2016 1st International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia

Plasmodium Falciparum Stages Classification on Red Blood Cell Image using Region Property

Febri Maspiyanti[1,a], Sri Rezeki C Nursari[1,b], Amir Murtako [1,c], Jullend Gatc[2,d]
[1]Department of Informatics Engineering, Pancasila University, Jakarta, Indonesia
[2]Department of Information System, Kalbis Institute, Jakarta, Indonesia
[a]febri.maspiyanti@univpancasila.ac.id, [b]sri.rezeki.candra.n@univpancasila.ac.id, [c]amir.murtako@univpancasila.ac.id, [d]jullend.gatc@kalbis.ac.id

***Abstract*—Determining the stage of Plasmodium falciparum in malaria disease diagnosis is crucial step to define the appropriate medical treatments so as to reduce the risk of death on patient. In this research, we propose plasmodium falciparum phases classification system using regionprops computation i.e. bounding box, centroid, extent, and extrema features. By using thin blood smear image data, our proposed method successfully classify falciparum data into three classes, namely, trophozoite, schizont and gametocyte with accuracy of 96%.**

Keywords—Classification, Falciparum Stages, Regionprops, Thin blood smear.

I. INTRODUCTION

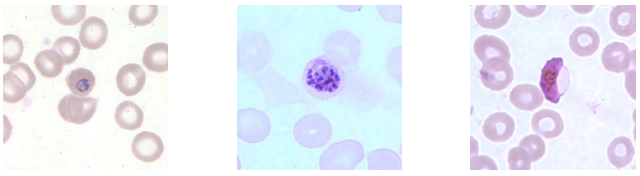
Plasmodium parasite that causes malaria consists of five types, namely: falciparum, vivax, malarie, ovale, and knowlesi [1]. According to World Health Organization country profiles of Indonesia 2015 report [2], Plasmodium falciparum is the most dominant among the malaria case in Indonesia. Based on the WHO report, total 57% of cases of malaria patients infected by the parasite Plasmodium falciparum type. According to the National Center for Biotechnolgy Information (NCBI) [3], Plasmodium falciparum appears to be the most common Plasmodium species in Indonesia, 58% of the total cases of malaria that occurred in Indonesia infected by P. Falciparum. Falciparum malaria may be fatal if treatment is delayed beyond 24 hours after the onset of clinical symptoms [4].

According to CDC [1], falciparum has three phases, namely: tropozoite, schizont and gametocit. Tropozoite is an initial phase of the parasite age, schizont is a mature phase of the parasite that is ready to spread to infect the healthy blood cells, while gametocyte is a mature phase that is ready to infect the mosquitoes through a mosquitoes bite. The third phase of falciparum can be seen in fig. 1

The clinical presentation on malaria patient can vary substantially depending on the infecting species, the level of parasitemia, and the immune status of the patient. Infections caused by *P. falciparum* are the most likely to progress to severe, potentially fatal forms with central nervous system involvement (cerebral malaria), acute renal failure, severe anemia, or acute respiratory distress syndrome [1]. Therefore, early detection and determination of falciparum phase is a crucial step to reduce the risk of death on patients.

According to CDC [1] on Treatment Information section, malaria can be a severe, potentially fatal disease (especially when caused by *Plasmodium falciparum*) and treatment should be initiated as soon as possible. Early detection on the presence of parasite in human blood is expected to be able to diagnose the severity of the infection hence the drug treatment will be more efficient. Conventional method for the detection of parasite commonly use microscope observation done by medics. Despite of the fact that good performance of microscopy can be difficult to maintain because of the requirements of adequate training and supervision of laboratory staff to ensure competence in malaria diagnosis [5].

In order to reduce human error, neither lack of human resource on microscope observation by medics, many researches on the early detection of malaria parasite have been done not only in effort to improve the quality of microscope observation but also to reduce mortality caused by malaria such as researches by Das [6], Purnama [7], Nugroho [8], and Anggraini [9]. In this research, we propose the development of plasmodium falciparum phases classification system using shape-based features in order to classify each parasite based on it life cycle stage hence the drug treatment for the patient will be more efficient.



