

Pothole Region Extraction Based On Similarity Evaluation Scale Classification Using Image Processing

Min-Jeong Kim¹, Hye-Jeong Kwon², Ji-Won Baek³, Hoill Jung⁴, Kyungyong Chung^{*5}

¹ Student, Division of Computer Science and Engineering, Kyonggi University, Suwon 16227, Gyeonggi, Korea

² Student, Division of Computer Science and Engineering, Kyonggi University, Suwon 16227, Gyeonggi, Korea

³ Student, Department of Computer Science, Kyonggi University, Suwon 16227, Gyeonggi, Korea

⁴ Professor, Department of Computer Software, Daelim University College, Anyang, 13916, Gyeonggi, Korea

^{*5} Professor, Division of Computer Science and Engineering, Kyonggi University, Suwon 16227, Gyeonggi,

Koreaminjeog0513@kyonggi.ac.kr¹, mjsgpwjd121@kyonggi.ac.kr², jwbaek@kyonggi.ac.kr³,

hijung@daelim.ac.kr⁴, dragonhci@gmail.com^{*5}

Corresponding author*: mobile Phone: +82-010-6362-4555

Article History: Received: 11 november 2020; Accepted: 27 December 2020; Published online: 05 April 2021

Abstract: Pavement deterioration and abnormal climate induced by global warming lead to a constant rise in the number of potholes. Accordingly, the loss cost for maintenance and accidents also increases. Therefore, it is necessary to develop a method of classifying pavement potholes and detecting their locations. This study proposes the pothole region extraction based on similarity evaluation scale classification using image processing. The proposed technique sets up a classification threshold appropriately by considering the structure, brightness, and other factors of the grayscale-converted image through SSIM (Structural Similarity Index Measure). It binarizes pothole images classified according to the threshold, and then extracts pothole regions through the threshold based segmentation. A conventional image classification method utilizes the rules found in objects or the label selected by a user. The proposed method can take into account detailed factors by comparing image similarity in the unit of pixel. According to the performance evaluation, the proposed classification method's F1-score is 0.83, and its accuracy of pothole region extraction is 0.851. Therefore, with the proposed technique, it is possible to make classification in consideration of similarity between images. In addition, the proposed method makes it possible to detect the regions similar to actual potholes accurately.

Keywords: Pothole detection, SSIM, Image processing, Segmentation, Binarization, Thresholding.

1. Introduction

With the development of transportation means, more people use pavements. Due to the sharp rise in pavement users and climate changes, the frequency of road damage is on the increase. Pavement damage includes potholes, spalling, and cracks. Potholes are generated by small cracks on pavements. If they fail to be managed timely, they can become larger, causing a big accident. Therefore, the importance of maintaining existing roads is increasing. According to the statistics of Ministry of Land, Infrastructure, and Transport, pothole cases that had occurred from 2016 to 2018 numbered 657,993, and thereby damage compensations and pavement repairs cost about 110 billion won [1]. Potholes can bring about big losses in terms of social and economic aspects and can cause human damage beyond physical damage. For this reason, it is necessary to develop a method of detecting potholes accurately in order to deal with them. Therefore, many studies on image classification for detecting potholes have been conducted [2]. For instance, images are classified with the uses of data, such as rules and labels. However, such image data classification can be different depending on models, labels, or scale criteria. Moreover, data can be distorted while image data is preprocessed, and consequently, accuracy can reduce. For accurate classification, it is necessary to take into account visual factors that help to judge the similarity between images, such as structure and brightness. Unlike different image evaluation indexes, SSIM (Structural Similarity Index Measure) can be used to find a structural difference recognized by humans, rather than a difference between pixels. Therefore, this study proposes the pothole region extraction based on similarity evaluation scale classification using image processing. The proposed technique sets up a threshold appropriately through SSIM which is an image quality scale and makes binary classification into two groups (pavement images with potholes and normal pavement images). The potholes images classified are binarized through Otsu threshold, and any random noise is removed by the Median filter. In addition, with the uses of dilation and erosion of Morphological processing, an object's unnecessary regions are removed, and the pixels of significant pothole region is expanded. In this way, it is possible to extract a region similar to a real pothole. Based on the extracted segmentation images, a location of a pothole can be detected more accurately. The performance of the proposed technique is evaluated in two aspects. The first performance assessment evaluates accuracy through the classification of normal road images and pothole images based on SSIM. The second performance assessment evaluates the matching rate of pixels between the segmentation image based on the Threshold and the original image.

