Rika Rosnelly

by LPPM Universitas Potensi Utama

Submission date: 05-Nov-2018 10:03AM (UTC+0700)
Submission ID: 1032882688
File name: 483-14776543209-13-IRES-2016.pdf (389.34K)
Word count: 2663
Character count: 14601

CLASSFICATION OF MALARIAL PARASITE AND ITS LIFE-CYCLE-STAGES IN BLOOD SMEAR

[1]SRI HARTATI, [2]AGUS HARJOKO, [3]RIKA ROSNELLY, [4]IKA CANDRADEWI

[1,2,3]Department of Computer Science and Electronics, Universitas Gadjah Mada
E-mail: shartati@ugm.ac.id, aharjoko@ugm.ac.id, rikarosnelly and ika.candra.kc@gmail.com

Abstract- A method to classify plasmodium of malaria disease along with its life stage is presented. The geometry and texture features are used as plasmodium features for classification. The geometry features are area and perimeters. The texture features are computed from GLCM matrices. The support vector machine (SVM) classifier is employed for classifying the plasmodium and its life stage into 12 classes. Experiments were conducted using 600 images of blood samples. The SVM with linear kernel gives the accuracy of 57% whereas SVM with RBF kernel yields an accuracy of 99.1%.

Index Terms- Malaria, geometry, texture, GLCM, SVM, RBF.

I. INTRODUCTION

Malaria is a highly hazardous disease to humans because it can cause death. Malaria is caused by parasites which is transmitted by the female Anopheles mosquito. These mosquitoes bite infected plasmodium from a person previously infected with the parasite. Plasmodium are divided into 4 types: plasmodium ovale, plasmodium malariae, plasmodium falciparum, and plasmodium vivax. Plasmodium vivax is often found in patients with malaria disease. Plasmodium falciparum is the cause of deaths of nearly most of the patients with malaria disease in the world.

Microscopic examination is required to determine the parasite plasmodium visually by identifying directly at the patient's blood dosage. The microscopic examination result is highly dependent on the expertise of the laboratory worker (health analyst) that identifies the parasite plasmodium. The microscopic examination technique is the gold standard in the diagnosis of malaria. Some techniques used for malaria diagnosis are the peripheral blood smear (PBS), quantitative buffy coat (QBC), rapid diagnosis test (RDT), Polymerase Chain Reaction (PCR), and Third Harmonic Generation (THG) [1],[2]. Among them, the PBS technique is the most widely used malaria diagnosis even though it has limitations of human resistance due to time requirement.

To diagnose malaria parasite, a manual calculation process that uses microscopic examination of Giemsa-stained thick and thin blood smears is carried out. This process requires a long time and is a tedious process. It is very susceptible to the capabilities and skills of technicians. Its potential for mistakes made by humans is significant [3]. As an illustration, a trained technician requires about 15 minutes to count 100 cells. Worldwide, technicians have to deal with millions of patients every year [4] To overcome a long and tedious process, several studies have been conducted to develop automated microscopic blood cell analysis. Some early studies showed a low performance which lead to inaccurate diagnoses [5]. Similar studies continued with various methods to improve the accuracy of identification of parasite infection, particularly for identifying 2-4 plasmodium parasite that can infect humans [6] without specifying the life stage of the malaria parasite. Each parasite has three different life stages. They are trophozoites, schizocytes, and gametocytes [3].

Therefore, studies on the classification of the life stages of the malaria parasite are a challenge [7]. The objective of this research is to detect Plasmodium and its life stage directly from a microscopic image of blood samples. The plasmodium that can cause human infection has four species: falciparum, vivax, ovale, and malaria. Each species is differentiated into four different life-cycle-stages, which are generally morphologically distinguishable: ring, trophozoite, schizont, and gametocyte. Therefore, we propose a method to classify plasmodium into twelve classes.

II. DATA COLLECTION

A total of 600 malaria image data of Giemsa - stained thin blood smears are obtained from Bina Medical Support Services (BPPM) in Jakarta. The sample malaria image data size is 2560 x 1920 pixels. The manual plasmodium classification is carried out by laboratory workers of the parasitology Health Laboratory of the North Sumatra Province, Indonesia, which provide the ground truth for the proposed method. Each image is given a label associated with the name of the parasite, i.e., plasmodium malaria, plasmodium falciparum, plasmodium vivax) along with its life-cycle stage (ring, trophozoite, schizont, or gametocyte).

