Hyperspectral Imaging Applied for Pixel-Level Crack Detection with Background Interferences

SIYI CHEN, YOUWU WANG and YIQING NI

ABSTRACT

Cracks in civil infrastructures are an important sign of structural degradation and may indicate the inception of catastrophic failure. Existing image-based crack detection techniques face challenges when it comes to the complex background scenes. These irrelevant background interferences are common in practice and may trigger false alarms in crack detection. To eliminate their influence, hyperspectral imaging is employed in this study, which captures hundreds of spectral reflectance values in a pixel in the visible and near-infrared region. Compared with the conventional greyscale/RGB images which are limited to one/three wide spectral bands (red, green, blue), hyperspectral imaging can therefore provide more rich spectral information for crack detection/distinguish cracks from other background interferences. Due to the high correlations in hyperspectral image data, this study proposed a hyperspectral crack detection method using the low rank representation-based algorithm. Moreover, a locality constraint together with the dictionary learning process is incorporated into the proposed method to train a multi-class classifier. The built classification model is tested based on a real-world hyperspectral imaging dataset, which contains eight different surface objects in total. The trained classifier achieves an overall accuracy of 92.1%. The results show that the proposed method can predict cracks and other materials under complex scenes.

Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
National Rail Transit Electrification and Automation Engineering Technology Research Center (Hong Kong Branch), The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China

INTRODUCTION

Cracks are a common issue in civil infrastructures and can cause material discontinuities, reduce local stiffness, and indicate potential structural failure [1]. While vibration-based methods are widely used for damage detection [2], they are not effective for identifying local damage such as cracks. To manage the local cracks, traditional visual inspection by experts is commonly adopted. The visual inspection completely depends on the knowledge and experience of trained personnel; hence, the whole process is often time-consuming, labor-intensive, and prone to human errors. In addition, in some regions of the target structure, it is difficult for human beings to gain access [3, 4]. Therefore, timely and accurate crack monitoring is crucial for ensuring structural safety.

To solve these problems, image-based crack detection using computer vision algorithms has gained popularity due to its low cost, high efficiency, and flexibility [5, 6]. Moreover, implementing robots or unmanned aerial vehicles (UAVs) as inspection platforms further automates the image-capturing process [7]. Up to now, a number of image-based crack detection methods have been developed, which can be broadly classified into two categories, namely image processing techniques and machine learning methods.

For image processing techniques, they include edge detection methods [8, 9], image binarization methods [10], percolation models , and mathematical morphology-based methods [11]. These methods magnify the visibility of cracks within images. Abdel et al. [12] compared four edge detection algorithms (fast Haar transform (FHT), fast Fourier transform (FHT), Sobel and Canny edge detection) and found that the fast Haar transform (FHT) was the most reliable for crack detection. Although image processing techniques have shown good performance in comparison to manual crack detection, they may not perform well in complex environments with noise interference and complex backgrounds, leading to lower detection accuracy.

Another kind of autonomous method for crack detection is the latest machine learning methods, particularly deep learning (DL) methods [13, 14]. These techniques rely on a training set of images to learn and identify the distinguishing characteristics of cracks. Among them, Cha et al. [15] first developed a deep convolutional neural network (D-CNN) model trained on 40,000 manually annotated images of concrete surfaces with dimensions of $256 \times 256 \times 3$ pixels to detect the presence or absence of cracks. Chen and Jahanshahi [16] trained a D-CNN augmented with a Naive Bayes classifier on 5,326 manually annotated images of size $120 \times 120 \times 3$ pixels that were extensively augmented to increase the dataset size to over 250,000 images. In these methods, the detected cracks are highlighted with bounding boxes. Further geometric information, such as the length, width and the area of the detected cracks, are not provided. Subsequently, some DL-based pixel-level crack detection methods have been developed [17, 18]. By using use deep CNNs, these methods can segment cracks in images or videos at the pixel level. The segmentation results are more precise, allowing for the quantification of parameters such as crack length and width.

The above studies have made great progress in the crack detection using conventional grayscale or RGB images. However, they may still face challenges under real-world conditions. In practice, complex surface scenes and background interferences such as water stains, vegetation, parking lines, and oil stains may cause false alarms and

pseudo cracks. This leads to a generalization issue for classification methods as a well-trained classifier can overfit and fail to generalize to other image scenes.

