

A high precision crack classification system using multi-layered image processing and deep belief learning

Author

Jo, Jun, Jadidi, Zahra

Published

2019

Journal Title

Structure and Infrastructure Engineering

Version

Accepted Manuscript (AM)

DOI

[10.1080/15732479.2019.1655068](https://doi.org/10.1080/15732479.2019.1655068)

Rights statement

This is an Author's Accepted Manuscript of an article published in Structure and Infrastructure Engineering, 16 (2), pp. 297-305, 21 Aug 2019, copyright Taylor & Francis, available online at: <https://doi.org/10.1080/15732479.2019.1655068>

Downloaded from

<http://hdl.handle.net/10072/392244>

Griffith Research Online

<https://research-repository.griffith.edu.au>

A High Precision Crack Classification System using Multi-layered Image Processing and Deep Belief Learning

Jun Jo¹, Zahra Jadidi²

¹ School of Information and Communication Technology, Griffith University, Gold Coast campus, 4215, Australia.

² Queensland University of Technology (QUT), Brisbane, Australia

A High Precision Crack Classification System using Multi-layered Image Processing and Deep Belief Learning

Road surfaces experience fatigue stress and loading, which often lead to cracks on the surface. The cracks might cause serious damage, and therefore, early detection can reduce the road maintenance cost. Traditional inspection methods are carried out by humans and are slow, costly and hazardous. To improve accuracy and reduce the hazards of current crack detection methods, this paper proposes a new autonomous crack detection system (ACDS) that can be used in any autonomous vehicles (UAVs). ACDS consists of three stages: image acquisition, image processing, and classification. The image processing stage consists of five parallel filtering methods, which remove noise and extract features from the images. In the classification stage, five deep belief network (DBN) classifiers separately analyse the images to detect cracks. The dataset used in this paper contains 15000 RGB and infrared images, with or without cracks. The results show the high precision of the proposed system.

Keywords: crack detection; image processing; Hessian filter; Gabor filter; Otsu filter; Retinex filter; Median filter; deep belief network; classification;

1. Introduction

Traditional human-based inspection methods are slow, costly and often dangerous. Many attempts have been made to automate crack detection (Hinton, 2009; Jo, Jadidi, & Stantic, 2017; Mohan & Poobal, 2017; Shi & Malik, 2000). The difficulties in detecting cracks are due to their random and irregular shapes, the sizes of cracks, and the background noises in images (Bu et al., 2015; Wang & Huang, 2010). These issues make it difficult to achieve high accuracy in the crack detection task. A number of automated crack detection techniques have been proposed to improve the detection rate using RGB and infrared images (Cha, Choi, & Büyüköztürk, 2017; Fujita & Hamamoto, 2011; Nishikawa, Yoshida, Sugiyama, & Fujino, 2012; Prasanna et al., 2016; Sham, Chen, & Long, 2008; Su, Li, Fang, & Wen, 2017; Zhang, Yang, Zhang, & Zhu, 2016).

A threshold-based automated crack detector, composed of pre-processing, image segmentation and feature extraction was proposed by (Yiyang, 2014). The threshold method is used after smoothing the input image during the pre-processing. This approach calculates the area and perimeter of the roundness index to analyse the image and finds cracks through comparison. Gabor filtering has also been widely used to detect cracks (Salman, Mathavan, Kamal, & Rahman, 2013). Another automatic approach uses median filtering to remove shading and detect cracks in noisy concrete surfaces (Fujita & Hamamoto, 2011). This method uses a multiscale line filter with a Hessian matrix to emphasise the cracks. Finally, an adaptive threshold algorithm is used to detect cracks.

Image processing techniques are responsible for extracting features such as crack shapes, then these rich features should be analysed and identified. Machine learning methods are utilised to analyse and classify the extracted features. For example, support vector machine (SVM) and Wavelet analysis methods are used in (Huo, Zhang, Zhou, & Huang, 2017) to classify crack features in the rotating shaft. In recent years, deep learning has been deployed by many studies to reduce the dimensionality of data and learn the data features in a hierarchical method. Deep learning can achieve higher accuracy due to its larger number of layers. Crack detection methods using a deep convolutional neural network (CNN) method are proposed by (Cha et al., 2017; Chen & Jahanshahi, 2018; Ding et al., 2018). The CNN method (Kang & Cha, 2018) was optimised in another paper (Cha, Choi, Suh, Mahmoudkhani, & Büyüköztürk, 2018) and a region-based CNN (R-CNN) was used for surface damage detection. The results confirmed the high efficiency of deep learning in crack detection. This study evaluates the performance of a deep belief network (DBN) in crack detection and compares its outputs with other methods.

An infrared image-based method is also deployed in (Rodríguez-Martín, Lagüela, González-Aguilera, & Martínez, 2016). The paper proposes an algorithm for the

extraction of isotherms, in order to detect cracks and the geometric characters. In addition, it detects the orientation of the crack and hence, it could predict the direction of propagation of the crack. Another infrared image processing method based on super-pixel (Xu, Xie, Chen, & Huang, 2014) uses a fuzzy c-means clustering for segmentation. In this method, the super-pixels are selected from raw grey images and high pass filtered images.

In this study, RGB and infrared crack images were obtained from road surfaces and used to perform crack detection. This paper proposes an autonomous crack detection system (ACDS) that provides high-quality processing of input images. This system adopts a parallel image processing method and a deep learning module that extracts and trains features of cracks. This paper studies the classification of the cracks, and the quantification of crack patterns is not investigated.

Cracks on the road surface disturb the current temperature profile in the crack area. An infrared camera can be employed to visualise these temperature changes and therefore combining infrared and RGB features of cracks will be beneficial to extract all significant features of cracks and will improve the accuracy. This study proposes a deep learning based crack detection in which the features are learned from RGB and IR images. Deep learning based detectors analyses the outputs of filter-based detectors to mitigate the impact of background noise on the performance. To the best of our knowledge, this is the first paper that combines IR and RGB images to improve the detection rate of deep learning based classifiers. Two low-cost cameras were used for the data collection. The ACDS was evaluated in two separate experiments. It was initially tested with RGB dataset. Afterward, the efficiency of the hybrid images (with infrared+RGB features) in improving the precision was evaluated using the infrared+RGB dataset. The results showed that the hybrid dataset led the ACDS to provide higher precision.

