

Computer vision-based classification of cracks on concrete bridges using machine learning techniques

Yang Yu¹, Maria Rashidi¹, Bijan Samali¹, Masoud Mohammadi¹, Andy Nguyen²

¹Centre for Infrastructure Engineering (CIE), Western Sydney University, Penrith, NSW 2751, Australia

²School of Civil Engineering and Surveying, University of Southern Queensland, Springfield, QLD 4300, Australia

email: y.yu3@westernsydney.edu.au, m.rashidi@westernsydney.edu.au, b.samali@westernsydney.edu.au, m.mohammadi@westernsydney.edu.au, andy.nguyen@usq.edu.au

ABSTRACT: Concrete crack is a significant indicator related to the durability and serviceability of concrete civil infrastructure such as dams, bridges and tunnels. Current inspection of concrete structures is based on manual visual operation, which is not effective in safety, cost and reliability. This research aims to address the problems in traditional inspection of concrete structures by proposing a novel automatic crack identification approach, which intelligently integrates both image processing and machine learning techniques. Through the crack-sensitive feature extraction and model self-learning, the proposed method has higher identification accuracy than conventional inspection method, which has been proved by the experimental verification.

KEY WORDS: Concrete crack; Image processing; Feature extraction; Machine learning.

1 INTRODUCTION

Cracks are always observed on the surface of civil infrastructures such as buildings, highways, dams, tunnels and bridges [1]. Some cracks are dangerous to the structures, while the others are not considered as the defect due to the small size. However, if they are neglected, the potential danger can lead to huge disaster. Accordingly, it is crucial to monitor and inspect civil structures to ensure safety and reduce the maintenance cost. Recently, with the rapid development of unmanned aerial vehicles (UAVs), health monitoring and condition diagnosis of large-scale civil infrastructure has been propelled to advance in the direction of automatic inspection and remote monitoring. Compared with traditional inspection methods with drawbacks of time-consuming and high cost, the UAVs are capable to conduct the quick assessment of structure which is not conveniently accessible, with fewer labour and cost [2,3]. As a consequence, it has been regarded as a promising approach for infrastructure assessment.

However, the main challenge for the application of this innovative technology is data analysis, that is how to translate the captured images to the assessment results. To solve this problem, a large number of studies have been carried out in crack identification and segmentation via image processing and pattern recognition. Kim et al. put forward a crack diagnostic method based on the fuzzy set theory for reinforced concrete structures [4]. Nnolim proposed an automated crack identification approach based on using partial differential equation [5]. Tsai et al. demonstrated the application of geodesic minimal path algorithm in the generation of the crack pattern, which can be used for the process of route planning [6]. Kim et al. presented a novel concrete crack assessment approach based on the integration of RAP technique and hybrid image processing, in which the binarization method was used to evaluate the crack width [7]. On the other hand, machine learning (ML) techniques have been introduced for image processing, and several studies have been reported in application of ML in crack classification and segmentation accordingly. Lee et al. presented an approach based on back-

propagation (BP) neural networks to identify and analyse the crack on the concrete surface [8]. Chun et al. employed gradient boost decision tree to design the crack detection algorithm, the inputs of which are colour, gradient and texture characteristics of crack [9]. In spite of the progressive advances of ML technique, further investigation is necessary to be conducted to solve the problems of complex network architecture, hyper-parameter setting and limited training data.

To solve the current challenges, this paper proposes a novel hybrid method to automate the crack detection of concrete structures in real-time. The proposed approach combines the image processing and machine learning techniques, including gabor filter, local binary pattern, kernel principal component analysis, *k*-nearest neighbours (*k*-NN), naïve Bayes (NB) and support vector machine (SVM). To start with, the two-dimensional (2D) gabor filter is adopted to filter the original image of concrete surface for eliminating the noise influence. Then, the local gabor binary pattern histograms are extracted from the filtered images to make up the feature vector. To decrease the dimension of crack-sensitive features, the kernel principal component analysis (KPCA) is conducted on the extracted feature vector and the first several principal components are used as the inputs of machine learning models for concrete crack identification. Finally, three machine learning models are developed and compared in terms of identification accuracy of concrete crack. The performance of the proposed method is evaluated by the images of a concrete bridge taken in the field with satisfactory results.

