Crack Detection in Masonry Structures using Convolutional Neural Networks and Support Vector Machines

Krisada Chaiyasarn[a], Wasif Khan[b], Luqman Ali[b], Mayank Sharma[b], Daniel Brackenbury[c] and Matthew DeJong[c]

[a]Department of Civil Engineering, Thammasat University, Thailand
[b]Department of Electrical and Computer Engineering, Thammasat University, Thailand
[c]Department of Engineering, University of Cambridge, United Kingdom
E-mail: ckrisada@engr.tu.ac.th, wasifuetm@yahoo.com, luqman.a@ku.th, macx.sharma@gmail.com, db559@cam.ac.uk, mjd97@cam.ac.uk

Abstract –

Masonry structures in historical sites are deteriorating due to ageing and man-made activities. Regular inspection and maintenance work is required to ensure the structural integrity of historic structures. The inspection work is typically carried out by visual inspection, which is costly and laborious, and yields to subjective results. In this study, an automatic image-based crack detection system for masonry structures is proposed to aid the inspection procedure. Previous crack detection systems generally involve the extraction of hand crafted features, which are classified by classification algorithms. Such approach relies heavily on feature vectors and may fail as some hidden features may not be extracted. In this study, we propose a crack detection system which combines deep Convolutional Neural Networks (CNN) and Support Vector Machines (SVM). CNN is used in extracting features from RGB images and SVM is used as an alternative classifier to a softmax layer to enhance the classification ability. A dataset containing images of cracks from masonry structures was created using a digital camera and an unmanned aerial vehicle from historical sites. The images were used for training and validating the proposed system. It is shown that the combined CNN and SVM model performs better than the model using CNN alone with the detection accuracy of approximately 86% in the validation images. It is also shown that the system can be used to detect cracks automatically for the images of masonry structures, which is useful for inspection of heritage structures.

Keywords –

Convolutional Neural Network; Support Vector Machine; Masonry structures; Crack detection, Computer vision

1 Introduction

Maintenance and condition assessment of historic structures are vital for Thailand as they are the country's cultural heritage. Cracks are considered to be the primary concern for the durability and safety of masonry structures. Therefore, crack detection is very important for the maintenance of structures and must be detected at the earliest stage to avoid unwanted situations, such as damage to buildings, or the collapse of structure due to severe cracks.

Visual inspection is a common procedure that is used in the examination and assessment of the current state of historical buildings. However, this procedure is laborious and time-consuming as it requires the experience and specialised knowledge of inspectors to assess structural conditions based on the visual appearance of structures. Furthermore, the procedure cannot be conducted frequently due to high labor cost and prone to human-error, and often the sites cannot be easily inspected due to inaccessibility. Figure 1 shows an example picture of a temple in Ayutthaya province in Thailand, where the top of the stupa is extremely high and is not accessible for human inspection. Figure 2 shows a picture of the Chapel viaduct, which is a masonry structure, and cracks and other damages in the structures need to be detected and monitored. In this paper, we propose an image-based crack detection system to inspect these masonry structures. Figure 3 shows example images of cracks that are found from various locations around the temple and the viaduct.

Cracks in images can be detected by either using techniques related to handcrafted feature extraction or by using automatic feature learning, i.e. a deep learning, which is a technique proposed in this paper. Crack detection systems based on handcrafted feature extraction techniques generally consist of two main steps. The first step is to extract relevant features, such as edges [1], features based on a percolation model [3-4] and multi-features [6] from input images. Then, in the second step, classifiers such as Support Vector Machines (SVM) [5] and Neural Networks [8] are trained to determine if extracted features are crack or non-crack. Since cracks can have various types and forms, this makes the features extraction task difficult. Therefore, deep learning algorithms are better in learning features automatically from raw images and complex high-level features can be built from low level features [12]. Deep learning has been applied in many problem as it proves to be a better technique in the classification task [10].