Another contribution of this study is the employment of a multi-layer method to extract the most efficient features from the crack images. The ACDS then uses the DBNs to accurately analyse features and detect cracks.

2. ACDS Architecture

The proposed deep learning-based ACDS is composed of three consecutive stages as shown in Figure 1: image acquisition, image processing, and classification. In image acquisition, both infrared and RGB images are collected. A FlirOne camera (<http://www.flir.com.au/flirone/ios/>) was used to capture infrared images from road surfaces. The RGB images were obtained from two resources: images manually collected from the road surface using an RGB camera; and an online CrackForest dataset (<https://github.com/cuilimeng/CrackForest-dataset>). Both the RGB and infrared images were processed through the multi-layered image processing module.

The image processing technique contains five parallel methods: Hessian-based method, Gabor filter, Otsu, Retinex filter, and Median filter. The output features of each method are used to train the dedicated DBN and the ACDS detects cracks on the road surfaces based on the trained model. The deep learning process in this system classifies the images into crack and non-crack categories. The average of the DBN results is used to analyse the final output. Figure 1 shows the architecture of ACDS system. This study combines deep learning with well-known image processing techniques, which helps to boost the power of deep learning by analysing filtered and rich features. The DBNs in ACDS are used as classifiers responsible for the classification of the input features into crack and non-crack groups.

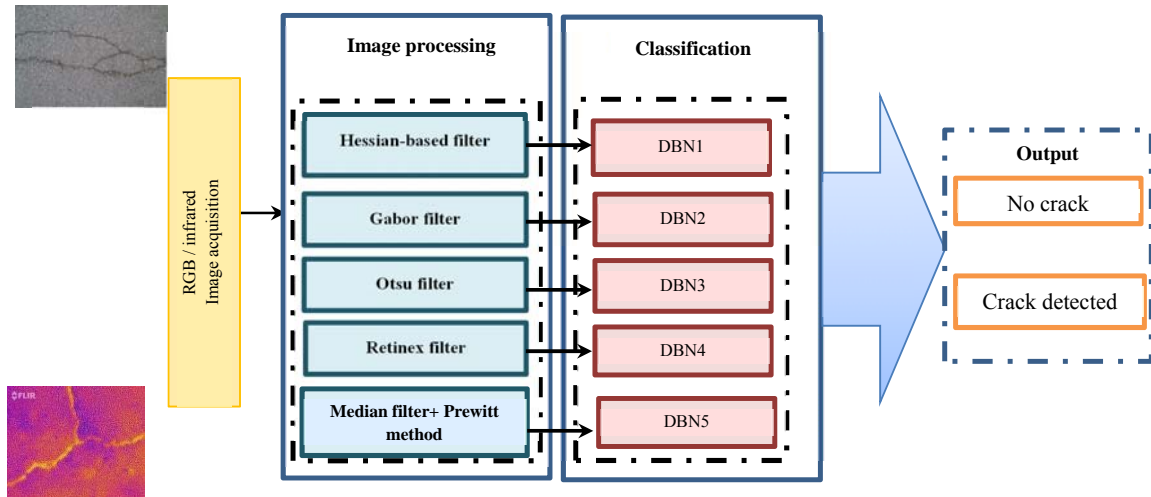


Figure 1: The system architecture of ACDS

3. Materials and Methods

3.1 Image acquisition and dataset

The infrared images contain thermal features that reflect information hidden in RGB images. The infrared images are useful for detecting small cracks when a crack surface creates heat difference (for example in sunlight). Invisible cracks can often be detected by infrared images (Sham et al., 2008). RGB images are helpful when there is enough contrast between the cracks and the surrounding areas. This research deals with a combination of these two image modes. 920 RGB and infrared images were collected in this stage.

Figure 2 shows a number of images from the training dataset, that contain various cracks in different illumination conditions. To collect the training dataset, 311 RGB and 480 infrared images were captured from different roads. In addition, a collection of 130 road crack images from the CrackForest dataset (Ding et al., 2018), was used to train the crack detector. The resolution of images in CrackForest dataset was 480x320 pixels.

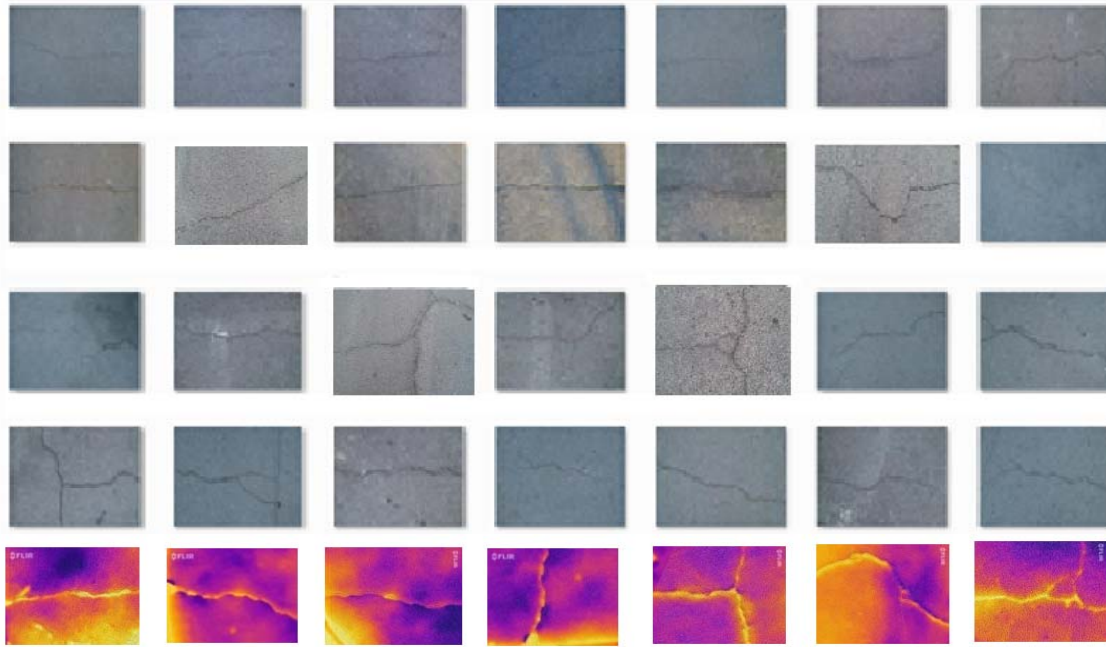


Figure 2: Road crack images in the dataset

The following cameras were used in this study:

