Rika Rosnelly

by Lppm Universitas Potensi Utama

An Overview of Malaria Identification Techniques for Microscope Blood Images

Ary Kusuma Ningsih
Departement of Computer Science and Electronics
Universitas Gadjah Mada
Yogyakarta, Indonesia
ary.kusuma.n[at]mail.ugm.ac.id

Sri Hartati
Departement of Computer Science and Electronics
Universitas Gadjah Mada
Yogyakarta, Indonesia
tati.ugm[at]gmail.com

Rika Rosnelly
Departement of Computer Science and Electronics
Universitas Gadjah Mada
Yogyakarta, Indonesia
rikarosnelly[at]gmail.com

***Abstract*— Malaria is a deathly disease caused by parasites that are transmitted to human through the bite of infected Anopheles mosquitoes. One of these parasites, *P. Falciparum* can progress to severe illness and often lead to death if not treated within 24 hours. Thus, early diagnosis of parasites in the red blood cell is crucial to decrease the number of malaria victim. Manual diagnosing require a well trained medical staff and several tools. This can be optimized using digital image processing to reliably detect or classify the presence of the parasites in red blood digital image. By doing microscopic digital image processing, the expected image of microscopic blood clots can be analyzed more effective, efficient and accurate. This paper provide an overview of the previous study of malaria identification techniques. We categorize differrent technique of microscopic digital image processing from pre-processing step to identification step and provide a clear comparison between them in term of pros and cons and identification accuracy.**

Keywords—malaria, image processing, identification, classification

I. Introduction

Malaria is a deathly disease caused by parasites that are transmitted to human through the bite of infected Anopheles mosquitoes. In 2015, there were 95 countries and areas affected by malaria. A total of 3.2 billion people are at risk of malaria disease; the number is nearly half of the world's population. The spread of malaria on a global scale is so vast that it includes 100 countries tropical or sub-tropical climates. Most cases of malaria deaths occur in sub-Saharan Africa. But Asia, Latin America, and the Middle East are also at high risk for malaria. [1]

There are five parasites that cause malaria in human, *Plasmodium falciparum*, *Plasmodium vivax, Plasmodium ovale, Plasmodium malariae* and *Plasmodium knowlesi*. The life-cycle of malaria parasites can be divided into four stages, which are ring, trophozoite, schizont and gametocyte.[2] One of these parasites, P. Falciparum can progress to severe illness and often lead to death if not treated within 24 hours.[1] Thus, early diagnosis of parasites in the red blood cell is crucial to decrease the number of malaria victim. Manual diagnosing require a well trained medical experts or microscopist and high quality medical facilities. Diagnosis process of malaria can be optimized using digital image processing to reliably detect or classify the presence of the parasites in red blood digital image. By doing microscopic digital image processing, the expected image of microscopic Red Blood Cell (RBC) can be analyzed more effective, efficient and accurate.

Microscopic images of malaria parasites are difficult to distinguish from one another and often contain noise coming from small artifacts such as bacteria, and so forth and so we need an experienced medical personnel to be able to provide an accurate diagnosis of the patient with symptoms of malaria morbidity where the diagnosis is done manually. Figure 2.1 shows a microscopic image of malaria.

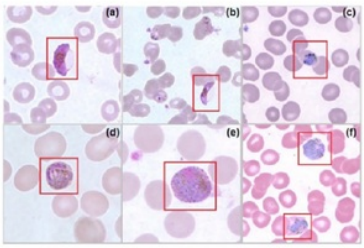


Figure 2.1 The image of the microscopic parasite Malaria (marked with a red box from left to right) are: (a) and (b) of *P. falciparum* (c) *P. knowlesi* (d) *P. malariae* (e) *P. ovale* (f) *P . vivax* [3]

A microsopic digital image processing of malaria blood images require several steps that are crucial to diagnose the malaria illness based on parasite type and stages. These steps include pre-processing, segmentation, feature extraction and the final step are identification or classification. There are many approaches of algorithm used in each step that lead to differrent accuracy.