2 PROPOSED HYBRID APPROACH FOR AUTOMATIC CRACK IDENTIFICATION

In this study, on basis of structural surface photos captured by the UAVs, an automatic crack identification approach is proposed, which consists of four steps. In the first step, the original image is processed by 2D Gabor filter. Then, the local gabor binary pattern histogram is extracted from filtered image for feature extraction. In the next step, the KPCA is employed to reduce the feature dimension, based on which three machine

learning models are developed to realize the crack prediction via pattern recognition. The details of these methods are presented in the following sub-sections.

2.1 Image pre-processing based on gabor filter

In this paper, gabor filter is applied to process the original image for feature extraction. The principle of gabor filter is summarised as follows.

2D gabor kernel function is defined as follows.

$$g(x, y) = s(x, y)w_r(x, y) \quad (1)$$

where $s(x, y)$ denotes the 2D complex wave, defined in Eq. (2):

$$s(x, y) = e^{i(2\pi(u_0x+v_0y)+P)} \quad (2)$$

where (x, y) denotes the spatial coordinate and P denotes the phase and its value is generally set to 0. (u_0, v_0) denotes the coordinates in the frequency domain, and corresponding expression of polar coordinates is shown in Eq. (3):

$$\begin{aligned} u_0 &= F_0 \cos \omega_0 \\ v_0 &= F_0 \sin \omega_0 \end{aligned} \quad (3)$$

where F_0 and ω_0 are expressed as follows:

$$\begin{aligned} F_0 &= \sqrt{u_0^2 + v_0^2} \\ \omega_0 &= \tan^{-1}\left(\frac{v_0}{u_0}\right) \end{aligned} \quad (4)$$

$\omega_r(x, y)$ is a 2D Gaussian function with rotation parameter, defined as

$$\omega_r(x, y) = Ke^{-\pi\left(\frac{(x-x_0)^2}{\delta_x^2} + \frac{(y-y_0)^2}{\delta_y^2}\right)} \quad (5)$$

where K is a constant; (x_0, y_0) denotes the central coordinates of Gaussian kernel; δ_x^2 and δ_y^2 are standard deviations, which are used to control the distributions in x and y directions. Hence, the specific form of 2D gabor kernel is provided in Eq. (6)

$$g(x, y) = Ke^{-\pi\left(\frac{(x-x_0)^2}{\delta_x^2} + \frac{(y-y_0)^2}{\delta_y^2}\right)} e^{i(2\pi F_0(x \cos \omega_0 + y \sin \omega_0) + P)} \quad (6)$$

In this paper, the value of P is set to 0, and the following 2D gabor filter (Eq. (7)) is selected for processing the original image of concrete crack.

$$\varphi_{\Pi(f, \theta, \gamma, \eta)}(x, y) = \frac{f^2}{\pi\gamma\eta} e^{-(\alpha^2 x'^2 + \beta^2 y'^2)} e^{i(2\pi f x')} \quad (7)$$

2.2 Feature extraction based on local gabor binary pattern histogram (LGBPH)

In this paper, the local binary pattern method is employed to encode the features of image after gabor filtering, which has higher recognition accuracy compared to the encoding of original images. Finally, the local gabor binary pattern histogram is adopted as the final features to indicate the concrete surface crack. Figure demonstrate the procedure of the proposed feature extraction method, the detailed procedure of which is summarised as follows.

Step 1. Gabor features are extracted from a raw image of concrete surface with or without crack.

Step 2. Uniform local gabor binary pattern method is employed to encode the images extracted from gabor filter.