*Corresponding author: Kyungyong Chung

Professor, Division of Computer Science and Engineering, Kyonggi University, Suwon 16227, Gyeonggi, dragonhci@gmail.com^{*5}

2. Relation Work

2.1. Image quality assessment

Image quality evaluation is a technique of evaluating the quality of a converted image in the way of calculating loss and similarity values between an original image and its converted (or compressed) image. Through the comparison in the unit of pixel, it is possible to draw a loss that people hardly recognize visually. Typically, such quality evaluation scales include MSE (Mean Squared Error) [3], PSNR (Peak Signal-to-Noise Ratio) [4], and SSIM (Structural Similarity Index Measure) [5]. MSE is a method of measuring a difference between images. It first measures the pixel values of an original image and its converted image and then presents how the two images are not similar on average. Since the method is the simplest to compare two objects, it is applied to a variety of loss measurement areas as well as image loss measurement. However, in the scale method, due to the square of a difference, a value less than ‘1’ can become smaller, and a value more than ‘1’ can become larger. As such, value distortion can occur. In addition, even if a MSE value is equal, each object’s original value can be different. For example, the MSE value of the predictive model of the object value 10000 is 500, and that of the objective value is 500. The two objects have the same MSE value, but a different loss actually. Equation (1) presents MSE. I represent the grayscale image of the size $A \times B$, and K means the distorted image of I.

$$\text{MSE} = \frac{1}{AB} \sum_{i=0}^{A-1} \sum_{j=0}^{B-1} [I(i, j) - K(i, j)]^2 \quad (1)$$

To overcome the limitation, PSNR, which is the evaluation scale with extended MSE, was developed. PSNR represents an object’s loss ratio for its maximum signals. With the scale, it is possible to compare a different-sized object’s loss. The higher PSNR, the lower loss. The scale is mainly used when loss information of image quality in compression is evaluated. PSNR represents the distortion ratio of the image with extended MSE. Just like MSE, PSNR evaluates image quality with the use of a numerical difference between image pixels. For this reason, PSNR can calculate the loss value which is inconsistent with the distortion recognized by human visually. Equation (2) presents PSNR. Max_I means the maximum value of an image pixel.

$$\text{PSNR} = 10 \log \frac{\text{Max}_I^2}{\text{MSE}} \quad (2)$$

To extract the image quality recognized by human actually, Structural Similarity (SSIM), which has a different approach from the two scale methods aforementioned, was developed. SSIM is used to evaluate structural similarity between images in the way of a human’s visual system to draw information from an image. SSIM is based on the comparison of luminance, contrast, and structure. It is used to evaluate a similarity between an original image and its converted image. Compared to PSNR and MSE, SSIM is based on a human’s visual cognitive factors [5]. Equation (3) presents SSIM. μ_a and μ_b mean the internal mean value of pixels in each one of images, respectively. σ_a and σ_b mean the internal standard deviation of pixels in each one of images, respectively. And σ_{ab} represents the internal covariance of pixels in image. SSIM has a value between – and +1. The closer a SSIM value is to +1, the more the two images are similar.

$$\text{SSIM}(a, b) = \frac{(2\mu_a\mu_b + C_1) + (2\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)(\sigma_a^2 + \sigma_b^2 + C_2)} \quad (3)$$

