# Texton Based Segmentation for Road Defect Detection from Aerial Imagery

Adhi Prahara[a,1,*], Son Ali Akbar[b,c,2], Ahmad Azhari[a,3]

*[a] Department of Informatics, Universitas Ahmad Dahlan, Yogyakarta, Indonesia*
*[b] Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta, Indonesia*
*[c] Faculty of Electrical and Electronics Engineering, Universiti Malaysia Pahang, Pahang, Malaysia*
*[1] adhi.prahara@tif.uad.ac.id*; [2]sonali@ee.uad.ac.id; [3]ahmad..azhari@tif.uad.ac.id*
* corresponding author*

**ABSTRACT**

Road defect such as potholes and road cracks, became a problem that arose every year in Indonesia. It could endanger drivers and damage the vehicles. It also obstructed the goods distribution via land transportation that had major impact to the economy. To handle this problem, the government released an online complaints system that utilized information system and GPS technology. To follow up the complaints especially road defect problem, a survey was conducted to assess the damage. Manual survey became less effective for large road area and might disturb the traffic. Therefore, we used road aerial imagery captured by Unmanned Aerial Vehicle (UAV). The proposed method used texton combined with K-Nearest Neighbor (K-NN) to segment the road area and Support Vector Machine (SVM) to detect the road defect. Morphological operation followed by blob analysis was performed to locate, measure, and determine the type of defect. The experiment showed that the proposed method able to segment the road area and detect road defect from aerial imagery with good Boundary F1 score.

## I. Introduction

Road is an important route that connect two places to allow travel by vehicles for transporting people or distributing goods. Damaged road can endanger drivers, damage the vehicles, and delaying distribution that can affect the economy. Indonesia has more than 47,000 km length of national roads and has spent around IDR 18.7 trillion to fix road defect and road maintenance in 2017.

Road defect especially potholes and road cracks is caused by insufficient pavement thickness, insufficient drainage, failures at utility trenches and castings, and pavement defects and cracks that left unmaintained and unsealed [1]. Potholes grow from cracks on the road surface into several inches of widths or depths. Road defect can damage vehicles tires, wheels, and suspensions. In a highway, it can cause serious accidents. Therefore, the government releases an online complaint system that utilizes information system and GPS technology.

The online complaints application allows the government to identify road defect faster because it takes advantage of community participation. User submits a complaint through the system and wait for verification. The response time depends on the level of report. After verification, the corresponding agencies will process the report and perform a survey on the road defect location to assess the level of damage. However, manual survey become less effective if the road area is large and may need to block the traffic to measure the defect. To overcome this problems, researchers develop methods to detect road defect from road observation image.

Based on the road observation result, researchers use image processing method [2-5], depth sensor and stereo vision [6-8], and accelerometer [9] or combination from all data to detect road defect. Koch and Brilakis [2] use several steps to detect road defect from image. Input image is segmented to extract

the area of defect using threshold on histogram based shape. From the segmentation result, they estimate the shape of defect using morphological operation and elliptic regression. Texture in the area of defect is compared to the surrounding area to determine the defect area. The experiment result shows good accuracy. Azhar *et al.* [3] detect potholes using shape based method. Histogram of Oriented Gradients (HOG) is used to extract feature descriptors from input image. Naïve Bayes is used to train and classify the feature to detect potholes. To localize potholes area after classification, a segmentation method using graph-cut is performed. The result shows 90% detection accuracy. Wang *et al.* [4] detect potholes into two steps. First, they extract wavelet energy from image to detect potholes using morphological operation and geometry criteria. Second, the detected potholes are segmented using Markov Random Field (MRF). Edge is extracted using edge detection method. From the experiment, the method shows 86.7% accuracy for various road condition.

Zhang *et al.* [6] detects potholes using stereo vision. They use disparity map and distance measurement from the road surface. The proposed system gives information of potholes position, size, and volume that helps corresponding agencies to perform analysis of damage. Kamal *et al.* [7] use Kinect sensor to detect potholes. Kinect sensor is used to visualize the depth of road surface and detect a curvature that assume as potential potholes. They use surface information and extract metrology to measure the level of damage.

Google in 2014 has patented a potholes detection system [9]. System is implemented on the vehicles by installing some sensors such as GPS and vertical accelerometer which sent back to Google server if vehicles passing through a pothole. Received data will be processed and the result is included in the Google map and navigation system so the user is informed and can choose safer route.

