

Improved wolf algorithm on document images detection using optimum mean technique

Wan Azani Mustafa¹, Mohamed Mydin M. Abdul Kader², Zahereel Ishwar Abdul Khalib³

^{1,2}Faculty of Engineering Technology, Universiti Malaysia Perlis, UniCITI Alam Campus, Sungai Chuchuh, 02100 Padang Besar, Perlis, Malaysia

³Embedded, Network and Advanced Computing Research School of Computer and Communication Engineering University Malaysia Perlis, 02600 Arau, Malaysia

Article Info

Article history:

Received Dec 1, 2018

Revised Jan 31, 2019

Accepted Mar 1, 2019

Keywords:

Binarization

Document

Mean

Optimum

Wolf

ABSTRACT

Detection text from handwriting in historical documents provides high-level features for the challenging problem of handwriting recognition. Such handwriting often contains noise, faint or incomplete strokes, strokes with gaps, and competing lines when embedded in a table or form, making it unsuitable for local line following algorithms or associated binarization schemes. In this paper, a proposed method based on the optimum threshold value and namely as the Optimum Mean method was presented. Besides, Wolf method unsuccessful in order to detect the thin text in the non-uniform input image. However, the proposed method was suggested to overcome the Wolf method problem by suggesting a maximum threshold value using optimum mean. Based on the calculation, the proposed method obtained a higher F-measure (74.53), PSNR (14.77) and lowest NRM (0.11) compared to the Wolf method. In conclusion, the proposed method successful and effective to solve the wolf problem by producing a high-quality output image.

Copyright © 2019 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Wan Azani Mustafa,

Department of Electrical Engineering Technology,

Faculty of Engineering Technology, Universiti Malaysia Perlis,

UniCITI Alam Campus, Sungai Chuchuh,

02100 Padang Besar, Perlis, Malaysia.

Email: wanazani@unimap.edu.my

1. INTRODUCTION

Recently, Document Image Binarization (DIB) is one of the first stages of the Document Image Analysis (DIA) and the further recognition pipeline. It consists mainly of classifying the pixels of the image into the background and foreground pixels [1-4]. At this stage, “binarization” will imply some cleaning and enhancement, since detecting the foreground pixel will remove, for instance, ink spots, degradations or ink-bleed through, besides it will recover lost strokes or other foreground information. The degraded image quality such as contrast variation and illumination give a big effect on the binarization process [5, 6]. The research areas of document binarization from old, degraded documents attract researchers because of the challenges of segmentation of the foreground text from degraded background [7-12]. Binarization is also known in the literature as thresholding, since a threshold value is applied to the pixels to saturate them in one of the classes. Image binarization takes into account any images, but DIB is meant for documents where text (and other graphics) are considered foreground and the background is the “common background of the image [13, 14].

The general image binarization (or thresholding algorithms) could be divided into global, local thresholding and mixing strategies. Global thresholding establishes a fixed threshold for all pixels of the image, whereas local thresholding applies different thresholds for each pixel (using pixel local information). Hybrid approaches try to compose both strategies: they use global and local information for thresholding. Global thresholding techniques usually need less computation, and they work well in simple cases, though fail in complex images. The most popular global image threshold is an Otsu's method [15] which after computing the gray level histogram, takes the optimal value that minimizes the inner class variance. There are some refinements based on this method [16, 17]. Taking a closer look at document specific contributions, Ntirogiannis *et al.* [18] proposed a combination of the global and local threshold for handwritten text recognition. First, a background estimation is applied to binarize the image globally, then they work with Connected Components (CCs), the approach takes stroke characteristics and discard components that do not suit to the common stroke criteria. The above finding is consistent with the study by Moghaddam and Cheriet [19]. The work in a similar manner since it extracts features such as stroke width and line height. It adapts itself to the document structure because it combines a grid-based modeling and the estimated background map.