In this paper, we propose an image-based crack detection algorithm, using convolutional neural network (CNN) as a feature extractor and SVM as a classifier. The image data has been collected using an Unmanned Aerial Vehicle (UAV) and a handheld DSLR camera from heritage masonry structures in Thailand and from a masonry bridge in United Kingdom as shown in Figure 1 and 2. The rest of the paper is organized as follows, Section 2 presents related works about automatic crack detection. Section 3 and 4 describe the methodology of the proposed system and experiments. Discussion and conclusion are drawn in Section 5 and 6.

Figure 1: A masory structure from Ayutthaya, Thailand, which is inaccesible for manual inspection. Hence, images are acquired using a drone.

Figure 2: The Chapel viaduct, Essex, United Kingdom, a masonry structure many visible damages.

Figure 3 :Sample images of cracks in masonry structures

2 Related Works

Many automatic crack detection systems are based on extracting handcrafted features. Abdel-Qader et. al. [1] applied four different edge detection techniques, i.e. Fast Haar Transform (FHT), Fast Fourier Transform, Sobel and Canny detectors for concrete bridges. The FHT was the best one among other detectors in the study. The limitation of edge detection algorithms is generally due to noise. Liu et. al. [9] applied image intensity features and Support Vector Machine (SVM) for tunnel crack detection. This method is prone to error due to noise.

Prasanna et. al. [7] proposed a crack detection system based on histograms for a bridge deck. Principle

Component Analysis (PCA) techniques are used for automatic inspection for concrete bridge decks in [2]. Yamaguchi et. al. [3-4] applied a percolation model for crack detection in concrete surfaces. Fujita et. al. [12] proposed a concrete crack detection system for noisy data using different filters for preprocessing to remove noise, followed by probabilistic relaxation and adaptive thresholding for the detection of cracks. Similarly, Prasanna et. al. [6] used STRUM (Spatially Tuned Robust Multi-feature) classifier for automatic crack detection on concrete bridges. In Prasanna's work, the proposed system applied intensity-based and gradient-based features with the combinations of scale-space features as the crack features, which were then classified. They demonstrated that the efficiency of the STRUM classifier was better than other image-based approach for crack detection.

For different types of cracks and images containing noise, the techniques based on handcrafted features fail to perform. Hence, automatic feature extraction based on learning techniques such as deep learning can perform well when compared to the techniques based on handcrafted features. Zhang et. al. [11] applied deep convolutional neural network for road crack detection from images collected using a low cost smart phone. Cha et. al. [13] used deep convolutional neural network (DCNN) for automatic concrete crack detection and presented 98% accuracy. Ellenberg et. al. [14] discussed several algorithms, including percolation approach, fractal method and tensor voting for crack detection. The paper also conducted a study on masonry crack detection although the paper did not provide details on their results.

3 Methodology

The outline of the proposed system is shown in Figure 4. The proposed system consists of three modules. Firstly, images are acquired using a UAV and a DSLR camera, Then, images are classified by a crack detection system based on CNN as a feature extractor and SVM as a classifier. The results from the crack detection module can then be used to localize cracks in the final module. Each module in the proposed system is explained below.

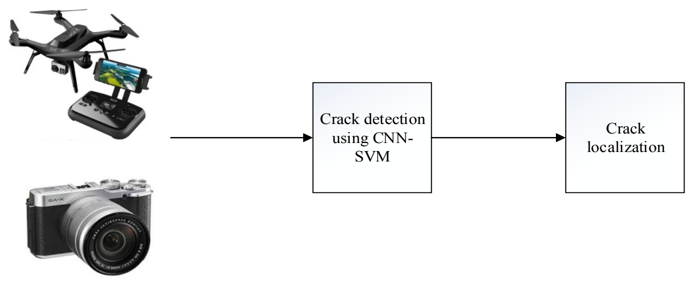

Figure 4: The outliine of the proposed methodology

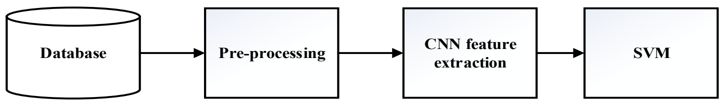

Figure 5: The outline of the proposed crack detection system.