This work focuses on detecting road defect such as road cracks and potholes from aerial imagery. Our main contribution is to perform texton based segmentation to segment road area, SVM to detect the road defect, and blob analysis to locate, measure, and determine the type of defect. We use UAV to capture the road aerial imagery. The rest of this paper is organized as follows: Section 2 presents the proposed road defect detection, Section 3 presents the results and analysis, and Section 4 describes the conclusion of this work.

## II. Methods

### A. *Road Defect Detection*

In this research, road defects are detected based on texton. Texton is used to segment road area using K-Nearest Neighbor (K-NN) algorithm then classify the road defect using Support Vector Machine (SVM). We use road aerial imagery taken from Unmanned Aerial Vehicle (UAV) to cover large area. Fig. 1 shows the general procedure of the proposed method.

The proposed method consists of three steps. At the first step, texton features are extracted from training images. The features are trained using K-NN to generate road area detector and trained using SVM to generate road defect detector. At the second step, the accumulative energy extracted from the road area detector is used to segment the road area. After we find the road area, edges are extracted using edge detection and morphological operation is performed. SVM is used to detect the road defect from the edges area. Finally, blob analysis is performed to locate the road defect, measure the defect size, and determine the defect type. The methods used in this research will be explained in the following sub section.
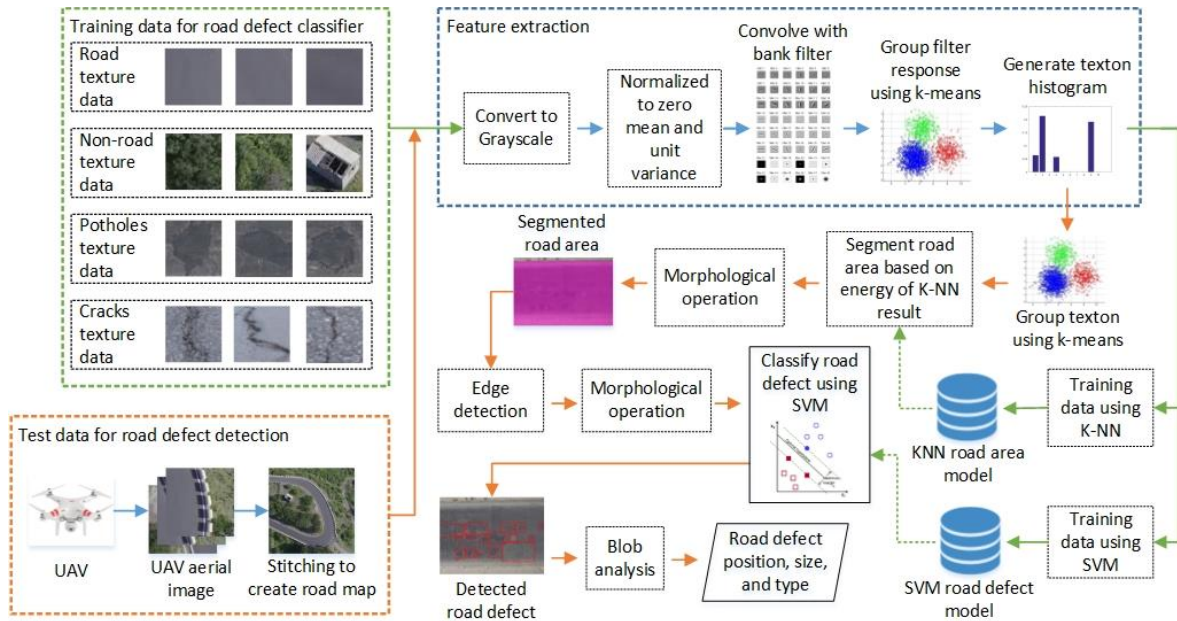
Fig. 1. General procedure of road defect detection based on texton.

