Malays. Appl. Biol. (2018) 47(2): 71–76

MALARIA PARASITES SEGMENTATION BASED ON SAUVOLA ALGORITHM MODIFICATION

WAN AZANI MUSTAFA[1*], AIMI SALIHAH ABDUL-NASIR[2] and ZEEHAIDA MOHAMED[3]

[1,2]Faculty of Engineering Technology, Universiti Malaysia Perlis, UniCITI Alam Campus, Sungai Chuchuh, 02100 Padang Besar, Perlis, Malaysia
[3]Department of Microbiology & Parasitology, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia
[*]E-mail: wanazani@unimap.edu.my

Accepted 25 April 2018, Published online 25 May 2018

ABSTRACT

Malaria is a serious disease caused by a blood parasite of the genus *Plasmodium* and becomes a leading cause of death in the world, particularly in Africa and South Asia. In general, the conventional malaria diagnosis based on manual microscopic observation under a light microscope will increase the chance of false detection and delay diagnosis process. As a result, many researchers have proposed automated malaria detection based on image processing approach in order to provide prompt detection of malaria parasite as well as increasing the accuracy of malaria diagnosis. This paper proposed a new method based on algorithm modification that has been inspired by the Sauvola's segmentation method. The objective of the proposed method is to improve the Sauvola method and achieve better segmentation results compared to the Feng method, Bradley method, and Nick method. Overall, the results of the numerical simulation indicate that the proposed method is the most effective and efficient (specificity = 99.94% and accuracy = 98.04%) compared to other methods. Hence, the implications of this image analysis would give future research directions for the researchers.

Key words: Algortihm, Malaria, Plasmodium, Parasites, Segmentation, Sauvola

INTRODUCTION

In the recent decades, malaria disease has been one of the major interesting research subjects due to several novel automated detection approaches. Segmentation of medical images such as blood image is a challenging task in the image processing field and analysis due to the appearance of the noise, shadow, random background, overlapping objects, and illumination problem (Mustafa *et al.,* 2014; Beliz-Osorio *et al.,* 2011; Genctav *et al.,* 2012; Mustafa & Yazid, 2018; Mustafa & Yazid, 2017). In the past decades, many researchers have been primarily concentrating and focusing on automated segmentation technique as it is more accurate and effective compared to the conventional methods (Mustafa & Yazid, 2016b; Mustafa & Yazid, 2016a; Abdul-nasir *et al.,* 2013; Abdul-nasir *et al.,* 2015; Mustafa *et al.,* 2017). The sensitivity and correct diagnosis information are very important in order to help doctors/pathologist to diagnose malaria cases. Furthermore, the automated system is a faster process to identify the disease compared to the conventional technique by using the microscope procedure (Kurer & Gejji, 2014; Thung & Suwardi, 2011).

Recently, researchers have shown an increase of interest in segmenting the malaria parasite by using an unsupervised technique rather than a supervised technique (Purwar *et al.*, 2011; Abdul-nasir *et al*., 2015). In general, the unsupervised technique is more practical and efficient for real environment/ condition during analysis of an image without the need of a positive sample. Fang *et al.* (2011) presented an unsupervised method using quaternion Fourier transform (QFT) spectrum. QFT spectrum was used to obtain and construct the image as well as to detect the malaria parasite location. This study is supported by Purwar *et al.* (2011) which suggested the use of an unsupervised technique for detection of malaria parasites in microscopic images. This method was inspired by Giemsa-stained idea which focusing on two stages; (1) enumeration and (2) identification.

* To whom correspondence should be addressed.

Most of the previous work applied Otsu's thresholding for segmentation of malaria image. However, this method heavily depends on image quality and fails when the histogram does not have distinct valleys. In order to overcome the issue, Ghosh *et al*. (2011) employed fuzzy divergence minimization approach based on Cauchy membership function to perform the threshold selection for segmenting the malaria parasite from the infected RBC and the background regions. The method was applied for segmentation of trophozoite, schizont and gametocyte regions in *P. vivax* images. This method aims to assist the thresholding technique by locating the deep valleys of the histogram, especially in the case when the histogram of the image is not crisp and the regions are not well defined. A number of researchers made a noticeable number of computer vision studies on segmenting the malaria images. The recent trends in segmenting the malaria images include Zack thresholding (Chakrabortya *et al.,* 2015), Active Contour based segmentation (Arnau & García, 2015), clustering (Abdul-nasir *et al.*, 2015; Abdul-nasir *et al.*, 2013). Hence, the current study will utilize the potential image segmentation technique via morphological approach for obtaining fully segmented malaria parasites.