Step 3. Each encoded image is divided into $k \times k$ sub-blocks, and the histogram of LBPH is calculated for each sub-block. Subsequently, the histograms of different sub-blocks are connected to make up the LGBPH features of the whole image.

2.3 Feature compression based on kernel principal component analysis (KPCA)

The feature dimension of LGBPHs extracted by the above method is high and redundant. If it is not compressed, it will greatly increase the complexity of classification decision and affect the speed of the classifier. Therefore, it is necessary to reduce the dimension of the original features to improve the classification efficiency. In this research, to fix this problem, a type of nonlinear dimension reducing method called kernel PCA (KPCA) is introduced. In KPCA, the kernel function is used to map the data in the input space to a high-dimensional feature space, which makes the input data have better separability. Then, linear PCA is performed on the transformed data of the new feature space, and the nonlinear principal component of the original data is obtained effectively, and the dimension compressed features with better separability are obtained. The basic principle and calculate process of using KPCA to compress the features of images for concrete crack identification are provided as follows.

First of all, the feature vector x is transformed into a new transformation space F , that is $\Phi: R^N \rightarrow F, x \rightarrow X$. Meanwhile, the condition $\sum_{k=1}^N \Phi(x_k) = 0$ is satisfied. The covariance matrix of new space can be written as $\tilde{C} = \frac{1}{N} \sum_{i=1}^N \Phi(x_i) \Phi(x_i)^T$. Then, the eigenvalue λ and corresponding eigenvector v are calculated, meeting the condition $\lambda v = \tilde{C} v$. Accordingly, we can get $\lambda \Phi(x_k) v = \Phi(x_k) \tilde{C} v$, and there exists $\alpha = [\alpha_1, \dots, \alpha_N]$ to meet the condition $v = \sum_{i=1}^N \alpha_i \Phi(x_i)$. Then, we define $K_{ij} = \Phi(x_i) \cdot \Phi(x_j)$, and get $N \lambda \alpha = K \alpha$. For the solution λ_k and α^k , the normalisation operation is needed, which is expressed as follows:

$$\begin{aligned} 1 &= \sum_{i,j=1}^N \alpha_i^k \alpha_j^k (\Phi(x_i) \cdot \Phi(x_j)) = (\alpha^k \cdot K \alpha^k) \\ &= \lambda_k (\alpha^k \cdot \alpha^k) \end{aligned} \quad (8)$$

For the data in the original space, the projection of x in the principal component direction of the transformation space is equivalent to the projection of $\phi(x)$ in the principal component direction v , that is $v \Phi(x) = \sum_{i=1}^N \alpha_i K(x_i, x)$. Here, denotes the kernel function. Generally, there are several commonly used kernel functions to choose, including linear function, polynomial function, radius basis function (RBF), etc. In this study, the RBF is selected due to its excellent nonlinear mapping capacity.

Through nonlinear dimensionality reduction by KPCA, the data dimension can be effectively reduced, and the nonlinear features in the original data can be better retained, which can greatly improve the accuracy and efficiency of the developed model for rapid and accurate identification of concrete cracks.

2.4 Crack identification using machine learning (ML) techniques

The final target of this research is to develop predictive models to automate the crack identification from concrete images. To achieve this target, three machine learning approaches are adopted and compared based on the compressed features in Section 2.3. The fundamentals of three ML methods are described in detail as follows.

- k -nearest neighbours (k -NN)

The k -NN algorithm was developed to classify the unknown data based on existing data with known labels.³² The essence of k -NN is analogy learning, that is learning by comparing the given test sample with training samples. Suppose that in the last section, l PCs are selected from feature vector of IMF energy ratios as the inputs of the ML model, which indicates that all the training samples should be in this l -dimensional space. When an unknown test sample is given, the k -NN algorithm searches the l -dimensional pattern space and finds k training samples that are closest to the unknown test sample. These k training samples can be regarded as k nearest neighbours of unknown sample.