### B. Texton

Texton is a representation of textures generated from the result of clustering the filter responses into a small set of prototype response vectors [10]. Filter responses can be calculated by convolving filter banks on image [11,12]. In this research, we use Leung-Malik (LM) filter banks [11]. LM-filter banks are formed from 48 filters with multi-scale variation and multi-orientation. It consists of first and second derivatives of Gaussian filters at 6 orientations on 3 scales, 8 Laplacian of Gaussian (LoG) filters, and 4 Gaussian filters. An example of LM filter bank is shown in Fig. 2.
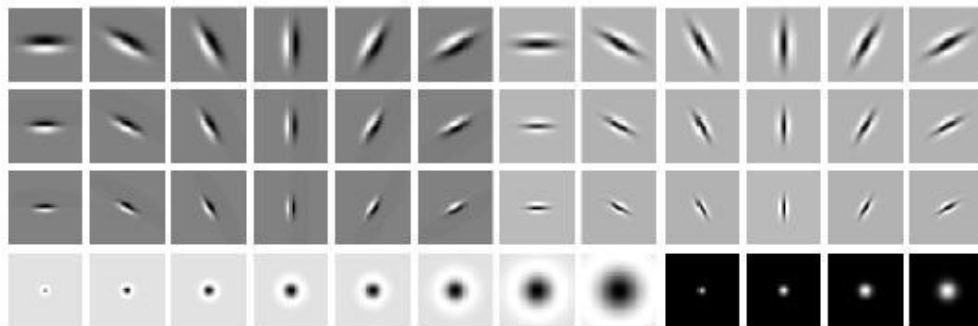


Fig. 2. Leung-Malik filter banks.

LM-filter banks generate 48 filter responses that will be grouped using k-means clustering method. K-means clustering has been widely used on image segmentation task, one of them in [13]. The procedure of k-means clustering is explained below:

1. Initialize $K$ number of cluster and cluster centers.

2. Calculate distance for each data to each cluster centers using distance metric. In this research, we use squared Euclidean distance metric and chi-squared distance metric.

3. Assign each data to the nearest cluster centers using (1) for squared Euclidean distance or (2) for chi-squared distance where $c_{(i)}$ is the *i-th* distance data, $x_i$ is the *i-th* data, $w$ is the weights associate with each dimension, and $\mu_j$ is the *j-th* cluster center.

$$c_{(i)} = \arg \min_{1 \leq j \leq K} \left(x_i - \mu_j\right)^2 \tag{1}$$

$$c_{(i)} = \arg \min_{1 \le j \le K} \sqrt{w(x_i - \mu_j)^2} \qquad (2)$$

4.  After all data assign to its nearest cluster, calculate new cluster centers.

5.  Repeat step 2 to 4 until converge or all cluster members are not changing label.

6.  Get the final cluster centers and cluster labels.

Texton histogram is calculated from the cluster labels. For each pixel, the corresponding cluster label on that pixel and all cluster labels in its neighborhood within certain radius are used to compute texton histogram. The number of bin histogram is equal to the number of cluster. Texton features are depended on the order of cluster labels. Because k-means always return different cluster labels when using random initialization of cluster centers, texton should be computed using fixed initial centers. We sort the cluster labels based on the number of member within cluster on ascending order. This method will keep the cluster labels in order even if the initialization is random. Before extracting texton from grayscale image, the grayscale image need to be normalized to zero mean and unit variance.

## *C. Segmentation based on Texton*

Image segmentation can be done based on color or texture. In this research, a texton based segmentation is used to segment road area. The procedure of road area segmentation is explained below:

1.  Image is convolved using LM-filter banks to get filter responses.

2.  Perform k-means clustering on filter responses with $K_1$ number of cluster and using squared Euclidean distance metric.

3.  Sort the centroid based on the number of member within cluster on ascending order.

4.  Set energy on each pixel to zero. Higher energy means higher possibilities as road area.

5.  Train K-NN model using texton features extracted from training images. The model consists of several pairs of $K_1$ centroid of filter responses called $C_1$ and its corresponding pairs of $K_2$ centroid of texton called $C_2$ also the labels for each centroid.

6.  Determine the number of sample for energy calculation. The samples are chosen among the $C_1$ trained data that gives minimum distance from the centroid of test data in step 3.

7.  For each samples do:

    a.  The filter responses on step 1 are grouped using k-means and $C_1$ as initial centroid with 1 iteration.

    b.  For each pixel, the corresponding cluster label on that pixel and all cluster labels in its neighborhood within certain radius are used to compute texton histogram.

    c.  Texton histograms are grouped using k-means with $K_2$ number of cluster and chi-squared distance metric. Chi-squared distance is used because it is good when dealing with histogram

    d.  Centroid from step 7 (c) is classified using K-NN model in step 5 with $C_2$ trained data to be labelled as road area or non-road area.

    e.  If the centroid belong to the road area, then the energy of each member within that cluster is added by one.