2.2. Trend of pothole detection research

Pavement deterioration and ground collapse caused by rapid climate changes lead to various types of road damage, such as cracks and potholes. Accordingly, the cost of vehicle and road damage caused by traffic accidents is on the rise. Conventional ways to detect potholes are people’s complaints or patrol cars’ detection, and repair work is conducted. In such ways, it is possible to detect damage in an accurate location, but immediate responses fail to be done. A pothole begins from a small crack in the pavement so that initial repair work is important. For this reason, it is necessary to detect road damage in order for maintenance. For pothole detection, manual and auto-detection systems are applied, including impact detection sensor, Ground Penetrating Radar (GPR), and image processing [6]. For pothole detection, and impact detection sensor detects an impact at the time of passing a pothole. However, the sensor detects what a vehicle passes only so that it mistakes an impact by a different road structure for a pothole. Its accuracy is low. Using broadband electromagnetic waves, GPR analyzes a sign reflected by an internal structure or the ground surface. Although the system can easily find the status and structure of pavement, it costs high. Image processing detects an object from an image by using deep neural networks. It can make detect in a real-time image so as to find pavement status fast. Therefore, studies on a variety of image processing algorithms for pothole detection are actively conducted. Koch, C. [7] developed the system that detects potholes with the use of the texture of asphalt pavement. By comparing images with the use of the color distribution histogram of pixels, it splits an image into a road damage region and into a normal road region. In addition, to detect a pothole, it extracts the texture of a damaged region and that of a normal region and then compares them. Accordingly, based on pavement damage, it can detect a pothole. Akagic, A. [8] proposed the unsupervised pothole detection method based on RGB color space segmentation. It sets up a global asphalt region point by using RGB and image standard deviation and then establishes a point more accurately by using the set-point and its four adjacent windows. In addition, with the use of Ostu threshold, the proposed method segments pothole regions and remove boundaries. By narrowing the range of pothole regions, it is possible to detect a pothole accurately.

3. Pothole region extraction based on similarity evaluation scale classification using image processing

3.1. Pothole image classification using SSIM

Potholes have structural characteristics that people can recognize. Accordingly, SSIM, which does not use similarity between images on the basis of pixel distance and is similar to the human visual method, is used to calculate a similarity with a normal road. As a data set, the Global Road Damage Detection Challenge 2020 is used. Based on the data labeled as a normal road, data sets in a mixture of normal and abnormal data are used for similarity calculation. A normal asphalt image converted to the grayscale image has the even distribution of '0' and '1'. On the contrary, pothole pavement is a large contrast to neighboring pavements according to shadow and hole depth. By reflecting the feature, and transforming the contrast and brightness that influence an SSIM score, it is possible to calculate the SSIM score of a pothole road and that of a normal road. Min-max Scaling is applied to the calculated whole similarity, and based on the threshold value drawn from the process, road images with or without potholes are classified. The classification process is shown in figure 1.

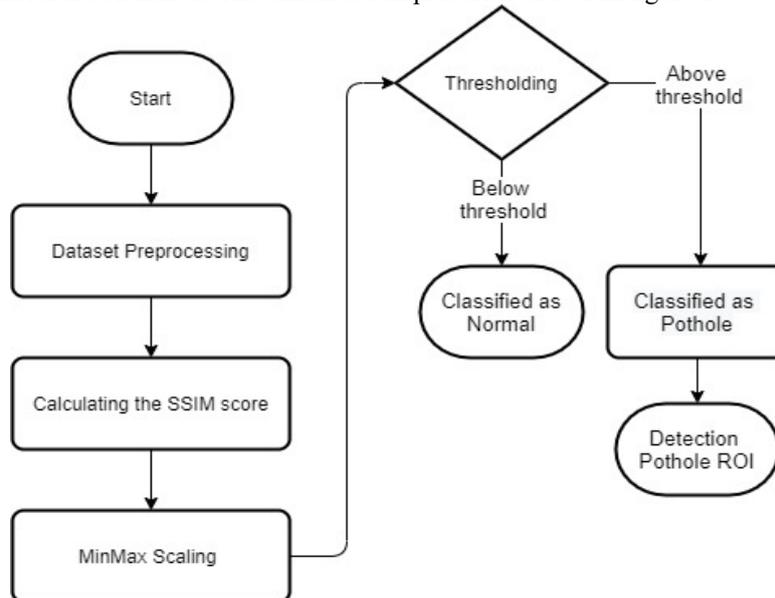


Figure 1. Process of Pothole Detection using Structural Similarity of Image

Figure 2 shows the preprocessed data and SSIM map. Fig. (a) and (b) are the images after SSIM feature values are adjusted for clear image differentiation. Fig. (a) is the converted normal road image. Fig. (b) is the converted pothole road image. Fig. (c) is the SSIM map after the comparison between the two converted images. Fig. (d) and (e) are the grayscale images without adjustment of feature values that original images have figure 2. When the local SSIM score of the two images is larger, mapping to a brighter pixel occurs. Therefore, through SSIM map, it is possible to find a pothole region [9].