The implementation of k -NN is based on the following assumptions: (1) all the data and labels belong to the numeric types; (2) the smaller the distance between two samples, the more similar these two samples; (3) each attribute (PC) has equal weight; (4) attribute values are normalized; (5) if one attribute value of sample x_1 and (or) sample x_2 are lost, their distance are supposed to be maximum possible distance. Here, Euclidean distance is employed to measure the similarity between test sample S_i and training sample S_j , the expression of which is shown in Eq. (9):

$$dist(S_i, S_j) = \sqrt{\sum_{k=1}^l (S_{i,k} - S_{j,k})^2} \quad (9)$$

The membership degree of test sample S_i belonging to the class C_j is defined as follows:

$$MD(S_i, C_t) = \sum_{S_j \in kNN(S_i)} dist(S_i, S_j) \delta(S_j, C_t) \quad (10)$$

where $kNN(S_i)$ denotes the nearest neighbour set; $\delta(S_j, C_t)$ denotes the class attribute of S_j belonging to C_t , i.e. $\delta(S_j, C_t) = \begin{cases} 1, & S_j \in C_t \\ 0, & \text{other} \end{cases}$.

Therefore, identifying the class of S_i is considered as calculating an optimization problem as follows:

$$C \text{ arg max } MD(S_i, C_t) \quad (11)$$

The class C with maximum membership degree should be the class of S_i .

- Naïve Bayes (NB)

Naïve Bayes method is a commonly used algorithm in the area of data mining and machine learning. Mainly applied in the classification area of machine learning, naïve Bayes method has a broad range of engineering applications in terms of public opinion analysis, text analysis, user preference analysis, biomedical diagnosis, etc. It has the benefits of highly efficient,

simple, stable and strong theoretical basis. The principle of Naïve Bayes method can be briefly summarised as follows.

Given a training data set (X, Y) , in which each sample X contains n -dimensional features and Y includes k categories as class label. If a new sample x is provided, the total probability equation can be expressed by:

$$P(y_k|x) = \frac{P(x|y_k)P(y_k)}{\sum_k P(x|y_k)P(y_k)} \quad (12)$$

where the conditional probability $P(x|y_k)$ can be transformed into:

$$P(x|y_k) = P(x_1, x_2, \dots, x_n|y_k) = \prod_{i=1}^n P(x_i|y_k) \quad (13)$$

Accordingly, the naïve Bayes classifier can be expressed in Eq. (14):

$$f(x) = \text{argmax} P(y_k) \prod_{i=1}^n P(x_i|y_k) \quad (14)$$

- Support vector machine (SVM)

SVM is a commonly used machine learning method based on statistical learning theory, which has the advantages for solving the nonlinear pattern recognition problems with small sample.³³ The mechanism of SVM is to learn and improve the generalization capacity by search for the structural risk minimization. Essentially, it can be transformed into a convex quadratic optimization problem. Suppose there is a data set $T = \{(x_1, y_1), \dots, (x_i, y_i), \dots, (x_l, y_l)\}$. Based on the nonlinear function, the input data can be mapped into the high-dimensional space for regression, shown in Eq. (15):

$$f(x_i) = w \cdot \varphi(x_i) + b \quad (15)$$

where l denotes the sample number; b denotes the bias; w denotes the weight vector; $\varphi(x_i)$ denotes the kernel function, which is used to transform the linear problem into nonlinear problem. In this work, the radial basis function (RBF) is chosen as the kernel with the following expression:

$$\varphi(x_i) = e^{-\frac{\|x_i - x\|^2}{2\sigma^2}} \quad (16)$$

where σ denotes the kernel parameter. The target of the SVM is to find an optimal hyperplane to separate two categories of data, which can be transformed into the following optimization problem with the constraints.

$$\min \phi(w) = \frac{1}{2} \|w\|^2 \quad \text{s. t. } y_i [(w \cdot \varphi(x_i) + b)] \geq 1 \quad (i = 1, 2, \dots, l) \quad (17)$$

