Plasmodium detection methods in thick blood smear images for diagnosing Malaria : A review

Chyntia Raras Ajeng Widiawati¹, Hanung Adi Nugroho², Igi Ardiyanto³
Department of Electrical Engineering and Information Technology
Universitas Gadjah Mada
Yogyakarta, Indonesia
¹chyntia.mti15@mail.ugm.ac.id, ²adinugroho@ugm.ac.id, ³igi@ugm.ac.id

Abstract—Malaria is one of the serious diseases in the world which often causes death. Based on the data from World Health Organisation (WHO) in 2015, Africa contributed 90% of death by malaria all over the world, followed by Southeast Asia and the Eastern Mediterranean. Early detection of plasmodium parasite is necessary to help the diagnose process of malaria. Nowadays, digital image processing is one of the methods used to help doctors in diagnosing a disease. In general, several stages in image processing consist of preprocessing, segmentation, feature extraction and classification. In this paper, a comparison of some image processing methods will be reviewed to determine the best method applied in this field. The confirmation is done by looking under microscope at 200 fields of view of the thick blood smear or counting of 200 to 500 of white blood cells to check the presence of the parasite. The confirmation from this research will affect the image classification result and affect the diagnosis. For the future work, look for the best method in distinguish the parasite and artifacts specially white blood cell in thick blood smear is really needed. So, can be used to count parasite density.

Keywords—Malaria, plasmodium, preprocessing, segmentation, feature extraction, classification, thick blood smear.

I. INTRODUCTION

Malaria is a serious disease caused by a blood parasite transmitted through the bite of a female Anopheles mosquito, which urgently need a medical treatment. Malaria has been staying as a global health problem in tropical and subtropical regions [1]. Based on data from the World Health Organization (WHO) most of the deaths due to malaria by 2015 found in Africa around 90% from world cases, followed by Southeast Asia and Eastern Mediterranean. Three countries in the Asia region contributed for 96% of malaria cases occur in the world: India 70%, Indonesia 16% and Myanmar 10%. About 1.4 billion people are at risk of malaria in the 10 Southeast Asia countries. [2].

There are various of laboratory techniques to validate the presence of malaria parasites in blood samples from thin and thick types of Peripheral Blood Smears (PBS), Quantitative Buffy Coat (QBC), and Rapid Diagnostic Test (RDT) [3]. In PBS technique, microscopic diagnosis method of malaria is conducted by staining blood sample with two types, thick or thin, on glass slides to check the availability of malaria parasites visually. Microscopic analysis now become the gold standard in the diagnosis of malaria, because of its sensitivity to check the presence of parasites and low cost [4][5][6]. One drawback of microscopic analysis is the blood smear condition is depend on lighting and imaging condition. In addition this technique is time consuming and the accuracy is depending on the expertise level of paramedics, especially when the diagnosis is made on the early stage of infection [3]. QBC smears the parasite Deoxyribonucleic Acid (DNA) in microhematocrit tube with a fluorescent stain and the next detection uses the microscope epi-fluorescent. QBC is simple, reliable and user-friendly, but it requires a particular instrumentation and more expensive compare with the conventional light microscope. It also has weaknesses in species determination and the number of parasites. In RDT techniques the same principles are used by detecting malaria antigen in the blood which flows along the membrane that contains the antibody of anti-malaria, so that the RDT method does not need any lab’s equipment. RDT is a malaria rapid diagnosis but it should be used along with other methods to confirm the results, the characteristic of infection and the monitoring treatment [3].

There are two microscopic techniques for detecting malaria in the blood smear. First is perform on thick blood smear which consist of a large number of hemolyzed red blood cells to know whether the patient is infected by the parasite Plasmodium or not. In the thick blood smear which is a drop of blood on a glass slide is used to detect the presence of the parasite. A paramedic should see the 200 fields of view of the thick blood smear or counting 200 to 500 of white blood cells to check the presence of the parasite [5]. It only shows white blood cells, platelets and parasites are to detect the infection and estimate the parasitaemia [7]. The second technique is thin blood smear which consists of one layer of red blood cells to identify the species of malaria, measure the parasitaemia and recognize the shape of the plasmodium stages that infected red blood cells. The main objective of this technique is to ensure the availability of the malaria parasite in red blood cells, and to assess the size of infected red blood cells compare with the uninfected cells.