8.  If the final energy is higher than threshold, then that pixel is belong to the road area.

Remove small road area and repair the shape of the road using morphological operation.

### D. *Road Defect Classification*

Classification of road defect is done using SVM. SVM is machine learning method that use hyperplane. Hyperplane can be calculated by maxing the distance or margin between two objects from different class. In classification, SVM uses voting strategy to determine in which group the data belong to. To perform classification using SVM, the procedure is explained as follows.

1. Compute SVM kernel.

2. Compute decision function using (3) where $y_i$ is the class label, $\alpha_i$ is the value $0 \leq \alpha_i \leq C$ where $C$ is the regurization parameter, $K(x_i, x)$ is the kernel that has input $x_i$ which is the input data, $x$ is the support vector model and $b$ is the bias.

$$sgn\left(\sum_{i=1}^{l} y_i \alpha_i K(x_i, x) + b\right) \tag{3}$$

3. Repeat step 1 and step 2 for the other class.

4. Determine the member of class from class function that gave maximum votes corresponding to the data.

The procedure to perform road defect detection on road area is explained as follows:

1. Perform edge detection to extract edges and calculate the mean of grayscale pixels on the road area. Edges are important to find the road defect because road cracks and potholes both tend to have strong edges.

2. Apply morphological operation dilation to the edges to get the potential defect area.

3. Remove road lines by thresholding the grayscale pixels using mean of road area. This step is used to minimize the false detection that caused by non-defect edges from road lines.

4. Train SVM model using texton features extracted from training images.

5. Perform sliding window over potential defect area to classify the area into road defect or normal road using SVM with previously trained road defect model in step 4.

Perform blob analysis to determine the defect location, to measure the defect size, and to determine the type of defect. Defect type is categorized as road cracks if the ratio of defect and its bounding box less than threshold otherwise it is categorized as potholes.

### E. *Performance Evaluation*

In order to measure the performance of the proposed method, we use contour matching score or boundary F1 score (BF score) [14]. BF score is said to be a good measure for semantic segmentation. BF score measures how close the predicted boundary of an object matches the ground truth boundary using (4). The score range from 0 to 1 when score 1 indicates a perfect match.

$$BF\ score = 2 \cdot precision \cdot \frac{recall}{(recall + precision)} \tag{4}$$

## III. Results and Analysis

The proposed texton based segmentation for road defect detection on road aerial image is written using Matlab and runs on laptop with processor Intel Core i-7 6700K, 16 GB of RAM, and NVidia GTX 1050. We use KITTI road dataset [15] as training data to segment road area and grab road defect image (potholes and road cracks) from the internet as training data to detect road defect.

### A. *Training Process*

The training process produces road area detector and road defect detector. Road area detector is trained using K-NN with K = 5 and consists of several pairs of centroid from filter responses and

several pairs of centroid from texton. Road defect detector is trained using linear-SVM. The clustering method to extract texton features for both models uses k-means with K = 40. The size of LM-filter banks kernel is 21x21 and the size of texton kernel is 5x5. Smaller kernel tends to produce smooth segments. Fig. 3 shows the road defect samples that used in this research.