In this paper, a new method based on Sauvola algorithm modification is proposed. The proposed method has been tested on 50 malaria parasites of *P. vivax* species. The input image was selected randomly from 4 slides dataset. The objective of this paper was to propose and improve the segmentation result compared to the conventional segmentation methods. As a comparison, a few selected local segmentation approaches such as local adaptive, Feng, Bradley, Sauvola and Nick and methods have been tested on the malaria image dataset. The descriptions of each segmentation method are summarized in Table 1.

Table 1. Comparison of Various Segmentation Methods

Method	Description
Feng	Based on adaptively exploiting the local image contrast. Normalize the contrast effect based on local mean and standard deviation (Feng & Tan, 2004; Khurshid *et al.*, 2009). $$T=(1-\alpha 1)\times m+\alpha 2\times(\frac{s}{R_s})\times(m-M)+\alpha 3\times M$$ R_s = dynamic range of gray value standard deviation, m = mean value, s = standard deviation, α = coefficient, M = minimum value of the gray levels. **Limitation**: Need to find the maximum coefficient and window size manually. Also proposed complicated algorithm.
Bradley	The key idea of the algorithm is that every image's pixel is set to black if its brightness is T percent lower than the average brightness of surrounding pixels in the window of the specified size, otherwise it is set to white (Bradley & Roth, 2011). k is the constant parameter. $$T=m\left(1-\frac{k}{100}\right)$$ m = mean value and k = 12 **Limitation**: Is not capable to solve the extreme illumination problem.
Sauvola	To solve the problem of black noise depending on the impact on the standard deviation value by using a range of gray-level values in the images (Bataineh *et al.*, 2011; Sauvola *et al.*, 1997; Sauvola & Pietikäinen, 2000). $$T=m\left(1-k\left(1-\frac{\sigma}{R}\right)\right)$$ R = gray-level (128), m = mean value, σ = standard deviation and k = 0.1 **Limitation**: If the contrast between the foreground and background is small or if the text is in thin pen stroke text.
Nick	To improve the Niblack method (black noise) and Sauvola method (low contrast) by shifting the thresholding value downward (Khurshid *et al.,* 2009; Bataineh *et al.,* 2011). $$T(x,y)=m+k\sqrt{\frac{(I^2-m^2)}{N}}$$ m = mean value, k = –0.13, I = pixel intensity and N = image size **Limitation**: Is unable to perform good segmentation result if the contrast is too small or the text is in thin pen stroke text.

MATERIALS AND METHODS

The proposed algorithm is inspired by the Sauvola method. In this paper, the modification of Sauvola algorithm by obtaining the binarization threshold which is likely to work better in order to improve the segmentation result especially for degraded and noisy sample images. In fact, the Sauvola method able to solve the problem of black noise depending on the impact on the standard deviation value by using a range of grey level values in the images (Bataineh *et al*., 2011; Sauvola & Pietikäinen 2000). However, the Sauvola method failed to segment if the contrast between the foreground and background is small or if the text is in thin pen stroke text. This method is proposed to overcome this problem by obtaining the maximum threshold value. The main advantage of the proposed method over Sauvola method is that it considerably improves binarization for "lost" detail images by shifting up the binarization threshold. The Sauvola algorithm is denoted as follows;

$$T = m\left(1 - k\left(1 - \frac{\sigma}{R}\right)\right) \quad (1)$$

Where, R is the grey level (128), m is the mean value, σ is the standard deviation, and k was set is 0.2 (default value). This method outperforms the Niblack approach in images where the text pixels have near 0 grey value and the background pixels have near 255 grey values. However, in images where the grey values of text and non-text pixels are close to each other, the results degrade significantly. Based on research, the mean value m will give a high effect on the threshold value as shown in Figure 1. According to Figure 1, example, if the mean value less than average (m – 20) the result was blurred and more information details were lost. Otherwise, if the mean value more than the average (m + 20) the result is better and improved. In this paper, the specific value was found to replace the normal mean value. The target is to determine the specific value that more than the normal average.

