Defect detection based on singular value decomposition and histogram thresholding

Xuan Tuyen Tran[1], Tran Hiep Dinh[1], Ha Vu Le[1], Qiuchen Zhu[2] and Quang Ha[2]

***Abstract*— This paper presents a novel method for defect detection based on singular value decomposition (SVD) and histogram thresholding. First, the input image is divided into blocks, where SVD is applied to determine if a region contains crack pixels. The detected crack blocks are then merged to construct a histogram to calculate the best binarization threshold by incoporating a recent technique for multiple peaks detection and Otsu algorithm. To validate the effectiveness and advantage of the proposed approach over related thresholding algorithms, experiments on images collected by an unmanned aerial vehicle have been conducted for surface crack detection. The obtained results have confirmed the merits of the proposed approach in terms of accuracy when using some well-known evaluation metrics.**

I. Introduction

Cracks in concrete surfaces are the initial indication of degradation of built infrastructure. These defects occur due to various reasons such as loading, chemical reactions or faulty construction, leading to a potential threat to human safety and asset damage. Therefore, regular inspection and monitoring of built infrastructure is essential to manage and maintain its serviceability and durability. Over the last decade, automatic inspection based on image processing techniques has received great interest from researchers due to its inexpensive and non-intrusive inspection process [1]–[3]. In processing of concerned images, there exists a significant difference between the intensity levels of pixels representing the region of interest and background, thresholding is hence widely applied due to its straightforwardness and effectiveness in object extraction. In [4], histogram thresholding for automatic binarization was employed in a vision-based automated manipulation system to pick up a single particle from a cluster of carbon nanotubes. In another intelligent system [5], thresholding plays an important role to extract the target from the image background for a more precise positioning.

In general, thresholding can be categorized into bi-level or multi-level techniques, where there is always an option to extend a bi-level technique into a multi-level one and vice versa. Among the binarization techniques, Otsu's method [6] is one of the most popular approach where an exhaustive search is employed to determine an optimized threshold that maximizes the inter-class variance between the object and background. As Otsu's algorithm is vulnerable to images with small objects, various extensions have been developed to improve its performance in defect detection by focusing on the contrast between the defect and background pixels. However, as discussed in [7], iterative approaches can be trapped into a non-convergent case, multiple convergence points or converging to a threshold value that leads to an invalid segmentation or increase in feature matching complexity [8]. Instead of calculating a global threshold for the whole image, alternative approaches [9], [10] have proposed to classify image pixels based on the local statistics or neighbourhood information. These approaches are limited in automation possibilities as user intervention is required to define the characteristics of the local window. On the other hand, a binarization problem can be solved by employing a multi-level thresholding approaches and setting the number of clusters to two. In [11], [12], spatial information and fuzzy membership functions are employed to generate a segmentation that is more robust to noise and artifacts. The segmentation result of these methods is based on various spatial constraints, leading to a difficulty to modify the algorithm for a specific application. In [13], [14], frequency and distribution of the histogram intensity values are utilized to calculate dominant peaks for thresholding purposes. While pre-defined parameters are essential in [13], a non-parametric approach has been developed in [14], where no prior knowledge about the number of histogram modes or distance between the modes in processing is required to obtain a desired segmentation.

Recently, machine learning and deep learning have been widely applied into computer vision due to the ability to accurately classify objects at pixel levels [15], [16]. However, the effectiveness of the approach is highly dependent on the data size and the accuracy level of the labeling phase.

According to our analysis, about 99 percent of the pixels of the surface images can be classified as background. Hence, the corresponding histograms also reflect this distribution of the intensity levels and usually appear

[1] University of Engineering and Technology, Vietnam National University, Hanoi

[2]Faculty of Engineering and Information Technology University of Technology Sydney, NSW 2000, Australia

as uni-modal. Therefore, to effectively solve a segmentation problem with thresholding, a pre-processing step is required to balance the number of crack and background pixels. Here, we propose to use singular value decomposition (SVD) to emphasize the crack features of the input image by filtering out the background pixels. First, the input image is divided into square blocks for local processing. Then, the singular value distribution, which presents the density of different components of the image, is obtained from the SVD. By evaluating the singular value energy decay rate, the background blocks and ones that contain crack pixels are classified. A histogram of the crack blocks is then constructed, where a combination of the Summit Navigator (SN) [14] and Otsu [6] is developed to determine the best binarization threshold. Experimental results have been taken to confirm the effectiveness of the proposed method in terms of incorporating a multilevel thresholding algorithm into a binarization problem, and improving the calculation of Otsu threshold to achieve a better defect detection.