(a) road cracks training images
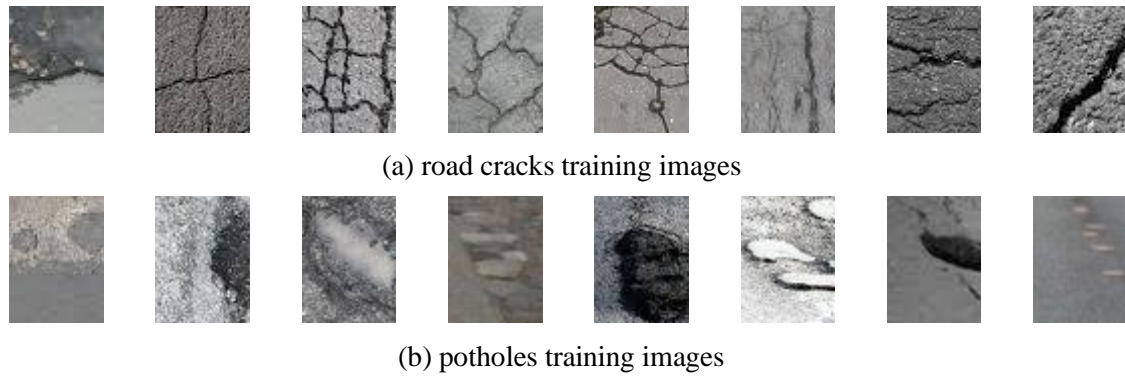


(b) potholes training images

Fig. 3. Samples of road defect training images

### B. Road Area Segmentation

Road area is segmented based on texture. Parameters that used to extract texton features are LM-filter banks size is 21x21, texton kernel size is 5x5, number of cluster (K) for k-means is 40, 5 samples of centroid set for energy calculation, and energy threshold is number of sample / 3. We remove small area which less than 32x32 pixels and apply morphological operation dilation with kernel size 32x32 to the potential road area. The largest area from potential road area is the final road area.

Fig. 4 shows the result of road area segmentation (a) road aerial image, (b) the energy result from K-NN classification (the scale from lowest to highest is indicated by blue – red color), and (c) the result of road area segmentation (magenta color marker). From Fig. 4(b), the road lines, grass and worn out asphalt have lower energy (blue color) because they have different texture than the road texture in the training images. From Fig. 4(c), the segmentation can cover most of the road area with 0.9884 BF score. The BF score is high because the road aerial image has a fine road texture with almost no defect on the road surface. On the road side, the rough grass texture is completely different with the road texture which makes the road shape is easier to be identified.



(a) Road aerial image

(b) Energy result from K-NN classification
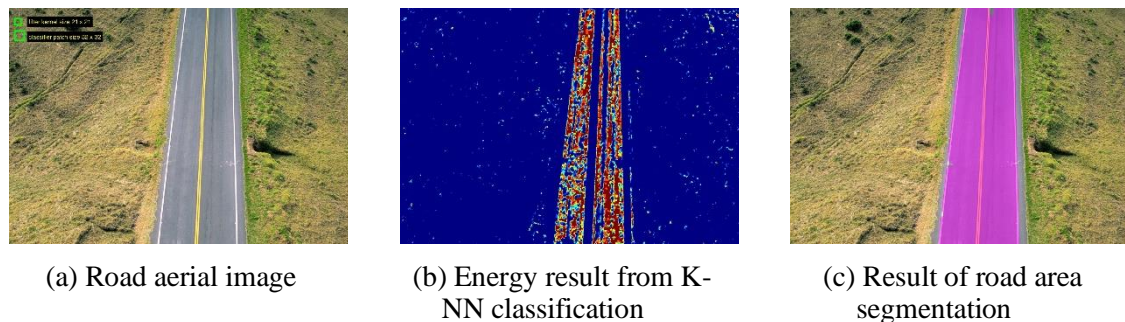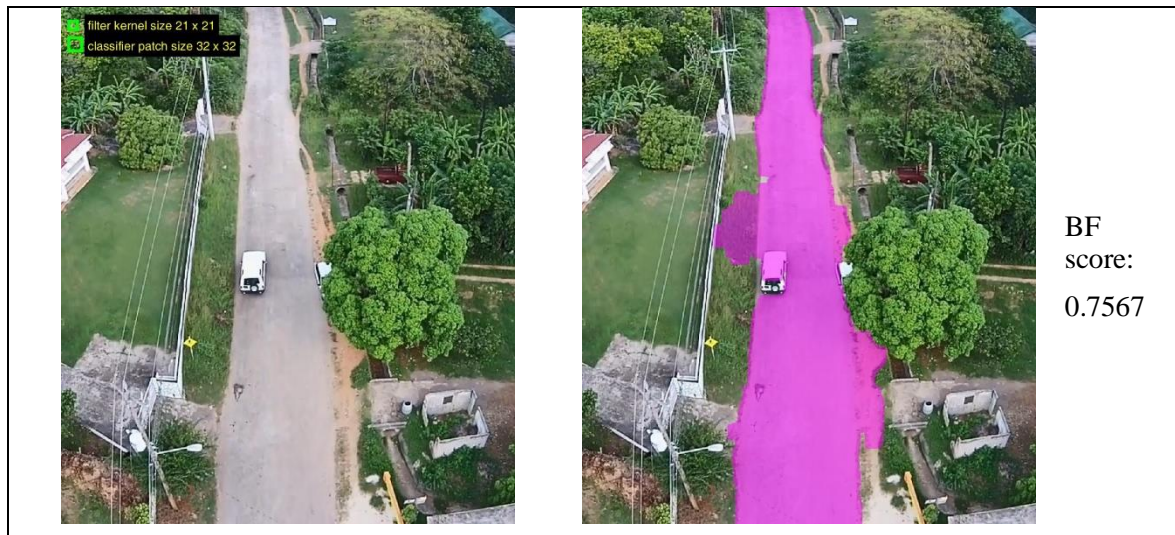
(c) Result of road area segmentation