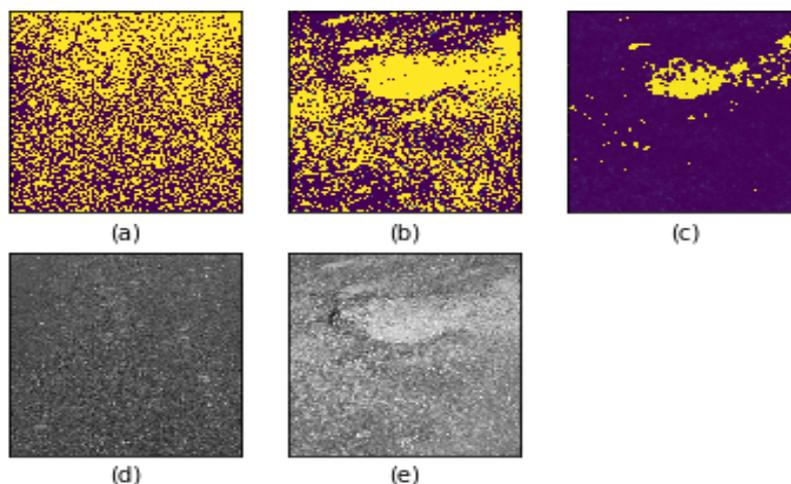


Figure 2. Preprocessed Data and SSIM Map

Figure 2 shows the preprocessed data and SSIM map. Fig. (a) and (b) are the images after SSIM feature values are adjusted for clear image differentiation. Fig. (a) is the converted normal road image. Fig. (b) is the

converted pothole road image. Fig. (c) is the SSIM map after the comparison between the two converted images. Fig. (d) and (e) are the grayscale images without adjustment of feature values that original images have figure 2. When the local SSIM score of the two images is larger, mapping to a brighter pixel occurs. Therefore, through SSIM map, it is possible to find a pothole region [9].

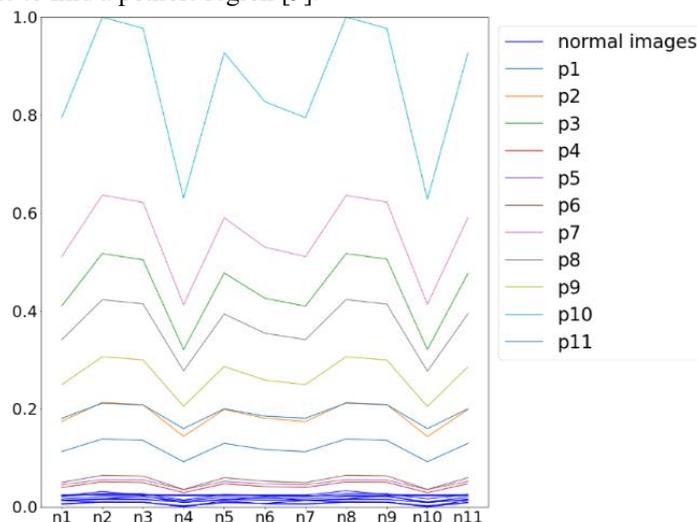


Figure 3. SSIM Score

Figure 3 shows the SSIM scores of the test data for the normal road images (n1, n2 ... , n11). X-axis presents the normal road images in comparison group. Y-axis presents the SSIM scores of the data. Data sets include both normal road images and pothole road images, and a blue line means the result value of a normal road image. Unlike PSNR and MSE, SSIM obtains the result similar to that of a human's perception ability. Therefore, it can show a clear difference between a pothole pavement and a normal pavement. In case of the data with the lowest similarity (n4 and n10), they were normal road images that used darker asphalt than other normal road images. Since they had less bright regions to compare, SSIM map was not clear at the time of comparing with pothole images. Nevertheless, it is possible to classify into a normal road image and pothole road image clearly. Accordingly, SSIM score is used to set up a threshold and classify images into a pothole road image and into a normal road image. In this study, a threshold was set to $\sigma=0.1$ for classification. SSIM based comparison between two images shows a clear difference, whereas PSNR and MSE, metrics that measure a distance between two different images, failed to present a clear difference between the two images. Figure 4 illustrates the MSE and PSNR results of the test data for a normal road image. X-axis is the same as in Figure 3. Y-axis presents the normalization of the MSE and PSNR scores for data.