In degraded documents, such historical records, traditional and general methods present poor performances. Degraded documents involve not only binarization, however, the restoration of document parts (mainly lost strokes) [20, 21]. In this purpose, usually, stroke characteristics are integrated into the binarization. It is possible to find a vast set of contributions addressing these issues [22-24]. For example, Moghaddam and Cheriet [25] treat binarization of degraded documents by combining different binarization scales. Their work aims to restore lost strokes and eliminate ink bleed-through. In contrast, the study by Nina *et al.* [26] indicated that a recursive Otsu's thresholding could be used for binarization of historical documents as well. In the same objective, another adaptive approach by the combination of image contrast, gradients, and text shapes are estimated by Canny's edge extractor was studied by Su *et al.* [27]. Besides, Valizadeh and Kabir [28] suggest to extract useful characteristics in order to classify the pixels in background or foreground; they use structural contrast taking into account information like the stroke width for feature extraction. With the selected features, space is partitioned into several regions (clustering) where the pixel within each region will be classified mixing the region information and the previous features extracted.

In this paper, a modification Wolf algorithm based on optimum mean was presented. The target of this study is to improve the detection accuracy and blurred problem on binarization result and at the same time to reduce the noise and any artefact by determining the optimum threshold value based on adjusting the mean. In terms of the analysis stage, a few image quality assessment (IQA) such as F-measure, Peak Signal Noise Ratio (PSNR) and Negative Rate Metric (NRM) will be obtained. This paper is organized into four sections. Section 2 explains the proposed method which focuses on Optimum Mean Technique, Section 3 shows the results performance in terms of the IQA. Section 4 describes the summary of the work.

2. RESEARCH METHOD

The optimum mean suggests based on a modification of Wolf method. Basically, the wolf method is to localize artificial text in images and videos using a measure of accumulated gradients and morphological processing. The quality of the localized text is improved by robust multiple frame integration [29]. However, wolf method is not working properly on the image which contents many structures such as text image and retinal image. The equation of wolf method as follows;

$$T = (1 - a)m + aM + a \frac{s}{R}(m - M) \quad (1)$$

m and s is a mean and standard deviation, while a is a gain parameter. Then, M represent image size minimum value of the gray levels and R is a maximum value of the standard deviations. Actually, the Wolf method successful to segment the text region, however failed when dealing with the non-uniform and low contrast input image. Besides, the output from the Wolf method also looks blurred, not clear and low quality especially when performed on thin text.

In this paper, the objective is to determine the optimum threshold value which is likely to work better for many types of low contrast document images. A slight modification of the Wolf threshold by using adjustment of mean value was proposed. The optimum mean method was proposed to overcome this problem by finding the maximum threshold value. The main advantage of the proposed method over Wolf is that it considerably improves the image quality, reduced the blurred image and eliminate the noise.

By referring to (1), the mean value m will give a high effect on the threshold value as shown in Figure 1. For example, according to Figure 1, if the mean value is less than average ($m - 15$) the result was blurred and more information details were lost. Otherwise, if the mean value is more than the average ($m + 15$) the result is better and improved. In this paper, the specific value (maximum mean) was found to replace the normal mean value. The aim is to find the specific value more than the normal average.