(a) (b) (c)

Figure 1. Plasmodium falciparum (a) Trophozoite, (b) Schizont, (c) Gametocyte [1].

II. RELATED WORK

Many research on detection and classification of the plasmodium parasite have been done widely, one of which is the research conducted by Das [6] using a rule-based segmentation techniques, Chan-Vese-based, and marker-controlled watershed to separate erythrocyt. After obtaining the segmented cell, use traditional Das F-statistic and information gain criteria for extracting intensity, texture, and

morphology level features from cells. Furthermore, Das uses Naive Bayes, Logistic Regression, Multilayer Perceptron, Radial Basis Function Network, Classification and Regression Tree to classify falciparum and vivax each into 3 phases (trophozoyte falciparum, schizont falciparum, gametocyte falciparum, trophozoyte vivax, schizont vivax, gametocyte vivax) managed to reach the level of accuracy of 96.84%

Purnama [7] use red-green-blue color histogram, hue channel HSV histogram, and hue channel HSI histrogram on their research as features extraction technique. Purnama also use Genetic Programming to detect and identify the type and stage of the parasites. Purnama divided their research into 2 different classification. First, Purnama classify the RBC into 2 classes: not-parasite (95,49% average accuracy) and parasite (95,58% average accuracy). Second, Purnama classify the RBC into 6 classes: Not-parasite (90,25% average accuracy), vivax tropozoite (82,25% average accuracy), vivax schizont (75,83% average accuracy), vivax gametocyte (81,75% average accuracy), falciparum thropozoite (90,75% average accuracy), and falciparum schizont (86,75% average accuracy).

Another research conduct by Nugroho [8] use contrast and noise filtering to enhance the image quality, K-means clustering algorithm to obtain parasite cells, morphological reconstruction algorithm to remove unwanted objects, histogram-based texture (mean, standard deviation, skewness, energy, entropy, smoothness, and kurtosis) to extract the features, and multilayer perceptron backpropagation algorithm to classify the falciparum parasites into 3 classes: trophozoite, schizont, and gametocyte. Nugroho research can reach accuracy about 87.8%.

Anggraini [9] use global thresholding dan connected component extraction to identify blood cell components, and then using Bayes Decision Theory to classify cell into infected and non-infected erythrocyt. Anggraini successfully obtained sensitivity of 92.59% and specificity of 99.65%. Anggraini also develop an expert system to identify cell conditions using thin blood smear data, namely: uninfected RBC, Erythrocyte and artifacts, Gametocyte stage of falciparum, Ring stage of falciparum, and unidentified infection [10]. By using the technique of grayscale conversion, noise reduction, contrast stretching, Otsu's thresholding, and multiple thresholding, Anggraini able to detect 10 out of the 12 data.

III. METHODOLOGY

A. Data

We use total 25 thin blood smear image data from CDC [11] and Atlas [12] with RGB 24 bit data format. The data distribution can be seen in Table 1 below:

Table 1. Data distribution

Data	Trophozoite	Schizont	Gametocyte
CDC	7	5	8
Atlas	3	2	0
	10	7	8

According to table 1, we use 20 data from CDC (7 trophozoite stage data, 5 schizont stage data, and 8 gametocyte stage data), and 5 data from Atlas (3 trophozoite stage data, and 2 schizont stage data).

B. Method

We propose plasmodium falcifarum parasite stage life classification which is an enhancement method proposed by Gatc [13] research. In this research, we use shape features to find the pattern of each stage life of falciparum. The method used in this research can be seen in fig.2.

The first step in this research start from the image data preprocessing using 3x3 windowing median filter in order to enhance and reduce noise in the data. In this step, the RGB image data will be converted into grayscale format to reduce complexity and dimension of the data.