- An iPhone SE with 12-megapixel RGB camera (resolution was 3246x2448 pixels)
- A thermal infrared camera, which has the capability of detecting fine temperature changes caused by cracks occluded by objects such as grass or moss

A Flir One Thermal Imaging Camera for IOS was used in this study to collect the infrared images in 480 x 640 pixels. This infrared camera plugs into the iPhone SE and is capable of measuring temperature differences as small as 0.1 °C. FlirOne has two camera lenses and can capture a single shot in two different modes, infrared and RGB. Both of these modes are shown in a single image. Therefore, the infrared images collected by FlirOne include both infrared and RGB features, and hence, these images are called infrared+RGB in this study. As the FlirOne shooting conditions, such as light intensity, temperature, and wind, affect the characteristics of the thermal images, all infrared+RGB images were collected in similar weather conditions. The images were captured on 15th April 2018 in the morning. The weather condition was: sunny, temperature 25 °C, humidity 45%, and no wind.

In this stage, the size of the collected dataset was not enough to prove the effectiveness of the proposed ACDS. Therefore, to provide an accurate evaluation of ACDS performance, each image was converted to 99x99 pixel image patches (Figure 3) (Zhang, Yang, Zhang, & Zhu, 2016). Out of the generated image patches, the duplicated patches were removed and then, 10000 samples were randomly selected for training the DBNs and 5000 samples were used as the testing set for the evaluation of the crack detection, as shown in Table 1. The numbers of crack and non-crack patches are equal in both training and testing datasets.

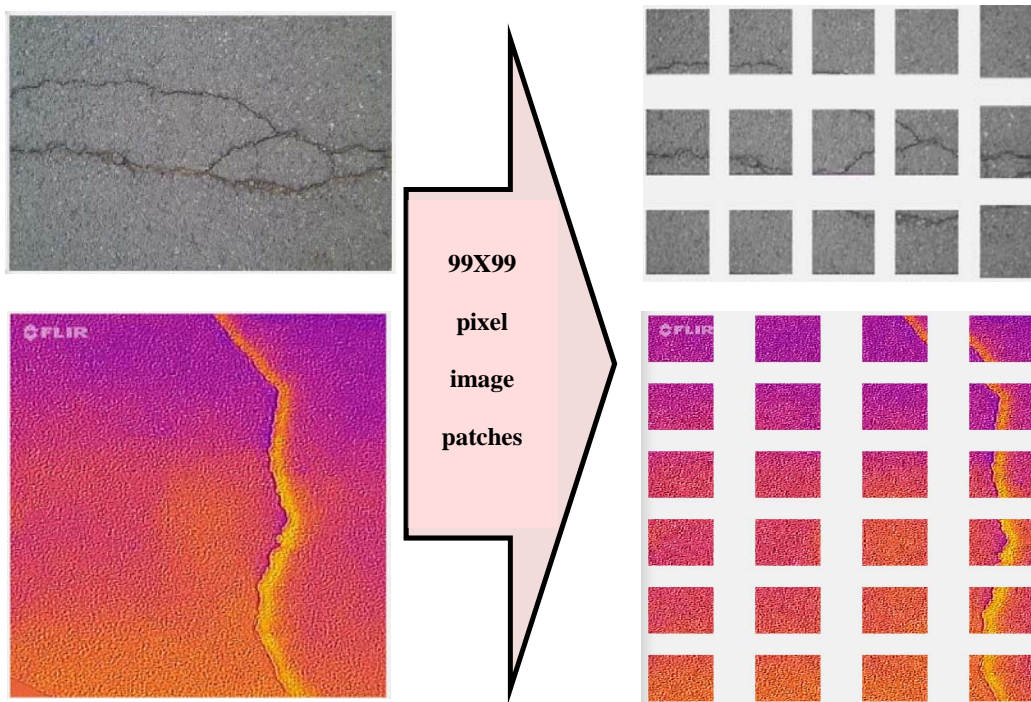


Figure 3: Images patches created from the original dataset

Table 1: The training and testing datasets

Image type	Original images	Generated images	Training images	Testing images	The ratio between crack and non-crack images
RGB	440 (collected images + 130 CrackForest)	7500	5000	2500	1
Infrared+RGB	480	7500	5000	2500	1
Total	920	15000	1000	5000	1

3.2 *Image processing*

The image processing stage includes: i) receiving an image from the input; ii) removing noise, and extracting features of segmentation, edges and background; and iii) sending the features to the classifiers.

Five parallel processing methods are used in the ACDS: Retinex, Hessian-based method, Gabor filter, Otsu and Median filter including edge detection. All of these methods process the input image and the outputs of this stage are transferred to the connected DBN classifiers. The average of the five DBNs is the final output.

Median filtering, Gabor filtering and Otsu's method used in ACDS are three well-known methods in crack detection (Mohan & Poobal, 2017). These three methods, as well as Hessian-based filtering and Retinex filtering, were used in parallel to improve the accuracy of crack detection in this study. The contribution of this study is the employment of a multi-layer method to extract the most efficient features from the crack images. The ACDS then uses the DBNs to accurately analyse features and detect cracks. The efficiency of the proposed ACDS method will be shown by comparing its performances with other methods. To the best of our knowledge, this is the first time that a layered image processing method has been applied to road crack detection.

