Rika Rosnelly

Diperikasa oleh LPPM Universitas Potensi Utama

Submission date: 05-Nov-2018 10:02AM (UTC+0700)
Submission ID: 1032881917
File name: 22Vol95No3_2-JATIT-2017.pdf (482.68K)
Word count: 6034
Character count: 32334

Journal of Theoretical and Applied Information Technology
15th February 2017. Vol.95. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IDENTIFICATION OF MALARIA DISEASE AND ITS STADIUM BASED ON DIGITAL IMAGE PROCESSING

[1]RIKA ROSNELLY, [2]SRI HARTATI, [3]EDI WINARKO, [4]SRI MULATSIH

[1,2,3]Department of Computer Science and Electronics, UGM
[4]Department of Pediatric, Faculty of Medicine, UGM/ Dr. Sardjito General Hospital
E-mail:[1]rikarosnelly@gmail.com, [2]shartati@ugm.ac.id, [3]ewinarko@ugm.ac.id, [4]smulat@gmail.com

ABSTRACT

Malaria is one of the public health problems that could cause death, especially in infants, toddlers, pregnant women. Malaria is still a health problem not only in Indonesia but also in some countries in the world. Microscopic testing is the gold standard to malaria disease diagnosis. However, the level of accuracy depends on the level of microbiological expertise and experience. Microscope testing also time consuming and requires extensive equipment. This study built a texture of feature extraction and morphology model that can be used to identify the type of malaria parasites along with the stadium on the image of a thin blood smear. Samples are a number of preparations that have been given indications of malaria and have given a Giemsa staining. Image acquisition process is done by using a digital microscope with 1000 times of magnification. The process of segmentation used thresholding method of Otsu. Selection feature used sequential forward selection (SFS). Classification technique used artificial neural network of learning vector quantization (LVQ) with K-fold cross validation to identify patterns of types of parasites and their life stages in order to get different types and stages of malaria disease. The results used a combination of texture with morphological traits to get the values of accuracy.

Keywords: *Malaria Disease, Texture Of Feature Extraction And Morphology, SFS, LVQ, K-Fold Validation*

1. INTRODUCTION

Malaria is a tropical disease type and many tropical countries have experienced it. It is threatening Asian countries, namely India, Indonesia, and other Asian countries [7]. WHO estimates malaria infects 300-500 million approximately and caused more than one million of deaths. Indonesia is one of the tropical country that highly prevalent types of tropical disease and still be a problem for public health. Figures for malaria disease are high enough for outside Java and Bali [14]. According to WHO reports that in 2012 the number of children decease decreased to fewer than 500,000. Overall, it predicted that there were 207 million cases of malaria in 2012 which caused 627,000 of deaths. According to the report, there were 102 countries with malaria transmission. The figure compares to 219 million cases and 660,000 of deaths in 2010 [11].

Microscope testing is the gold standard for malaria diagnosis, but the level of accuracy depends on the level of technician/ microbiologist expertise and experiences. Test use a microscope is time consuming and requires extensive equipment [16]. Another disadvantage of diagnosis uses manual microscope is requiring human intervention during the diagnostic process which may cause delays and misdiagnosis may occur. Microscopic needs training to gain expertise in the diagnosis, from the samples that need to be analyzed, this method is inconsistent and depend on blood smear, staining quality and the quality of the microscope [4]. For the examination of malaria disease has developed a new way that aim to confirm more quickly. The examination of malaria disease uses RDT (Rapid Diagnostic Test) has the ability to improve diagnostic system, but this method is not able to detect some of parasitic infection on samples with low concentrations of parasites. Accuracy results are affected by humidity and high ambient temperatures [16].

The problem in this research was determining the correct feature extraction model of texture and morphology in order to distinguish the type of malaria disease and its stages. The limit problems in this research are the images of the malaria parasite used here are the images of the malaria

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

parasite falciparum, malariae, vivax with stage ring, tropozoit, schizonts, and gametocytes. Types of malaria identified are malaria falciparum, malariae, vivax with stage ring, tropozoit, schizonts, gametocytes. Some studies which related to the features that are used to analyze the types of malaria parasites are using morphology feature [3] [12] [15] [22] [23] [24] [25], texture and morphology [5], color and statistic [2], color and shape [13], geometry, color and statistic [4], color, shape and texture [9][21]. According to previous studies of malaria parasites as aforesaid, mostly using morphological features. The study of malaria parasites that used a combination of texture and morphological feature which conducted by previous researchers are still rare. Therefore, they need to obtain the texture and morphology feature so it is more complete.