II. Literature Review

A. Pre-process

The first stage in the process of digital image processing is pre-processing that aims to improve the image

978-1-5090-1648-8/16/$31.00 ©2016 IEEE

quality and remove noise from the image being studied. [4] proposed a pre-processing stage with a 5x5 median filter that works well to eliminate any unexpected noise gray scale images without losing the detail of the parasite. Later, [6] created a system of automatic identification of malaria by P. Falciparum parasite research object. Pre-processing on the research was done by converting the RGB image to grayscale and then filtered using a median filter to remove noise in the form of small spots on the image. The image that was filtered and then subtracted to the image imposed morphological closing to obtain background. Last stages serves to eliminate uneven lighting in the image so that the resulting image is ready to be segmented.

Median filters are also used in research [5] to remove noise from gray scale images. The 3x3 median filter works by replacing each pixel value at the midpoint of the window 3x3 median of intensity neighbors. This process aims to reduce the sharpness of the image (blur) to obtain a noise-free image.

[3] proposed a method of pre-processing by converting an RGB image into a binary image and then removes noise using adaptive filtering method Weiner. Research [12] focused on improving image quality at the stage of pre-processing using Gamma Equalization (GE). The process was carried out as follows first, convert RGB image to grayscale to reduce the effect of the difference in color and was used for processing a single channel. After that, the calculation of the range of values for γ order of grayscale images was carried out. In this study produced that $\gamma = 0.8$ was the most excellent contrast, it was more advanced than some other methods are histogram equalization (HE), IM adjust (IA) and Contrast-limited adaptive histogram equalization (CLAHE).

Median filter 3x3 was also used in [11] and [8] research, as well as were done in previous studies by [4], [5] and [6], but in this study, [11] proposed median filter followed by a Laplacian filter to refine and improve the image of the edge of the image. While [8] proposes to perform contrast stretching and image conversion from RGB to HSV before being given a 3x3 median filter, this stage produces images with good contrast and noise reduction.

Convertion of image from RGB to HSV is considered better because it produces clearer image, it was stated by [10]. Furthermore, [10] proposed the application of morphological methods morphological opening followed by closing the image that has been converted into HSV. This proccess to eliminate noise in the form of artifacts that contain small objects that are not required. Table 1 show study of differrent preprocessing techniques for malaria blood images.

B. Segmentation

The next stage is the stage of segmentation which aims to separate the foreground object (erythrocytes, blood components, and parasites) on a background. Research [4] converted the image into a binary image preprocessing by setting the threshold value that separates the image into two classes based intensity. This process was followed by applying the Otsu method to set a threshold level that maximizes the variance between the two classes histogram automatically. Red blood cells have a biconcave shape that causes cavities contained on foreground objects. Therefore, [4] designed process of filling a cavity (hole filling) based on the extraction of connected components. Hole filling process started by creating a complementary image later labeling of connected components. Furthermore, the reconstructed image by adjusting the image intensity at a value of 0 on the object that has been labeled and value of 255 on objects that are not labeled. This process concludes with the complement of the image of the image that has been reconstructed. Results from hole filling process were treated to the next connected component extraction step to obtain binary markers of individual cells.[5] was using this method in 2013.

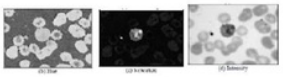
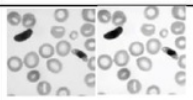
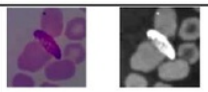
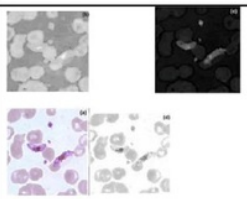
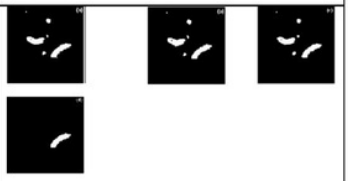
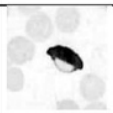
According to [7], several methods that were proposed previously for segmentation which can work well if there were a striking difference in color between the area of blood infected by the parasite and background. Therefore, [7] focused on the detection of malaria based segmentation approach that utilized the potential of color image segmentation approach using HSI color space and moving algorithm k-means clustering. This study aimed to get the full segmentation of malaria-infected blood by a thin smear of blood image. The segmentation was done by applying a clustering method of MSM in the S component of the image obtained blood cells infected with segmented full.