3.2.1 *Retinex filtering*

The non-local Retinex method (Zosso, Tran, & Osher, 2013) used in this study is an illumination correction method. This method combines previous variational models of Retinex and focuses on the reflectance function to improve the performance. The Retinex model receives the multiplication of the illumination and the true underlying reflectance of the object as an input (Zosso et al., 2013), then searches for a filtered gradient and calculates the reflectance gradient of the observed image.

This combined Retinex, which takes into consideration the sparsity and fidelity of the reflectance, can be used for texture-preserving shadow removal and colour and hyperspectral image enhancement. The proposed ACDS method uses an adaptive smoothing, based on the Retinex theory to remove illumination effects.

3.2.2 Hessian-based filter

The Hessian-based filter has been widely applied to cerebral angiograms for enhancing the visualization of intracranial aneurysms (Frangi, Niessen, Vincken, & Viergever, 1998; Jerman, Pernuš, Likar, & Špiclin, 2016a, 2016b; Li & Sone, 2003; Wiemker et al., 2013; Zosso et al., 2013). A novel blob enhancement filter based on a modified volume ratio of Hessian eigenvalues was proposed by (Jerman et al., 2016b). This modified method showed a more uniform response inside the blob-like structures compared to other methods. The results also proved that this modified filter is independent of the size and intensity of structures, and it is capable of detecting small blob-like structures such as aneurysms.

Due to the high performance of this modified filter, it was used as one layer of ACDS in this study to enhance the visualization of cracks in the images captured from road surfaces and to improve the performance of crack detection. The Hessian filter has been considered as a method of crack detection in a number of studies (Fujita & Hamamoto, 2011; Fujita, Mitani, & Hamamoto, 2006).

3.2.3 Gabor filter

This filter is used to distinguish cracks from the background in a crack image. A Gabor filter-bank proposed in (Jain & Farrokhnia, 1991) was implemented in this study. Gabor filtering can be used to estimate the mean and standard deviation of the energy of the filtered image. A Gabor impulse response has a sinusoidal plane wave of orientation and

frequency. This can be modulated by a two-dimensional Gaussian envelope as defined by (El-Tarhouni, Boubchir, Al-Maadeed, Elbendak, & Bouridane, 2016; Jain & Farrokhnia, 1991):

$$h(x, y) = \exp\left\{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right\} \cos(2\pi U_x + \varphi) \quad (1)$$

U_x and φ show the frequency and phase of the sinusoidal plane wave. σ_x and σ_y represent the space constants of the Gaussian envelop. The Gabor filter-bank used in this research has Gabor filters with a Gaussian kernel function, which is based on sinusoidal plane waves of various orientations from the same Gabor-root filter:

$$g_{m,n}(x, y) = a^{-m} h(x', y'), a > 1 \quad (2)$$

Where x , y and θ are defined as follows:

$$x' = x \cos \theta + y \sin \theta \quad (3)$$

$$y' = -x \sin \theta + y \cos \theta$$

$$\theta = \frac{n\pi}{k} (k = \text{total orientation}, n = 0, 1, \dots, k - 1, m = 0, 1, \dots, s - 1) \quad (4)$$

For the image $I_E(r, c)$ with size $H \times W$, the Gabor filtered output is as:

$$I_g(r, c) = \sum_{s,t} I_E(r - s, c - t) g_{m,n}(s, t) \quad (5)$$

Based on this convolution, the energy of the filtered image is used to predict the mean and standard deviation, which are used as features. In this study, the Gabor filter is used for texture segmentation

3.2.4 Otsu

Otsu is an automatic threshold selection providing with high contrast segmentation (Otsu, 1979). This method can be used to select a threshold by analysing the distribution of grey values of an image. Otsu is based on grey histogram and it uses maximum inter-class variance between target and background to find the best segmentation threshold (Garcia,

2010; Hoang, 2018). This method has shown high performance in crack detection (Mohan & Poobal, 2017) and is used as a layer of the image processing phase of ACDS.

3.2.5 Median filter

A Median filter is an image processing method with high accuracy in crack detection (Mohan & Poobal, 2017; Kang & Cha, 2018). In ACDS, the output of this method is fed into an edge detector, the Prewitt method (Senthilkumaran & Rajesh, 2009), to show the edges of the crack.

In this study, the Median filter was used to improve the capability of detecting a crack in noisy images (Kang & Cha, 2018). Gabor filtering was chosen because it is a potential technique for multidirectional crack detection. Therefore, it can help the proposed ACDS to detect cracks aligned to different directions (Mohan & Poobal, 2017). Otsu has been widely used in crack detection due to its computational efficiency, which is due to the calculation of the threshold of grey intensity by the interclass separability. Due to this advantage of Otsu, this thresholding algorithm is used in ACDS (Hoang, 2018). Hessian-based multi-scale line filter is another method used to remove the background noise. Retinex is an image enhancement method, which synthesises a high local contrast using a non-linear spatial transform, and it helps to remove shadow (Hines, Rahman, Jobson, & Woodell, 2004).

Each of the image processing methods in this study has shown high accuracy in feature extraction. However, the performance of these methods depends on the image quality and the background surface. Due to the high complexity of the background noise in a road surface, a layered image processing method can utilise the advantages of all methods and reduce false negative alarms.

3.3 Deep Belief Network

The DBN (Bengio, 2009; Carreira-Perpinan & Hinton, 2005; Chen, Zhao, & Jia, 2015; Hinton, Osindero, & Teh, 2006; Larochelle, Erhan, Courville, Bergstra, & Bengio, 2007; Senthilkumaran & Rajesh, 2009) has become one of the most popular methods of deep learning. In this method, a generative model is employed for pre-training, and back-propagation is used in the fine-tuning stage. This fast learning algorithm can find the optimal parameters very quickly. The DBN is a powerful hierarchical neural network beneficial for feature extraction. The input layer receives images as inputs. The features of these images will be transferred through multiple hidden layers and new features will be extracted. Then, the output layer is responsible for detecting the class of features using a linear classifier. The DBN provides an accurate classifier, which can detect cracks based on the optimised images received from the image processing stage (Hinton, 2009).