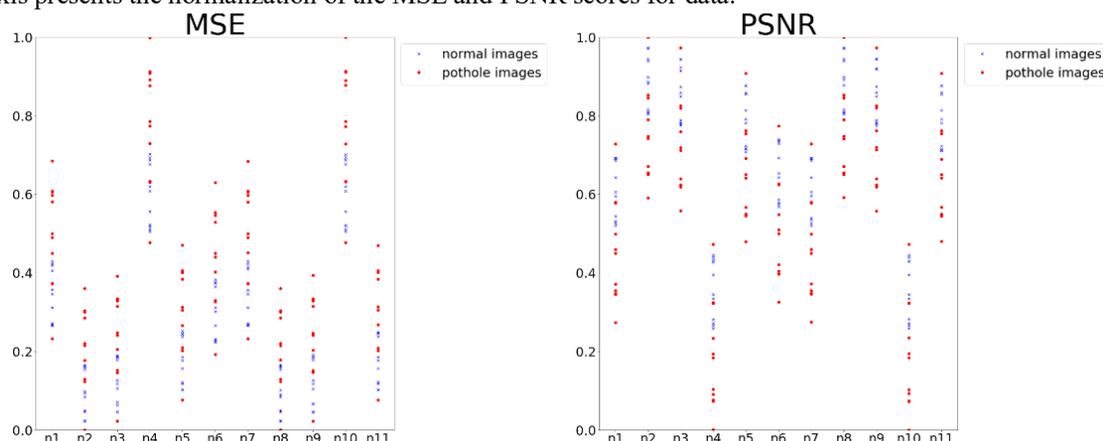


Figure 4. PSNR and MSE score

3.2. Threshold based pothole region extraction

To detect a pothole from the image classified by SSIM, preprocessing is performed. In order to separate a pothole from a road image, OTSU [10] based binarization is applied. OTSU detects a valley region from an image histogram and sets it up as a threshold and then finds an intensity value for the best segmentation with binary class. In this way, it is possible to separate an object with a background. An asphalt pavement image generates a lot of random noises in the binarization due to its curved surface. Since it is hard to process a noise only with binarization, Median filtering [11,12] is performed. Median filter as a nonlinear filter is used to remove a random noise by keeping a border. Morphological processing is conducted to add a pixel layer into a

region in order to fill small holes. It also removes unclear detailed objects [13,14,15]. As such, morphological dilation and erosion are applied to remove a small object and expand an object in order for clear detection. Figure 5 illustrates the image after binarization and morphological processing.