Actually, the Sauvola method failed to binarize the low contrast region because the threshold value is low. Therefore, the proposed method tends to increase the threshold value to segment the information in the low contrast region. However, if the threshold value is higher, it will introduce noise and artefact on the resulting image. Oppositely, if the low threshold, the output image will lose. Thus, the specific and the maximum threshold value needs to be proposed. In this paper, the maximum mean is calculated in order to replace the original mean. The maximum mean equation is depicted as follows;

$$m_{\max} = \frac{\max(x, y) + mean}{2} \quad (2)$$

Where, $\max(x, y)$ is the maximum intensity of input image and *mean* is the original mean for the whole image. The average between the highest intensity and mean image was calculated. The main target is to improve the lost details on binarization result and at the same time to reduce the noise and any artefact. The final proposed algorithm is;

$$T = \frac{\max(x, y) + mean}{2}\left[1 - k\left(1 - \frac{\sigma}{R}\right)\right] \quad (3)$$

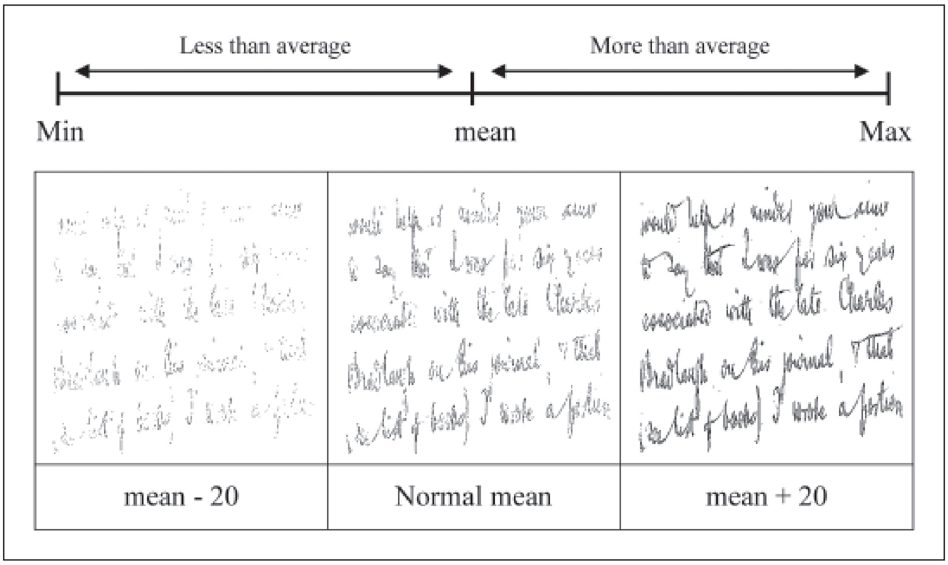

Fig. 1. Binarization effect after applying the different mean value.

where, k and R value used a default value form Sauvola method. From this algorithm, the low contrast problem can be solved and automatically increased the binarization result. In order to evaluate the proposed method and compare the results with a few local methods, 50 malaria images were tested and the results are given in the following section.

RESULTS AND DISCUSSION

In this experiment, 50 sample malaria images with non-uniform image background intensity due to uneven illumination were tested. All the processed images are in grayscale images and the size of each image is 808 × 608 pixels, 96 dpi, and 8-bit depth. Firstly, the original colour image was separated into three channels which are red, green, and blue in order to compare for the best contrast. Based on the observation among these three colour channels, the green channel was selected as the input image for segmentation process as the target region is mostly highlighted compared to other channels. The result of images after the segmentation process is shown in Figure 2. As shown in Figure 2, all segmentation approachable and satisfied in order to detect the malaria parasite, however, the segmentation methods complicated to deal with the contrast problem cause the intensity level of the target region approximate to the intensity level of background images. The low contrast will effect on the segmented result.