In the thick blood smear it is necessary to recognise the object, which is a parasite and artifacts such as white blood cells or platelets. Hence, the object pattern should be determined initially. In this paper, the discussion will be focused on the detection method of parasites and white blood cells.
II. CHARACTERISTICS OF MALARIA

A. Life Cycle and Morphology

An initial infection by Anopheles mosquito named, sporozoites gets into the human blood stream, then within an hour the sporozoites enters hepatocytes phases and begin to divide into merozoites exoerythrocytic (tissue schizogony) [8]. After the merozoites leaves the liver, they invade red cells and start to grow depend on the plasmodium parasite’s development.

At this stage the trophozoites parasite-shaped like a ring with one minor point or two cytoplasms. While schizontes has a circular shape with a lot of dots inside the cytoplasm and has a small number of chromatin. Gametocytes is the division of the parasite into male and female gametocyte which shaped more likely to a banana. The size of the male gametocyte is smaller than the female gametocyte.

B. Signs and Symptoms

The common symptoms that often occur from a malaria patient are headache, chills, dizziness, abdominal pain, nausea, vomiting, mild diarrhea and a dry cough [8].

C. Effects

The severe stage of the clinical condition in patient can develop around three to seven days after the fever. This can lead to several complications such as neurological complications, pulmonary complications, hypoglycemia, hypotension and shock, or hematological abnormalities [8].

D. Methods of Diagnosis

Conventional microscopy diagnosis has become a malaria diagnosis standard. In the diagnostic process a blood smear check is carried out, either thick blood smear or thin blood smear.

III. CHARACTERISTICS OF WHITE BLOOD CELLS

White blood cells (WBC) have a nucleus of cells without hemoglobin. They have a duty to prevent the spread of the pathogens [9]. The parasites can be identified by estimating the number of parasites presence in the blood 1μl which is the standard value of white blood cells number (WBC 8000 / ml). The accounted number of parasites in 200 white blood cells then multiplied by 40, hence the number of parasites is count of parasites per microliter of blood. If the exactly number of the white blood cell is known, it can be used to provide more accurate result with the correct conformation of the multiplication factor [7].

Based on its function, the white blood cells consist of two characteristics, they are [10]:

a. Phagocyte: is responsible to produce the antibody and serve the pathogens or bacteria acceptance.

• Neutrophils: 10 to 15 micron sized core with two to five lobes. In dry condition, it colored a pink-purple granules, shown in Figure 1 [9].

b. Lymphocytes: helps the immune system which has two characteristics. First characteristic have a size of 6 to 9 micron to the red blood cell where the bottom line is almost full, but not full by the cytoplasm, and the second characteristic is large lymphocytes with 10 to 17 microns in size with cytoplasmatic that has many small lymphocytes, shown in Figure 5 [9].
conducted. The image result should be good, hence, it can facilitate and support the detection process. Most of the research in plasmodium detection are implemented the preprocessing stage.

Elter et al. [11] changed the original image into monochrom image, hence, the object which contain of chromatin will be dark color and the objects without chromatin will be colored by the light gray. The proposed technique is quite discriminatory because it focused on the object with the chromatin.

Hanif et al. [12] stretched the contrast to deploy a range of values of the color image so the image contrast is increased. Generally, this technique refers to the illumination difference between the object and the background. In his study, the dark stretching technique is used for the image enhancement process, where dark stretching is a linear mapping function to improve the image brightness and the contrast levels. The process will stretch the image value range which less than the threshold value, so the image with the greater contrast need to be compressed beforehand.

May et al. [13] is converting the image from RGB to L * a * b for the image intensity values extraction. Image sharpening collaborated with negative Laplacian filter to sharpen the image and eliminate the blur effect from the original image. In further, median filter is applied for image filtering to reduce the salt and pepper noise and help to preserve the edges of objects. The last preprocessing step is a histogram stretching to increase the contrast and the image intensity.

Purnama et al. [14] used a combination of three color space on his research, the Red, Green and Blue (RGB), Hue Saturation and Value (HSV) and Hue, Saturation and Intensity (HSI) color space. They compare the contrast of Red, Green and Blue channel from the RGB image. Hue, Saturation and Value channel from the HSV image. Hue, Intensity and Saturation from the HSI image. Then, they chose the Red, Green, and Blue channel from the RGB image to combine with the Hue channel from HSI image.