In order to overcome this problem, this research built a model with a combination of textures and morphological feature extraction. The result of the combination was the correct identification of malaria disease.

In this research, the type of malaria disease that had been researched was falciparum ring stadium, falciparum tropozoit stadium, falciparum schizon stadium, falciparum gametocytes stadium; malariae ring stadium, malariae tropozoit stadium, malariae schizon stadium, malariae gametocytes stadium; and vivax ring stadium, vivax tropozoit stadium, vivax schizon stadium, vivax gametocytes stadium. The features that had been used on the texture were energy, entropy, homogeneity, contrast, and correlation. Meanwhile, the features that had been used on morphology were perimeter and area.

In this research, the classification method gave impact the identification result. The system to classify malaria parasite can be improved by using aneural network strategy that imitated from feed forward back propagation [9][15][17]. Another classifier system used machine learning [21][23][26].

Based on the study [6] of reviews for the identification of malaria parasites computers based, there are 112 articles that describe computer-based malaria diagnosis using microscopic images ranging from 2000 to 2015. The classification method in [6] mostly using Support Vector Machine (SVM). Learning vector quantization (LVQ) based classification to classify 12 classes of plasmodium malaria and their life-cycle stages has not been developed. This method developed by Teuvo Kohonen has many advantages over other types. The advantages of Learning Vector quantization are easy to classify multiclass problems and adapted to the complexity of the algorithm and the needs during training [27]. LVQ training will produce a final weight, and then it will be used for data validation.

Data with limited numbers usually contained different value of accuracy in the validation process. The value of the training data accuracy can reach 100% but at validation data was lower than training data [19]. One way to overcome this issue is using K-Fold Validation method. This research used LVQ with K-Fold validation to identify patterns of malaria parasites along with the stages. The results of pattern recognition are measured by using the value of accuracy, sensitivity, and specificity.

The rest of the paper is organized as follows : The second part presents a review of the literature. The third section presents the methodology of the study. The fourth section presents the experimental results. The fifth part is the conclusion.

2. LITERATURE REVIEW

Some of the previous research used highly advanced computerized imaging method for malaria disease detection, classification and automatic identification. Each study has different characteristic of image types. Therefore the method that has been used for segmentation was varying as well. The result of segmentation is stated well if the image can be differentiated objects on the image of Region of Interest (ROI). The object is malaria parasite.

In general, the process of segmentation is performed by image enhancement and thresholding. The study [3] was using global thresholding method for automatic status segmentation of microscopic images which obtained from thin blood smear for malaria plasmodium falciparum.

The study [20] used watershed algorithms edge detection for blood cells segmentation. This research separated blood cell components from the background and identified the elements of parasite, namely nucleus and cytoplasm in infected erythrocytes. Although the threshold had been selected automatically, this method is very dependent on the image quality.

The stages of the segmentation process that had been done in research [16] is different from [3] where the study [16] used Otsu method to search for local boundary so that the image after preprocessing stage can be converted into a binary image. The main principle of otsu method is classifying the objects and the background of malaria parasite.

The study [16] was only to classify samples of malaria parasite type of vivax tropozoit stadium. The characteristic of vivax malaria has larger than red blood cells. There are red blood cells in human blood, red blood cells is infected by malaria parasites, white blood cells that is larger than the malaria parasite, platelet and artifacts.

The stage of feature extraction is an important part of the classifier because it affects the work of the classifier. The purpose of feature extraction is to identify and extract relevant information in the image that will be recognized by other objects. Features that have been extracted then used as parameters that can distinguish one object with another object in the identification stage [21].

The selection of features has an important role in identify the type of malaria parasite. Types of malaria parasites can be distinguished by a combination of the features of shapes, textures and colors that had been taken from across the red blood cel. The number of images that had been used was 68 images of falciparum plasmodium and vivax plasmodium with ring stages, tropozoit and gametocytes. The identification of falciparum plasmodium and vivax plasmodium were very difficult to distinguish on ring stage [21]. On ring stadium, the form of feature is not able to distinguish falciparum plasmodium and vivax plasmodium because of the similar shape [15].

The system to classify malaria parasite can be improved by using an imitation of neural network strategy of feed forward back propagation [9] [15] [17]. The result of the previous research indicated vary of accuracy, sensitivity and specificity depend on the number of parasite species to be identified, the type of features and classification methods that had been used.