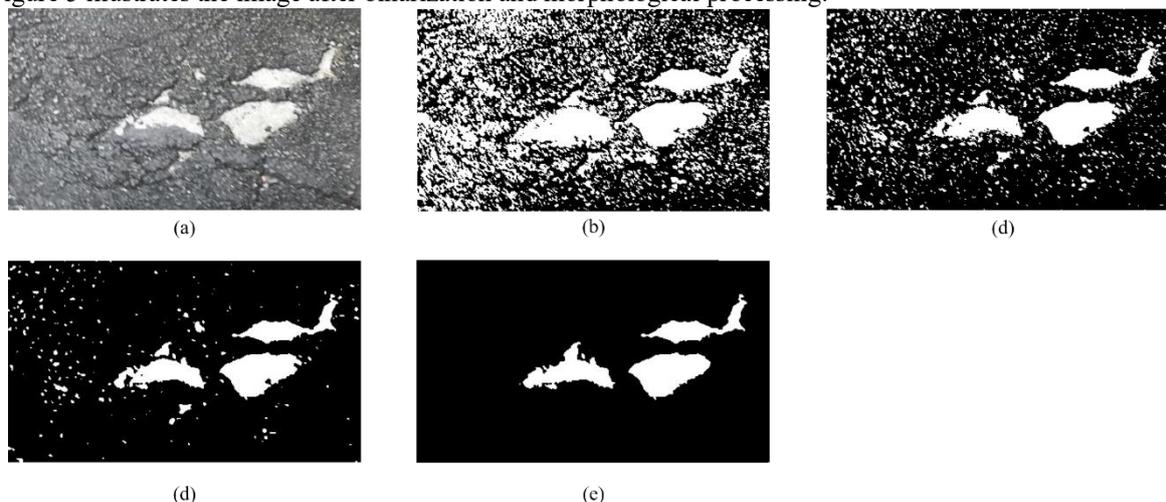


Figure 5. Process of threshold based pothole segmentation

In Figure 5, (a) is the original image of pothole; (b) is the binarized image through global threshold; (c) is the image after OTSU algorithm is applied; (d) is the image after OTSU algorithm is applied to Median-filtered image; (e) is the image after Morphological processing. As such, through image binarization, filtering, and morphological processing, it is possible to remove unnecessary small objects, and extract a pothole region for clear detection.

4. Evaluation

As a data set, Global Road Damage Detection Challenge 2020 is used. The specifications of operating system and hardware are as follows: Window10, Intel(R) Core(TM) i7-9700K CPU 3.60GHz, 16GB RAM. As software, Python 3.6.0 is used. Performance evaluation is conducted in two ways. The data used in this study are eleven pothole road images and eleven normal road images, both of which have 512x512 in size. For the detection of the structural features of potholes, the brightness and contrast of these images, which are the features used in SSIM, were adjusted and preprocessed [16,17]. As for eleven normal road images, a total of twenty-one images were compared, and therefore comparative tests were conducted two hundreds forty-two times. In the first performance evaluation, the accuracy of the proposed SSIM based image classification was evaluated, in the second evaluation, image segmentation regions were compared with real pothole regions in terms of a consistency ratio.

In the first evaluation, the accuracy of the proposed SSIM based image classification is evaluated through confusion matrix. Table 1 presents the SSIM based pothole image classification confusion matrix.

Table 1: Confusion Matrix of Pothole Image Classification

		Classification	
		Normal Image	Pothole Image
Ground Truth	Normal Image	121	0
	Pothole Image	35	86

The proposed method's precision for pothole detection was 0.71; its recall was 1.00; its f1-score was 0.83. The reason why pothole road images are less distinguished than normal road images is that caves in some pothole images are smooth and thus there is no big difference in SSIM map. In addition, these images have different circumstances and areas so that their brightness and contrast are different depending on weather conditions, asphalt colors, and other factors when they are shot. These factors not only influence image features but detection results. It is possible to compare pothole road images with normal road images significantly through a human's visual cognitive way, rather than the distance and noise ratios in the unit of pixel. Therefore, unlike PSNR and MSE, SSIM makes it possible to draw a clear structural similarity ratio and make classification on the basis of a threshold.

In the second evaluation, image segmentation regions were compared with real pothole regions in terms of a consistency ratio. Based on the pixels of consistency between actual pothole regions and segmentation regions, accuracy, precision, and recall are evaluated. Figure 6 illustrates the results of the comparison between original image regions and segmentation regions.