Fig. 4. The result of road area segmentation

Fig. 5 shows the other samples of road segmentation and its corresponding BF score. From Fig. 5(a), the road aerial image has rough road texture with many cracks and rough grass texture on the road side. The different textures make the road shape easier to be identified which resulted in 0.9108 BF score. From Fig. 5(b), the road aerial image has rough road textures with some defects, various grass textures on the left side, and soft soil texture on the right side. The result of road area segmentation has 0.7567 BF score. The moderate score is caused by false segmentation of similar textures, even though they have different color, that present on image such as soft grass texture on the left and soft soil texture on the right side of the road.

(a)

(b)

Fig. 5. Performance score of road area segmentation

## C. Road Defect Detection

Road defect is detected based on texture using linear-SVM. Assume that road defect has different texture than normal road then there is exist a strong edge in the road area. Therefore, we perform edge detection on road area using Canny edge detection. Because of the thinning characteristic of Canny edge detector, a morphological operation dilation with kernel size 21x21 is applied to expand the potential defect area. SVM is used to classify the potential defect area into defect or normal category. Small area is removed and blob analysis is performed to find the defect location, measure the defect size, and calculate the ratio of defect area and its bounding box to determine the defect type. The defect type is categorized as road cracks if the ratio is lower than 0.5 otherwise it is categorized as potholes.

Fig. 6 shows the result of road defect detection. Fig. 6(a) is the road aerial image, (b) is the result of road area segmentation (shown by magenta color marker), (c) is the result of morphological operation on edge (shown by cyan marker), and (d) is the result of road defect segmentation along with information of defect location, defect size, and defect type (shown by cyan marker and green text and rectangle). Fig. 6(d) shows that the result of road defect detection on the image is several road cracks.

From Fig. 6(b), the proposed method is failed to cover all of the road area and resulted in 0.6217 of BF score. The moderate score is mainly caused by large damage on the road surface (shown by red circle marker). The damage is too large to be covered by morphological operation dilation with kernel size 21x21. Similar texture of road and gravels in the road side also contributes to lower the score. This poor segmentation of road area will affect the road defect detection in the next step.

From Fig. 6(c), the edge detection followed by morphological operation dilation with kernel size 15x15 is almost covers all the road cracks although there are some non-defect edges that caused by