None of the 600 image data consist of plasmodium ovale. Therefore the 600 image data consist of 12 classes. Fig. 1 shows different plasmodium, and their life-stages.

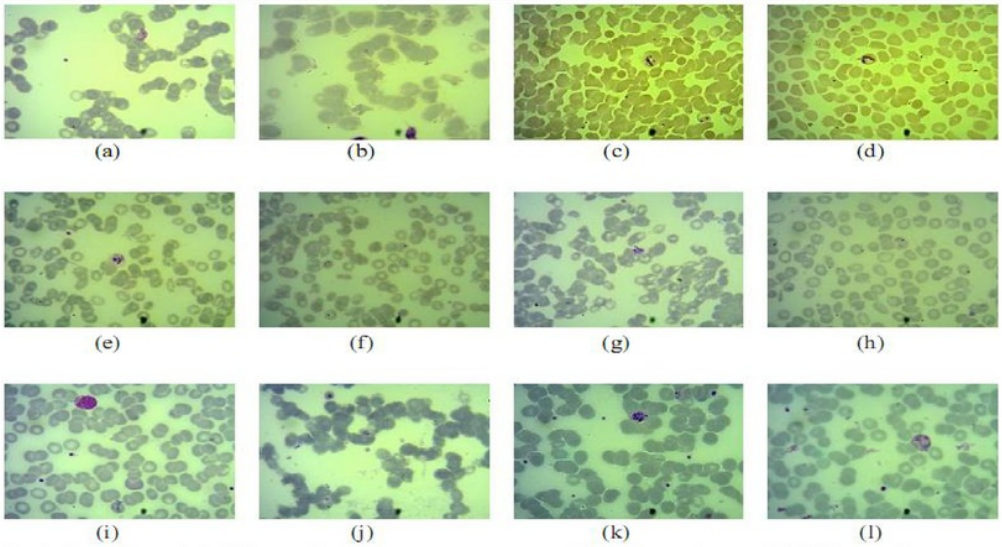


Fig. 1. Plasmodium and their life stage. (a) Falcifarum, gametocyte stage (b) Falcifarum, ring stage (c) Falcifarum, schizont stage (d) Falcifarum, trophozoite stage (e) Malariae, gametocyte stage (f) Malariae, ring stage (g) Malariae, schizont stage (h) Malariae, trophozoite stage (i) Vivax, gametocyte stage (j) Vivax, ring stage (k) Vivax, schizont stage (l) Vivax, trophozoite stage

III. METHOD

The classification process of the malaria parasite is shown in Fig. 2. A blood smear is performed on the blood sample. The region of interest (ROI) is then determined to locate the area, which contains parasite. Next, three basic image processing steps are carried out, that is, preprocessing, segmentation, and feature extraction. Following that, the image classification and detection of infected red blood cells (RBC), that is called parasitemia, are carried out. In this work, the malaria images consist of three types of plasmodium and each has four different life stages, i.e., ring, schizont, trophozoite, and gametocyte stages.

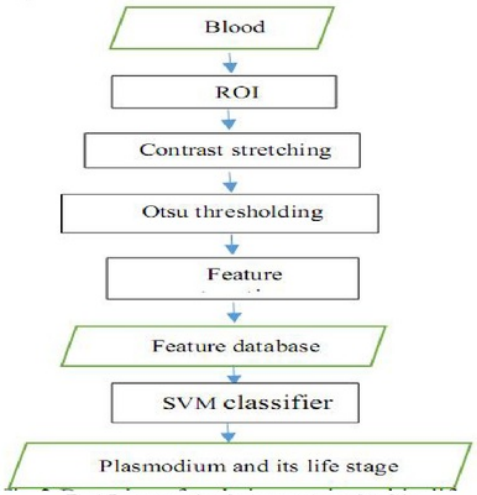


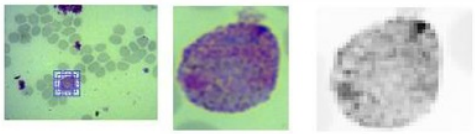
Fig.2 Detection of malaria parasite and its life stage