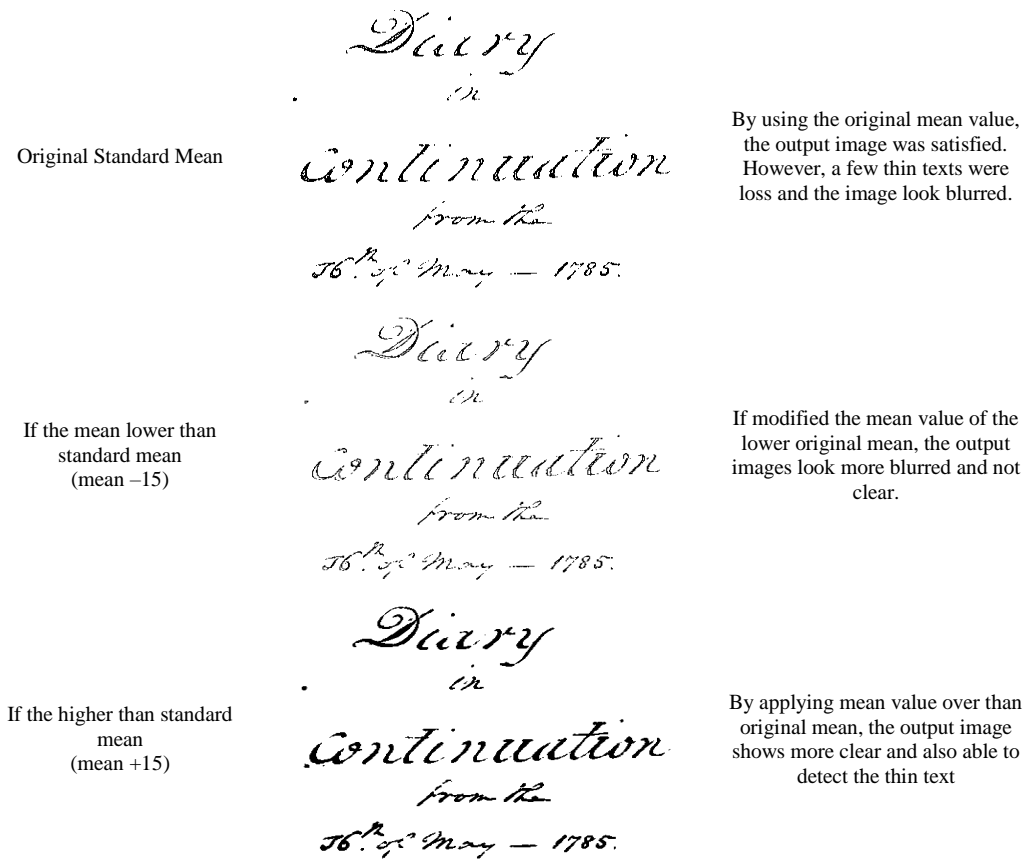


Figure 1. Comparison of resulting image with different mean value

Actually, the Wolf method failed to binarize the text image in low contrast condition because the threshold value is not accurate. Therefore, the proposed method tends to improve the threshold value based on an optimum mean calculation and applied to the low contrast image. The specific and maximum threshold value was suggested caused if the threshold value is too high, it's will introduce noise and artefact on the resulting image. In this paper, the maximum mean is calculated in order to replace the original mean. The maximum mean equation is depicted as follows;

$$O_{mean} = \max [mean(x, y)] \tag{2}$$

where, $\max(x, y)$ is the maximum intensity of the overall mean image and $mean$ is the standard mean. Normally, the method was processed on 3x3 windowing size and contain a low of the mean value in the whole images. So, we select the maximum mean value to put into the final algorithm. The final proposed algorithm is;

$$T = (1 - a)O_{mean} + aM + a \frac{S}{R} (O_{mean} - M) \tag{3}$$

Based on the modified algorithm, the blur problem and unsegment thin text can be solved. Then, it will increase the binarization result and produced high image quality. In order to evaluate the effectiveness of

the proposed method and compared to the Wolf methods, 14 document images experimented and the results are explained in the following section.

3. RESULTS AND ANALYSIS

In this study, 14 document images with non-uniform image background intensity due to uneven illumination were tested [30]. The document dataset is publicly image provided by <http://utopia.duth.gr/~ipratika/HDIBCO2012/benchmark>. Every input image shows the different pixel size and the major problem images is caused by the fade ink and degrade of background. In this experiment, all images were processed in grayscale images. The proposed process was performed on the default windowing size, which is 3x3 [31, 32]. In binarization stage, a small windowing size will give a better result compared to the larger size [33]. A few mathematical statistics such as variance and the standard deviation was obtained before applying the final proposed equation. Figure 2 illustrated the comparison resulting image for Wolf method and the proposed method. Based on observation, the wolf method failed in order to detect the thin text, especially on illumination and low contrast problem input image. However, after applying the modification proposed method was increased the resultant performance. The resulting image shows that more clearly and improved the image quality.