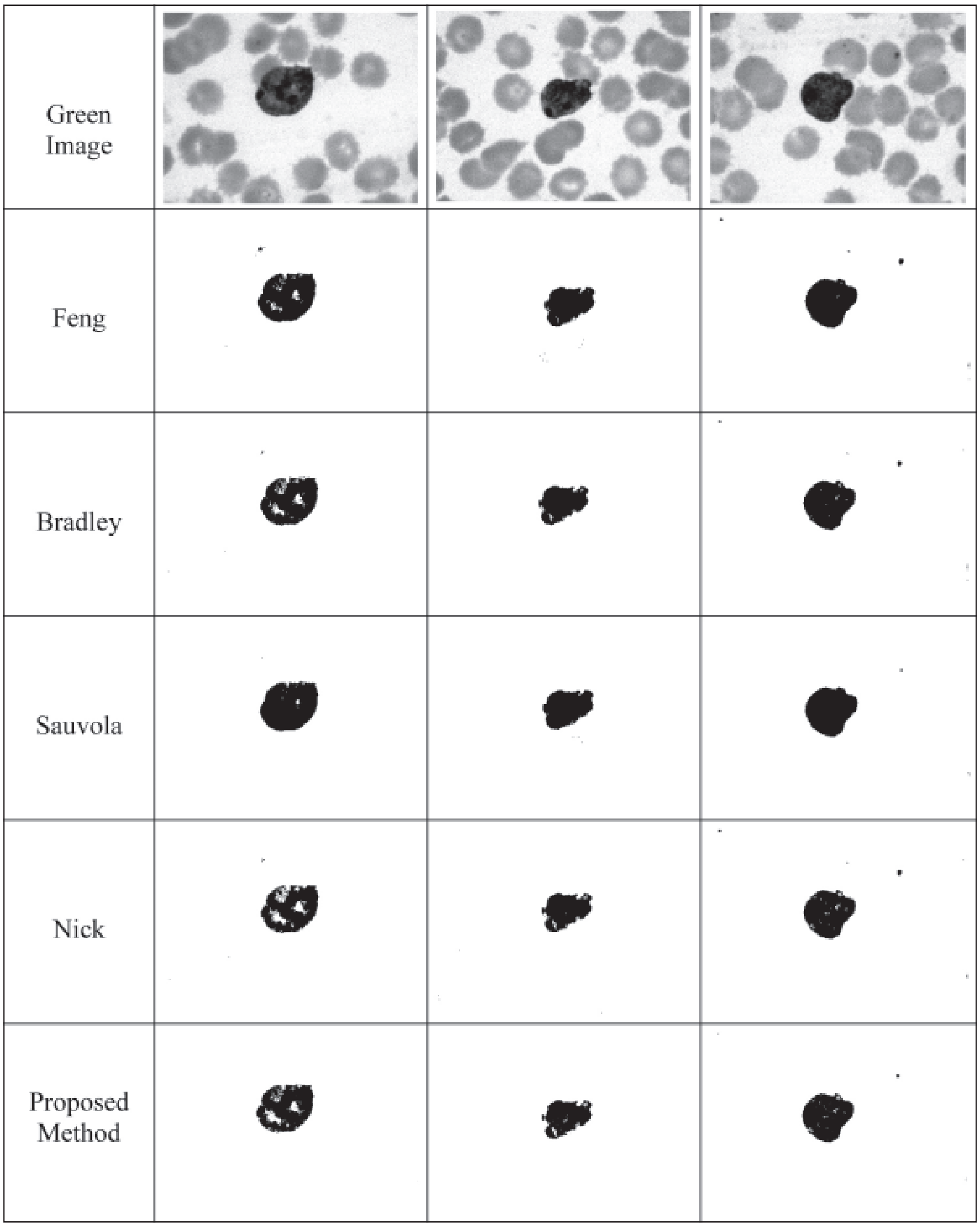

Fig. 2. Comparison of different segmentation techniques on malaria image.

