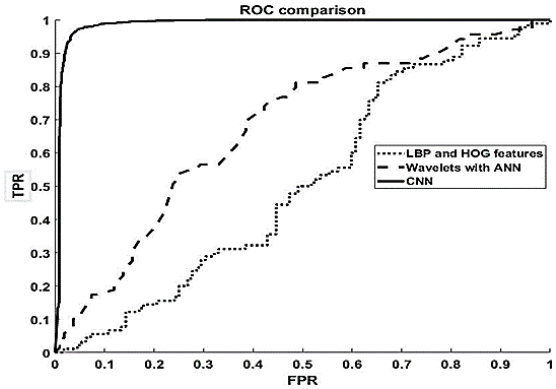

Fig. 11. The ROC curves between the proposed method and other handcrafted methods.

5. Discussion

It can be seen from the results in Table 4 that damages can be detected using images and are useful for inspection. The performance of the proposed system using CNN is significantly improved as compared to other traditional handcrafted features. The CNN architecture used in this study are only simple as the complex architecture such as those by He et al. [29] and Krizhevsky et al. [13], may require a large amount of training data for the system to work well, and they are generally computationally complex. A large training dataset can be difficult to create as collecting images of damages to use as training samples were not straightforward since the interested site did not have many visible damages. Furthermore, experienced inspectors were required to verify damage and non-damage tiles for creating the training datasets. Manually labelling data for training are a time-consuming process as discussed in Cha et al. [15], Krizhevsky et al. [13], and Zhang et al. [14]. The accuracy of the system can be increased by a larger number of training data and the quality of the training data can be improved by using verification by experienced inspectors. The process of creating data is in fact a significant problem in most machine learning systems.

CNN can extract discriminative features from many images without any pre-processing. A large amount of training images is required, which is a disadvantage. Since the tile patterns employed in this work are generally complex even for the human inspectors, therefore; it can be difficult to identify damaged tiles for making training dataset. To overcome this issue and to increase the training samples, data augmentation and transfer learning can be applied, which is beyond the scope of this study.

Conclusions

The experiments show that CNN can provide good results in damage detection in ceramic tiles. The proposed method is successfully applied to classify damages in tile images in both validation and testing datasets, although with better and larger datasets, the performance of the system can be improved further.

The traditional approach in detecting tile damages based on the feature extraction technique clearly performed worse than the CNN-based technique as it automatically learns features which aggregates essential information from images to improve the

classification accuracy. The capability of the proposed system is a significant achievement towards automation and automatic inspection for heritage structures.

Acknowledgement

The authors would like to thank the grant from the Research Division, Department of Engineering, and the Research Department of Thammasat University to allow this research work to be successfully completed.

Nomenclatures

L	SoftMax loss function
n	Number of image patches
P	Probability of label $b^{(i)}$ of $a^{(i)}$
(b^i)	Class label
(k_j^i)	Output of unit j in the last layer of $a^{(i)}$

Abbreviations

ANN	Artificial neural network.
CNN	Convolutional neural networks.
FN	False negative.
FP	False positive.
FPR	False positive rate.
HOG	Histogram of oriented gradients.
IBR	Image-based rendering.
LBP	Local binary patterns.
ROC	Receiver operating characteristic.
RIMLV	Rotation invariant of local variance.
ReLu	Rectified liner unit.
SGD	Stochastic gradient descent.
SVM	Support vector machine.
TN	True negative.
TP	True positive.
TPR	True positive rate.
UAVs	Unmanned aerial vehicles.

References

1. Hocenski, Z.; Vasilic, S.; and Hocenski, V. (2006). Improved canny edge detector in ceramic tiles defect detection. *Proceedings of the* 32*nd Annual Conference on IEEE Industrial Electronics*. Paris, France, 3328-3331.
2. Jacob, G.; Shenbagavalli, R.; and Karthika, S. (2016). Detection of surface defects on ceramic tiles based on morphological techniques. *Computer Vision and Pattern Recognition*. arXiv:1607.06676v1.
3. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; and Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and their applications*. 13(4), 18-28.
4. Liaw, A. and Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18-22.