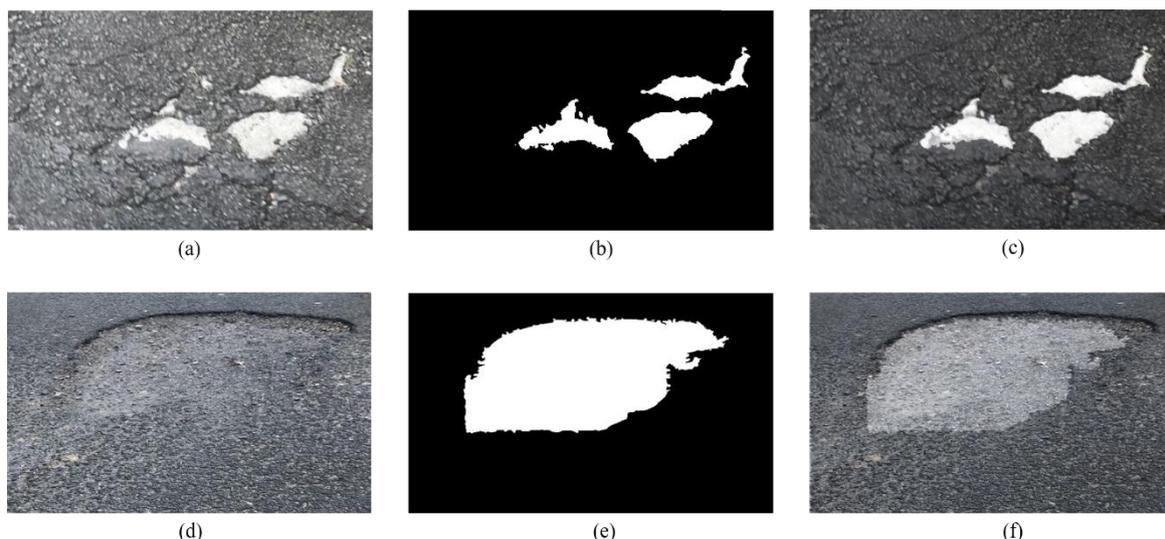


Figure 6. Comparison between original image and segmentation

Fig 6. (a) and (d) show original images; (b) and (e) are segmentation images; (c) and (f) are overlaid images of original and segmentation images. In this way, it is possible to find how much pothole regions are consistent with segmentation regions. Accuracy is used to judge how much segmentation regions are consistent with pothole regions. Recall is used to judge how much segmentation regions are incorporated into real pothole regions. Precision is used to judge how much the detected pothole regions are real pothole regions. Figure 7 shows the results of Accuracy, Recall, and Precision for pothole detection.

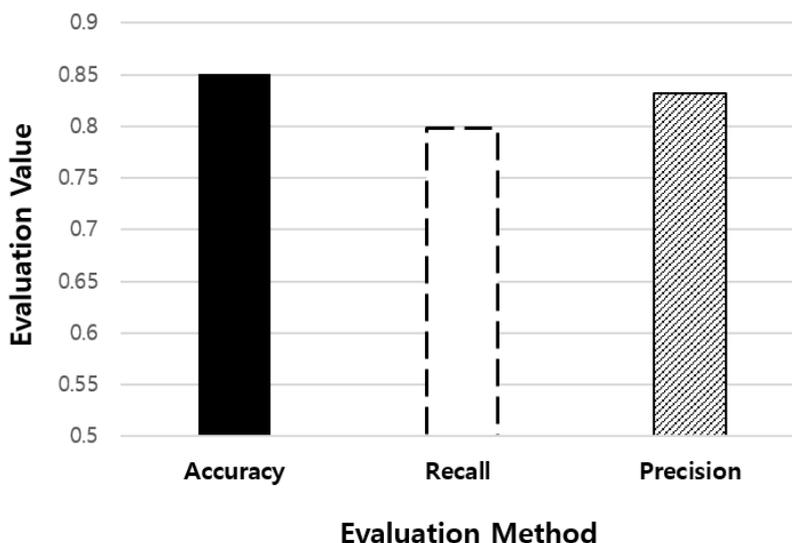


Figure 7. Result of accuracy, recall, precision according to Pothole detection

As shown in Fig. 7, Accuracy was 0.851; Recall was 0.798; Precision was 0.832. Given the results, the method removed unnecessary noises through image preprocessing and extracted pothole regions through pixel expansion, and thereby produced a good performance. In most images, it is possible to make similar detection just as in real pothole regions. However, in pothole images with thin holes and unclear boundaries, image binarization was not performed well. In this case, the performance of the method was not good. Therefore, potholes can be detected through SSIM based significant image classification and threshold-based segmentation.