3.1 Image acquisition

In recent years, UAVs have been utilized in surveying as an alternative to conventional surveying methods since they are faster, simpler and cheaper. UAVs can also be utilized to collect images for inspection as shown in this paper. We collected images from a stupa from Wat Chai Watthanaram, which is located in Ayutthaya in Thailand. The images were collected using a DJI Phantom 4 drone. The drone was programmed to fly around the stupa to collect images at two different heights. At the ground level, a DSLR camera were used to collect images of the stupa near the ground. Sample images taken from the drone are shown in Figure 6. As shown in this figure, the top of the stupa cannot be easily reach by human inspectors, and utilizing the drone to collect images around this area clearly provides great benefits for data collection.

In addition, with the collaboration from the Cambridge University team, more sample of masonry structures were also collected from the Chapel viaduct. This is a railway viaduct that crosses the River Colne in the Colne Valley in Essex, UK. The images were collected using a Sony DSLR camera from various locations around the viaduct. The sample images are shown in Figure 7.

Figure 6: Sample images acquired using a UAV

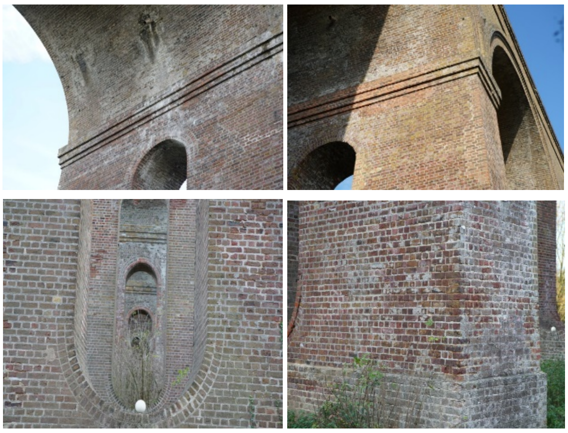

Figure 7 :Samples images of the Chapel viaduct

The images collected using the drone were converted into patches of size 96x96 using a Matlab algorithm for training and classifying by CNN. Only patches belonging to masonry areas were selected and used for training, the patches containing surrounding objects (e.g. trees) were ignored. The total number of image patches used in this work were 6002, which were manually separated into cracks and non-cracks images. Out of all patches, 3162 patches were used for training and validation, and the remaining 2840 patches were used for testing the proposed system. The crack and non-crack patches were manually labelled as either 0 or 1. The example of crack and non-crack patches are shown in Figure 8 and 9.

Figure 8: An example of crack patches

Figure 9: An example of non-crack patches

3.2 Crack Detection

In the proposed crack detection system, we applied Convolutional Neural Network (CNN) as it has ability in solving many real-world problems efficiently. The architecture of CNN used in this paper is shown in Figure 10.

The multilevel deep feature extractor and a classifier are the two main tools of the CNN architecture. The role of multilevel deep feature extractor is to retrieve discriminant features from image pixel intensity values presented in the RGB colour channels and SVM is used as a classifier for the purpose of classification. In our proposed CNN architecture, Keras sequential model was used, which was composed of convolutional, activation and max pooling layers. The first convolutional layers consist of 32 filters of size 3x3 pixels as shown in Figure 10. After filtering, the max pooling operation was activated with a ratio of 2.