Nazlibilek et al. [15] in his research conducted the mapping of intensity value. Then, convert the RGB images to grayscale and then do the calculation complement the image. After that, the Otsu method is used to convert automatically from the grayscale to the binary image.

Pinkaw et al. [16] generated the sub-image of size 140 by 140 pixels with slight modification. The input RGB image is converted to the HSI image to create a mask that will identify the location of chromatin. The Intensity and Hue channel from HSI image is used to detect a large number of white blood cells. While to detect a small number of white blood cells, Intensity channel -from the HSI image is used.

Elter et al. [11] separated the plasmodium from the object which contain chromatin based on the characteristic shape and the color intensity. Black-top-hat morphological operation is used in this study and followed by a threshold calculation operation. Afterwards, the circular morphological dilatation is applied, and flat structuring element to combine the blobs in the binary image is generated from operating threshold. Further extraction of candidate positions of plasmodium-connected components are labeled by using a simple algorithm to extract the objects from the binary image. Centroid of the object is extracted and considered as the candidate for the plasmodium position.

Huang et al. [17] used the Otsu method in the segmentation process to obtain the right threshold values automatically. Otsu thresholding will split the image into two classes. The study resulted 19.8 of Relative Distance Error (RDE) and 0.965 of dice coefficient in this process. Besides, these techniques also done in May et al. [13] and Nazlibilek et al. researches [15]. But, the Otsu thresholding can’t be used in all situations, Otsu can’t give the right threshold value if the data variance is very different with another data.

C. Feature Extraction

Feature extraction is a process of making the unique characteristics of an object that will be analyzed in the next stage.

Elter et al. [11] initially made a small cuts on the Region of Interest (ROI) with a size of 80x80 pixels. The feature extraction consist of statistical features, textures and color analysis features. The used statistical features contain of four parameters: mean, variance, skewness and kurtosis. All features are normalized to obtain the mean value of 0 and standard deviation of 1. At first, invariant ranked is applied so only the top 60 features are used. The chosen features then selected by using Genetic Algorithm (GA) for the automatic selection features of the selected section.

Huang et al. [17] used the feature shape and size of the erythrocytes. In this research, 85 features are employed. The 80 texture features and five shape features are used, but the computation found very hard and generate some problems in the computation process. Furthermore, the selected features of each core are obtained. Principal Component Analysis (PCA) is used to determine the important feature, and to minimize the dimension of data with maintain the characteristic of dataset.

B. Segmentation

Segmentation is a process of dividing the image into several homogeneous regions based on certain criteria. The main focus at this stage is the white blood cells and plasmodium. The proper segmentation technique will be very influential for the next stages and give an effect to the diagnosis accuracy.
Purnama et al. [14] used the statistical analysis features of the five components of the three color spaces, which are Red, Green and Blue in RGB. As well as Hue on HSV and HSI. The parameters used in the statistical analysis are mean, standard deviation, kurtosis, skewness and entropy. These statistical features, are used to identify the six classes.

Pinkaew et al. [16] used the statistical analysis features of four color components, which are Green of RGB, Intensity of HSI, Saturation of HSV and Value of HSV. The used statistical analysis parameters are mean, standard deviation, kurtosis, skewness and entropy. The statistical analysis, of kurtosis can show the asymmetry of the skewness from the histogram around the mean sample. Furthermore, the entropy is used to characterize the number of image information.

D. Classification

The obtained features are then classified in the classification process, whether the object from the image is white blood cells or plasmodium. Several studies of plasmodium classification already done before.

Elter et al. [11] used Support Vector Machine (SVM) classifier to classify two classes of the ROI. They used the SVM with Radial Basic Function (RBF). Basic construction can minimize the feature dimension vector by projecting them onto low dimension with maximum variance between feature dot’s, so the classification process can be more efficient. This method is also used in Pinkaew et al.’s [16] research.

Purnama et al. [14] used a Genetic Algorithm (GA) to classify two classes and six classes. The result of the experiment then compared, and obtained the classification into two classes in the thick blood smear are more accurate than the classification into six classes.

Nazlibilek et al. [15] used the classification method of Multilayer Perceptron (MLP) and Principal Component Analysis (PCA) technique. PCA is the oldest transformation technique used for multivariable analysis. In this technique, the input dimension of the dataset is minimized. Mathematically, PCA is the orthogonal linear transformation to change the data to the new coordinate system, so the largest variance from the data projection is located in second coordinate.