poor road segmentation. From Fig. 6(d), the result of road defect detection is only 0.6024 of BF score. The moderate score is caused by large road damage shown by red circle marker that does not belong to road area hence not considered as road defect by the proposed method. Fig. 7 shows the other samples of road defect detection on road aerial image. Fig. 7(a) shows that the result of road defect detection is several road cracks and has 0.7379 of BF score while Fig. 7(b) shows that the result is potholes and has 1.0000 of BF score.
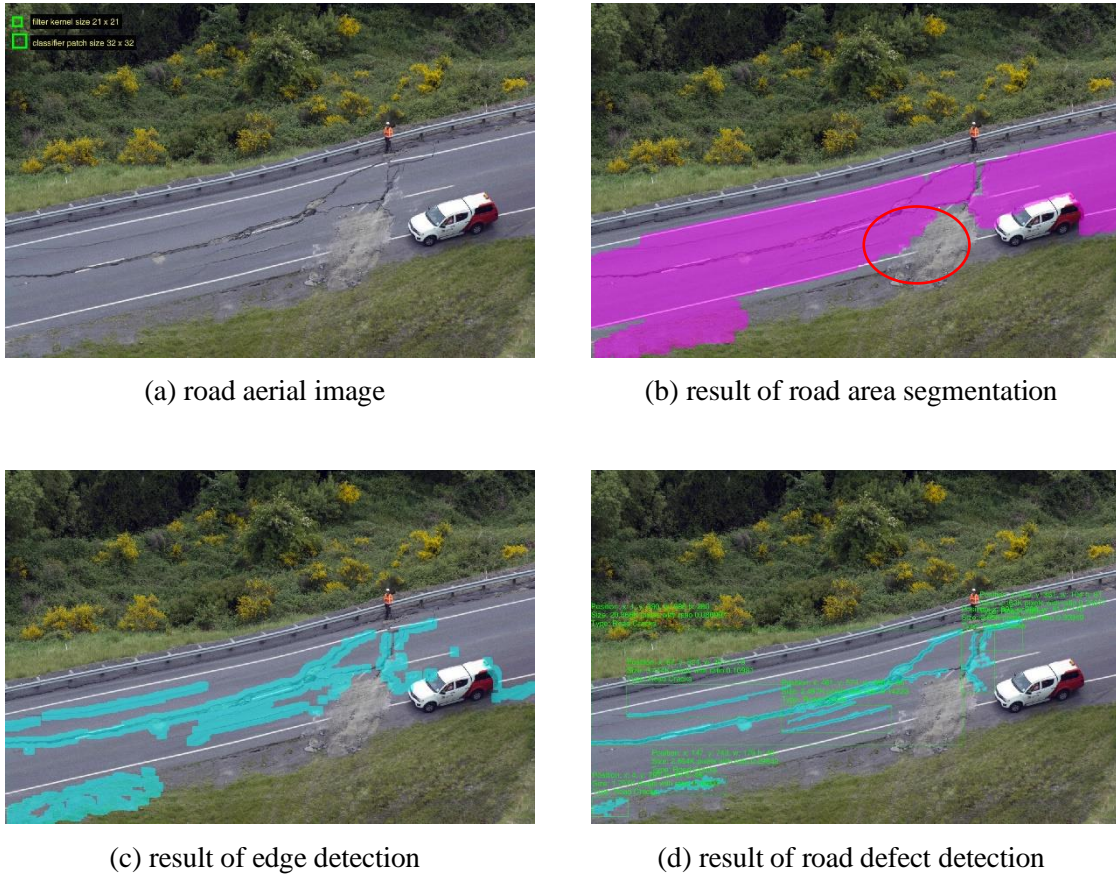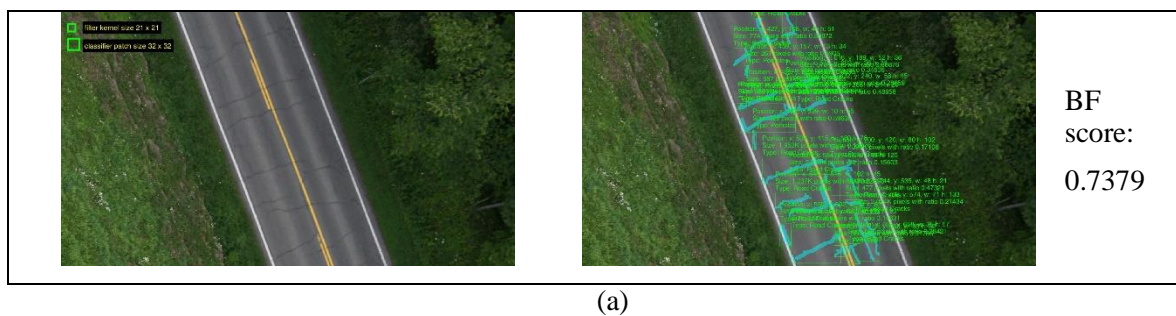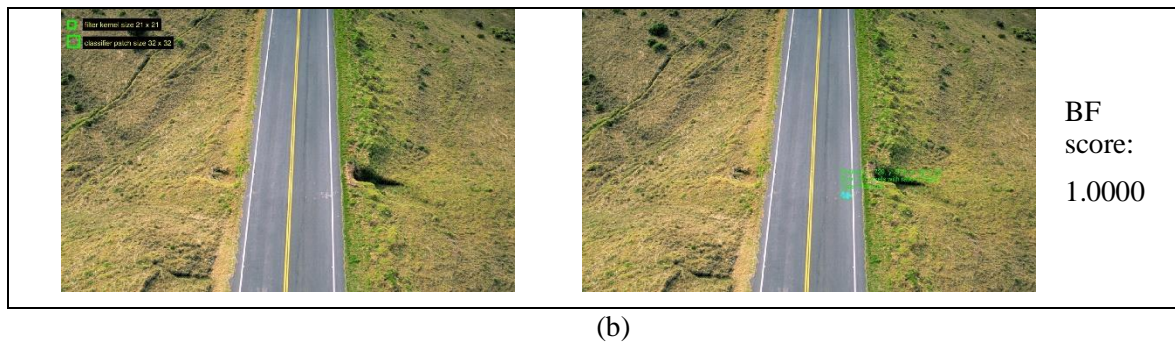