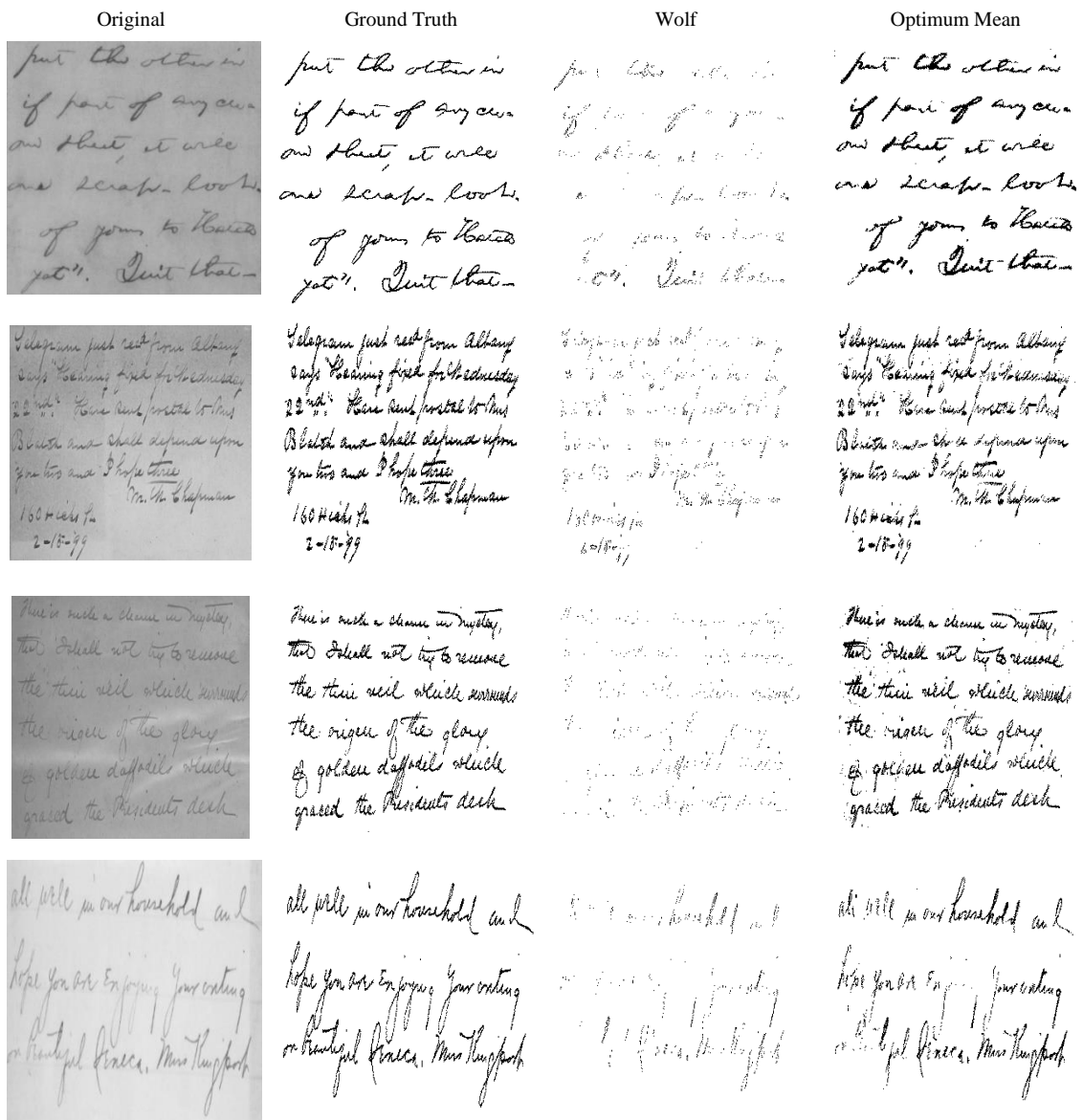


Figure 2. Comparison of Wolf method and Optimum mean method on the document image

In order to prove the performance method, a few selected Image Quality Assessment (IQA) such as an F-measure, PSNR and NRM was calculated. The quality of the segmented image is determined based on the pixels similarity of the resultant segmented image against the manually segmented image. The mathematical for F-measure, PSNR and NRM as follows;

$$F - measure = \frac{2 \times \left(\frac{TP}{TP+FN} \right) \times \left(\frac{TP}{TP+FP} \right)}{\left(\frac{TP}{TP+FN} \right) + \left(\frac{TP}{TP+FP} \right)} \quad (4)$$

$$NRM = \frac{\left(\frac{FN}{TP+FN} \right) + \left(\frac{FP}{FP+TN} \right)}{2} \quad (5)$$

$$PSNR = 10 \log_{10} \frac{S^2}{MSE} \quad (6)$$

TP, FP, FN and TN denote the True Positive, False Positive, False Negative and True Negative values, respectively [34, 35]. Mathematically, the highest of F-measure and PSNR show the good binarization result [36, 37]. Oppositely, the lowest of NRM should be present the good image quality. Table 1 presents the comparison result between Wolf method and Optimum Mean method based on IQA.