The propose research will enhance research [9] in order to determine the type of malaria plasmodium and its species from the blood smear image. In the process of feature extraction, the feature simply grouped into four classes according to the four life stages of plasmodium [9]. Research [1] showed that vivax malaria had consistent levels based on the pixel values. In this case is the area feature was used as a main feature in developing a classification system. Research [2] used color features and statistics to feature extraction, that is phase of the image, mean value of green plane, skewness, kurtosis, standard deviations, and energy, only to diagnose malaria parasites or not exposed to malaria by digital image processing.

3. RESEARCH METHODOLOGY

Sampling was carried out in the Central Health Laboratory North Sumatera Province, parasitology department where the sample was malaria preparation from Bina Pelayanan Penunjang Medik (BPPM) Jakarta. Malaria preparation was reexamined by laboratory worker and labeled it to determine the type of parasite and its stadium. There were 3 types of malaria parasites that exist in the Central Health Laboratory North Sumatra Province, namely falciparum, malariae and vivax with four stages, namely ring, tropozoit, schizont and gametocyte.

Identification model of malaria including image acquisition, selection of ROI (Region of Interest), image enhancement, segmentation, feature extraction of texture and shape, feature selection, identification and diagnostic tests (Figure 1).

Training Examining

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

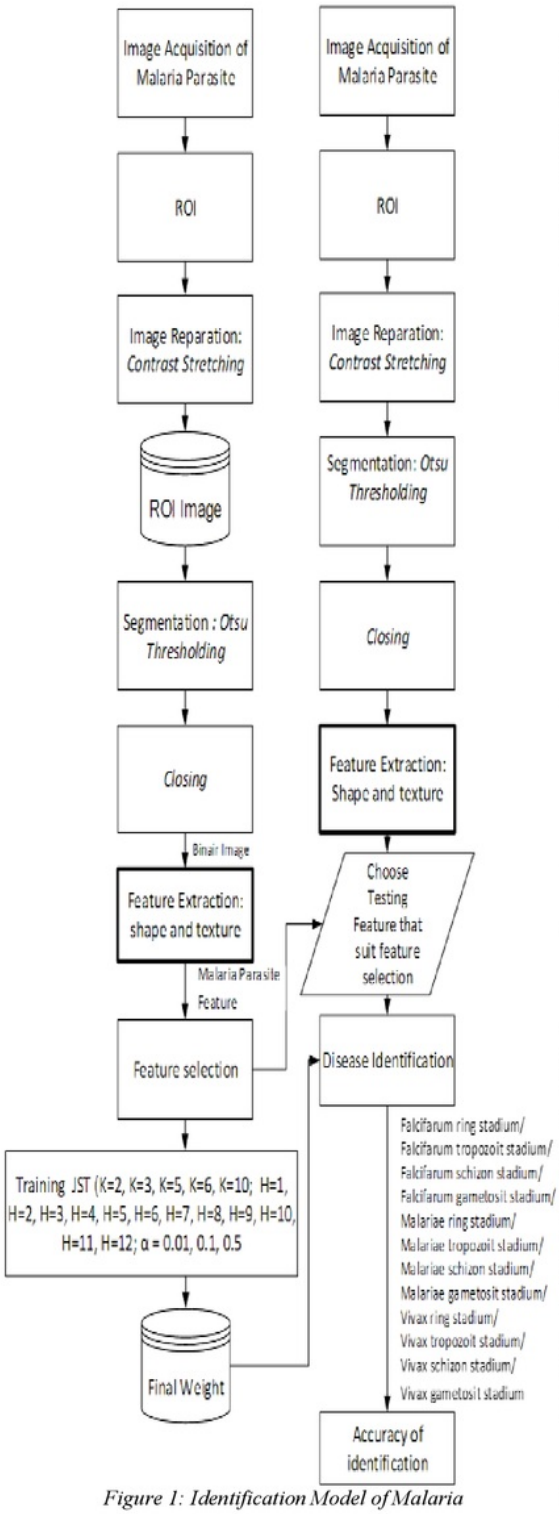

Figure 1: Identification Model of Malaria

In the image acquisition process of the malaria parasite, at Central Health Laboratory North Sumatera Province were already contained Giemsa stain in the preparation. Before the inspection was done, preparation dripped by immersion oil, then the image acquisition process is done by using a digital microscope. Magnification on the image acquisition process is 1000 times. Furthermore, the image can be viewed on screen at which the reexamination process was accompanied by medical laboratory workers to identify the type of malaria parasite. In all three species of plasmodium, each has 4 stages namely stage of rings, trophozoites, schizonts and gametocytes. The image of plasmodium type can be seen in Figure 2.