To eliminate their influence, hyperspectral imaging (HSI) is introduced in this study, which captures hundreds of spectral reflectance values in a pixel in the visible and near-infrared region. Compared with the conventional greyscale/RGB images which are limited to one/three wide spectral bands (red, green, blue), hyperspectral imaging can therefore provide more rich spectral information to distinguish cracks from other background interferences. Due to the high correlations in hyperspectral image data, this study proposed a hyperspectral crack detection method using the low rank representation-based algorithm. As a result, an eight-class classification model could be built based on the proposed method, which includes "Concrete", "Asphalt", "Crack", "Oil stains", "Artificial marking", "Green vegetation", "Water stains", and "Dry vegetation".

METHODOLOGY

Low-Rank Representation in HSI

Given an HSI image, the 3D HSI data $\mathbf{I} \in \mathbb{R}^{B \times n_1 \times n_2}$ can be built, where B denotes the number of spectral bands and $n_1 \times n_2$ denotes the number of pixels in the spatial domain. Then, the 3D HSI data $\mathbf{I}$ is rearranged into a matrix $\mathbf{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_N] \in \mathbb{R}^{B \times N}$, where $N = n_1 \times n_2$ and $\boldsymbol{y}_i \in \mathbb{R}^{B \times 1}$ denotes reflectance spectrum of the i^{th} test pixel. With an overcomplete spectral dictionary $\mathbf{D}$, each test pixel $\boldsymbol{y}_i$ can be represented as a linear combination of the atoms in the dictionary. As such, $\mathbf{Y}$ could be expressed as

$$\mathbf{Y} = \mathbf{DZ} + \mathbf{E} \tag{1}$$

where $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \cdots, \mathbf{D}_K] \in \mathbb{R}^{B \times m}$ denotes a spectral dictionary formed by the HSI pixels from all K classes, and $\mathbf{D}_i$ is sub-dictionary corresponding to the class i; $\mathbf{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_N]$ denotes the representation matrix, and $\mathbf{E}$ represents the sparse noise component. If the test pixel $\boldsymbol{y}_i$ belongs to class c, it is assumed that $\boldsymbol{y}_i$ can only be described by the sub-dictionary $\mathbf{D}_c$ of class c. As a result, the non-zero elements in $\boldsymbol{z}_i$ are those corresponding to the category $\boldsymbol{y}_i$ belongs to. Therefore, the resulting representation matrix $\mathbf{Z}$ reveals the critical class information, which could be used for class prediction.

To address Eq. (1), more constraints are needed for characterizing matrices $\mathbf{Z}$ and $\mathbf{E}$. Ideally, if atoms in dictionary $\mathbf{D}$ are rearranged according to the order of class, the optimal representation matrix $\mathbf{Z}$ would have a classwise block-diagonal structure. Then Eq. (1) could be solved by the following low-rank representation (LRR) problem:

$$\begin{aligned} &\min_{\mathbf{Z},\mathbf{E}} \ \|\mathbf{Z}\|_* + \lambda \|\mathbf{E}\|_1 \\ &\text{s.t. } \mathbf{Y} = \mathbf{DZ} + \mathbf{E} \end{aligned} \tag{2}$$

where $\|\cdot\|_*$ denotes the nuclear norm of a matrix, which is a good surrogate of the rank minimization problem, and $\|\cdot\|_1$ denotes the ℓ_1 norm; $\lambda > 0$ is the trade-off parameter that balances the weights of the rank term and the noise component.

Dictionary Learning

In LRR, dictionary quality is of great importance for classification, and it will influence the discriminative power of the representation $\mathbf{Z}$. Instead of employing the whole training data set as the dictionary, this study learns a dictionary behaving well in the training set. The mathematical model for dictionary learning can be described as follows:

$$\begin{aligned} &\min_{\mathbf{Z},\mathbf{D},\mathbf{E}} \ \|\mathbf{Z}\|_* + \lambda\|\mathbf{E}\|_1 + \frac{\gamma}{2}\|\mathbf{D}\|_F^2 \\ &\text{s.t. } \mathbf{X} = \mathbf{DZ} + \mathbf{E} \end{aligned} \tag{3}$$

where $\mathbf{X}$ denotes the training data set, and $\frac{\gamma}{2}\|\mathbf{D}\|_F^2$ is to avoid scale change during the dictionary learning process.