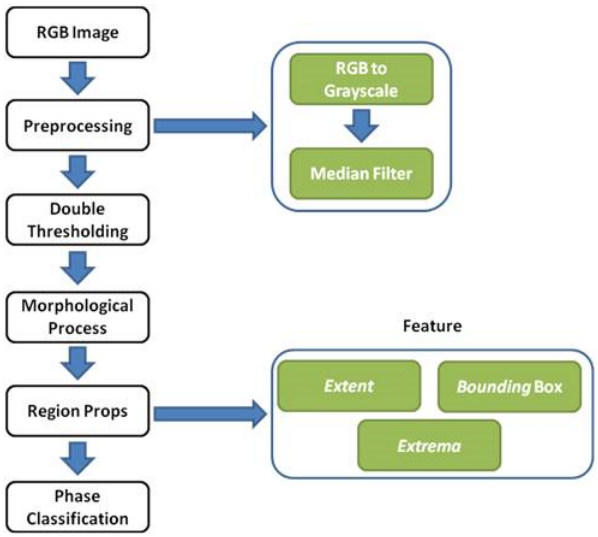
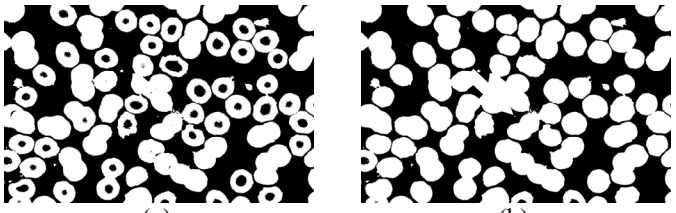


Figure 2. Plasmodium falciparum stage classification method

1. Double Thresholding Segmentation.

The following step, we use double thresholding method proposed by Gatc et al [13] as image segmentation process to eliminate background from foreground (blood cells). Thin blood cell image segmentation results can be seen in fig. 2.



(a) (b)

Figure 3. Image segmentation result using Double Thresholding method

Double Thresholding is simple segmentation process to separate image foreground and background based on image histogram into binary image classification. We adopted Gatc et.al method which combines standard threshoding and Otsu method. Otsu method will maximize the variable to divide foreground object and background the equation:

$$\sigma_B^2(k^*) = \max_{1 \le k < L} \sigma_B^2(k) \quad (1)$$

With function:

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

Where omega ω is the Zeroth cumulative moment value, $\mu(k)$ is the first cumulative moment and μ_T is the mean probability of each pixel [14].

2. Feature Extraction using Region Props

Once the background has been eliminated from the red blood cells, each red blood cell will be divided into two classes, namely infected and uninfected red blood cells. We use thresholding technique on image data that only consist of red blood cell (RBC). Parasite on infected RBC will have greater intensity compared to uninfected RBC. Furthermore, once the uninfected RBC have been eliminated, the infected RBC will be extracted using one of the Measurement-Based [15] features extraction method, namely, regionprops function.

Region properties (regionprops) [15] is one of the functions used to measure a set of properties from each region of the image that has been labeled in matrix S, where the positive integer that is an element of S correspond to the corresponding region. Region Properties able to extract many of the properties of an image, but in this research, we only use four types of properties to classify each stage i.e. Bounding Box, Centroid, Extent, and Extrema. The ilustration of the bounding box, centroid, extent, and extrema can be seen in fig. 4.

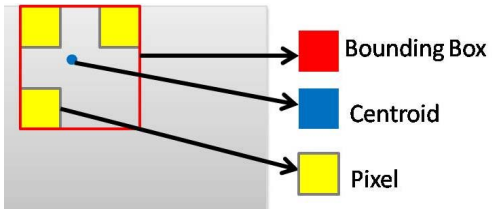


Fig 4. Regionprops on an image data ilustration

Bounding box is the smallest rectangle containing the region, centroid is the center of mass of the region, extent is the ratio of pixels in the region to pixels in the total bounding box, and extrema specifies the outer pixels coordinates (x,y) inside the region.