[6] proposed the use of Zack thresholding to determine boundaries between the two peaks and extract the foreground of the image. Zack thresholding method was performed iterative least three times to remove the clumps apart RBC (Red blood cell), eliminating artifacts, unwanted fill the gap and to increase RBC shape and sharpen image edges do morphological operations. No image segmentation results of this process were then used for the analysis of stained pixels.

[3] used Sobel Operator method for estimating the absolute distance gradient at each point of input gray scale image. Furthermore, [3] used Harris corner detection to detect the position of pixels to obtain a value converted to the matrix so that the detection of malaria can be done.

[11] proposed stages of segmentation that begins by sharpening the image using a Laplacian filter. Result from filter and then subtracted from the original image binary image of blood cells obtained using the method of Otsu, this image is a binary mask 1. The Laplacian filter is applied back to the green channel of the image and the result is reduced by the original image. Contrast enhanced image using histogram equalization. Otsu method was

TABLE 1 REPRESENTATIVE WORKS OF MALARIA IDENTIFICATION METHODS (PRE-PROCESS)

Author	Methods	Result	
		Original Image	Pre-Processing Image
Anggraini (2011)	a. Gray-scale conversion b. Noise reduction-5x5 median filter		
Nasir (2012)	a. Partial constrast stretching b. RGB → HSV		Hue Saturation Intensity 
Permata (2012)	Not Available	Not Available	Not Available
Mushabe (2013)	a. RGB → grayscale b. Median filter c. Morphogical closing	Not Available	Not Available
Gatc (2013)	a. RGB → grayscale b. Median filter	Not Available	Not Available
Somasekar (2014)	a. RGB→grayscale b. Gamma Equalization		
Rakshit (2014)	a. RGB→biner b. Complement binary image c. Adaptive Filtering		
Nugroho (2015)	a. Contrast Stretching b. RGB→ HSV c. 3x3 median filter		
Ravendran (2015)	a. RGB→ HSV b. *Morphological opening* c. *Morphological closing*		
Savkare (2015)	a. RGB → grayscale b. Median filter c. Laplacian filter		

implemented to obtain a binary image; this image is a binary mask 2. To obtain the final binary image of the binary mask of blood cells 1 and 2 is added. An object that has area less than the average area of the red blood cells removed using

morphological opening with a disc-shaped element structure.

The morphological reconstruction algorithm is also used by the [8] in the process of image segmentation. The initial stage of the K-means algorithm is used to calculate the distance from the cluster of data to the midpoint of the cluster. Data included in one cluster can be determined by looking at the minimum distance from the cluster. Furthermore, morphological reconstruction algorithm used by the method of dilation and erosion to remove the objects that are not needed. Table 2 shows the study of methods of segmentation.

C. *Feature Extraction*

Feature extraction is an essential process in digital image processing that aims to identify and obtain relevant information from images or called with features. [4] and [5] developed two groups of features.

The first group's feature based on the difference in the intensity distribution of the infected red blood cells, healthy or an area of the object identified. In the gray scale, image of the infected red blood cells had a few pixels with intensity values near zero because the nucleus of the malaria parasite. The same applies to the red blood cells that contain artifacts, but this condition does not apply to healthy red blood cells.

Another part that plays an important role in identifying malaria parasites in red blood cells is cytoplasm. In contrast to the nucleus, cytoplasm has a lighter color. [4] proposed a second group features views of the area that may be infected and is divided into three areas; nucleus, cytoplasm and blood cells. The second feature of the group obtained from dark area ratio (black; nucleus) with areas of light (white; cytoplasm) so that the infected red blood cells are classified based on the ratio of black and white pixels. While the research conducted [5] propose a second feature group is determined by the size of the area that has been segmented.

In the study of [4], features white & black pixel ratio has not managed to identify red blood cells infected with malaria, when the parasite is in the early stages of trophozoites, which the cytoplasm has not yet appeared. While at [5] study, the proposed method can detect 70 of the 76 cell parasite infected cells.