A. Pre-processing

The aim of preprocessing step is to obtain images with lower noise and higher contrast than the original images for further processing. Blood smear images might be affected by illumination and color distribution of blood images due to the camera setting and staining variability. Most of microscopes yield blood cells with quite similar colors. Therefore, image enhancement and noise reduction operations are required. Low intensities of light might decrease the contrast of blood image [8], therefore the contrast image has to be improved using a contrast enhancement method. After image enhancement is performed, the region of interest (ROI) is carried out by manually cropping the infected RBC, because the image not only contains infected red blood cells but also normal red blood cells, white blood cells, platelets, and artifacts. Experts validate the process of determining ROI. Experience indicates that the appropriate size for ROI is 256 x 256 pixels. These preprocessing produces an image with good contrast.

B. Segmentation

Segmentation attempts to subdivide an image into sub images or segments such that each segment fulfills a certain characteristics. In this case, as the malaria parasite affects the red blood cells, the segmentation is carried out to separate red blood cells from the rest, and the results are the red blood cells of the microscopic images of blood sample. Initially the RGB image of ROI is converted into the gray image since the red blood cells can be distinguished from the rest from its gray level value. In this research, Otsu's thresholding method is used for its ability to determine threshold automatically. An example is depicted in Figure 3.



(a) (b) (c)

Fig. 3 (a) Initial image, (b) Region of Interest (ROI) (c) gray scale ROI

After thresholding, the morphological closing and opening are performed to extract the hole inside the infected cell and eliminate the unwanted artifacts [8]. These segmented cells are further processed and then the infected red blood cells are identified.

C. Features Extraction

Many studies concerning the analysis of red blood cells recently use texture features [5], [9], and color features[10],[11], to differentiate normal cells and infected cells. In this research texture and geometry, features are used. Geometry features are selected for analyzing blood since hematologist uses these features. The selected geometric features are area and perimeter. The area is the number of pixels of the object that indicates the size of the object, and is calculated using as

$$Area = \sum_{x}\sum_{y} f\,(x,y). \quad (1)$$

The perimeter is the continuous line forming the boundary of a closed geometric object. It can be calculated as

$$Perimeter = \sum_{x}\sum_{y} f\,(x,y), \quad x,y \in Boundary\ region \quad (2)$$

The texture features are computed from the Gray-Level Co-occurrence Matrix (GLCM) of the ROI image. The GLCM is used to calculate the co-occurrence of a pair of pixels with gray-level value *i* and *j* in a particular direction. A GLCM element $P\theta,d(i, j)$ is the joint probability of the gray level pairs i and j in a given direction θ separated by a distance of d units. In this research, the GLCM features are extracted using one distance (d = {1}), and three directions(θ = {45^0, 90^0,135^0}). These texture based features can be calculated as follows:

1. Contrast is the measure of intensity contrast between a pixel and the neighboring pixel over the complete image.

$$\sum_{i,j=0}^{N-1} p_{i,j}(i-j)^2 \quad (3)$$

2. *Entropy*

Entropy is the measure of the complexity of the image, and it represents the amount information contained in data distribution. The higher the entropy value, the higher the complexity of the image.

$$\sum_{i,j=0}^{N-1} p_{i,j}(-\ln p_{i,j}) \quad (4)$$

3. ***Energy***

Energy is a measure of the pixel intensities in gray scale value. Energy is computed by summing all squared elements in the GLCM matrix,

$$\sum_{i,j=0}^{N-1} p_{i,j^2} \quad (5)$$

4. Homogeneity is the measure of the homogeneity of a particular region. This value is high when all pixels have the same values or uniform.

$$\sum_{i,j=0}^{N-1} \frac{p_{i,j}}{1+(i-j)^2} \quad (6)$$

5. Correlation indicates how a pixel is correlated with the neighboring pixels in a particular area

$$\sum_{i,j=0}^{N-1} p_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right] \quad (7)$$

1 Classification

The support vector machine (SVM) is selected to classify the Plasmodium type along with its life stage. There are 12 possible classes since there are three types of Plasmodium and four life stages. Two different kernels are implemented, and their performances are compared.