Sometimes the samples cannot be well classified and the classification errors exist in the developed SVM models. In this case, a slack variable is added into the constraints to solve the error problem, shown as:

$$y_i [(w \cdot \varphi(x_i) + b)] \geq 1 - \xi_i \quad (i = 1, 2, \dots, l) \quad (18)$$

When $0 < \xi_i < 1$, all the samples can be correctly classified. When $\xi_i \geq 1$, x_i will be misclassified. To avoid this problem, the penalty term $C \sum_{i=1}^l \xi_i$ is added to the minimization target and the fitness function can be written as follows:

$$\phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \quad (19)$$

Therefore, combine Eq. (18) and Eq. (19), and the optimization problem with the constraints can be expressed as:

$$\min \phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \quad s.t. \quad y_i [(w \cdot \varphi(x_i) + b)] \geq 1 - \xi_i \quad (i = 1, 2, \dots, l) \quad (20)$$

where $\sum_{i=1}^l \xi_i$ denotes the upper bound of the misclassified sample number, which is used to measure the deviation degree of the data from ideal partition condition. C denotes the penalty coefficient. To calculate this optimization problem, the Lagrange function is employed to transform the optimal classification problem into its dual form, shown as:

$$\max L(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \varphi(x_i, x_j) \quad (21)$$

$$s.t. \quad 0 \leq \alpha_i, \alpha_j \leq C \quad (i, j = 1, 2, \dots, l), \quad \sum_{i=1}^l y_i \alpha_i = 0$$

Hence, the corresponding classification decision function of SVM can be expressed in Eq. (22).

$$f(x) = \text{sgn} \left(\sum_{i=1}^l \alpha_i y_i \varphi(x_i, x) + b \right) \quad (22)$$

In SVM, the penalty coefficient C and kernel factor are two main parameters that are capable of remarkably affecting the performance of the model. In this research, the values of C and σ^2 are set as 30 and 2, respectively, according to the suggestions in [10].

3 EXPERIMENTAL VALIDATION

In this study, the images, with the resolution of 2048×2000, are taken from a bridge are used for evaluating the proposed hybrid approach. Then, the images are cut into more images with smaller size of 256×256 to make up the dataset, including 10,000 images of intact surface and 10,000 images of concrete crack. For the purpose of model training and validation, we randomly select 80% of images of each label as training samples, while the rest of images are used as validation samples to evaluate the capacities of trained machine learning models. Figure 1 provides an example of intact surface image and cracked surface image, respectively.

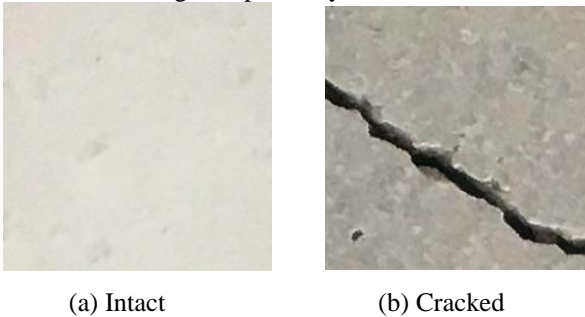


Figure 1. Example of images of concrete surface without and with crack.

Subsequently, the 2D gabor filter is employed to process the original images. Figure 2 shows an example of filtering results of both intact surface and cracked surface images. It is clearly

seen that the proposed filter is capable of effectively distinguish different types of images (intact and cracked), even though background brightness may influence the filtering results. Based on the 2D gabor filter, all the images are filtered, which will be used for crack-sensitive feature extraction.

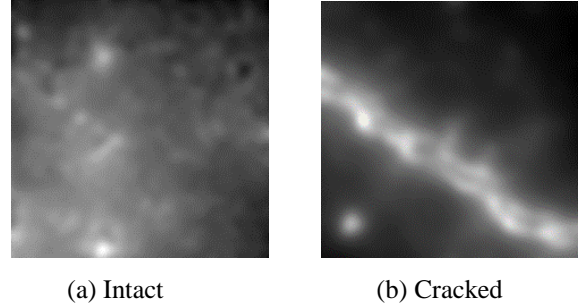


Figure 2. Processed images using gabor filter.