5. Rowley, H.A.; Baluja, S.; and Kanade, T. (1998). Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1), 23-38.
6. Yao, W.; Li, Z.; and Qi, M. (2000). Damage process detection of a ceramic tile system by acoustic emission. *Experimental mechanics*, 40, 265-270.
7. Ar, I.; and Akgul, Y.S. (2008). A generic system for the classification of marble tiles using Gabor filters. *Proceedings of the* 2008 23*rd International Symposium on Computer and Information Sciences*. Istanbul, Turkey, 1-6.
8. Najafabadi, F.S.; and Pourghassem, H. (2011). Corner defect detection based on dot product in ceramic tile images. *Proceedings of the* 2011 *IEEE* 7*th International Colloquium on Signal Processing and its Applications*. Penang, Malaysia, 293-297.
9. Ghazvini, M.; Monadjemi, S.A.; Movahhedinia, N.; and Jamshidi, K. (2009). Defect detection of tiles using 2D-wavelet transform and statistical features. *World Academy of Science, Engineering and Technology*, 49, 901-904.
10. Hanzaei, S.H.; Afshar, A.; and Barazandeh, F. (2017). Automatic detection and classification of the ceramic tiles' surface defects. *Pattern Recognition*, 66, 174-189.
11. Thomas, A.S.; Hassan, M.F.; Ibrahim, M.; Rahman, M.N.A.; Sapuan, S.Z.; and Ahmad, M.F. (2019). A study on close-range photogrammetry in image based modelling and rendering (IMBR) approaches and post-processing analysis. *Journal of Engineering Science and Technology* (*JESTEC*), 14(4), 1912-1923.
12. Samarawickrama, Y.C.; and Wickramasinghe, C.D. (2017). Matlab based automated surface defect detection system for ceremic tiles using image processing. *Proceedings of the* 2017 6*th National Conference on Technology and Management* . Malabe, Sri Lanka, 34-39.
13. Krizhevsky, A.; Sutskever, I.; and Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks. *Proceedings of the Advances in Neural Information Processing Systems*. New York, United States, 1097-1105.
14. Zhang, L.; Yang, F.; Zhang, Y.D.; and Zhu, Y.J. (2016). Road crack detection using deep convolutional neural network. *Proceedings of the* 2016 *International Conference on Image Processing* . Phoenix, AZ, USA, 3708-3712.
15. Cha, Y.J.; Choi, W.; and Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 32(5), 361-378.
16. Wang, N.; Zhao, X.; Zou, Z.; Zhao, P.; and Qi, F. (2020). Autonomous damage segmentation and measurement of glazed tiles in historic buildings via deep learning. *Computer-Aided Civil and Infrastructure Engineering*, 35(3), 277-291.
17. Huang, Y.; Qiu, C.; and Yuan, K. (2020). Surface defect saliency of magnetic tile. *The Visual Computer*, 36(1), 85-96.
18. Wang, N.; Zhao X.; Zhao, P.; Zhang, Y.; Zou, Z.; and Ou, J. (2019). Automatic damage detection of historic masonry buildings based on mobile deep learning. *Automation in Construction*, 103, 53-66.
19. Alipour, M.; Harris, D. K.; and Miller, G. R. (2019). Robust pixel-level crack detection using deep fully convolutional neural networks. *Journal of Computing in Civil Engineering*, 33(6): 04019040.

20. Dung, C.V.; and Anh, L.D. (2019). Autonomous concrete crack detection using deep fully convolutional neural network. *Automation in Construction*, 99, 52-58.
21. Kalfarisi, R.; Wu, Z.Y.; and Soh, K. (2020). Crack detection and segmentation using deep learning with 3D reality mesh model for quantitative assessment and integrated visualization. *Journal of Computing in Civil Engineering*, 34(3): 04020010.
22. Chollet, F. (2015). Keras-Deep learning library. Retrieved 2015, from https://github.com/keras-team/keras.
23. Scherer, D.; Müller, A.; and Behnke, S. (2010). Evaluation of pooling operations in convolutional architectures for object recognition. *Proceedings of the International conference on Artificial Neural Networks*. Berlin, Heidelberg, 92-101.
24. Nair, V.; and Hinton, G.E. (2010). Rectified linear units improve restricted boltzmann machines. *Proceedings of the* 27*th international conference on machine learning* ,*M*adison, United States, 807-814.
25. LeCun, Y.A.; Bottou, L.; Orr, G.B.; and Müller, K.R. (2012). Efficient backprop. *Neural networks: Tricks of the trade*. Lecture Notes in Computer Science, 7700, Berlin, Heidelberg, 9-48.
26. Turtinen, M.; Mäenpää, T.; and Pietikäinen, M. (2003). Texture classification by combining local binary pattern features and a self-organizing map. *Proceedings of the Scandinavian Conference on Image Analysis*. Berlin, Heidelberg, 1162-1169.
27. Ojala, T.; Pietikainen, M.; and Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971-987.
28. Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110.
29. He, K.; Zhang, X.; Ren, S.; and Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA, 770-778.