Before the SVM is used for classification, it is trained using training data. In the process of training, the SVM uses feature matrix, as the training input, which is obtained in the features extraction process. The training data classification process is to seek support vector and bias of input data. The following is the training algorithm for each binary SVM:

Input: Z is a matrix of Plasmodium features obtained from feature extraction process

Output: Strain vector as a target. Ytrain vector is a column vector for classification of the first class, where all images of blood preparations of the first class will be symbolized by number 1, all images of blood smears from other classes with number -1. In this study, a Gaussian kernel function with variance (σ) = one is used. The next step is to calculate Hessian matrix, i.e., multiplication of a Gaussian kernel with Ytrain. Ytrainis a vector that contains values of 1 and -1. Hessian matrix is later used as input variables in quadratic programming. The training steps are described as follows:

1. Determine input (Z = Xtrain) and Target (Ytrain) as a pair of training from two classes.
2. Calculating Gaussian kernel

$$K(Z,Z_i) = \exp\left(\frac{-|Z-Z_i|^2}{2\sigma^2}\right) \quad (8)$$

3. Calculate Hessian matrix

$$H = K(Z,Z_i) * Y * Y_T. \quad (9)$$

4. Assign c and epsilon. The term c is a constant in Lagrangian multipliers and epsilon (cost parameter) is the upper limit value of α, which serves to control classification error. This study used value of c = 100000 and epsilon = 1x10-7.

5. Assign vector e as a unit vector which has the same dimension with the dimension of Y.

6. Calculating quadratic programming solution

$$L(\alpha) = \frac{1}{2}\alpha^{+} H\alpha + e^{T}\alpha \quad (10)$$

,with $Y_\alpha^{\ T} = 0$ and $a \le L \le c$

In testing process, data that have never been used for training are used. Results of this process are an index value of the largest decision function, stating the class of the testing data. If a class in the classification test match the test data classes, classification is stated to be correct. The final result of classification is an image of blood that matches with an index value of decision function using SVM one against all.

Having an input data feature vector T for test data (w,x,b), and k number of classes, the input data then is used for the testing process. The input is generated in the process of feature extraction, The process of testing is as follows:

1. Calculate Kernel Gaussian

$$K(T, x_i) = \exp(-\frac{|T - x_i|^2}{2\sigma^2}) \quad (11)$$

2. Calculate

$$f_i = K(T, x_i)w_i + b_i \quad (12)$$

3. Repeat steps 1,2 for i = 1 to k.

4. Determining the maximum value of f_i

5. A class i is a class from T which has the largest value of f_i

The performance of the proposed method is measured regarding accuracy, sensitivity, and specificity. The True positive (TP) shows the image of blood smears correctly identified. False positive (FP) is the image of Plasmodium classified incorrectly. The true negative (TN) indicates the number of image that is not a member of a class and is correctly identified as not a member of class (NV). False negative (FN) shows the number image of blood smears that should not be members of class but identified as a member of class.

Accuracy=(TP+TN)/(TP+TN+FP+FN), (11)
Sensitivity=TP/(TP+FN),
Specificity=TN/(FP+TN)

Table 1. Experimental results for SVM classifier with linear kernel.

	K=1	K=2	K=3	K=4	K=5
Accuracy (%)	53,0	52,0	62,0	60,0	56,0
Precision (%)	33,0	34,4	45,3	37,8	44,0
Sensitivity (%)	38,9	37,4	45,4	44,1	43,4
Specificity (%)	95,4	95,6	95,5	96,3	95,7

Table 2. Experimental results for SVM classifier with RBF kernel.

	K=1	K=2	K=3	K=4	K =5
Accuracy (%)	100	98,0	100	100	98,0
Precision (%)	100	96,2	100	100	97,9
Sensitivity (%)	100	99,1	100	100	97,0
Specificity (%)	100	99,8	100	100	99,8

CONCLUSION

A method to classify plasmodium of malaria disease along with its life stage is presented. The geometry and texture features are used for classification. The texture features are computed from GLCM matrices. The SVM classifier is employed for classifying the plasmodium and its life stage into 12 classes. The SVM with linear kernel gives the accuracy of 57% whereas SVM with RBF kernel yields an accuracy of 99.1%.