Table 2. Comparison segmentation result based on IQA

Methods	Image Quality Assessment (IQA)		
	Sensitivity (%)	Specificity (%)	Accuracy (%)
Feng	88.65	96.56	96.28
Bradley	85.15	97.75	97.31
Sauvola	91.09	97.81	97.58
Nick	82.15	98.03	97.47
Proposed Method	84.54	99.94	98.04

Next, a few objective evaluations were calculated in order to assess the overall performance of different segmentation techniques such as accuracy, sensitivity, and specificity. The quality of the segmented image is determined based on the pixels similarity of the resultant segmented image against the manually segmented image. The details explanation about the accuracy, sensitivity, and specificity can refer on (Mustafa *et al.*, 2014). The result of sensitivity, specificity, and accuracy should be higher to show a well-segmented image and otherwise (Yazid & Arof, 2013; Pratikakis *et al.,* 2011). The result performance is presented in Table 2. From the data in Table 2, it is apparent that all selected method gave the satisfied result. The Sauvola method achieves the highest result in terms of sensitivity with 91.09%, followed by Feng method which is 88.65% and others. Based on the specificity and accuracy, the proposed method shows the effectiveness with achieved the highest result (specificity = 99.94 and accuracy = 98.04) compared to the other methods. Mathematically, the specificity will be high since all malaria parasites which are detected similarly to the ground truth image. However, the sensitivity will be lower since most of the detected black region is not correctly identifies as the background.

CONCLUSION

Malaria is an endemic, global and life-threatening disease. Development of automated malaria detection techniques is still a field of interest. Automated detection is faster and high accuracy compared to the traditional technique using microscopy. In this paper, a new method was presented based on a modification of Sauvola algorithm. The aim of modification is to find the best maximum threshold value, especially for the malaria images. The proposed method was tested on 50 sample malaria images and compared with the Feng method, Sauvola method, Bradley method, and Nick method. After applied the objective evaluations, the proposed method achieved; sensitivity = 84.54%, specificity = 99.94%, accuracy = 98.04%. Overall, the proposed method is more effective and successful in order to detect the malaria parasite compared to the other methods.

ACKNOWLEDGMENT

This work was supported by Short Term Grant Scheme (Grant No. 9001-00514).

REFERENCES

Abdul-nasir, A.S. Mashor, M.Y., Halim, N.H.A. & Mohamed, Z. 2015. The Cascaded Moving k-Means and Fuzzy c-Means Clustering Algorithms for Unsupervised Segmentation of Malaria Images. *AIP Conference Proceedings*. pp. 1–11.

Abdul-nasir, A.S., Mashor, M.Y. & Mohamed, Z. 2013. Colour Image Segmentation Approach for Detection of Malaria Parasites Using Various Colour Models and k-Means Clustering. *WSEAS Transactions on Biology And Biomedicine*, **10(1)**: 41-55.

Arnau, X.G. & García, S.H. 2015. Erythrocyte shape classification using integral – geometry – based methods. *Medical & Biological Engineering & Computing*, pp.623-633. Available at: http://dx.doi.org/10.1007/s11517-015-1267-x.

Bataineh, B., Abdullah, S.N.H.S. & Omar, K. 2011. An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows. *Pattern Recognition Letters*, **32**: 1805-1813.

Beliz-Osorio, N., Crespo, J., García-Rojo, M., Muñoz, A. & Azpiazu, J. 2011. Cytology Imaging Segmentation Using the the Locally Constrained Watershed Transform. In *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing*. pp. 429-438.

Bradley, D. & Roth, G. 2011. Adaptive Thresholding Using the Integral Image. *Journal of Graphics, GPU, and Game Tools*, **12(2)**: 13-21.

Chakrabortya, K., Chattopadhyayb, A., Chakrabarti, A., Acharyad, T. & Dasguptae, A.K. 2015. A Combined Algorithm for Malaria Detection from Thick Smear Blood Slides. *Journal Health Medical Informatics*, **6(1)**: 1-6.