The bounding box of a region R is the minimal axis-parallel rectangle that enclose all points of R,

$$BoundingBox(R) = \langle u_{min}, u_{max}, v_{min}, v_{max} \rangle$$

Where $u_{min}, u_{max}, v_{min}$ and v_{max} are the minimal and maximal coordinate values of all points $(u_i, v_i) \in R$ in the x and y direction, respectively. Extrema ilustration can be seen in fig.5 below.

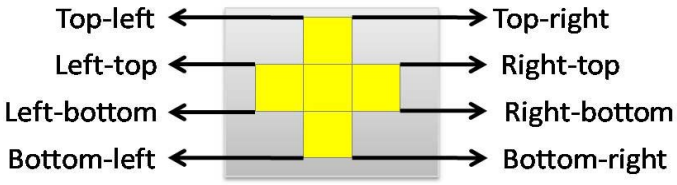


Fig.5. Extrema ilustration

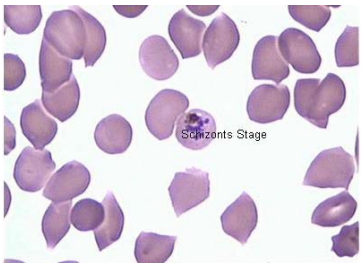
In this research, we studied about the morphology characteristics of each falciparum stage then use regionprops to compute the properties of each stage. In general, the Plasmodium parasite that infected human RBC have three major stages, namely trophozoite stage, gametocyte stage, and schizont stage [1]. Each stage have specific characteristic that is able to distinguish one stage to other stage. Thropozoite phase is characterized by small dot ring shaped in the cytoplasm. Trophozoite stage is the initial stage of the falciparum life cycle. From trophozoite stage, the parasite will develop into two other stage i.e. gametocyte and schizont. Gametocyte stage has oval dark pigments that resemble the shape of banana. The other stage is schizont that similar to trophozoite stage but has more dark colored granules that almost cover the entire of RBC.

Tropozoite stage is easy to identify due to the single chromatin dot of the parasite. In contrast, the dark chromatin dots on schizont stage will produce more point coordinates (x, y). The number of coordinate correspond to the number of centroid. Furthermore, we compute each member of the extrema. While on gametocyte stage will be classified easily due to it shape resemble to banana shape. In addition, gametocyte area recognize by the area that greater than trophozoite stage. Each area that has been extracted and calculated will be labeled correspond to the region member and will be used as training data.

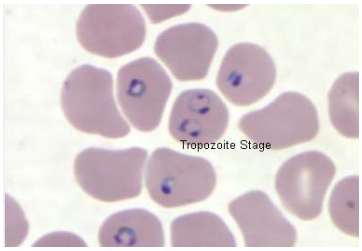
IV. RESULT

The classification of plasmodium falciparum parasites stage on 25 thin blood smear image data using our propose method succesfully achieved accuracy of 96%. The examples of classification results can be seen in fig. 6, and the confusion matrix can be seen in table 2 below:

(a) Gametocyte Stage



(b) Schizont Stage



(c) Tropozoite Stage

Fig. 6. Classification result

Table 2. Confusion Matrix

Data	Falciparum Stage Class		
	Gametocyte	***Schizont***	***Tropozoite***
Gametocyte	8	0	0
Schizont	0	7	0
Tropozoite	0	1	9

V. DISCUSSION

Plasmodium falciparum stages classification can be easily done manually by experts. In contrast to build automated plasmodium falcifarum stage classification system, there are several important factors that need to be considered such as artifacts on the camera lens whether from the non-standard collecting procedure of blood smear will affect the diagnostic system. We consider that the preprocessing stage is one of the crucial steps in order to reduce the diagnostic errors. Experiment without preprocessing stage has been carried out misclassification due to the form of artifacts that have similar intensity to parasite.

Our research use three major processes in developing the plasmodium falciparum stages classification system i.e. data preprocessing to enhance the image quality, segmentation to eliminate RBC from background, and feature extraction using the region property to extract the image information. This method is succesfully achieved accuracy of 96%.