Table 1. Result of Image Quality Assessment (IQA) in terms of F-measure PSNR, NRM and MPM

Image	F-measure		PSNR		NRM	
	Wolf	Optimum Mean	Wolf	Optimum Mean	Wolf	Optimum Mean
1	70.81	78.12	15.79	15.66	0.22	0.08
2	28.90	76.28	11.49	14.19	0.42	0.14
3	26.86	81.47	12.68	15.69	0.42	0.04
4	78.95	82.46	17.68	17.29	0.17	0.05
5	72.54	65.52	15.83	12.07	0.21	0.03
6	59.42	69.90	14.12	12.59	0.28	0.04
7	27.06	60.98	12.51	13.76	0.42	0.25
8	81.31	71.35	16.40	12.44	0.15	0.03
9	37.71	84.67	10.54	14.84	0.38	0.11
10	22.53	80.93	10.58	14.63	0.44	0.13
11	12.06	84.77	10.93	15.57	0.47	0.06
12	28.17	74.29	13.18	15.81	0.42	0.17
13	18.29	49.86	13.58	14.59	0.45	0.32
14	69.86	82.89	16.66	17.62	0.23	0.05
Average	45.32	74.53	13.71	14.77	0.33	0.11

According to the Table 1, the F-measure and PSNR show the increment result after applying the proposed method. Based on the average value, the proposed method achieved an F-measure (74.53) and the PSNR (13.71) compared to the Wolf method which is an F-measure (45.32) and the PSNR (13.71). In terms of increment percentage, the proposed method scored 64.45 % (F-measure) and 7.73 % (PSNR). Besides, the proposed method also gave the lowest NRM (0.11) compared to the Wolf method (0.33). From this data, we can see that the proposed method was successful to improve the binarization performance of Wolf method

4. CONCLUSION

Nowadays, digital image processing is very important in order to analyse the information on the image. Thousands of valuable historical documents stored on the shelves of national libraries throughout the world are waiting to be scanned in order to facilitate access to the information it contains. In this paper, a new approach based on a modification of Wolf method was present. The main goal of the study was to determine the optimum threshold value using the maximum mean. The second objective is to solve the blur problem and improved the detection performance. Based on the numerical result, the proposed method obtained the highest PSNR (14.773), F-measure (74.539) and lowest NRM (0.111) compared to the Wolf method. Summary, the proposed method effective and successful to solve the Wolf problem and also increased the binarization result.

ACKNOWLEDGEMENTS

The author would like to acknowledge support from the Short Term Grant under a grant number of 9001-00576, Universiti Malaysia Perlis (UniMAP)