5. Conclusion

This study proposed the pothole region extraction based on similarity evaluation scale classification using image processing. The proposed method sets up a threshold by comparing images with the use of SSIM in order to classify images into pothole road and into the normal road. In addition, it binarizes a pothole image and removes noises through filtering. After that, it performs a more improved segmentation process through morphological processing. In this way, it is possible to extract and detect pothole regions. Performance evaluation was conducted in two ways. Firstly, the accuracy of the SSIM based classification model is evaluated. Secondly, a pixel consistency ratio between the pothole regions extracted by threshold-based

segmentation and real pothole regions is evaluated. According to the performance evaluation of the proposed SSIM based classification model, its f1-score was 0.83, the accuracy for pothole region extraction was 0.851. As a result, the proposed model showed good performance. Therefore, with the use of the proposed method, it is possible to make a classification in consideration of similarity between pothole road image and normal road image. In addition, by removing unnecessary noises and reducing the loss of pothole regions that are removed through binarization, it is possible to extract regions accurately. In this way, it is possible to classify images, extract significant regions, and detect potholes. Nevertheless, when images include shades or different objects, the proposed model showed low performance in terms of classification and pothole region extraction. Therefore, it is planned to research in-depth image processing to remove shades and unnecessary objects.

Acknowledgment

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2018-0-01405) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation)

References

1. Information on <http://www.molit.go.kr>
2. Pramestya, R. H., Sulistyaningrum, D. R., Setiyono, B., Mukhlash, I and Firdaus, Z. (2018) Road Defect Classification Using Gray Level Co-Occurrence Matrix (GLCM) and Radial Basis Function (RBF). In 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE) 285-289.
3. Wang, Z and Bovik, A. C. (2009) Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE signal processing magazine* 26(1), 98-117.
4. Huynh-Thu, Q and Ghanbari, M. (2008) Scope of validity of PSNR in image/video quality assessment. *Electronics letters* 44(13), 800-801.
5. Wang, Z., Bovik, A. C., Sheikh, H. R and Simoncelli, E. P. (2004) Image quality assessment: from error measurement to structural similarity. *IEEE transactions on image processing*, 13(1).
6. Jo, Y and Ryu, S. (2016) Real Time Pothole Detection System based on Video Data for Automatic Maintenance of Road Surface Distress. *KIISE Transactions on Computing Practices* 22(1), 8-19.
7. Koch, C and Brilakis, I. (2011) Pothole detection in asphalt pavement images. *Advanced Engineering Informatics* 25(3), 507-515.
8. Ouma, Y. O and Hahn, M. (2017) Pothole detection on asphalt pavements from 2D-colour pothole images using fuzzy c-means clustering and morphological reconstruction. *Automation in Construction* 83, 196-211.
9. Sun, Y., Tistarelli, M and Maltoni, D. (2013) Structural similarity based image quality map for face recognition across plastic surgery. In 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS) 1-8.
10. Zhang, J and Hu, J. (2008) Image segmentation based on 2D Otsu method with histogram analysis. In 2008 International Conference on Computer Science and Software Engineering 6, 105-108.
11. Shan, W., Yi, Y., Qiu J and Yin, A. (2019) Robust median filtering forensics using image deblocking and filtered residual fusion. *IEEE Access* 7, 17174-17183.
12. Rudnitskii, A. G and Rudnytska, M. A. (2017) Segmentation and denoising of phase contrast MRI image of the aortic lumen via fractal and morphological processing. In 2017 IEEE 37th International Conference on Electronics and Nanotechnology (ELNANO) 344-348.
13. Choi, S. Y., Chung, K. (2020) Knowledge Process of Health Big Data using MapReduce-based Associative Mining. *Personal and Ubiquitous Computing* 24(5), 571-581.
14. Kim, J. C., Chung, K. (2020) Discovery of Knowledge of Associative Relations using Opinion Mining Based on a Health Platform. *Personal and Ubiquitous Computing* 24(5), 583-593.
15. Shin, D. H., Chung, K., Park, R. C. (2020) Prediction of Traffic Congestion Based on LSTM through Correction of Missing Temporal and Spatial Data. *IEEE Access* 8, 150784-150796.
16. Kang, J. S., Baek, J. W., Chung, K. (2020) PrefixSpan based Pattern Mining using Time Sliding Weight for Streaming Data. *IEEE Access* 8, 124833-124844.
17. Kim, H. J., Baek, J. W., Chung, K. (2020) Optimization of Associative Knowledge Graph using TF-IDF based Ranking Score. *Applied Sciences* 10(13), 4590-4610.
18. Shin, D. H., Park, R. C., Chung, K. (2020) Decision Boundary-Based Anomaly Detection Model using Improved AnoGAN From ECG Data. *IEEE Access* 8, 108664-108674.