The DBN has several hidden layers, representing non-linear patterns in training data. It consists of a stack of Restricted Boltzmann Machines (RBMs) and uses a layer-wise pre-training phase. Each layer is trained by an unsupervised algorithm to fix the parameters of that layer, and then the next layer will be trained on top of the data that is received from the previous hidden layer. Once pre-training is completed, the output layer is added and the whole network will be trained using a supervised learning algorithm. This phase is called fine-tuning. The joint distribution of observable data x and hidden layers h^i is defined as Equation (6), where θ shows the network parameters, weights, and biases (Wu et al., 2017).

$$P(x, h^1, \dots, h^k | \theta) = P(h^{k-1}, h^k | \theta) \prod_{i=2}^{k-1} P(h^{i-1} | h^i, \theta) P(x | h^1, \theta) \quad (6)$$

DBN has a layer-by-layer learning system. In another word, the learned features of the first RBM are used as the input data for the second RBM (Le Roux & Bengio, 2010; Salakhutdinov & Larochelle, 2010). This method helps DBN to extract the deep

features of the input data. Then, pre-training in a hierarchical method is employed in the DBN to learn these deep features of the input. This research employed a two-layer DBN, in which different numbers of hidden layers were manually tested to find the most efficient structure. There were 800 nodes in each hidden layer and 2 nodes in the output layer. Figure 4 shows the structure of a two-layer DBN.

After training the RBMs, the whole pre-trained network is fine-tuned to integrate the DBN layers to classify inputs based on the learned features (Hecht-Nielsen,1992; Sutskever &Hinton,2008). To evaluate the performance of the proposed ACDS, precision (P) and recall (R) are defined as Equations (7) and (8) (Zhang et al.,2016). False negative shows the wrong detection of cracks, and false positive reflects the wrong detection of non-crack images. True positive shows the correct detection of non-crack images.

$$P = \frac{\text{true positive}}{\text{true positive} + \text{false positive}} \quad (7)$$

$$R = \frac{\text{true positive}}{\text{true positive} + \text{false negative}} \quad (8)$$

Then, the F1 score is expressed as:

$$F_1 = \frac{2PR}{P+R} \quad (9)$$

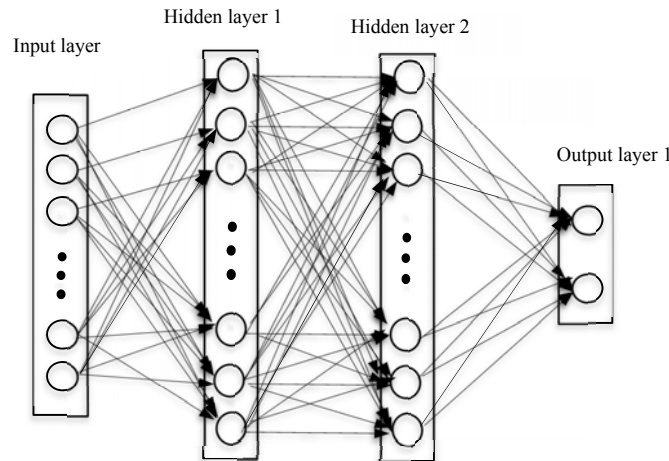


Figure 4: A DBN with two layers RBM

4. Results

DBNs with two hidden layers were used for classification. Each hidden layer had 800 nodes, and the output layer had 2 nodes. In this study, DBN-based classifiers were simulated in MATLAB R2017b. The results were saved after 500 iterations of the algorithm training. The data set comprised 15000 images, of which 10000 samples were used for training purposes. Then, the classifiers were separately tested using the RGB and infrared+RGB testing datasets. For each image, five image processing methods were employed to remove noise and extract features. The output of each image processing method is shown in Figure 5. The cracks in infrared samples became visible due to the heat difference. After the extracted features of training images were used to train the DBN classifiers, these classifiers became capable of detecting crack images in a new dataset. To evaluate the performance of the DBNs, they were tested with features extracted from testing images using five image processing methods.

The root-mean-square-error (RMSE) rates over different iterations are shown in Figure 6. RMSE value represents the difference between the predicted output and the desired output. Figure 6 shows the RMSE error for every 100 iterations and that the percentage of RMSE is reduced when the iteration number increases, and is stabilized after 100 iterations in all methods. Minimising RMSE is important because the predicted output should be very close to the actual output. Figure 6 shows the results for the infrared+RGB testing images. In this research, the classifiers were evaluated in two phases. Initially, each classifier was trained and evaluated using the same dataset to ensure that training appropriately adjusted the weight coefficients of the classifier. The system required the high performance of the classifiers in the training dataset. Secondly, the classifiers were evaluated with new samples, as they should be able to deal with the combination of known and new samples in the real world.