The paper is structured as follows: Section II provides a brief introduction about the property and implementation of SVD for crack blocks detection. An automatic thresholding method is also developed for calculation of the best binarization threshold. Experimental results will be discussed in Section III.

II. Methodology

A. Crack blocks detection based on SVD property

1) SVD basic and its property: Let $X \in \mathbb{R}^{M\times N}$ is an arbitrary rank n matrix, the theory of SVD [17] states that X can be decomposed into sum of n rank-1 matrices as:

$$X = U\Lambda V^T = \sum_{i=1}^{n} \alpha_i u_i v_i^T \tag{1}$$

where U and V are respectively an $M \times M$ and $N \times N$ orthogonal matrices, and $\Lambda = \text{diag}(\alpha_1, \alpha_2, \alpha_3, ...\alpha_n)$ is a $M \times N$ diagonal matrix of singular values α_i. The diagonal elements of Λ are arranged in a descending order and called the singular values (SVs) of X. Generally speaking, if we divide an image into square blocks and consider them as matrices, the employment of SVD allows decomposing each block into several rank-1 matrices, $\alpha_i u_i v_i^T$ representing linearly independent components of the block. The magnitude of α_i would illustrate the contribution of component i to the original matrix . If an image region contains only background, the energy would concentrate mostly in the first singular value α_1, while the magnitudes of the following SVs are negligible. In contrast, the existence of both crack and background components in a block will result in more than one significant SVs. It has been confirmed in

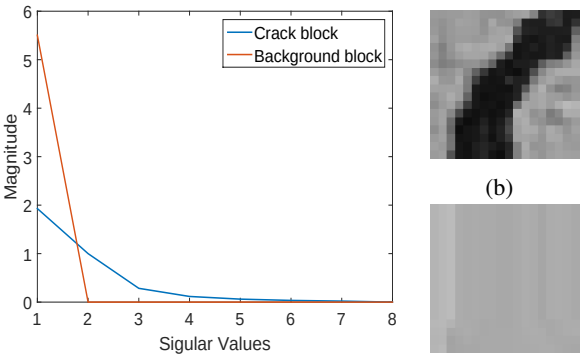


Fig. 1: Illustration of the difference between a crack and non-crack block: (a) Distribution of the singular value gaps of two blocks, (b) a crack block, and (c) a background block.

[18] that the singular values (SVs) of smoothed images have a higher decaying rate compared with those from a random ones. Therefore, the difference between the calculated SVs could be a reliable metric to detect the degree of appearance of different components in the concerned defect image. Fig. 1 illustrates an example of a crack and background blocks. While there is a significant difference between the first and second SVs of the background block (red line), the gap between these two values in the crack block (blue line) is significantly smaller.

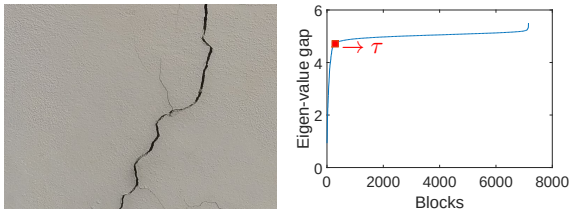


Fig. 2: Example of the singular value gap of a crack image: (a) original image, (b) the corresponding eigen-value gap distribution.