To extract the effective features from the filtered images, the local binary patterns (LBP) are considered in this research. Figure 3 gives an example of LBP histogram of two types of images (intact and cracked). Obviously, the LBP can transform the filtered image into a limited number of features, which can be used as inputs of machine learning model. It is noticeable in Figure 3 that the LBP features were effectively extracted to distinguish non-cracked and cracked images. However, it is hard to say that this type of feature can be effective in all the images in the dataset. To solve this problem, KPCA is used to compress the features, which makes the feature separable. In this study, the RBF kernel is selected and its parameter is set to 2.5 according to the suggestion in [11]. Here, t-distributed stochastic neighbour embedding (t-SNE) is adopted to visualize the features extracted from filtered images. Figure 5 shows a comparison of features before and after KPCA operation using the former three principal components (PCs). It can be observed from Figure 4 (a) that before KPCA operation, some non-cracked and cracked images are hard to distinguish due to the corresponding red and blue points mixed together in the figure. However, after the KPCA operation, it is apparent that using 3 PCs the cracked cases can be well distinguished from the non-cracked cases. Accordingly, if we can use more PCs as the input variables, the developed machine learning models will have higher prediction accuracy.

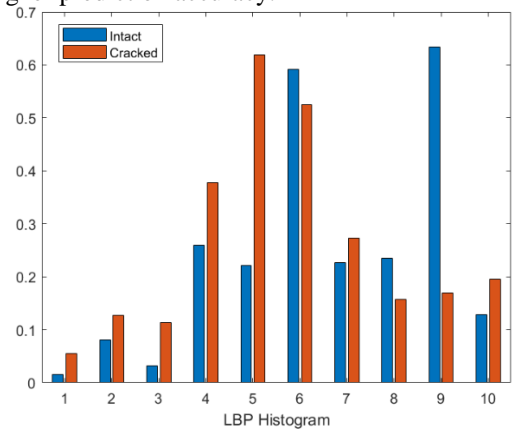
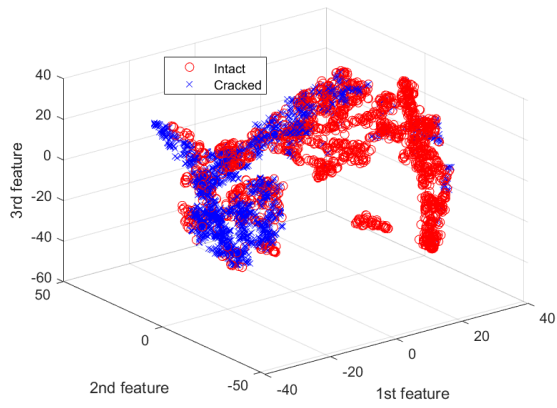
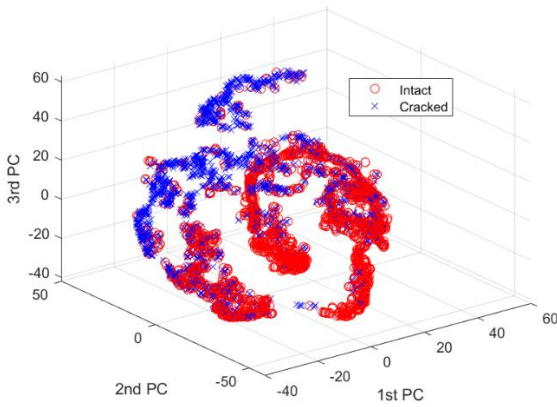


Figure 3. Feature extraction using LBP.