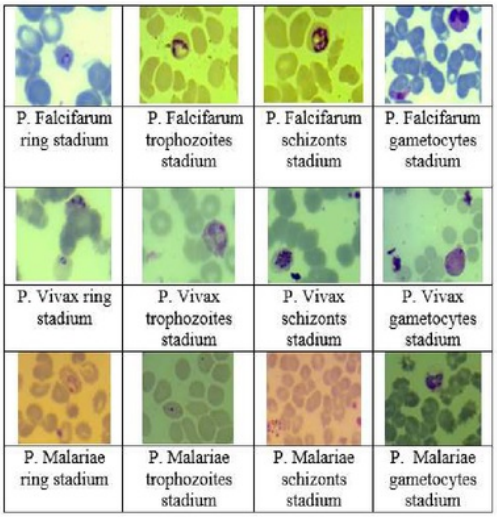

Figure 2: Type of Malaria Parasites

After getting the image of the type of malaria parasites along with the stage turns to the image of the malaria parasite has some objects consisted of normal red blood cells, red blood cells exposed to the malaria parasite, platelets, white blood cells and artifacts. In order to overcome the problem, the ROI was done. In the image of malaria parasite turns out the position or location of the malaria parasite in the image was random. Furthermore, ROI was done automatically based on the color; turns to the image of malaria parasites contained other objects that had similar color as the object of malaria parasite, such as platelets and white blood cells. Based on the problem above, this study did ROI process manually. The input in this process was the image of malaria parasite.

When the image acquisition process of the malaria parasite was done, the quality of the image

was not bright well. The aim of image reparation was to improve the image quality of image display for humans view or converting an image to have a better quality so that the image becomes easier to be processed by computer.

Segmentation method used Otsu thresholding. Otsu method is based on the intensity values of the pixels in the image and to find the optimum threshold. Optimum means the way it works that maximizes the variance interclass (between-class variance).

Feature extraction process is then performed to acquire the feature of texture and morphology of malaria parasites in each ROI. The proposed method was texture feature and shape features to identify the type of malaria disease consists of energy 0^0, energy 45^0, energy 90^0, energy 135^0, entropy 0^0, entropy 45^0, entropy 90^0, entropy 135^0, contrast 0^0, contrast 45^0, contrast 90^0, contrast 135^0, homogeneity 0^0, homogeneity 45^0, homogeneity 90^0, homogeneity 135^0, correlation 0^0, correlation 45^0, correlation 90^0, correlation 135^0, área and perimeter.

Thus, the study has a data set consisted of 22 features and they were grouped into 12 classes which consist of three types of malaria parasites along with its stadium where the features clustering bécame one of the contributions to this study.

It was important to select the features from the 22 features that obtained. This is because not all features that had been gained in feature extraction played an important role in the examination of malaria. In this study, the method used in the features selection was sequential forward selection (SFS) in Matlab. On Sequential Forward Selection (SFS), the most significant feature would be selected gradually. At the beginning each feature would be tested of how big its contribution to the identification results. Feature with the best accuracy would be selected, and then it would add to the following feature so there were 2 features to be tested. The combination of two features with the highest yields have subsequently tested by adding one following feature and so on until the combination of several features with the highest accuracy were found. Each feature on histogram represented malaria parasites feature that played an important role in the next process which was the training and testing.

The input on classification was the result texture and shape feature selection. The amount of data used in pattern recognition was 600 data with the number of features as the result of feature selection. Animitation neural networks founded very useful in many medical applications. In the learning process of supervisión pattern recognition, the recommended method is animitation neural network because the training process was relatively fast, provided relatively good accuracy and able to model non linear systems. This study used Learning Vector Quantization (LVQ).

Figure 3 shows the LVQ architecture design, the input was input vector which indicated the number of features that would be trained. Input vectors would be grouped into 12 classes. The parameters that was used were the value of learning rate, the decrease in the rate of learning, learning epoch and the minimum rate to the extent of minimum value and maximum rate of learning epoch has met. At the stage of training, LVQ algorithm processed the input to receive input vector and description of the class number, then the calculated distance vectors against the entire weight vector for existing k classes. Input vector was then classified into classes with a distance weight vector closest to the input vector.