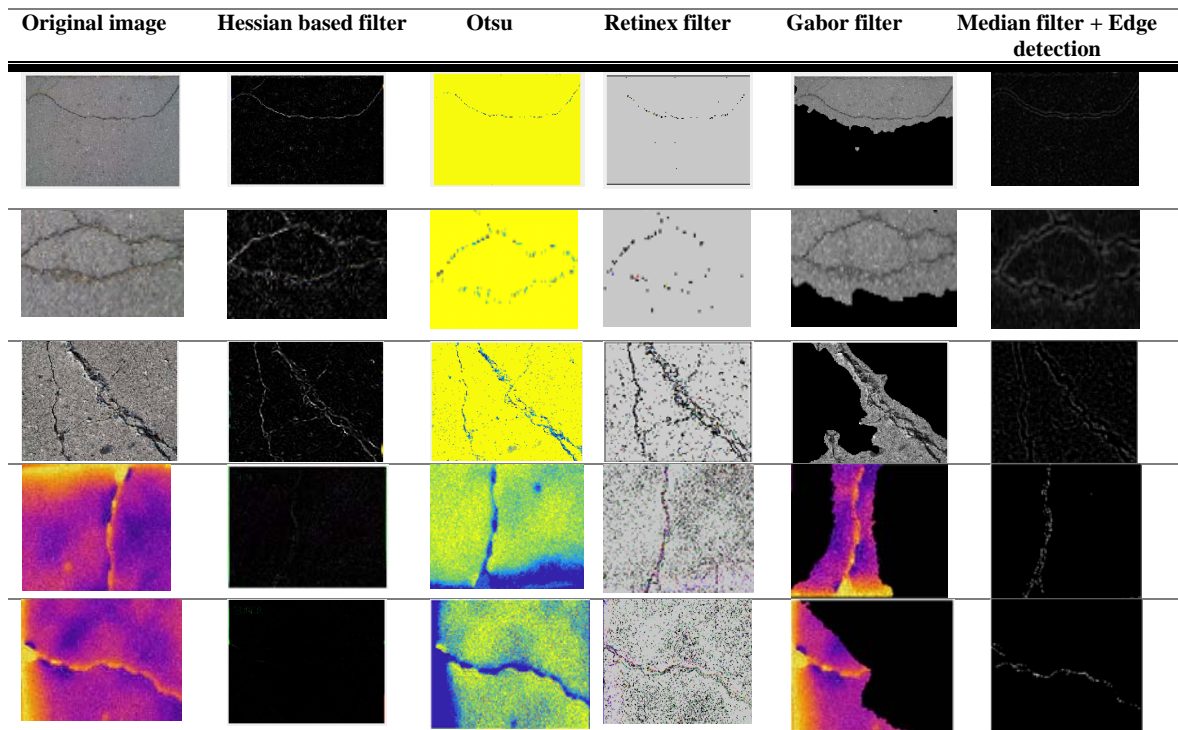


Figure 5: Comparison of the outputs of different image processing methods for crack images (from left to right: Original image, Hessian-based filter, Otsu, Retinex filter, Gabor filter, and Median filter + Edge detection)

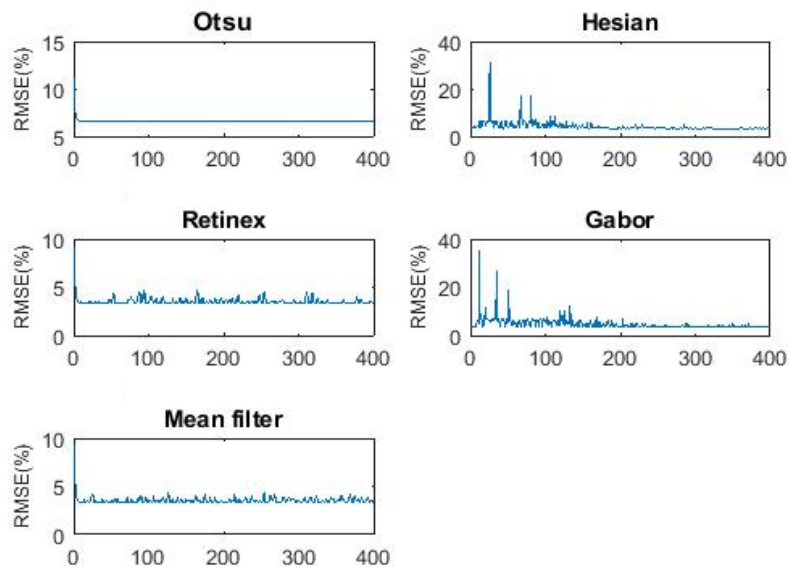


Figure 6: RMSE versus fine-tuning iterations for DBNs trained and tested with different image processing methods

The maximum, minimum and average precision results of the ACDS over 10 experiments are shown in Table 2. The evaluation results are provided for both training and testing datasets. The results show the high performance of ACDS in both known and new images. As can be seen in Table 2, the ACDS trained by infrared+RGB outperforms the ACDS trained by RGB only images, because of the combination of RGB and infrared features in infrared+RGB images.

For further evaluation, the averaged performance of ACDS is compared with other studies in Table 3. As the results show, the precision of ADCS with infrared+RGB is superior to the other methods'. Crack detection using infrared+RGB images helps the system extract features of thermal and RGB information together, and therefore, ACDS can detect even thin cracks via the existence of heat difference.

Table 2. The precision of ACDS in detecting cracks in training and testing datasets

Type of image		Testing dataset = Training dataset	Testing dataset (new samples)
RGB	Minimum Precision	92.41	88.11
	Average Precision	95.12	90.18
	Maximum Precision	98.21	95.13
Infrared+RGB	Minimum Precision	92.26	90.63
	Average Precision	97.63	91.98
	Maximum Precision	98.91	95.36

Table 3. The precision of ACDS in comparison with other methods

Method	Reference	Precision	Recall	F1 Score
SVM	(Zhang et al., 2016)	0.81	0.67	0.74
Boosting	(Zhang et al., 2016)	0.74	0.76	0.75
ConvNets	(Zhang et al., 2016)	0.87	0.92	0.89
ACDS (RGB)	0.90	0.88	0.87
ACDS (Infrared+RGB)	0.92	0.93	0.91

SVM and Boosting methods (Zhang et al., 2016) were used to evaluate images of smaller size (300X300) compared with the ACDS (Table 3). For these methods, the features were automatically generated by manually annotated image patches. The SVM and Boosting methods were trained with features, which were based on the colour and texture of patches. The patches had binary labels representing crack and non-crack groups. As Table 3 shows, the ACDS still outperforms these methods. Compared with the ConvNets used in (Zhang et al., 2016), ACDS provided better performance especially when it was trained with hybrid images.

All ACDS experiments were performed using an Intel(R) Core™ i5-3230M CPU @2.60 GHZ with 8GB RAM. As shown in Table 4, the recorded training speed for ACDS was about 63 minutes in which the averaged feature extraction time was about 20 minutes.

Table 4. The processing time of ACDS

Method	Average processing time (min)	Max. processing time (min)	Min. processing time (min)
ACDS (RGB)	61	64	59
ACDS (Infrared+RGB)	63	65	60

A key challenge for UAVs is power consumption. However, novel machine learning algorithms introduce high computational workload and affect the battery and flight time of the UAV. Convolutional Neural Networks should use Graphics Processing Units (GPUs) to process inputs and reach high performance. However, GPU-based analysis generates a high overhead which is inappropriate for low-power UAVs (Kyrkou et al., 2015).