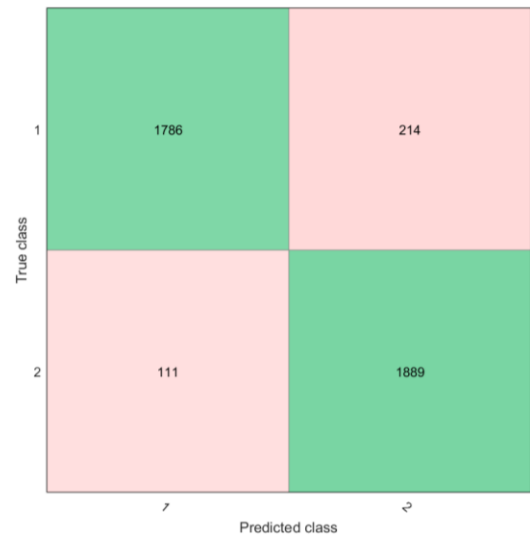
(a) Before PCA



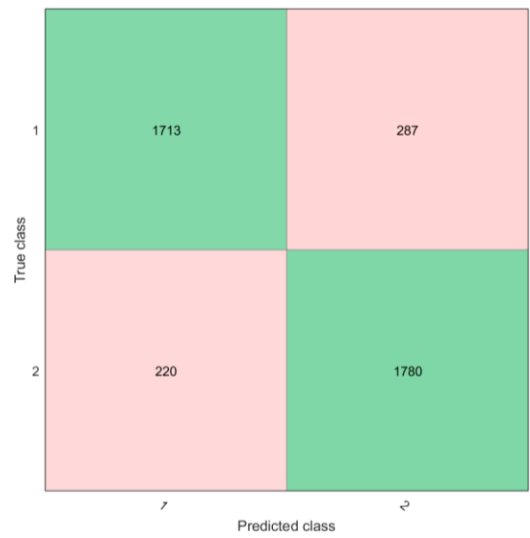
(b) After PCA

Figure 4. Feature comparison before and after PCA.

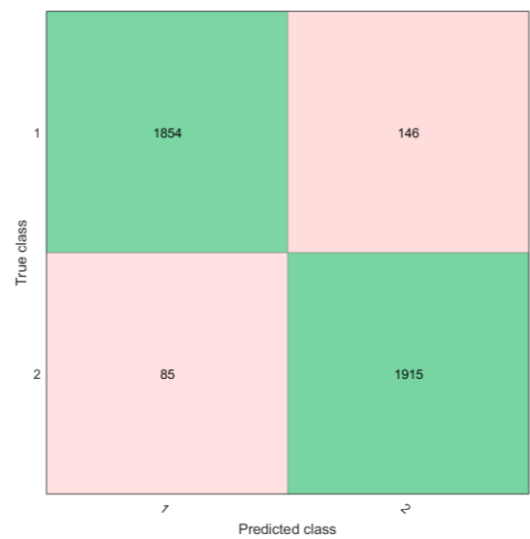
To automate the crack identification in the field, the machine learning models are built up, including naïve Bayes, *k*-NN and SVM. It is well known that a training process is necessary to get three predictive models ready, which may result in the trained models over-fitted. To avoid this phenomenon, 5-fold cross validation is considered during the model learning, in which all the training data are randomly divided into five groups. For each time, one group are used as the test data while the other four groups are used as training samples. The whole procedure is repeated five times and five results are averaged as the objective function to tune the model parameters. Figure 5 compares the training results of three models in terms of confusion matrix. Obviously, the SVM has the best training results with prediction accuracy of 94.2%, while the training accuracies of naïve Bayes and *k*-NN models are 91.9% and 87.3%, respectively. Then, the testing images are sent to the trained models for performance evaluation. The corresponding results are displayed in Table 1. It is noted that the testing accuracies of three models are a little lower than training accuracies. Similarly, the SVM has the best prediction performance (90.7%) among all three models, which can be considered as the optimal solution to the crack identification of concrete structures in the field.