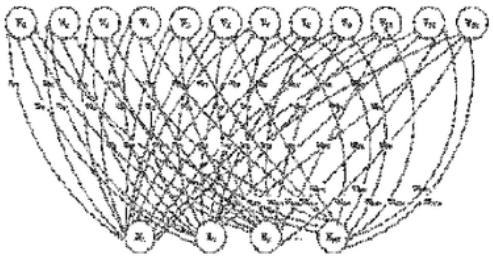

Figure 3: Architectural model of LVQ development

In the process of training and testing, LVQ used vector operations. LVQ had weight vector in each unit of output which was often called a codebook. During the training process, LVQ did an update on each of the weight vector from unit that representing each class. Once the training process was completed, we conducted the testing process by calling the final weight and the data to be tested [8]. Figure 3 shows that LVQ in this research, $X_1 \ldots X_{22}$ were the 1 st to 22 nd feature, $W_{11} \ldots W_{22\ 12}$ were the weight that related input to output. The 1 weight for 1 class to the 22 weight for 12 classes.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. EXPERIMENTAL EVALUATION

The process of determining ROI was done by manual cropping on malaria parasite object. Then two types of ROI size had been taken, namely 400 x 400 pixels and 256 x 256 pixels. These two types of measurements were consulted to laboratory worker and did the validation. This process was done to ensure that ROI was taken only from the area of malaria parasite and it was not contained with other objects such as platelets, leukocytes, or artifact. The tests showed that the appropriate size for ROI is 256 x 256 pixels. ROI is then stored in the image file in bitmap format.

Image enhancement method used here was a contrast stretching, which performed to enhance the quality of the original image. The result will be used for the segmentation process.

The segmentation process was used to separate the object of malaria parasite with red blood cells. Through this process, the boundary between the object and the malaria parasite of red blood cells can be determined. The excess used of Otsu thresholding was the selection of the threshold value with the discriminant analysis. The image histogram was divided into two groups and the threshold value was determined when the variance between the two groups was maximum [18].

The analysis result showed the differentiation of morphology identification class of malaria parasites along with its stadium. The result of class morphology analysis of falciparum ring stadium (FR), falciparum tropozoit stadium (FT), falciparum schizon stadium (FS), falciparum gametocytes stadium (GF), vivax ring stadium (VR), vivax tropozoit stadium (VT), vivax schizon stadium (VS), vivax gametocytes stadium (VG), malariae ring stadium (MR), malariae tropozoit stadium (MT), malariae schizon stadium (MS) and malariae gametocytes stadium (MG) can be seen in Figure 4.

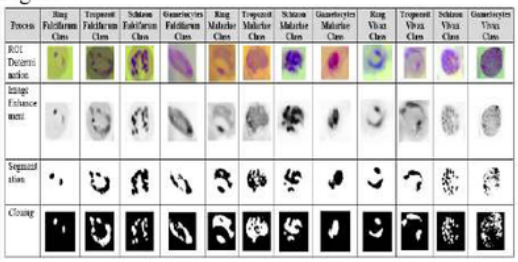

Figure 4: Morphology Analysis Classes FR, FT, FS, GF, MR, MT, MS, MG, VR, VT, VS, VG

Then, feature extraction process conducted on the result of image segmentation in order to get texture and shape features from malaria parasite in each ROI. Texture feature extraction for training phase used Gray Level Co-occurrence Matrix (GLCM), calculated by multiplying the probability of adjacency relationship between two pixels whithin a certain angle and orientation. In this study, there were four GLCM that had been used to determine the features of the image that was GLCM with one spatial distance with the angle 0^0, 45^0, 90^0, and 1350. The stages to search for feature extraction used GLCM can be seen in Figure 5.

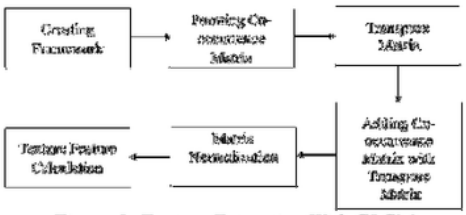

Figure 5: Feature Extraction With GLCM

Figure 5 in this research created the image with a size 256 x 256 pixels and a framework formed with a size of 256 x 256. Then, co-occurrence matrix was made by entering the number of spatial relationships that exist in the matrix. Co-occurrence matrix that had been obtained then transposed in order to obtain a symmetrical angle which was 180^0 angle. The next step was summed up co-occurrence matrix and transposed matrix where this step would position the symmetrical between 0^0 and 180^0 angle. The normalization was obtained by added up all symmetric matrix, and then used as dividers for all pixels contained in the symmetric matrix. The last step was done by calculating feature texture which consists of energy, entropy, contrast, homogeneity, and correlation. The calculation for the texture features of one image data can be seen on formula 1 to 5 [10].