In this paper, cloud-based processing is required to integrate the proposed ACDS into UAVs with two cameras. The UAV will transfer its collected images to a cloud server

for analysis. Therefore, image processing and classification tasks can be performed using a CPU processor on the cloud server. By off-loading the computational operations, energy consumption will be reduced. Therefore, the proposed ACDS can be implemented in a small and low-price UAV.

5. Conclusions

This paper proposed an autonomous crack detection system to reduce the maintenance cost for civil infrastructures. The ACDS could improve the performance of crack detection on road surfaces. Two types of images, RGB and infrared, were used for analysis by the image processing stage of the ACDS. Infrared images can represent features hidden in RGB images and this is useful for detecting small cracks with visible thermal features. In this study, the ACDS was evaluated separately with RGB and the hybrid images (with infrared+RGB features) to evaluate the impact of infrared features on the detection rate of deep learning based classifiers. This was the contribution of this study.

The proposed image processing stage of ACDS included five parallel feature extraction methods. These features were utilised to train the deep learning-based classifiers. The performance parameters of the proposed ACDS were compared with other studies and the results showed improvement in the performance. The precision of the ACDS in detecting cracks in RGB images was 90.18%, about 10% more than the SVM and Boosting methods in another study. This precision increased to 91.98% when the ACDS was trained with both infrared and RGB features. Another contribution of this study is the deployment of a multi-layer image processing method to improve the precision and minimise the number of false negatives in crack detection. The features

created by the layered image processing stage were classified by deep learning into the crack and non-crack groups.

This proposed ACDS may also be applied to other structures such as bridges, buildings, etc. This study aimed to improve the precision of crack detection and examine the impact of thermal features on the performance. However, the processing time of the proposed ACDS is still higher than for a single layer crack detector. This problem will be addressed in a future study.

6. References

- Bengio, Y. (2009). *Learning deep architectures for AI*. Foundations and trends in Machine Learning, 2(1), 1-127.
- Bu, G., Chanda, S., Guan, H., Jo, J., Blumenstein, M., & Loo, Y. C. (2015). *Crack detection using a texture analysis-based technique for visual bridge inspection*. Electronic Journal of Structural Engineering, 14(1), 41-48.
- Carreira-Perpinan, M. A., & Hinton, G. E. (2005). *On contrastive divergence learning*. Tenth International Workshop on Artificial Intelligence and Statistics (Aistats), Barbados.
- Cha, Y. J., Choi, W., & Büyüköztürk, O. (2017). *Deep learning-based crack damage detection using convolutional neural networks*. Computer-Aided Civil and Infrastructure Engineering, 32(5), 361-378.
- Cha, Y. J., Choi, W., Suh, G., Mahmoudkhani, S., & Büyüköztürk, O. (2018). *Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types*. Computer-Aided Civil and Infrastructure Engineering, 33(9), 731-747.
- Chen, F.-C., & Jahanshahi, M. R. (2018). *NB-CNN: deep learning-based crack detection using convolutional neural network and naive Bayes data fusion*. IEEE Transactions on Industrial Electronics, 65(5), 4392-4400.
- Chen, Y., Zhao, X., & Jia, X. (2015). *Spectral-spatial classification of hyperspectral data based on deep belief network*. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(6), 2381-2392.
- Ding, L., Fang, W., Luo, H., Love, P. E., Zhong, B., & Ouyang, X. (2018). *A deep hybrid learning model to detect unsafe behavior: integrating convolution neural networks and long short-term memory*. Automation in Construction, 86, 118-124.
- El-Tarhouni, W., Boubchir, L., Al-Maadeed, N., Elbendak, M., & Bouridane, A. (2016). *Multispectral palmprint recognition based on local binary pattern histogram fourier features and gabor filter*. 6th European Workshop on Visual Information Processing (EUVIP 2016), (pp. 1-6), Marseille, France.
- Frangi, A. F., Niessen, W. J., Vincken, K. L., & Viergever, M. A. (1998). *Multiscale vessel enhancement filtering*. International Conference on Medical Image Computing and Computer-Assisted Intervention, (pp. 130-137), Cambridge, MA, USA.