REFERENCES

- [1] B. M. Singh, R. Sharma, D. Ghosh, and A. Mittal, "Adaptive binarization of severely degraded and non-uniformly illuminated documents," *Int. J. Doc. Anal. Recognit.*, pp. 393–412, 2014.
- [2] E. T. Zemouri, Y. Chibani, and Y. Brik, "Restoration based Contourlet Transform for historical document image binarization," *Int. Conf. Multimed. Comput. Syst. -Proceedings*, vol. 0, pp. 309–313, 2014.
- [3] H. Z. Nafchi, R. F. Moghaddam, and M. Cheriet, "Phase-based binarization of ancient document images: Model and applications," *IEEE Trans. Image Process.*, vol. 23, no. 7, pp. 2916–2930, 2014.
- [4] A. Sehad, Y. Chibani, and M. Cheriet, "Gabor Filters for Degraded Document Image Binarization," *Proc. Int. Conf. Front. Handwrit. Recognition, ICFHR*, vol. 2014-Decem, pp. 702–707, 2014.
- [5] W. A. Mustafa and H. Yazid, "Image Enhancement Technique on Contrast Variation : A Comprehensive Review," *J. Telecommun. Electron. Comput. Eng.*, vol. 9, no. 3, pp. 199–204, 2017.
- [6] W. A. Mustafa, H. Yazid, and S. Yaacob, "A Review : Comparison Between Different Type of Filtering Methods on the Contrast Variation Retinal Images," in *IEEE International Conference on Control System, Computing and Engineering*, 2014, pp. 542–546.
- [7] W. A. Mustafa, H. Yazid, and M. Jaafar, "An Improved Sauvola Approach on Document Images Binarization," *J. Telecommun. Electron. Comput. Eng.*, vol. 10, no. 2, pp. 43–50, 2018.
- [8] E. Ahmadi, Z. Azimifar, M. Shams, M. Famouri, and M. J. Shafiee, "Document image binarization using a discriminative structural classifier," *Pattern Recognit. Lett.*, vol. 63, pp. 36–42, 2015.
- [9] N. Mitianoudis and N. Papamarkos, "Document image binarization using local features and Gaussian mixture modeling," *Image Vis. Comput.*, vol. 38, pp. 33–51, 2015.
- [10] W. A. Mustafa and H. Yazid, "Illumination and Contrast Correction Strategy using Bilateral Filtering and Binarization Comparison," *J. Telecommun. Electron. Comput. Eng.*, vol. 8, no. 1, pp. 67–73, 2016.
- [11] W. A. Mustafa and H. Yazid, "Background Correction using Average Filtering and Gradient Based Thresholding," *J. Telecommun. Electron. Comput. Eng.*, vol. 8, no. 5, pp. 81–88, 2016.
- [12] W. A. Mustafa and H. Yazid, "Contrast and Luminosity Correction Based On Statistical Region Information," *Adv. Sci. Lett.*, vol. 23, no. 6, pp. 5383–5386, 2017.
- [13] W. A. Mustafa and M. M. M. A. Kader, "Binarization of Document Images: A Comprehensive Review," *J. Phys. Conf. Ser.*, vol. 1019, no. 012023, pp. 1–9, 2018.
- [14] W. A. Mustafa and M. M. M. A. Kader, "Binarization of Document Image Using Optimum Threshold Modification," *J. Phys. Conf. Ser.*, vol. 1019, no. 012022, pp. 1–8, 2018.
- [15] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," in *IEEE Transactions On Systems, Man, And Cybernetics*, 1979, vol. 20, no. 1, pp. 62–66.
- [16] A. D. Brink, "Thresholding of digital images using two-dimensional entropies," *Pattern Recognit.*, vol. 25, no. 8, pp. 803–808, 1992.
- [17] J. Kittler and J. Illingworth, "On threshold selection using clustering criteria," *IEEE Trans. Syst. Man. Cybern.*, vol. SMC-15, no. 5, pp. 652–655, 1985.
- [18] K. Ntirogiannis, B. Gatos, and I. Pratikakis, "A combined approach for the binarization of handwritten document images," *Pattern Recognit. Lett.*, vol. 35, pp. 3–15, Jan. 2014.
- [19] Reza Farrahi Moghaddam and M. Cheriet, "AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization," *Pattern Recognit.*, vol. 45, pp. 2419–2431, Jun. 2012.
- [20] W. A. Mustafa, H. Aziz, W. Khairunizam, Z. Ibrahim, S. Ab, and Z. M. Razlan, "Review of Different Binarization Approaches on Degraded Document Images," in *IEEE International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, 2018, pp. 1–8.
- [21] W. A. Mustafa, W. Khairunizam, Z. Ibrahim, S. Ab, and Z. M. Razlan, "Improved Feng Binarization Based on Max-Mean Technique on Document Image," in *IEEE International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, 2018, pp. 1–6.
- [22] J. Amudha, N. Pradeepa, and R. Sudhakar, "A survey on digital image restoration," in *Procedia Engineering*, 2012, vol. 38, pp. 2378–2382.
- [23] C. C. Fung and R. Chamchong, "A review of evaluation of optimal binarization technique for character segmentation in historical manuscripts," in *3rd International Conference on Knowledge Discovery and Data Mining, WKDD 2010*, 2010, pp. 236–240.
- [24] J. Kaur and R. Mahajan, "A Review of Degraded Document Image Binarization Techniques," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 3, no. 5, 2014.
- [25] R. Farrahi Moghaddam and M. Cheriet, "A multi-scale framework for adaptive binarization of degraded document images," *Pattern Recognit.*, vol. 43, no. 6, pp. 2186–2198, Jun. 2010.
- [26] O. Nina, B. Morse, and W. Barrett, "A Recursive Otsu Thresholding Method for Scanned Document Binarization," in *IEEE Workshop on Applications of Computer Vision (WACV)*, 2010, pp. 307–314.
- [27] B. Su, S. Lu, and C. L. Tan, "Robust document image binarization technique for degraded document images," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1408–1417, 2013.
- [28] M. Valizadeh and E. Kabir, "Binarization of degraded document image based on feature space partitioning and