V. DISCUSSION

Elter et al. [11] proposed a study by using ground-truth segmentation to identify positive and True-False-positive. Sensitivity was calculated and will be acquired True-positive rate (TPR), which will be used to calculate the average number of false-positive detection per image (FPI). The proposed method is accurate for cases of plasmodium density of less than five per image. Better features are needed to reduce of False-positive and to discriminate the plasmodium or artifacts.

Hanif et al. [12] in their research, produced images with white and dark areas. Dark areas called the parasite while the unwanted area is converted to white areas. Dark stretching utilized for enhancement and thresholding technique and generated clearly visible parasite result. Dark stretching is potential for enhancement and segmentation process by adjusting the threshold. The threshold value must be smaller than stretching dark factor.

Huang et al. [17] proposed a segmentation method by using Otsu multilevel thresholding. Texture features and forms of the erythrocytes core in the feature extraction process used as the PCA before the detection process. The study resulted 19.8 of Relative Distance Error (RDE) and 0.965 of dice coefficient on image segmentation process. By using the feature selection, the study succeeded in reducing the number of features significantly from 80 to 7 features.

May et al. [13] proposed a method of research of preprocessing, segmentation and classification. The study resulted 99.72% of sensitivity, 99.94% of specificity and Positive Predictive Value (PPV) of 98.90%. The used of Otsu in the segmentation process resulted a high accuracy segmentation in the different picture’s conditions. The Otsu method can obtain the right threshold automatically. This technique is effective and can make the process faster because the determination of manual threshold is not necessary. In case of significant different contrast this technique is recommended. But, for the image with low different contrast this technique can abolish the detected object.

Purnama et al. [14] showed that the statistical analysis features successfully utilized for the feature extraction, especially if implemented on each component in RGB and Hue color components of HSV and HSI. RGB is the color space with Red, Green and Blue channel. HSV is the color space with Hue, Saturation and Value channel, while the HSI is the color space with Hue, Saturation and Intensity channel. The statistical analysis is appropriate to use in the feature extraction, because we can see the asymmetry of the histogram distribution of the image. Furthermore, the uninfected image value parameter in statistical analysis can be used to compare with the infected image value.

Nazlibilek et al. [15] proposed a method of image enhancement, segmentation and classification. Initially the image passed the enhancement stage to get a better image quality as the input of segmentation process. Afterward, the image is segmented and calculated. The counted segmentation image then grouped into five types of white blood cells by each characteristic. The number of each type obtained from the input images. However, misclassification occurred due to the variations in the intensity. The uninfected cells which have the darker colors are considered as an infected cell and the morphological operations are not be able to remove it. The overlapped cell also caused an inaccuracy.

Sarrafzadeh et al. [18] proposed a method of segmentation of nuclei stage, cytoplasm segmentation and White Blood Cells’s (WBC) extraction. In the nuclei segmentation K-means algorithm and morphological operations are used. Afterwards in the WBC’s extraction phase the features of white blood cells are taken. In this case, segmentation by observing the color information is not a good method except for the typical color variations images.

Pinkaew et al. [16] proposed a method with preprocessing, feature extraction, basic construction, and classification. In the image preprocessing stage the image is converted from RGB
into HSI. Furthermore, in feature extraction process a statistical analysis on four colour components is used. Afterwards, the feature dimension on the basic stages of construction is reduced. In the last stage, the classification process is implemented by using SVM classifier. The basic construction of the classification can be more efficient. However, SVM classifier can only classify into two classes. It is suitable to use in the thick blood smear case, because it classifies into positive and negative classes.