ACKNOWLEDGMENT

The authors would like to thank the Directorate General of Higher Education, the Ministry of Research and Higher Education of the Republic of Indonesia for sponsoring this research. The authors would like also to thank the parasitology Health Laboratory of the North Sumatra Province and Bina Medical Support Services (BPPM), Jakarta, for supporting this research.

REFERENCES

[1] Guide, World Health Organization, Geneve, Switzerland, 2nd edition, 2010.

[2] P. Jain, B. Chakma, S. Patra and P. Goswami, "Potential biomarkers and their applications for rapid and reliable detection of malaria," BioMed Research International, vol.2014, pp.201-221, 2014.

[3] F.E. McKenzie, "Dependence Of Malaria Detection And Species Diagnosis By Microscopy On Parasite Density," Am. Soc. Trop. Med. Hyg., 2008.

[4] FB. Tek, A.G Dempster, I. Kale, "Malaria Parasite Detection in Peripheral Blood Images," Proceeding of British Machine Vision Conference 347-356, 2006.

[5] H.A Nugroho, S.A Akbar, E.E.H Muhandarwari, "Feature Extraction and Classification for Detection Malaria Parasites in Thin Blood Smear", Conference on Information Technology, Computer and Electrical Engineering (ICITACEE), Indonesia, pp.198- 201, 2015.

[6] E. Komagal, K. S. Kumar, A.Vigneswaran, Recognition And Classification Of Malaria Plasmodium Diagnosis,

International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 1, p.1-4, 2013.

[7] N.E Ross, C.J. Pittchard, D.M. Rubbin, A.G. Duse, "Automated image processing method for the diagnosis and classification of malaria on thin blood smears," Medical and Biological Engineering and Computing, vol. 44, no. 5, pp. 427–436, 2006.

[8] K. M. Khatri, V. R. Ratnaparkhe, S. S. Agrawal, A. S. Bhalchandra,"Image processing approach for malaria parasite identification", National Conference on Growth of Technologies in Electronics, Telecom and Computers—India's Perception(GTETC—IP '14), IJCA Proceedings, pp. 5–7, 2014.

[9] A. Kumar, A.Choudhary, P. U. TembhareC. R. Pote, "Enhanced identification of malarial infected objects using otsu algorithm from thin smear digital images". International Journal of Latest Research in Science and Technology, vol 1, no, 159 pp. 2278-5299, 2012.

[10] N. Ahirwar, S. Pattnaik and B. Acharya, "Advanced image analysis based system for automatic detection and classification of malaria parasite in blood mages." International Journal of Information Technology and Knowledge Management, vol. 5, No. 1, pp. 59-64, 2012.

[11] T.Chen, Y.Zhang, C.Wang, Z.Ou, F.Wang, T.S.Mahmood, "Complex local phase based subjective surfaces (CLAPSS) and its application to DIC red blood cell image segmentation", Journal of Neurocomputing, vol.99, pp.98-110,2013.

★★★

Bu Rika 4

ORIGINALITY REPORT

16%	11%	5%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	ijmas.iraj.in Internet Source	10%
2	Hanung Adi Nugroho, Son Ali Akbar, E. Elsa Herdiana Murhandarwati. "Feature extraction and classification for detection malaria parasites in thin blood smear", 2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), 2015 Publication	3%
3	Hanung Adi Nugroho, Wahyu Andi Saputra, Adhistya Erna Permanasari, E. Elsa Herdiana Murhandarwati. "Automated determination of Plasmodium region of interest on thin blood smear images", 2017 International Seminar on Intelligent Technology and Its Applications (ISITIA), 2017 Publication	1%
4	iasir.net Internet Source	1%

5	S. T. Khot, R. K. Prasad. "Chapter 9 Optimal Computer Based Analysis for Detecting Malarial Parasites", Springer Nature, 2015 Publication	1%

Exclude quotes	On	Exclude matches	< 25 words
Exclude bibliography	On		