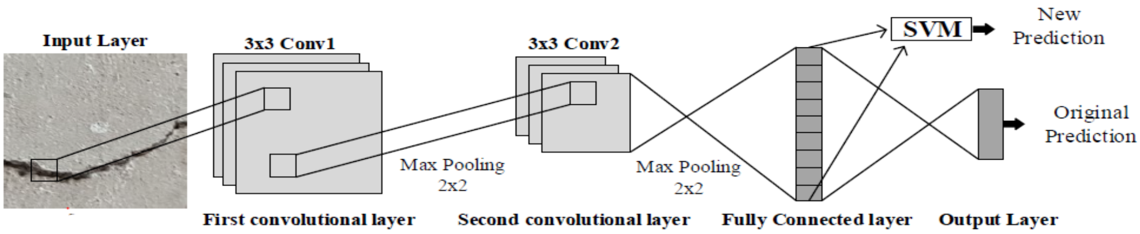

Figure 10: The architecture of CNN used in the proposed system

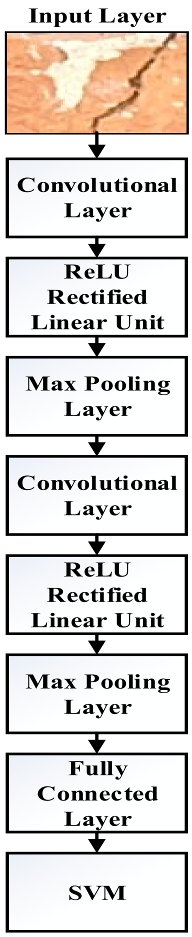

Figure 11: The architecture of CNN-SVM technique.

Figure 11 shows a more detail architecture of CNN used in this study. As shown in the figure, the key role of the convolutional layers is to detect the local connections of features from the previous layer. The output result of the feature maps is then passed to the activation layer ReLU. The max-pooling operation was used in the vision systems for two reasons, (1) to eliminate non-maximal values, which helps to reduce computation time for layers, and (2) to down-sample operations for 2x2 sub-regions to reduce the dimension of intermediate feature vectors .The filters are stacked together and fully connected layers can then be used in computing the class scores. In the proposed system, the output from the fully connected layer become the input feature vectors of size 2352 for the SVM classifier as depicted in Figure 10 and 11.

An SVM classifier was applied in the final stage of the proposed crack detection system. The SVM classifier was used to replace the softmax layer in CNN. The main objective in SVM is to find a hyperplane that separates the largest fraction of a labelled dataset for binary classification. The training data is a set of training samples pairs $\{(x_1, y_1), \ldots, (x_i, y_i)\}$, where x_i is the observation or input feature for the i^{th} sample and $y_i \in \{1, 0\}$ is the associated class label .The SVM classifier is the discriminant function that maps an input feature space x_i into a class label y_i. An interested reader is referred to read [16] for the detail of Support Vector Machines.

4 Experiments and Results

The evaluation of the proposed crack detection system was conducted on validation and testing datasets. The system was evaluated against our own ground truth data, which was manually labelled, to estimate inaccuracy that can occur from the proposed system. As mentioned in Section 3.1, a total of 6002 image patches were labelled either crack or non-crack patches in our experiment, 3162 patches were used in training and validation. The 2840 patches were used for testing. The Receiver Operating Characteristic curve (ROC) analysis, confusion matrix and classification report were used to evaluate the performance of the proposed system.

For SVM, the Radial Basis Function (RBF) was used as a kernel, hence a cross-validation technique was employed to obtained the optimal values for the kernel. Table 1 shows a parametric study for SVM, where a different combination of C and gamma values were tried in the validation dataset to obtain the maximum accuracy. As shown in the table, the best accuracy occurred when C = 4 and gamma = 1.

Table 1: Parametric study for SVM

C	gamma	Accuracy
1	0.5	0.770
1	1	0.72
2	1	0.72
3	1	0.73
4	1	0.73
5	1	0.71

The classification report and ROC curves were obtained based on the confusion matrix, which can be explained as shown in Table 2.

Table 2: Confusion matrix for class classification

Ground Truth Label	Predicted Label	
	Positive (Crack)	Negative (Non-crack)
Positive (Crack)	True Positive (TP)	False Negative (FN)
Negative (Non-crack)	False Positive (FP)	True Negative (TN)

The following equations are used for the classification analysis in the classification report.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1score = \frac{2xPrecxRecall}{Prec + Recall} \quad (4)$$

We conducted experiment by comparing the results between the method with the CNN alone and the proposed system, where we combined CNN and SVM. Table 3 shows the results of the proposed crack detection system on the validation dataset. It can be seen that the combined method CNN-SVM offers the better accuracy of 85.94 than the CNN method with the accuracy of 82.94. Other metrics, including Precision, Recall and F1Score are also better for the CNN-SVM method.