1. Energy
   Measure the uniformity of texture. The maximum value of the energy is 1. The value of energy will be high when the distribution of pixels in a state of constant or periodic.
   $$\sum_{i,j=0}^{N-1} P_{i,j}^2 \quad (1)$$
   P (i, j) is the probability of co-occurence.
2. Entropy
   Measure images randome. Entropy would be high value when images were not uniform. Entropy was inversely related to energy.
   $$\sum_{i,j=0}^{N-1} P_{i,j}\left(-\ln P_{i,j}\right) \quad (2)$$
   P (i, j) is the probability of co-occurence

3. Contrast
   Measure the spatial frequency of an image as well as the difference between high and low of a pixel.
   $\sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2$ (3)
   P (i, j) is the probability of co-occurence
   Requirement: when i = j, then the contrast filled 0.
4. Homogeneity
   Measure the homogeneity of an image.
   $\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$ (4)
   P (i, j) is the probability of co-occurence
5. Correlation
   Measure the linierity of an image. The correlation measure can easily to be used with other feature.
   $\sum_{i,j=0}^{N-1} P_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right]$ (5)
   P (i, j) is the probability of co-occurence.

Shape feature extraction for training step used area and perimeter [5].

a. Area is the number of pixels of the object that indicates its size. Área formula can be seen on formula 6: $Area = \sum_x \sum_y f(x,y)$ (6)
   Where f(x,y) is malaria parasite object binary image.
b. Perimeter is the length of the resulting framework. In the chain code, perimeter formula can be seen on formula 7:
   Perimeter =
   $\sum_x \sum_y f(x,y).x, y \in Boundary\ region$ (7)

By using 600 data consisted of 510 training data and 90 test data, the feature selection of malaria parasites was done by using Matlab software. The data had 22 features or attributes as the result of feature extraction of texture and shape. In feature selection of malaria parasite, produced 5 important features, namely: area, perimeter, energy 0^0, energy 90^0 and contrast 0^0. The result of feature selection can be seen in Figure 6.

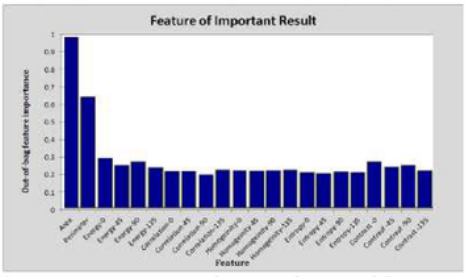

Figure 6: Histogram on feature selection of the type of malaria parasites

From the five features of the type of malaria parasites, the most important feature was area feature to examine the type of malaria while the perimeter, energy 0^0, energy 90^0, and contrast 0^0 was a feature considered to examine the type of malaria.

Feature of energy 45^0, energy 135^0, correlation 0^0, correlation 45^0, correlation 90^0, correlation to 135^0, homogeneity 0^0, homogeneity 45^0, homogeneity 90^0, homogeneity 135^0, entropy 0^0, entropy 45^0, entropy 90^0, entropy 135^0, contrast 45^0, contrast 90^0, contrast 135^0 based on the histogram was a feature that did not play a role to examine the type of malaria disease. It can be eliminated and was not used in the LVQ training process. The result of feature selection of the malaria parasite then can be used as input in training on LVQ.

Training used to obtain the best value of learning rate (α), the decrease learning rate (dec α), minimum learning rate (min α), and the final weight that will be used for examination. The parameters that will be used are listed in Table 1.

Table 1: Parameters LVQ for training hallmark characteristic of malaria parasites by selection

Total input	The initial weight	*K-Fold*	Value (α)	Dec Value (α)	Min Value (α)
5	Random	2,3,5,6,10	0.01	0.001	0.0001
		2,3,5,6,10	0.1	0.01	0.001
		2,3,5,6,10	0.5	0.1	0.05

In the training process, the initialization of weights used random numbers. Each parameter had been tested by training and validation data in accordance with K-fold validation. K value was 2, 3, 5, 6, and 10. The value of learning rate (α) was 0.01, 0.1, and 0.5. The condition was stopped at the minimum value of the learning rate (min α) was 0.0001, 0.001 and 0.05. Once the training was done, the result of accuracy value, epoh, mean square error (MSE), and timer were obtained. The result showed that the average of training data accuracy on all parameters had 100% accuracy value. In order to get the appropriate parameter, the highest accuracy value had been chosen from validation data. If we get the same accuracy value, then the parameter selection conducted based on minimum epoch on the accuracy value. Graph of average percentage of accuracy values on learning rate value 0.01, 0.1, 0.5, training validation data can be seen in Figure 7.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