2) Crack blocks detection: To apply the aforementioned SVD property, we consider an input image as a matrix $X \in \mathbb{R}^{M\times N}$ where M and N are respectively the height and width of the image. The original image is initially divided into $\frac{MN}{w^2}$ small blocks of size $w \times w$, where w is empirically selected as 8 to provide the best result in terms of accuracy and computation time. Let us consider these blocks as sub-matrices X_{ij} for $i = 1, 2, ...\frac{M}{w}$, $j = 1, 2, ...\frac{N}{w}$. First, the diagonal matrix Λ_{ij} containing the singular values of X_{ij} is obtained from Equation (1). Then, λ_{ij} is a vector extracted from

the diagonal of the matrix Λ_{ij}

With the assumption that the image background is uniform, we consider that there are two meaningful components in each block, which are the crack and background. The detection of crack component could be achieved through estimating the distance between two largest eigenvalues $\lambda_{ij}^{(1)}$ and $\lambda_{ij}^{(2)}$. If a block has background pixels as the principal component, the energy will concentrate almost in the first eigen value, and the value for the other is considerably smaller, leading to an increase in the gap between the first and second eigenvalue. Let D be an array that contains sigular value gaps sorted in an increasing order of all blocks in image:

$$D_{ij} = |\lambda_{ij}^{(1)} - \lambda_{ij}^{(2)}| \tag{2}$$

Due to the large difference between the eigengap of crack and non-crack blocks, D would have an L-shape as shown in Fig. 2(b). The corner of this L-shape is considered as a transition, from which a threshold τ is selected to separate the crack blocks from the background ones. If the difference between crack and background pixels is not clear enough, a heuristic factor is employed to determine τ. Based on our analysis on collected crack images, τ should be set to 0.05 if the eigen-value gap distribution does not appear as a L-shape. Let C be a function to check whether a concerned block X_{ij} is background or contains crack pixels, C can be formulated as:

$$C(X_{ij}) = \begin{cases} 1 & \text{if} \quad D_{ij} \leq \tau \\ 0 & \text{if} \quad D_{ij} > \tau. \end{cases} \tag{3}$$

B. Binarization using Summit Navigator and Otsu

The blocks containing potential crack pixels determined in Equation 3 are then employed to construct a histogram where the number of background pixels is drastically decreased compared to the one from the original image. Since there is a better balance between the number of crack and background pixels, the distribution of the generated histogram becomes bi-modal. Fig. 3 demonstrates an example of a surface image, the histogram of which is unimodal and the crack emphasized image where only the pixels of the crack blocks are considered. Here, SN and Otsu are employed to calculate the best threshold for binarization of the crack blocks. SN has been developed in [14] to precisely identify true peaks from multi-modal gray-scale histograms of images. Inspired by the advance of SN in background removal applications, the algorithm is employed in this work to aid with the peak selection step. Nevertheless, as an approach to determine an optimized threshold is not discussed in [14], we utilize Otsu for the best threshold calculation. The flowchart of the proposed method is presented in Fig. 4. Let $h = (h_k)_{k=0..L-1}$ be the discrete histogram of the crack blocks extracted from the input image the pixels of which contributed into L bins. The probability of the intensity level k is then evaluated as:

$$p_k = \frac{h_k}{A}, \quad p_k \geq 0, \quad \sum_{k=0}^{L-1} p_k = 1, \tag{4}$$

where A is the total pixel number from the extracted crack blocks. It follows that:

$$\sum_{k=0}^{L-1} h_k = A. \tag{5}$$

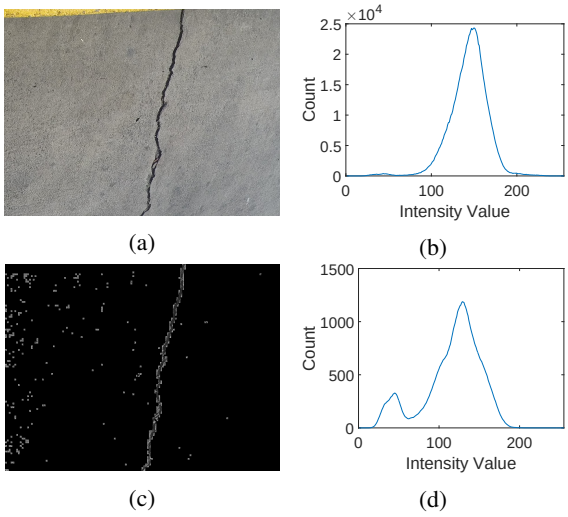


Fig. 3: Illustration of the crack emphasis process: (a) and (b) original image and its unimodal histogram, (c) and (d) image of crack block and its constructed bi-modal histogram.