(a) road aerial image

(b) result of road area segmentation



(c) result of edge detection

(d) result of road defect detection

Fig. 6. The result of road defect detection



BF score: 0.7379

(a)

(b)

Fig. 7. Performance score of road defect detection

## IV. Conclusion

In this research, a method to detect road defect from aerial imagery is proposed. From the experiment, we can conclude that the proposed method can detect road defect if the defect is on early stage such as road cracks and small size potholes. The proposed method works well in detecting road defect inside the road area. Potholes can only be detected if the proposed method found closed contours and will fail otherwise (see Fig. 6(b)). From several tests, false detection can also be caused by shadow and object on the road. The present of object on the road can caused a hole in the road area segmentation and messed up the edge. Therefore, the acquisition of road aerial image should be done in less traffic and captured multiple times to accommodate the hole with other same image location but with different acquisition time. Shadow can be removed with certain algorithm. Overall, the proposed method can detect road cracks and potholes on road aerial image and gives information of defect location, defect size, and defect type. This method is effective because it covers large road area, do not disturb the traffic, and produce good BF score.

### Acknowledgment

### References

[1] R. A. Eaton, R. H. Joubert, and E. A. Wright, "Pothole Primer: A Public Administrator's Guide to Understanding and Managing the Pothole Problem (Rev. Dec. 1989.)," *Special Report*, US Army Corps of Engineers-Cold Regions Research & Engineering Laboratory, p. 34, 1989.

[2] C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," *Adv. Eng. Informatics*, vol. 25, no. 3, pp. 507–515, Aug. 2011.

[3] K. Azhar, F. Murtaza, M. H. Yousaf, and H. A. Habib, "Computer vision based detection and localization of potholes in asphalt pavement images," in *2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2016, pp. 1–5.

[4] P. Wang, Y. Hu, Y. Dai, and M. Tian, "Asphalt Pavement Pothole Detection and Segmentation Based on Wavelet Energy Field," *Math. Probl. Eng.*, vol. 2017, pp. 1–13, Feb. 2017.

[5] Y. Jo, and S. Ryu, "Pothole Detection System Using a Black-box Camera," *Sensors*, vol. 15, no. 11, pp. 29316–29331, Nov. 2015.

[6] Z. Zhang, X. Ai, C. K. Chan, and N. Dahnoun, "An efficient algorithm for pothole detection using stereo vision," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 564–568.

[7] K. Kamal *et al.*, "Performance assessment of Kinect as a sensor for pothole imaging and metrology," *Int. J. Pavement Eng.*, vol. 19, no. 7, pp. 565–576, Jul. 2018.

[8] M. Bellone and G. Reina, "Pavement distress detection and avoidance for intelligent vehicles," *Int. J. Veh. Auton. Syst.*, vol. 13, no. 2, p. 152, 2016.

[9] J. Bridgers and T. Chiang, "Mobile pothole detection system and method," *United States patent*-US 9,365,217, Booz Allen Hamilton Inc, 14 June 2016.

[10] J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and Texture Analysis for Image Segmentation," *Int. J. Comput. Vis.*, vol. 43, no. 1, pp. 7–27, 2001.

[11] T. Leung and J. Malik, "Representing and Recognizing the Visual Appearance of Materials using Three-dimensional Textons," *Int. J. Comput. Vis.*, vol. 43, no. 1, pp. 29–44, 2001.

[12] C. Schmid, "Constructing models for content-based image retrieval," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 2, p. II-39-II-45.

[13] A. Prahara, I. T. R. Yanto, and T. Herawan, "Histogram Thresholding for Automatic Color Segmentation Based on k-means Clustering," Springer, Cham, 2017, pp. 344–354.

[14] G. Csurka, D. Larlus, and F. Perronnin, "What is a good evaluation measure for semantic segmentation?," in *Procedings of the British Machine Vision Conference 2013*, 2013, p. 32.1-32.11.

[15] J. Fritsch, T. Kuhnl, and A. Geiger, "A new performance measure and evaluation benchmark for road detection algorithms," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 2013, pp. 1693–1700.