Although LRR has a powerful ability to capture the global structure of the given HSI data, the local structure between two pixels, which is also helpful in classifying HSI, is ignored by LRR [19, 20]. Therefore, a locality constraint term, which explicitly utilizes the spectral similarity of pixels, is employed in this study to improve the performance of the original LRR model. The locality constraint as a penalty term for LRR is introduced as

$$\|\mathbf{M} \odot \mathbf{Z}\|_1 \tag{4}$$

where $\odot$ denotes the Hadamard product, and $M_{i,j}$ is the distance between the i^{th} and j^{th} pixels. Finally, the locality constraint term is integrated into Eq. (3) to learn a discriminative dictionary and representation as follows,

$$\begin{aligned} &\min_{\mathbf{Z},\mathbf{D},\mathbf{E}} \ \|\mathbf{Z}\|_* + \lambda\|\mathbf{E}\|_1 + \alpha\|\mathbf{M} \odot \mathbf{Z}\|_1 + \frac{\gamma}{2}\|\mathbf{D}\|_F^2 \\ &\text{s.t. } \mathbf{X} = \mathbf{DZ} + \mathbf{E} \end{aligned} \tag{5}$$

where α controls the contribution of the locality constraint term. Associating locality constraint in the training process, a compact and discriminative dictionary $\mathbf{D}$ can be learned from all training data by Eq. (5).

Classification and Crack Detection

Using the dictionary learned from Eq. (5), the low-rank representation matrix $\mathbf{Z}_{Test}$ of test HSI data $\mathbf{Y}$ could be obtained by solving Eq. (2).

Based on the representation $\mathbf{Z}_{Train}$ of training data $\mathbf{X}$ and class label matrix $\mathbf{H}$, multivariate ridge regression model is designed to obtain a linear classifier $\mathbf{L}^*$:

$$\mathbf{L}^* = \arg\min_{\mathbf{M}} \left\|\mathbf{H} - \mathbf{L}\mathbf{Z}_{Train}\right\|_F^2 + \eta\|\mathbf{L}\|_F^2 \tag{6}$$

where η is the weight of the regularization term.

Then the class label for the test pixel i is determined by

$$k^* = \arg\max_k \mathbf{L}^* \mathbf{z}_i^{Test} \quad (7)$$

where k^* is the predicted label corresponding to the classifier with the largest output.

RESULTS

Datasets

The hyperspectral images of pavement cracks are obtained from a public crack dataset [21]. This public dataset consists of 50 hyperspectral images of cracks in the concrete or asphalt surface, captured by a UAV-based HSI system (Figure 1).

The HSI camera captures 139 narrow spectral bands ($B = 139$) in wavelengths ranging from 450 nm to 950 nm (most of the VNIR region). The spectral resolution is 8 nm in each band, and the spatial resolution is 50 × 50 pixels. After calibration, a calibrated 3D reflectance hyperspectral cube is built for each image. The 3D HSI data cube provides both spatial and spectral information, and holds some information that human eyes cannot see without the hyperspectral image. It comprises 139 spectral bands, 50 pixels in length and 50 pixels in width. Each pixel in HSI image contains a sampled spectrum that could be used to recognize and classify different materials. Consequently, the size of hyperspectral cube is 50 × 50 × 139.

In this study, a total of 50 hyperspectral images are used. All images contain hairline or apparent cracks. Among them, 34 crack images are for concrete surfaces, and 16 are for asphalt surfaces. Considering the complex background interference in the real environment, these 50 crack images are disturbed by multiple types of noise, including parking lines, water stains, oil stains, green and dry vegetation.

To train a supervised classification model, a “hyperspectral reflectance library” of different material is required first. After feature extraction, a total of eight different classes are considered during the pixel extraction process, including "Asphalt” "Crack", "Oil stains", "Artificial marking", "Green vegetation", "Water stains", "Dry vegetation" and "Concrete", respectively.

For each class, 25% of labeled HSI pixels from the dataset $\boldsymbol{A}$ are randomly chosen to be the training set, and the remaining 75% are used for testing. That is, each class consists of 50 data points for training and 150 data points for testing. The number of dictionary atoms for each class is set to be 20.

(a) (b)

Figure 1. (a) the Cubert FireflEYE S185 snapshot camera for HSI; (b) the UAV-based HSI system [22].

Classification Performance

After the training process is completed, the performance of the trained model in accurately predicting class labels is assessed by the testing set. By comparing the predicted class labels and the true class labels of the testing set, a confusion matrix C can be generated, as Figure 2 shows.

From the confusion matrix, true positive (TP), true negative (TN), false positive (FP) and false negative (FN) can be determined for each class. Then several commonly used metrics can be generated to evaluate the performance of the classification model, which are defined as

$$A_c = \frac{\sum_{i=1}^{K} C_{ii}}{\sum_{i=1}^{K}\sum_{j=1}^{K} C_{ij}} \times 100\% \qquad (8)$$
$$p = \frac{TP}{TP+FP};\ r = \frac{TP}{TP+FN};\ F_1 = \frac{2\times p\times r}{p+r}$$

where A_c represents the overall accuracy. p and r represent the precision and recall, respectively. Since the precision competes against the recall in any model, F1 score, acting as a comprehensive index, tends to better gauge the classification performance.