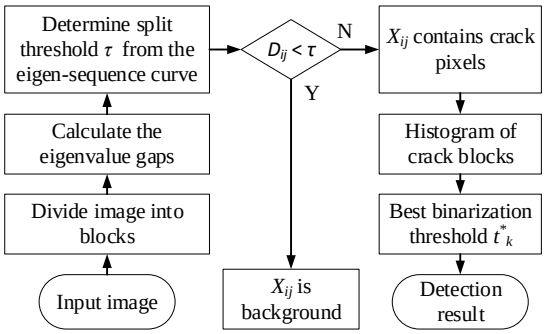
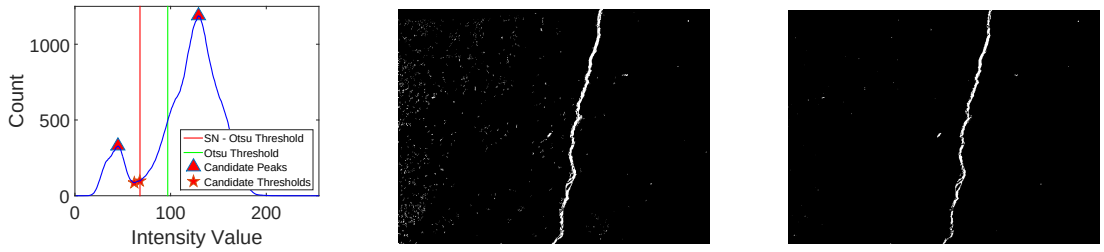


Fig. 4: Flowchart of the proposed algorithm.

The frequency at each intensity level are then compared with its two nearest neighbors to calculate initial peaks and valleys. Let $\mathbf{S}$ be a set of intensity levels of initial peaks s_k corresponding to frequency h_k as per



(a) (b) (c)

Fig. 5: Result comparison between Otsu and the combination of SN and Otsu: (a) thresholds returned by two approaches, (b) segmentation by Otsu, and (c) segmentation by SN-Otsu.

Algorithm 1 Crack blocks detection

1: Divide image into $\frac{M \times N}{w^2}$ blocks X_{ij}
2: ▷ Form the eigenvalue gap distribution of all blocks in image
3: **for** $i \leftarrow 1, \frac{M}{w}$ **do**
4: **for** $j \leftarrow 1, \frac{N}{w}$ **do**
5: $\lambda_{ij} \leftarrow$ eigenvalues calculated from SVD
6: Store $|\lambda_{ij}^{(1)} - \lambda_{ij}^{(2)}|$ in D in decreasing order
7: **end for**
8: **end for**
9: $\tau \leftarrow$ L-shape corner detection of D
10: ▷ Detect crack blocks and set background block to zero
11: Set any block X_{ij} that fulfil $|\lambda_{ij}^{(1)} - \lambda_{ij}^{(2)}| \geq \tau$ to zero
12: Apply Summit Navigator and Otsu algorithms on the remaining blocks for binarization
13: Overwrite input non-zero blocks by binarized blocks

the following condition:

$$\vec{S} = \{s_k | h_k \geq h_{k-1} \text{ AND } h_k \geq h_{k+1}\}. \quad (6)$$

Similarly, a set of intensity levels of initial valleys t_k corresponding to frequency h_k is determined as:

$$\vec{T} = \{t_k | h_k \leq h_{k-1} \text{ AND } h_k \leq h_{k+1}\}. \quad (7)$$