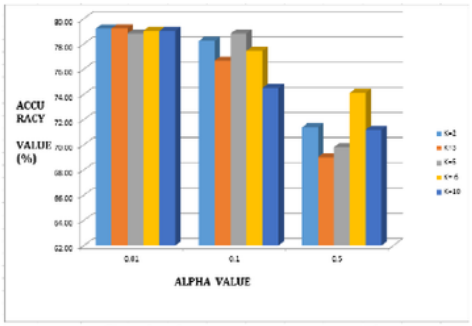

Figure 7: Graph of average percentage of accuracy value of validation data on malaria parasite feature as a result of feature selection

According to Figure 7, it shows that average value of the highest training accuracy 79.22% was achieved in 3-fold with α value was 0.01. In the study also conducted a training of complete malaria parasites feature (without feature selection) using LVQ. The training parameters with this method can be seen in Table 2, while the results of training to get the value of 100% accuracy on all K.

Table 2: Parameters LVQ for training complete malaria parasite feature

Total input	Initial weight	*K-Fold*	Value (α)	Dec Value (α')	Min Value (α')
22	Random	2,3,5,6,10	0.01	0.001	0.0001
		2,3,5,6,10	0.1	0.01	0.001
		2,3,5,6,10	0.5	0.1	0.05

The result of average accuracy of data validation on complete malaria parasite features based on parameters in Table 2 can be seen in Figure 8.

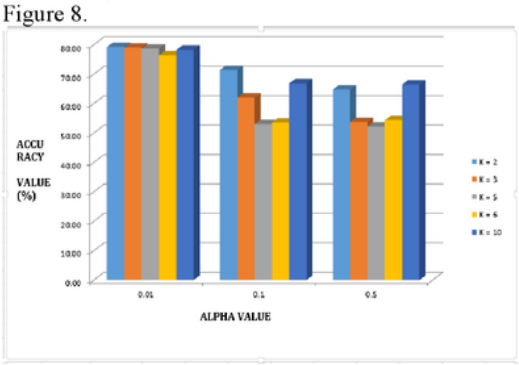

Figure 8: Graph of average percentage of accuracy value of validation data on complete malaria parasite feature

Figure 8 shows that, the highest value of the training accuracy was 79.21%. This value was reached in the case of 2-fold. Alpha value was 0.01, alpha decrease was 0.001, alpha minimum was 0.0001 and epoch value was 4603. Validation process and testing on malaria parasite features was done by using learning vector quantization (LVQ) method with k-fold.

In the testing process, 90 data that represent 12 classes were used. The result of testing data using selected features reached the average of accuracy of 77.78%. By using all features, not selected, the experimental testing showed that the average value of accuracy was 74.44%. Having these two results showed that the testing with selected feature resulted in better accuracy that is 77.78%. Table 3 showed the results of testing data.

Table 3: Accuracy Value on Training Data, Validation and Testing by Using LVQ

Method	Data	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)
LVQ	Feature of malaria parasite with feature selection	100 at all k-*fold*	79,22 at 3-*fold*	77,78
LVQ	Complete Feature of malaria parasite (without feature selection)	100 at all k-*fold*	79,21 at 2-*fold*	74,44

The testing process on malaria parasite feature was carried out by using LVQ with k-fold. It also was carried out with and without any feature selection. In order to determine the best validation results, we used the average accuracy value of validation data at each fold value. The best parameter was selected from the highest average value. The final weight and the best α value would be saved. Then they were used in testing process.

The result of testing process then be used to evaluate the accuracy, sensitivity and specificity, using confusion matrix, such as shown in Table 4.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 4: Confusion matrix of LVQ testing on malaria parasites feature as feature selection results

Actual	Prediction												
	FR	FS	FT	GF	MG	MR	MS	MT	VG	VR	VS	VT	Total
FR	5	2	0	0	0	2	0	1	0	4	0	0	5
FS	0	2	0	0	0	2	1	0	0	0	0	0	5
FT	0	0	5	1	0	0	0	0	0	0	0	0	5
GF	0	0	0	15	0	0	0	0	0	0	0	0	15
MG	0	1	0	1	3	0	0	0	0	0	0	1	5
MR	0	0	0	1	0	1	1	0	0	0	0	0	3
MS	0	0	0	0	0	0	13	0	0	0	0	0	13
MT	0	0	0	0	0	0	2	3	0	1	0	0	6
VG	0	0	0	0	0	0	0	0	9	0	0	0	9
VR	0	0	0	0	0	0	0	0	0	15	0	0	15
VS	0	0	0	0	0	0	1	0	0	0	2	0	3
VT	0	0	0	2	0	0	0	0	0	0	0	4	6
Total	5	5	5	18	3	5	28	5	9	20	2	5	90

Table 4 shows that there were 70 from 90 data which can be identified correctly. The test results can be calculated as well by sensitivity and specificity values. Those values can be seen on Table 5.