Table 2 shows results on the testing dataset. From Table 2, it can be seen that, again, the CNN-SVM method is more accurate than the CNN method. The combined method improved the accuracy by 7.4%.

Table 3: Results of the proposed system on validation dataset

Method	Validation Accuracy	Precision	Recall	F1 score
CNN	82.94	0.83	0.71	0.74
CNN-SVM	85.94	0.84	0.79	0.79

Table 4 :Results of the proposed system on testing dataset

Method	Accuracy	Precision	Recall	F1 score
CNN	67.5	0.80	0.68	0.73
CNN-SVM	74.9	0.82	0.78	0.78

The ROC curve is shown in Figure 12. The ROC is a plot between True Positive Rate (TPR) and False Positive Rate (FPR) for the probabilities values of the output as computed by comparing predicted labels to ground truth values. TPR and FPR are calculated by the following equations,

$$TPR = \frac{TP}{TP + FN} \quad (5)$$

$$FPR = \frac{FP}{FP + TN} \quad (6)$$

It can be seen that, from Figure 12, the CNN-SVM method is better than the CNN method as the ROC curve of the CNN-SVM method appears to approach more towards the top left corner of the graph.

To localize cracks, some sample images were used to show the result. We divided an image into grid and each grid was classified by the proposed crack detection system. If the grid is classified as crack, we highlight the grid as red. Figure 13 shows an example of crack localization. It can be seen that, on the top pair of images, most crack regions are correctly identified, although some misclassification can still be observed. However, the bottom pair of images contains many false negative areas, which may be due to the proposed system was confused with the grout lines as cracks. The bottom pair of images also has many false negative, especially around the grout lines. This suggests that the inaccuracy of the system may be due to these regions as the appearance of the grout lines is similar to cracks. Nevertheless, with more training dataset, the result should improve.

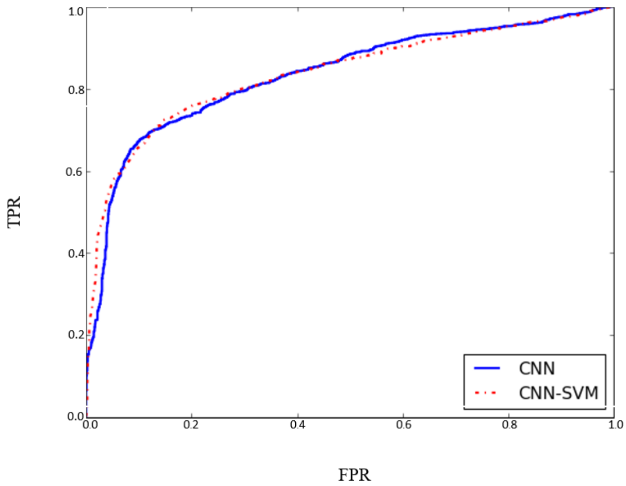

Figure 12: The ROC curves between the CNN technique, and the combined CNN-SVM model

Figure 13: Crack localization of selected sample images

5 Discussion

The combined model, i.e. CNN-SVM, shows an increase in accuracy as shown in Table 3 and 4. It can be seen from the results that CNN-SVM is extremely good at extracting features and classification, although it relies heavily on good datasets. Good dataset can be difficult to create as they rely on human as a gold standard and the database may need to be verified by multiple sources. In our database, we have many non-crack patches but fewer crack patches. To overcome the problem with a small dataset, transfer learning and data augmentation may be applied.

Crack detection on masonry structures are difficult as cracks cannot be easily identified in images. Cracks in masonry structures have similar appearance to grout lines, which can be mistaken as cracks, unlike the problems of crack detection on simple concrete surface shown in previous studies. Therefore, it can be difficult to create good datasets since this type of scene, i.e. the masonry surface, is complex and confusing, even for human inspectors themselves. Nevertheless, good datasets are still required for any CNN systems.