- Fujita, Y., & Hamamoto, Y. (2011). *A robust automatic crack detection method from noisy concrete surfaces*. *Machine Vision and Applications*, 22(2), 245-254.
- Fujita, Y., Mitani, Y., & Hamamoto, Y. (2006). *A method for crack detection on a concrete structure*. 18th International Conference on the Pattern Recognition (ICPR 2006), (Vol. 3, pp. 901-904), Hong Kong, China.
- Garcia, D. (2010). *Image segmentation using Otsu thresholding*. MATLAB File Exchange, 26532.
- Hecht-Nielsen, R. (1992). *Theory of the backpropagation neural network*. *Neural networks for perception* (pp. 65-93).
- Hines, G. D., Rahman, Z. U., Jobson, D. J., & Woodell, G. A. (2004). *DSP implementation of the retinex image enhancement algorithm*. *Visual Information Processing XIII*, (SPIE Vol. 5438, pp. 13-25), Orlando, USA.
- Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). *A fast learning algorithm for deep belief nets*. *Neural computation*, 18(7), 1527-1554.
- Hinton, G. E. (2009). *Deep belief networks*. *Scholarpedia*, 4(5), 5947.
- Hoang, N.-D. (2018). *Detection of surface crack in building structures using image processing technique with an improved Otsu method for image thresholding*. *Advances in Civil Engineering*, 2018.
- Huo, Z., Zhang, Y., Zhou, Z., & Huang, J. (2017). *Crack detection in rotating shafts using wavelet analysis, Shannon entropy and multi-class SVM*. 3rd International Conference on Industrial Networks and Intelligent Systems (INISCOM 2017), (pp. 332-346), Ho Chi Minh, Vietnam.
- Jain, A. K., & Farrokhnia, F. (1991). *Unsupervised texture segmentation using Gabor filters*. *Pattern recognition*, 24(12), 1167-1186.
- Jerman, T., Pernuš, F., Likar, B., & Špiclin, Ž. (2016a). *Enhancement of vascular structures in 3D and 2D angiographic images*. *IEEE Transactions on medical imaging*, 35(9), 2107-2118.
- Jerman, T., Pernuš, F., Likar, B., & Špiclin, Ž. (2016b). *Blob enhancement and visualization for improved intracranial aneurysm detection*. *IEEE Transactions on Visualization and Computer Graphics*, 22(6), 1705-1717.
- Jo, J., Jadidi, Z., & Stantic, B. (2017). *A Drone-Based Building Inspection System Using Software-Agents*. Paper presented at the International Symposium on Intelligent and Distributed Computing (pp. 115-121). Springer, Cham.
- Kang, D., & Cha, Y. J. (2018). *Autonomous UAVs for Structural Health Monitoring Using Deep Learning and an Ultrasonic Beacon System with Geo-Tagging*. *Computer-Aided Civil and Infrastructure Engineering*, 33(10), 885-902.
- Larochelle, H., Erhan, D., Courville, A., Bergstra, J., & Bengio, Y. (2007). *An empirical evaluation of deep architectures on problems with many factors of variation*. *Proceedings of the 24th international conference on Machine learning*, (pp. 473-480), Corvallis, OR, USA.
- Le Roux, N., & Bengio, Y. (2010). *Deep belief networks are compact universal approximators*. *Neural computation*, 22(8), 2192-2207.
- Li, Q., & Sone, S. (2003). *Selective enhancement filters for nodules, vessels, and airway walls in two-and three-dimensional CT scans*. *Medical Physics*, 30(8), 2040-2051.
- Mohan, A., & Poobal, S. (2017). *Crack detection using image processing: A critical review and analysis*. *Alexandria Engineering Journal*.
- Nishikawa, T., Yoshida, J., Sugiyama, T., & Fujino, Y. (2012). *Concrete crack detection by multiple sequential image filtering*. *Computer-Aided Civil and Infrastructure Engineering*, 27(1), 29-47.

- Otsu, N. (1979). *A threshold selection method from gray-level histograms*. IEEE Transaction on systems, man, and cybernetics, 9(1), 62-66.
- Prasanna, P., Dana, K. J., Gucunski, N., Basily, B. B., La, H. M., Lim, R. S., & Parvaneh, H. (2016). *Automated Crack Detection on Concrete Bridges*. IEEE Transaction on Automation Science, 13(2), 591-599.
- Rodríguez-Martín, M., Lagüela, S., González-Aguilera, D., & Martínez, J. (2016). *Thermographic test for the geometric characterization of cracks in welding using IR image rectification*. Automation in Confusion, 61, 58-65.
- Salakhutdinov, R., & Larochelle, H. (2010). *Efficient learning of deep Boltzmann machines*. Proceedings of the thirteenth international conference on artificial intelligence and statistics (pp. 693-700), Sardinia, Italy.
- Salman, M., Mathavan, S., Kamal, K., & Rahman, M. (2013). *Pavement crack detection using the Gabor filter*. 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), Hague, Netherlands.
- Senthilkumaran, N., & Rajesh, R. (2009). *Image segmentation-a survey of soft computing approaches*. International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom'09). (pp. 844-846). Kerala, India.
- Sham, F., Chen, N., & Long, L. (2008). *Surface crack detection by flash thermography on concrete surface*. Insight-Non-Destructive Testing and Condition Monitoring, 50(5), 240-243.
- Shi, J., & Malik, J. (2000). *Normalized cuts and image segmentation*. IEEE Transactions on pattern analysis and machine intelligence, 22(8), 888-905.
- Su, H., Li, X., Fang, B., & Wen, Z. (2017). *Crack detection in hydraulic concrete structures using bending loss data of optical fiber*. Journal of Intelligent Material Systems and Structures. 28(13), 1719-1733.
- Sutskever, I., & Hinton, G. E. (2008). *Deep, narrow sigmoid belief networks are universal approximators*. Neural computation, 20(11), 2629-2636.
- Wang, P., & Huang, H. (2010). *Comparison analysis on present image-based crack detection methods in concrete structures*. In Image and Signal Processing (CISP), 2010 3rd International Congress on (Vol. 5, pp. 2530-2533). IEEE.
- Wiemker, R., Klinder, T., Bergtholdt, M., Meetz, K., Carlsen, I. C., & Bulow, T. (2013). *A radial structure tensor and its use for shape-encoding medical visualization of tubular and nodular structures*. IEEE Transactions on visualization and computer graphics, 19(3), 353-366.
- Wu, F., Wang, Z., Lu, W., Li, X., Yang, Y., Luo, J., & Zhuang, Y. (2017). *Regularized Deep Belief Network for Image Attribute Detection*. IEEE Transactions on Circuits and Systems for Video Technology, 27(7), 1464-1477.
- Xu, C., Xie, J., Chen, G., & Huang, W. (2014). *An infrared thermal image processing framework based on superpixel algorithm to detect cracks on metal surface*. Infrared Physics & Technology, 67, 266-272.
- Yiyang, Z. (2014). *The design of glass crack detection system based on image preprocessing technology*. In Information Technology and Artificial Intelligence Conference (ITAIC), 2014 IEEE 7th Joint International (pp. 39-42). IEEE.
- Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2016). *Road crack detection using deep convolutional neural network*. In Image Processing (ICIP), 2016 IEEE International Conference on (pp. 3708-3712). IEEE.
- Zosso, D., Tran, G., & Osher, S. (2013). *A unifying retinex model based on non-local differential operators*. In Computational Imaging XI (Vol. 8657, p. 865702). International Society for Optics and Photonics.

Kyrkou, C., Plastiras, G., Theocharides, T., Venieris, S. I., & Bouganis, C. S. (2018). DroNet: Efficient convolutional neural network detector for real-time UAV applications. In *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2018* (pp. 967-972). IEEE.