Table 1. Method performance

<table>
<thead>
<tr>
<th>Paper</th>
<th>Year</th>
<th>Preprocessing, segmentation, feature extraction and classification</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elter et al. [11]</td>
<td>2011</td>
<td>black-top-hat morphological operation followed by a threshold calculation operation</td>
<td>97%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>circular morphological dilation is applied, and flat structuring element is generated from operating threshold</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>extraction of candidate positions of plasmodium-connected components are labeled by using a simple algorithm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>the feature extraction consist of statistical features, textures and color analysis features</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>the chosen features then selected by using Genetic Algorithm (GA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>used the SVM with Radial Basic Function (RBF) for classification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hanif et al. [12]</td>
<td>2011</td>
<td>stretched the contrast to deploy a range of values of the color image</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dark stretching technique is used for the image enhancement process</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May et al. [13]</td>
<td>2013</td>
<td>converting the image from RGB to L * a * b for the image intensity values extraction</td>
<td>99,72%</td>
<td>99,44%</td>
<td>98,90%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>image sharpening collaborated with negative Laplacian filter followed by median filter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>used the Otsu method, negative image transformation, image dilation, image erosion, clear border and hole filling in the segmentation process</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purnama et al. [14]</td>
<td>2013</td>
<td>used a combination of three color space, the Red, Green, and Blue channel from the RGB image combine with the Hue channel from HSV and HSI image</td>
<td>-</td>
<td>-</td>
<td>95%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>used the statistical analysis features of the five components of the three color spaces</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>used a Genetic Algorithm (GA) to classify two classes and six classes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pinkaew et al. [16]</td>
<td>2015</td>
<td>RGB image is converted to the HSI image</td>
<td>-</td>
<td>-</td>
<td>P.Falciparam : 85,71%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>used the statistical analysis features of four color components, which are Green of RGB, Intensity of HSI, Saturation of HSV and Value of HSV</td>
<td></td>
<td></td>
<td>P. Vivax : 78,72%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>used the SVM with Radial Basic Function (RBF)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paper</th>
<th>Year</th>
<th>Preprocessing, segmentation, feature extraction and classification</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huang et al. [17]</td>
<td>2012</td>
<td>used the Otsu method in the segmentation process</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>used the feature shape and size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>the selected features of each core are obtained. Principal Component Analysis (PCA) is used</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nazlibilek et al. [15]</td>
<td>2014</td>
<td>convert the RGB images to grayscale and then do the calculation complement the image</td>
<td>-</td>
<td>-</td>
<td>95%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>used the Otsu method in the segmentation process</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>used the classification method of Multilayer Perceptron (MLP) and Principal Component Analysis (PCA) technique</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sarrafaideh et al. [18]</td>
<td>2015</td>
<td>used the K-means algorithm in image segmentation by observing the color information</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>morphological operation is used to make the mask of object</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1 mentions about five plasmodium detection method and three white blood cells detection method. The table also shows the result comparison of all conducted methods. Unfortunately, not all research explain the quantitative result. The quantitative result explained in Elter et al. [11] for their research with 97% of sensitivity, May et al. [13] research with 99,72% of sensitivity, 99,94% of specificity and 98,90% of accuracy. I Ketut et al. [14] and Nazlibilek et al. [15] research obtained 95% of accuracy, while Pinkaew et al. [16] gained 85,71% of accuracy for plasmodium falciparum and 78,72% of accuracy for plasmodium vivax.

All researches result as discussed above are not confirmed by looking at 200 of the visual field and counting 200 to 500 of white blood cells.

VI. CONCLUSION

In some studies the general steps of image processing are not necessary and omitted. It can be concluded that the different type of data required different steps to produce a better accuracy. However, there is no rule to produce good accuracy and help the diagnosis of malaria.

Dark stretching technique is potential to stretch and segmented the image by observing the threshold value. To obtain the right value, Otsu thresholding can be used because this method will automatically determine the right threshold value. Otsu thresholding method is capable to produce a better accuracy of image segmentation in various image conditions.

The used of features selection is used for classification need to be considered in order to optimally reduce the occurred false-positive detection. Statistical analysis features are commonly used and generated good result in feature extraction process. Murer's dot shape features color also need to be evaluated. However, for the features color is not proper enough in some cases, because the obtained results did not have the distinctive color images. Hence, the feature selection stage is required to obtain the important and dominant features from the image.

In addition, SVM classifier gives an accurate result for plasmodium with image density less than five plasmodium cases. SVM classifier gives the results of 97% of sensitivity. In malaria screening cases, SVM classifier is very suitable for two-class of classification (negative and positive). However, to obtain a better result in the SVM classifier, more training data is necessary.

A research which focused on thick blood smear should be added that stage. The confirmation from this research will affect the image classification result and affect the diagnosis. Look for the best method in distinguishing the parasite and artifacts is really needed, so can be used to count parasite density.

References