Next, the SN algorithm is applied on $\mathbf{S}$ to determine the two most dominant peaks, s_1^* and s_2^*, corresponding to two distribution modes of crack and background pixels. Although Otsu technique can be applied directly on the crack blocks, it has been pointed out that the calculated threshold might lead to an invalid segmentation. To overcome this limitation, we proposed to use the between-class variance developed by Otsu to search for an optimized threshold among the valley points between two dominant peaks returned by SN. This approach ensures that the calculated threshold is located at the valley between two distributions and avoids an exhaustive search in the whole range of intensity of the constructed histogram h. Let $t_k \in \mathbf{T}$ be the threshold that separates the pixels into two classes (background and crack), the between-class variance can be expressed as:

$$\sigma_B^2(t_k) = \frac{[\mu_T \omega(t_k) - \mu(t_k)]^2}{\omega(t_k)[1 - \omega(t_k)]}, \quad (8)$$

where

$$\omega(t_k) = \sum_{k=0}^{t_k} p_k, \quad (9)$$

$$\mu(t_k) = \sum_{k=0}^{t_k} k p_k, \quad (10)$$

$$\mu_T = \mu(L) = \sum_{k=0}^{L-1} k p_k. \quad (11)$$

The optimal threshold t_b^* is then defined as:

$$\sigma_B^2(t_k^*) = \max_{s_1^* < t_k < s_2^*} \sigma_b^2(t_k). \quad (12)$$

The pseudo code of the proposed algorithm is presented in Algorithm 1. Fig. 5 presents a comparison of the binarization results returned by Otsu and SN-Otsu. It is significant to see that the segmentation by Otsu in Fig. 5(b) has more noise than that of SN-Otsu in Fig. 5(c) as the threshold calculated by Otsu was located on one side of a mode instead of at the valley between two peaks.

III. Results and Discussion

The effectiveness of the proposed method is evaluated on the crack images of the SYDCrack dataset collected by our UAVs [3]. Performance of this approach is also compared with the following relevant techniques: Otsu's method [6], Sauvola's adaptive thresholding technique [10], contrast iterative thresholding (CIT) [3], slope difference distribution (SDD) [13], and the superpixel-based fast fuzzy c-means clustering (SFFCM) [12].

In this experiment, five evaluation measures [19], [20], namely the F-measure (F_β), the probabilistic rand

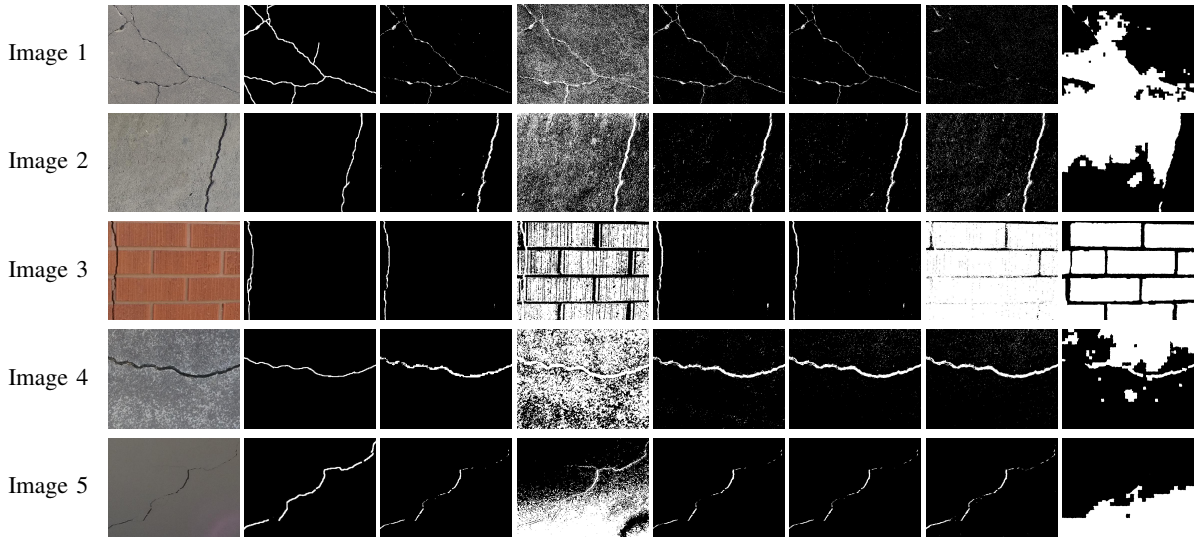


Fig. 6: Crack detection results: From left to right: Image name, original image, segmentation respectively by the proposed method, Otsu, Sauvola, CIT, SDD, and SFFCM.