From Table 2, we can see that our research can only classify 24 out of 25 data correctly. One misclassify data is tropozoite stage. Generally, the early phase of tropozoite stages has one chromatin dot with ring-shaped parasite's cytoplasm, and vacuoles are clearly visible. However, tropozoite late stages have two chromatin dot attached to a ring inside one parasite vacuoles. Thus condition similar to shcizont stage which have chromatin dot between 2 to 32, the difference is each chromatin dot in schizont have one vacuole it self (each chromatin dot seperated from others chromatin dot by vacuole). This vacuoles have low grey level intensity that easily distinguished by our method, hence misclassed the tropozoite late stage into schizont stage.

CONCLUSION

We propose plasmodium falciparum stages classification system using regionprops computation i.e. bounding box, centroid, extent, and extrema features. By using thin blood smear image data, our proposed method successfully classify falciparum data into three classes, namely, trophozoite, schizont, and gametocyte with accuracy of 96%.

REFERENCES

[1] Centers for Disease Control and Prevention. Malaria. CDC. Web. 1 May 2016. <http://www.cdc.gov/dpdx/malaria/index.html>

[2] World Health Organization. Malaria: Country Profile of Indonesia 2015. WHO. 2015.

[3] Elyazar, Iqbal R.F., Simon I. Hay, and J. Kevin Baird. "Malaria Distribution, Prevalence, Drug Resistance and Control in Indonesia." *Advances in parasitology* 74 (2011): 41–175. *PMC*. Web. 1 May 2016.

[4] World Health Organization. International Travel and Health: Malaria. WHO. Web. 1 May 2016. <http://www.who.int/ith/diseases/malaria/en/>

[5] World Health Organization. Guideline for the Treatment of Malaria Third Edition. WHO. 2015.

[6] Das, D.K., A.K. Maiti, C. Chakraborty. Automated System for Characterization and Classification of Malaria-infected Stages using Light Microscopic Image of Thin Blood Smears. Journal of Microscopy, vol. 257, Issue 3, pp 238-252. Royal Microscopical Society. 2015.

[7] Purnama, I.K, F. Z. Rahmanti, M. H. Purnomo. Malaria parasite identification on thick blood film using genetic programming. 3rd International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME). pp. 194-198. Bandung. IEEE. 2013.

[8] Nugroho, Hanung Adi, Son Ali Akbar, E. Elsa H. Murhandarwati. Feature Extraction and Classification for Detection Malaria Parasites in Thin Blood Smear. 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE). IEEE. 2015.

[9] Anggraini, D., et al. Automated Status Identification of Microscopic Images Obtained From Malaria Thin Blood Smears using Bayes Decision: A Study Case in Plasmodium Falciparum. International Conference on Advanced Computer Science and Information System (ICACSIS). IEEE. 2011.

[10] Anggraini, Dian, et al. Automated Status Identification of Microscopic Images Obtained from Malaria Thin Blood Smears. International Conference on Electrical Engineering and Informatics. IEEE. 2011.

[11] Centers for Disease Control and Prevention. Malaria. CDC. DPDx - Laboratory Identification of Parasitic Diseases of Public Health Concern. Web. <http://www.cdc.gov/dpdx/malaria/gallery.html#>

[12] Web Atlas of Medical Parasitology. Plasmodium falciparum. Web. <http://atlas.or.kr/atlas/alphabet_view.php?my_codeName=Plasmodium%20falciparum>.

[13] Gatc, Jullend, et al. Plasmodium Parasite Detection on Red Blood Cell Image for the Diagnosis of Malaria Using Double Thresholding. ICACSIS: IEEE. 2013.

[14] Liu, Dongju., JianYu. Otsu Method and K-means.International Conference on Hybrid Intelligent Systems. Department of Computer Science Beijing Jiaotong University Beijing. IEEE. China. 2009.

[15] Mathworks Documentation. Regionprops. Web. http://www.mathworks.com/help/images/ref/regionprops.htm