Fang, Y., Xiong, W., Lin, W. & Chen, Z. 2011. Unsupervised Malaria Parasite Detection Based on Phase Spectrum. In *Annual International Conference of the IEEE EMBS*. pp. 7997-8000.

Feng, M. & Tan, Y. 2004. Contrast adaptive binarization of low quality. *Electronics Express*, **1(16)**: 501-506.

Genctav, A., Aksoy, S. & Onder, S. 2012. Unsupervised segmentation and classification of cervical cell images. *Pattern Recognition*, **45**: 4151-4168.

Ghosh, M., Das, D. & Chakraborty, C. 2011. *Plasmodium vivax* segmentation using modified fuzzy divergence. In *International Conference on Image Information Processing (ICIIP 2011)*. pp. 10–12.

Khurshid, K., Siddiqi, I., Faure, C. & Vincent, N. 2009. Comparison of Niblack inspired Binarization methods for ancient documents. *Proceedings of SPIE-IS&T Electronic Imaging*, **7247**: 1-9.

Kurer, D.A. & Gejji, V.P. 2014. Detection of Malarial Parasites in Blood Images. *International Journal of Engineering Science and Innovative Technology (IJESIT)*, **3(3)**: 651-656.

Mustafa, W.A., Yazid, H., Yaacob, S.B. & Basah, S.N.B. 2014. Blood vessel extraction using morphological operation for diabetic retinopathy. *IEEE Region 10 Symposium*, **(3)**: 208-212. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6863027.

Mustafa, W.A. & Yazid, H. 2016a. Background Correction using Average Filtering and Gradient Based Thresholding. *Journal of Telecommunication, Electronic and Computer Engineering*, **8(5)**: 81-88.

Mustafa, W.A. & Yazid, H. 2017. Contrast and Luminosity Correction Based On Statistical Region Information. *Advanced Science Letters*, **23(6)**: 5383-5386.

Mustafa, W.A. & Yazid, H. 2018. Conversion of the Retinal Image Using Gray World Technique. *Journal of Biomimetics, Biomaterials and Biomedical Engineering*, **36**: 70-77.

Mustafa, W.A. & Yazid, H. 2016b. Illumination and Contrast Correction Strategy using Bilateral Filtering and Binarization Comparison. *Journal of Telecommunication, Electronic and Computer Engineering*, **8(1)**: 67-73.

Mustafa, W.A., Yazid, H. & Kamaruddin, W. 2017. Combination of Gray-Level and Moment Invariant for Automatic Blood Vessel Detection on Retinal Image. *Journal of Biomimetics, Biomaterials and Biomedical Engineering*, **34**: 10-19. Available at: http://www.scientific.net/JBBBE.34.10.

Pratikakis, I., Gatos, B. & Ntirogiannis, K. 2011. ICDAR 2011 Document Image Binarization Contest (DIBCO 2011). *International Conference on Document Analysis and Recognition*, pp.1506-1510. Available at: *http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6065249*.

Purwar, Y., Shah, S.L., Clarke, G., Almugairi, A. & Muehlenbachs, A. 2011. Automated and unsupervised detection of malarial parasites in microscopic images Automated and unsupervised detection of malarial parasites in microscopic images. *Malaria Journal*, **10(1)**: 364. Available at: *http://www.malariajournal.com/content/10/1/364.*

Sauvola, J., Seppanen, T., Haapakoski, S. & Pietikainen, M. 1997. Adaptive Document Binarization. *International Conference on Document Analysis and Recognition*. pp. 147-152.

Sauvola, J. & Pietikäinen, M. 2000. Adaptive document image binarization. *Pattern Recognition*, **33**: 225-236.

Thung, F. & Suwardi, I.S. 2011. Blood Parasite Identification using Feature Based Recognition. *International Conference on Electrical Engineering and Informatics*. pp. 1-4.

Yazid, H. & Arof, H. 2013. Gradient based adaptive thresholding. *Journal of Visual Communication and Image Representation*, **24(7)**: 926-936. Available at: *http://linkinghub.elsevier.com/retrieve/pii/S1047320313001144*.