index (PRI), the variation of information (VI), the global consistency error (GCE), and the boundary displacement error (BDE), are calculated to evaluate performance of participated algorithms against our human annotated segmentation. The PRI measures the similarity between two segmentations by calculating the fraction of pairs of pixels, the labels of which are consistent between the computed and ground-truth segmentation. The difference between two segmentations are also evaluated by calculating the average conditional entropy (VI), the degree of multual consistency (GCE) and the average displacement error of boundary pixels (GCE). A better segmentation should have higher F_β and PRI but lower VI, GCE, and BDE. The F-measure is calculated as:

$$F_\beta = \frac{(1+\beta^2) \times Precision \times Recall}{\beta^2 \times Precision + Recall}, \qquad (13)$$

where $Precision$ and $Recall$ represent the ratio of the correctly reported crack pixels among the predicted crack pixels and the correctly predicted crack and background pixels, and β^2 is the weight between $Precision$ and $Recall$. As discussed in [19], β^2 was selected to be 0.3 to emphasize precision over recall in defect detection.

Fig. 6 presents the segmentation results of the participated algorithms on some images from our collected UAV images. It is significant to see that the results returned by Otsu and SFFCM are not satisfying as a considerable number of background pixels are recognized as crack. On the other hand, the proposed method has provided a better result compared to Sauvola, CIT, and SDD with less noise in each segmentation. The average measures of the participated algorithms on 170 images of the SYDCrack dataset are reported in Table I, where our proposed method outperforms other algorithms in terms of F_β, PRI and BDE. The proposed method is also the second best among the participated algorithms in terms of VI and GCE.

TABLE I: Average performance of participated algorithms on the SYDcrack dataset

Algorithms	$F_\beta \uparrow$	PRI↑	VI↓	GCE↓	BDE↓
Proposed	**0.8102**	**0.9792**	0.1081	0.0107	**29.1941**
Otsu	0.5547	0.6778	0.7554	0.0217	80.8131
Sauvola	0.7879	0.9723	0.1356	0.0129	51.5721
CIT	0.7988	0.9773	**0.1046**	**0.0098**	77.3155
SDD	0.6541	0.9349	0.2367	0.0160	83.3373
SFFCM	0.5547	0.6285	0.8634	0.0253	83.8068

TABLE II: Average computation time in seconds

Proposed	Otsu	Sauvola	CIT	SDD	SFFCM
0.99	**0.03**	0.21	0.24	1.24	37.60

The experiment was executed by using MATLAB R2015a on an Intel(R) Core(TM) i5-5200U CPU @2.20 GHz with 64 bit Windows 10. The average computation time of participated algorithms is reported in Table II, where Otsu is the most computationally effective algorithm in the defect detection task. Although the proposed method is only faster than SDD and SFFCM, the result can be improved in future work as parallel

computation has not been applied on the crack blocks detection using SVD.

The experimental results obtained have indicated improved performance in terms of accuracy and consistency in combining advantages of the Summit Navigator and Otsu methods. Moreover, the simple implementation of the proposed technique makes it promising for vision-based health monitoring and fault diagnosis applications [21], [22].

IV. Conclusion

In this paper, a hybrid method integrating singular value decomposition into histogram thresholding has been proposed to deal with the defect detection problem using thresholding techniques. Based on the detected crack blocks resulted from the pre-processing step using SVD, a combination between SN and Otsu is developed for a better segmentation of crack pixels from the input image. The contribution of the research is twofold: First, the effectiveness of SVD for emphasizing crack pixels has been verified, where the constructed histogram from the crack blocks appears as a bi-modal distribution instead of a uni-modal from the input image. Then, the proposed SN-Otsu technique has improved the binarization result compared with other related thresholding techniques. Experimental results on our UAV collected images have confirmed the advantage of the proposed approach in terms of accuracy and consistency.