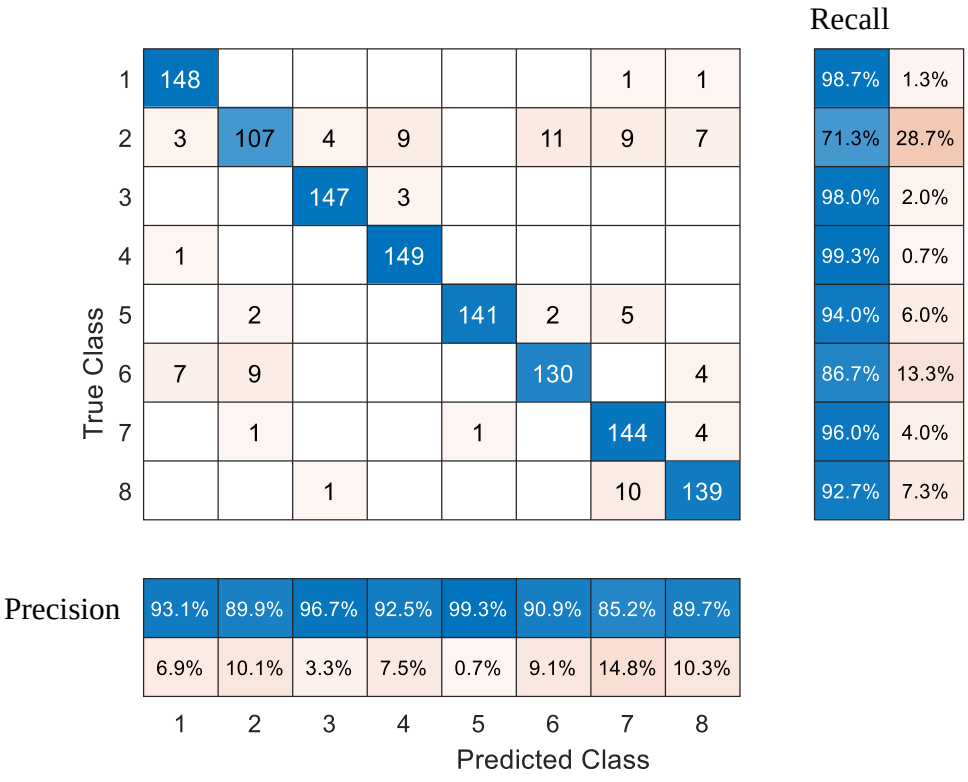

Figure 2. Confusion matrix of HSI pixels for eight classes. (Class 1: Asphalt; Class 2: Crack; Class 3: Oil; Class 4: Artificial marking; Class 5: Green vegetation; Class 6: Water; Class 7: Dry vegetation; Class 8: Concrete)

In this study, the trained classification model resulted in an overall accuracy of 92.1%, that is, 92.1% HSI pixels in testing set are correctly classified. The averaged precision, recall, and F1 scores for all 8 classes are 92.2%, 92.1%, and 91.9%, respectively. It reveals that, in general, the trained model achieves satisfactory classification performance on all eight different classes. In addition, the general equilibrium of average precision and recall also indicates that the trained model is balanced to both positive and negative samples for each class. In summary, the results show that the proposed method can predict cracks and other materials under complex scenes.

CONCLUSIONS

This study focuses on the application of HSI for crack detection and identification of surface materials under complex scenes. The proposed method uses a low-rank representation-based approach to build a multi-class classifier for the prediction of cracks and other surface materials. The method has some unique characteristics, including the use of dictionary learning process, which reduces redundancy and increases discriminative ability for HSI classification. The introduction of a locality constraint term encourages similar HSI pixels to have similar representations, leading to high classification accuracy. The trained classification model was tested on a real-world HSI crack dataset captured from the UAV. The results showed that the trained model achieved satisfactory classification performance and high efficiency, with an overall accuracy of 92.1%.