6 Conclusion

From the experiments, we can conclude that the emerging class of technologies known as deep CNN offers the possibility of automatic crack detection for masonry structures. The combined techniques known as CNN-SVM has been implemented in this work to automatically classify image patches and to localize crack regions on masonry images for inspection. The proposed method is successfully applied to classify image in our validation and testing datasets, although with better and larger datasets, our system performance can be improved further.

We also concluded that the efficiency of the model can be further improved by fine tuning the CNN architecture and its parameters, such as adding more layers to CNN. From the results shown in this work, it can be concluded that the combined model, namely CNN and SVM performs better than the method using CNN alone. As shown in this work, CNN is best to be used as a feature extractor, and these features can be classified by any classifiers. In the future work, different classifiers can be explored to see if the accuracy of the system can be improved.

7 Acknowledgment

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

8 References

[1] Abdel-Qader I., Osama A. and Kelly M. Analysis of edge-detection techniques for crack identification in bridges, *Journal of Computing in Civil Engineering*, 17 (October 2003), pp. 255-263

[2] Abdel-Qader I., Pashaie-Rad, S., Abudayyeh, O. and Yehia, S. (2006). PCA-based algorithm for unsupervised bridge crack detection. *Advances in Engineering Software*, *37*(12), 771-778.

[3] Yamaguchi T., Suzuki K., Hartono P. and Hashimoto S. Percolation approach to image-based crack detection. In *Proceedings of the 7th international conference on Quality Control by Artificial Vision*, vol. 11, pp. 291-296. 2005.

[4] Yamaguchi T., Shingo N. and Shuji H. An efficient crack detection method using percolation-based image processing.

In *Industrial Electronics and Applications, 2008. ICIEA 2008. 3rd IEEE Conference on*, pp. 1875-1880. IEEE, 2008.

[5] Suykens AK. and Joos V. Least squares support vector machine classifiers. *Neural processing letters*9, no. 3 (1999): 293-300.

[6] Prasanna P., Kristin J., Nenad G., Basily B., Hung M., Ronny L, and Hooman P. "Automated crack detection on concrete bridges." *IEEE Transactions on Automation Science and Engineering*13, no. 2 (2016): 591-599.

[7] Prasanna P., Kristin D., Nenad G. and Basily B. Computer-vision based crack detection and analysis. *SPIE Smart Structures and Materials and Nondestructive Evaluation and Health Monitoring* (2012).

[8] Rowley H., Shumeet B. and Takeo K. Neural network-based face detection. *IEEE Transactions on pattern analysis and machine intelligence* 20, no. 1 (1998): 23-38.

[9] Liu Z., Shahrel A., Takeshi O., and Toshiaki E. Tunnel crack detection and classification system based on image processing. In *Machine Vision Applications in Industrial Inspection X*, vol. 4664, pp. 145-153. International Society for Optics and Photonics, 2002.

[10] Gamboa J. and Borges C.. Deep Learning for Time-Series Analysis. *1701.01887* (2017).

[11] Zhang L., Fan Y., Yimin D. and Ying J.. Road crack detection using deep convolutional neural network. In *Image Processing (ICIP), 2016 IEEE International Conference on*, pp. 3708-3712. IEEE, 2016.

[12] Fujita Y. and Yoshihiko H. A robust automatic crack detection method from noisy concrete surfaces. *Machine Vision and Applications* 22, no. 2 (2011): 245-254.

[13] Cha Y., Wooram C. and Oral B. Deep Learning Based Crack Damage Detection Using Convolutional Neural Networks. *Computer Aided Civil and Infrastructure Engineering* 32, no. 5 (2017): 361-378.

[14] Ellenberg A., Kontsos A., Bartoli I. and Pradhan A. Masonry crack detection application of an unmanned aerial vehicle. In *Computing in Civil and Building Engineering (2014)*, pp. 1788-1795. 2014.

[15] Downey A., D'Alessandro A., Laflamme S. and Ubertini F. (2017). Smart bricks for strain sensing and crack detection in masonry structures. Smart Materials and Structures, 27(1), p.015009.

[16] Cortes C. and Vapnik V. (1995). Support vector machine. Machine learning, 20(3), 273-297.