(a) Naïve Bayes



(b) *k*-NN



(c) SVM

Figure 5. Training results of machine learning models. (“1” denotes intact; “2” denotes cracked)

Table 1. Prediction accuracy of test samples.

Model	Prediction accuracy
NB	90.7%
k-NN	85.8%
SVM	93.1%

- [11] Y. Yu, M. Rashidi, B. Samali, A.M. Yousefi, W. Wang, Multi-image-feature-based hierarchical concrete crack identification framework using optimized SVM multi-classifiers and D-S fusion algorithm for bridge structures, *Remote Sensing*, 2021, 13, 240.

4 CONCLUSIONS

This paper proposes a hybrid approach for automatic crack identification of concrete structures. First of all, the 2D gabor filter is used to deal with the original image of concrete surface for eliminating the noise influence. Then, local binary pattern histogram is used to extract features from filtered images as the feature vector. To compress the features and make them separable, kernel PCA is conducted on the feature vector and the transformed PCs are used as the inputs of machine learning models. Finally, three state-of-the-art machine learning methods, i.e. naïve Bayes, k-nearest neighbours (*k*-NN) and support vector machine (SVM) are developed to automate the identification process. The field images taken from a bridge are used to validate the effectiveness of the proposed method. The results show that SVM has optimal prediction performance among three models, with accuracy of 94.2% for training data and 93.1% for testing data. In the future work, the deep learning technology will be investigated to carry out the crack segmentation for the images, which are of great significance for the engineers to do onsite inspection of concrete bridges.

ACKNOWLEDGMENTS

The authors would like to acknowledge the technical and financial supports from Transport for New South Wales.

REFERENCES

- [1] Y. Liu, J.K.W. Yeoh, Robust pixel-wise concrete crack segmentation and properties retrieval using image patches, *Automation in Construction*, 2021, 123, 103535.
- [2] M. Rashidi, M. Mohammadi, S. Kivi, M. Abdolvand, L. Truong-Hong, B. Samali, A decade of modern bridge monitoring using terrestrial laser scanning: Review and future directions, *Remote Sensing*, 2020, 12, 3796.
- [3] M. Rashidi, C. Zhang, M. Ghodrati, S. Kempton, B. Samali, A. Akbarnezhad, L. Zhu, Bridge abutment movement and approach settlement : A case study and scenario analysis, *International Journal of Structural Stability and Dynamics*, 2018, 18, 1840011.
- [4] Y.M. Kim, C.K. Kim, G.H. Hong, Fuzzy set based crack diagnosis system for reinforced concrete structures, *Computers & Structures*, 2007, 85, 1828-1844.
- [5] U.A. Nnolim, Partial differential equation-based enhancement and crack detection, *Mathematical Problems in Engineering*, 2019, 2019, 8157205.
- [6] Y.C. Tsai, V. Kaul, C.A. Lettsome, Enhanced adaptive filter-bank-based automated pavement crack detection and segmentation system, *Journal of Electronic Imaging*, 2012, 24, 043008.
- [7] H. Kim, J. Lee, E. Ahn, S. Cho, M. Shin, S.H. Sim, Concrete crack identification using a UAV incorporating hybrid image processing, *Sensors*, 2017, 17, 2052.
- [8] B.Y. Lee, Y.Y. Kim, S.T. Yi, J.K. Kim, Automated image processing technique for detecting and analysing concrete surface cracks, *Structure and Infrastructure Engineering*, 2013, 9, 567-577.
- [9] P.J. Chun, S. Izumi, T. Yamane, Automatic detection method of cracks from concrete surface imagery using two-step light gradient boosting machine, *Computer-Aided Civil and Infrastructure Engineering*, 2021, 36, 61-72.
- [10] Y. Yu, U. Dackermann, J. Li, E. Niederleithinger, Wavelet packet energy-based damage identification of wood utility poles using support vector machine multi-classifier and evidence theory, *Structural Health Monitoring*, 2019, 18, 123-142.