Research on feature extraction to classify red blood cells infected and uninfected also done by [6]. [6] proposed to use the values of R, G and B of each pixel of the input image as a feature. These values are then used for the classification process. In 2015,[11] proposed parameters used for the extraction of features into two groups. A first group is a feature-based form that consists of area, perimeter, the major axis and a minor axis. The second group is based on textual features which include the standard deviation, momentum, and kurtosis. The binary image of red blood cells infected and uninfected analyzed to eliminate parasites that are found outside the red blood cells. The result of this process is the number of red blood cells in the image as well as the number of infected red blood cells. A total of 21 images used for training data and test data stained-pixel classifier. The final result of this study showed 99.8% sensitivity for classification using KNN (k = 3) and 99.5% for the classification using Linear Bayesian.

[8] conducted a study that focused on the extraction of texture-based features. The method used is a histogram-based texture consisting of the mean, standard deviation, energy, skewness, entropy, smoothness, and kurtosis. By using these features, [8] obtained an accuracy of 87.8%, 81.7% sensitivity and 90.8% specificity.

D. *Identification*

The identification process is the final process that aims to identify the object, group the features into groups that exist or detect an object in the image. It is useful to find the malaria parasite. On paper discussed, groupings do vary. However, there are some papers that identify parasite based on segmentation. The following table is a study method used for the identification, classification and detection techniques and comparison of the advantages and disadvantages of each method. Table 2 show study of differrent classification and/or identification techniques for malaria blood images.

III. DISCUSSION AND FUTURE DIRECTION

Eventhough there are vast discussion about malaria identification techniques, several challenges still remain. In this paper research that has been carried out according to literature survey are: 1) Identification of red blood cells: infected or not infected, 2) Classification of the malaria parasite life's phase Plasmodium Falciparum and Plasmodium vivax, 3) Detection of malaria parasite. The highest accuracy for malaria identification is 97,73 % aimed by [9].

Based on the exposure in this paper, there still left a research question. How to identify malaria parasites other that *P. Falciparum* and *P. Vivax*? The digital image processing research for malaria can still be developed. Future works are available for malaria identification methods for Plasmodium Ovale, Knowlesi and Malariae and each of these parasites life-cycle Trophozoite, Ring, and Schizon. According to [13] early ring-form trophozoites (rings) of *P. knowlesi* are similar to *P. falciparum*, as rings may show double chromatin dots. Future study are suggested to use Sobel Operator and Haris corner detector as used in [9]. The Sobel Opertaor executes a 2- D spatial gradient measurement on an image and emphasizes on regions of high spatial frequency that are related to edges, it is used to find approximate absolute gradien magnitude at each point of an input gray scale image. Whereas Haris Corner detector is based on the local auto correctiom function of a signal which measures the local changes of the

signal with patches shifted by a small amount in different directions [9]

This paper presents a study of the method of identification of malaria by microscopic image of red blood cells. Information obtained from this paper can be used for research related to malaria later. The conclusion of this paper can be reference about any known method that can be used for any digital image processing as well as the advantages and disadvantages of each method

REFERENCES

[1] Malaria Fact Sheet

http[colon]//www[dot]who[dot]int/mediacentre/factsheets/fs094/en/

[2] J. Frean, "Improving quantitation of malaria parasite burden with digital image analysis," *Transactions of the Royal Society of Tropical Medicine and Hygiene,* vol 102, pp. 1062-1063, Feb 2008.

[3] Nugroho, A. H. (2015). Feature Extraction and Classification for Detection Malaria Parasit es in Thin Blood Smear. *Conference on Information, Technology, Computer and Electrical Engineering (ICITACEE).* Indonesia: IEEE.

[4] Anggraini, D., Nugroho, A., Pratama, C., Rozi, I., Pragesjvara, V., & Gunawan, M. (2011). Automated Status Identification of Microscopic Images Obtained from Malaria Thin Blood Smears using Bayes Decision : A study case in Plasmodium Flaciparum. *ICACSIS*, 317-352.

[5] Gatc, J.,(2013). Plasmodium Parasite Detection on Red Blood Cell Image for the Diagnosis of Malaria Using Double Thresholding. *International Conferrence on Advanced Computer Science and Information System,* 381-385.

[6] Mushabe, M. (2013). Automated Detection of Malaria in Giemsa-Stained Thind Blood Smears. *35th Annual International Conferrence of the IEEE EMBS Osaka, Japan, 3-7 July*, 3698-3701.