REFERENCES

1. Aboudi, J. 1987. “Stiffness Reduction of Cracked Solids,” *Eng. Fract. Mech.,* 26(5):637-50.
2. Das, S., P. Saha, and S. Patro. 2016. “Vibration-based Damage Detection Techniques Used for Health Monitoring of Structures: A Review,” *J. Civ. Struct. Health,* 6(3):477-507.
3. Phares, B. M., G. A. Washer, D. D. Rolander, B. A. Graybeal, and M. Moore. 2004. “Routine Highway Bridge Inspection Condition Documentation Accuracy and Reliability,” *J. Bridge. Eng.,* 9(4):403-13.
4. Gallagher, R. P., P. Reasenberg, and C. D. Poland. 1999. *Earthquake Aftershocks: Entering Damaged Buildings: Applied Technology Council*.
5. Deng, J., A. Singh, Y. Zhou, Y. Lu, and V. C. S. Lee. 2022. “Review on Computer Vision-based Crack Detection and Quantification Methodologies for Civil Structures,” *Constr. Build. Mater.,* 356:129238.
6. Sun, L., Z. Shang, Y. Xia, S. Bhowmick, and S. Nagarajaiah. 2020. “Review of Bridge Structural Health Monitoring Aided by Big Data and Artificial Intelligence: From Condition Assessment to Damage Detection.” *J. Struct. Eng.,* 146(5):04020073.
7. Lee, A. J., W. Song, B. Yu, D. Choi, C. Tirtawardhana, and H. Myung. 2023. “Survey of Robotics Technologies for Civil Infrastructure Inspection,” *J. Infrastruct. Intell. Resil.,* 2(1):100018.
8. Kamaliardakani, M., L. Sun, and M. K. Ardakani. 2016. “Sealed-crack Detection Algorithm Using Heuristic Thresholding Approach,” *J. Comput. Civ. Eng.,* 30(1):04014110.
9. Dorafshan, S., R. J. Thomas, and M. Maguire. 2018. “Comparison of Deep Convolutional Neural Networks and Edge Detectors for Image-based Crack Detection in Concrete,” *Constr. Build. Mater.,* 186:1031-45.
10. Kim, H., E. Ahn, S. Cho, M. Shin, and S. H. Sim. 2017. “Comparative Analysis of Image Binarization Methods for Crack Identification in Concrete Structures,” *Cem. Concr. Res.,* 99:53-61.

11. Qu, Z., L. D. Lin, Y. Guo, and N. Wang. 2015. “An Improved Algorithm for Image Crack Detection Based on Percolation Model,” *IEEJ Trans. Electr. Electron. Eng.,* 10(2):214-21.
12. Abdel-Qader, I., O. Abudayyeh, and M. E. Kelly. 2003. “Analysis of Edge-Detection Techniques for Crack Identification in Bridges,” *J. Comput. Civ. Eng.*, 17(4):255-63.
13. Hsieh, Y. A., and Y. J. Tsai. 2020. “Machine Learning for Crack Detection: Review and Model Performance Comparison,” *J. Comput. Civ. Eng.,* 34(5):04020038.
14. Ali, R., J. H. Chuah, M. S. A. Talip, N. Mokhtar, and M. A. Shoaib. 2022. “Structural Crack Detection Using Deep Convolutional Neural Networks,” *Autom. Constr.,* 133:103989.
15. Cha, Y. J., W. Choi, and O. Büyüköztürk. 2017. “Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks,” *Comput.-Aided Civ. Infrastruct. Eng*., 32(5):361-78.
16. Chen, F. C, and M. R. Jahanshahi. 2018. “NB-CNN: Deep Learning-Based Crack Detection Using Convolutional Neural Network and Naïve Bayes Data Fusion,” *IEEE Trans. Ind. Electron.,* 65(5):4392-400.
17. Jiang, K, Q. Han, X. Du, and P. Ni. 2021. “A Decentralized Unsupervised Structural Condition Diagnosis Approach Using Deep Auto-encoders,” *Comput.-Aided Civ. Infrastruct. Eng.,* 36(6):711-32.
18. Bang, S., S. Park, H. Kim, and H. Kim. 2019. “Encoder-decoder Network for Pixel-level Road Crack Detection in Black-box Images,” *Comput.-Aided Civ. Infrastruct. Eng.,* 34(8):713-27.
19. Zheng, Y., X. Zhang, S. Yang, and L. Jiao. 2013. “Low-rank Representation with Local Constraint for Graph construction,” *Neurocomputing,* 122:398-405.
20. Hajiani, F., N. Parhizgar, and A. Keshavarz. 2021. “Hyperspectral Image Classification Using Cluster based Graph Regularized Low Rank Representation and Dictionary Learning,” *Neurocomputing,* 462:208-20.
21. Chen, Z. 2020. Hyperspectral Imagery for Material Surface Damage. Figshare2020.
22. Aryal, S., Z. Chen, and S. Tang. 2021. “Mobile Hyperspectral Imaging for Material Surface Damage Detection,” *J. Comput. Civ. Eng.*, 35(1).