References

[1] H. Oliveira and P. L. Correia, "Automatic road crack detection and characterization," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 1, pp. 155–168, March 2013.

[2] L. Wang and Z. Zhang, "Automatic detection of wind turbine blade surface cracks based on uav-taken images," *IEEE Trans. Ind. Electron.*, vol. 64, no. 9, pp. 7293–7303, Sep. 2017.

[3] V. T. Hoang, M. D. Phung, T. H. Dinh, and Q. P. Ha, "System architecture for real-time surface inspection using multiple uavs," *IEEE Syst. J.*, online 5 JUL 2019. [Online]. Available: http://dx.doi.org/10.1109/JSYST.2019.2922290

[4] Q. Shi, Z. Yang, Y. Guo, H. Wang, L. Sun, Q. Huang, and T. Fukuda, "A vision-based automated manipulation system for the pick-up of carbon nanotubes," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 2, pp. 845–854, April 2017.

[5] P. Wang, D. Li, S. Shen, and Y. Shen, "Automatic microwaveguide coupling based on hybrid position and light intensity feedback," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 3, pp. 1166–1175, June 2019.

[6] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62–66, Jan 1979.

[7] C. Leung and F. Lam, "Performance analysis for a class of iterative image thresholding algorithms," *Pattern Recognit.*, vol. 29, no. 9, pp. 1523 – 1530, 1996.

[8] N. M. Kwok, Q. P. Ha, and G. Fang, "Effect of color space on color image segmentation," in *Proc. 2009 2nd Int. Congress Image Signal Process.*, 2009, pp. 1–5.

[9] W. Niblack, *An Introduction to Digital Image Processing*. Prentice-Hall, 1986.

[10] J. J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," *Pattern Recognit.*, vol. 33, pp. 225–236, 2000.

[11] S. Aja-Fernández, A. H. Curiale, and G. Vegas-Sánchez-Ferrero, "A local fuzzy thresholding methodology for multiregion image segmentation," *Knowl.-Based Syst.*, vol. 83, pp. 1 – 12, 2015.

[12] T. Lei, X. Jia, Y. Zhang, S. Liu, H. Meng, and A. K. Nandi, "Superpixel-based fast fuzzy c-means clustering for color image segmentation," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 9, pp. 1753–1766, Sep. 2019.

[13] Z. Wang, J. Xiong, Y. Yang, and H. Li, "A flexible and robust threshold selection method," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 9, pp. 2220–2232, Sep. 2018.

[14] T. H. Dinh, M. D. Phung, and Q. P. Ha, "Summit navigator: A novel approach for local maxima extraction," *IEEE Trans. Image Process.*, vol. 29, pp. 551–564, 2020.

[15] Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, and S. Wang, "Deepcrack: Learning hierarchical convolutional features for crack detection," *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1498–1512, March 2019.

[16] Y. Fei, K. C. P. Wang, A. Zhang, C. Chen, J. Q. Li, Y. Liu, G. Yang, and B. Li, "Pixel-level cracking detection on 3d asphalt pavement images through deep-learning-based cracknet-v," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–12, 2019.

[17] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, July 2017.

[18] M. Bayat, M. Fatemi, and A. Alizad, "Background removal and vessel filtering of noncontrast ultrasound images of microvasculature," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 3, pp. 831–842, March 2019.

[19] Q. Hou, M. Cheng, X. Hu, A. Borji, Z. Tu, and P. H. S. Torr, "Deeply supervised salient object detection with short connections," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 4, pp. 815–828, April 2019.

[20] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.

[21] S. Permana, E. Grant, G. M. Walker, and J. A. Yoder, "A review of automated microinjection systems for single cells in the embryogenesis stage," *IEEE/ASME Trans. Mechatronics*, vol. 21, no. 5, pp. 2391–2404, Oct 2016.

[22] Q. Zhu, T. H. Dinh, V. Hoang, M. D. Phung, and Q. P. Ha, "Crack detection using enhanced thresholding on uav based collected images," in *Proc. 2018 Australasian Conf. on Autom. Robot.,*, Canterbury New Zealand, 4-6 DEC 2018, pp. 1–7.