- classification,” *Int. J. Doc. Anal. Recognit.*, vol. 15, no. 1, pp. 57–69, 2012.
- [29] C. Wolf, J. Jolion, R. F. V Rr-, I. De Lyon, and A. A. Einstein, “Extraction and Recognition of Artificial Text in Multimedia Documents,” *Pattern Anal. Appl.*, vol. 6, no. 4, pp. 309–326, 2003.
- [30] W. A. Mustafa and M. M. M. A. Kader, “Document Image Database (2009 - 2012): A Systematic Review,” *J. Phys. Conf. Ser.*, vol. 1019, no. 012024, pp. 1–7, 2018.
- [31] K. Saeed and M. Albakoor, “Region growing based segmentation algorithm for typewritten and handwritten text recognition,” *Appl. Soft Comput.*, vol. 9, no. 2, pp. 608–617, Mar. 2009.
- [32] J. Ruiz-del-Solar and J. Quinteros, “Illumination compensation and normalization in eigenspace-based face recognition: A comparative study of different pre-processing approaches,” *Pattern Recognit. Lett.*, vol. 29, pp. 1966–1979, 2008.
- [33] S. Y. W. Azani Mustafa, Haniza Yazid, “Illumination Normalization Of Non-Uniform Images Based On Double Mean Filtering,” in *IEEE International Conference on Control System, Computing and Engineering*, 2014, pp. 366–371.
- [34] R. Hedjam, H. Z. Nafchi, R. F. Moghaddam, M. Kalacska, and M. Cheriet, “ICDAR 2015 contest on MultiSpectral Text Extraction (MS-TE_X 2015),” *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, pp. 1181–1185, 2015.
- [35] I. Pratikakis, B. Gatos, and K. Ntirogiannis, “ICDAR 2013 document image binarization contest (DIBCO 2013),” in *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, 2013, pp. 1471–1476.
- [36] W. A. Mustafa and H. Yazid, “Conversion of the Retinal Image Using Gray World Technique,” *J. Biomimetics, Biomater. Biomed. Eng.*, vol. 36, pp. 70–77, 2018.
- [37] W. A. Mustafa and M. M. M. A. Kader, “A Comparative Study of Automated Segmentation Methods for Cell Nucleus Detection,” *Malaysian Appl. Biol.*, vol. 47, no. 2, pp. 125–129, 2018.

BIOGRAPHIES OF AUTHORS



Wan Azani Mustafa received degree in Biomedical Electronic Engineering (2013) and PhD in Mechatronic Engineering (2017) from University Malaysia Perlis (UniMAP). Registered on Board Engineer Malaysia (BEM) on 2014. Published more than 40 academic articles and 1 book. Now, working as a senior lecturer at University Malaysia perlis, Malaysia. Current interest includes image processing, biomechanics, intelligence system, and control system.



Mohamed Mydin M. Abdul Kader received Master of Science In Electrical Engineering from Universiti Sains Malaysia (USM). Recognition as a Professional Technologist member from Malaysia Board Of Technologists (MBOT) (2018). Now, working as a lecturer at University Malaysia perlis, Malaysia. Current interest includes image processing, control system, Robotic and Automation.



Zahereel Ishwar Abdul Khalib received Ph.D in Computer Engineering, UniMAP MSc. in Real Time Software Engineering, UTM B.Sc in Computer Engineering, California State University, Sacramento, USA. Now, working as a senior lecturer at School of Computer & Communication Engineering, University Malaysia perlis, Malaysia. Research interest Image and Signal Processing, Real Time, SystemMulticore Programming and Multimedia Streaming.