[7] Nasir, A., Mashor, M., & Mohamed, Z. (2012). Segmentation Based Approach fro Detection of Malaria Parasites Using Moving K-Means Clustering. *IEEE EMBS International Conference on Biomedical Engineering and Sciences,* 653-658.

[8] Permata, E. (2015). Klasifikasi Parasit Plasmodium Vivax pada Citra Sel Darah Merah menggunakan Metode Support Vector Machine One Against All. *Seminar Nasional Polinema.*

[9] Rakshit, P. (2013). Detection of Presence of Parasite in Human RBC in Case of Diagnosing Malaria using Image Processing. *Proceedings of the 2013 IEEE Second International Conference on Image Processing*, 329-334.

[10] Ravendran, A. (2015). Moment Invariant Features for Automatic Identification of Critical Malaria Paraites. *International Conference on Industrial and Information System, ICIIS.* Sri Lanka: IEEE.

[11] Savkare, S. (2015). Automated System for Malaria Parasite Identification. *International Conference on Communication, Information & Computing Technology (ICCICT).* Mumbay, India: IEEE.

[12] Somasekar, J. (2014). Contrast-enhanced Microscopic Imaging of Malaria Parasites. *IEEE International Conference on Computational Intelligence and Computing Research.*

ACKNOWLEDGMENT

The authors acknowledge the support of Smart System Laboratorium of Universitas Gadjah Mada.

TABLE 2 REPRESENTATIVE WORKS OF MALARIA IDENTIFICATION METHODS (CLASSIFICATION)

No	Author, year	Methods	Coments		Accuration
			Pros	Cons	
1	Anggraini, 2011	*Bayes Classifier*	High sensitivity and specificity	1. Only use one parasite as objects, namely P. falciparum 2. The proposed algorithm has not been able to identify the infected red blood cells when the parasites are in the early stages of the stadium trophozoite 3. Not to classify the stage of the parasite based on size and shape	Sensitivity 92.59% Specificity 99.65%
3	Rakshit, 2013	*Harris Corner Detection Algorithm*	1. Perform edge detection using Sobel operator to detect malaria parasites 2. It is the classification of the stage of the malaria parasite 3. Application of Harris Corner methods can be used to identify the stage of the malaria parasite by pixel position value computed into value matrix	Only use Plasmodium falciparum as a research object	97.73 %
4	Mushabe, 2013	*KNN Linear Bayesian*	1. It has been calculating the number of parasites in the red blood cells 2. The method of classification using KNN showed high accuracy	1. The classification of the stage and type of parasite 2. Only use Plasmodium falciparum as a research object 3. Calculation of estimates of the size of red blood cells has no effect on the segmentation process	Sensitivity KNN (k=3) 99.8 % Error 0.0012 Sensitivity Linear Bayesian 99.5 % Error 0.0055
5	Nugroho, 2013	*ANN-Multiperc epton Backprop agation*	Already the classification stage of the malaria parasite	Only use Plasmodium falciparum as a research object	87.8% Sensitivity 81.7% Specificity 90.8%
6	Savkare, 2015	*SVM Classifier*	1. Already classify stage and type of parasite 2. Classify stadium ring, trophozoites, and gametocytes 3. High segmentation's accuracy	Only use Plasmodium falciparum and Plasmodium vivax as a research object	Correct detection rate : 98.66% Sensitivity : 98.94% Specificity : 96.12%
7	Ravendran, 2015	*Naive Bayes KNN*	Features used invariant to translation, rotation and scale changes	1. Calculation of the complexity of malaria parasite detection remains low 2. Identification is only made for gametocytes stage of Plasmodium Falciparum	K-NN True positive rates 77.78% True negative rates 95.24% Naive Bayes True positive rates 88.89% True negative rates 80.95%

Bu Rika ICIC

ORIGINALITY REPORT

13%	10%	9%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

1%

★ Tatsuya Nonoyama, Lian Wen, Terry Rout. "Chapter 33 Current Challenges and Proposed Software Improvement Process for VSEs in Developing Countries", Springer Nature, 2016

Publication

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		