Table 5: Accuracy, sensitivity and specificity values of LVQ testing on malaria parasites feature with feature selection

Accuracy (%)	Sensitivity (%)		Specificity (%)	
77.78	FR	0.00	FR	82.35
	FS	40.00	FS	80.00
	FT	100.00	FT	76.47
	GF	100.00	GF	73.33
	MG	40.00	MG	80.00
	MR	20.00	MR	81.18
	MS	100.00	MS	74.36
	MT	50.00	MT	79.76
	VG	100.00	VG	75.31
	VR	100.00	VR	74.03
	VS	50.00	VS	79.07
	VT	100.00	VT	76.19
	Average	66.67	Average	77.67

The results of testing process showed that the average of sensitivity, specificity and the accuracy were 66.67%, 77.67% and 77.78% respectively.

Compared to the previous research [2], this research used color and texture features. While the previous research [2] features used, skewness, kurtosis, standard deviations, and energy as features. Both used learning vector quantification. LVQ network was used for identification classification of malaria parasite.

In this research, 600 data of malaria parasite images were used for training, 510 data and testing, 90 data. The results of all six features used LVQ and selection of malaria parasites feature can be seen in Table 6.

Table 6: LVQ parameter for training of malaria parasite feature with feature extraction

Feature	Method of identification	Data	Training accuracy value (%)	Validation accuracy value (%)	Calibration accuracy value (%)
Colour and texture (phase of image, mean of green plane, skewness, kurtosis, standard deviation and energy)	LVQ	Feature of Malaria parasite	100 in all *k-fold*	44.30 in *5-fold*	58.89%
Texture and shape (area, perimeter, energy 0^0, energy 90^0, contrast 0^0	LVQ	Feature of Malaria parasite	100 in all *k-fold*	79.22 in 3-fold	77.78

According to these results, it can be stated that texture and shape features that consisted of area, perimeter, energy 0^0, energy 90^0 and contrast 0^0 are appropriate for input identification using LVQ. It results in 77.78% accuracy. It is better result compared with previous research that gave 58.89% accuracy.

5. CONCLUSIONS

A method for identification of malaria disease and its stadium was presented. The test results showed that the accuracy of the system to identify the incidence of malaria and the stage reached 74.44% without feature selection. While the feature selection produces 77.78% of accuracy. These results indicate there is an increase of 18.89%, when compared with the results of previous studies of accuracy that only 58.89% by using the mean of the green plane, skewness, kurtosis, standard deviation and energy. LVQ and K-Fold cross validation, can be used to classify the features of the malaria parasite to identify the type of malaria in the process of testing the accuracy value 77.78%, 66.67% and a sensitivity value a specificity value of 77.67%.

This research has not been considered the determination of region of interest automatically so it would be included for the future research.

ACKNOWLEDGEMENT

The authors would like to thank Yayasan Potensi Utama for funding this research, also to Parasitology Health Laboratory of the North Sumatera Province and Bina Medical Support Services (BPPM), Jakarta, for supporting the required data for this research.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES

[1] Ajala, F.A., Fenwa, O.D. and Aku, M.A., Comparative Analysis of Different Types of Malaria Diseases using First Order Features, International Journal of Applied Information Systems (IJAIS), Volume 8, No.3, New York, USA, February 2015

[2] Annaldas, S. and Shirgan S.S., Automatic Diagnosis of Malaria Parasites Using Neural Network and Support Vector Machine, International Journal of Advance Foundation and Research in Computer (IJAFRC), Volume 2, January 2015

[3] Anggraini, D., Nugroho, A.S., Pratama, C., Rozi, I.E., Pragesjvara, V., Gunawan, M., "Automated Status Identification of Microscopic Images Obtained from Malaria Thin Blood Smears using Bayes Decision: A study case in Plasmodium Falciparum", *ICACSIS*, Jakarta, 17-18 Desember 2011, ISBN: 978-979-1421-11-9, page 347-352

[4] Chavan, S., Nagmode, M., "Malaria Disease Identification and Analysis Using Image Processing", *International Journal of Computing and Technology (IJCAT)*, Volume 1, July 6th 2014

[5] Das, D., Ghosh, M., Chakraborty C., Maiti, A.K. and Pal, M., "Probabilistic Prediction of Malaria using Morphological and Textural Information", *International Conference on Image Information Processing (ICIIP), IEEE,* 2011

[6] Das, D.K., Mukherjee, R., Chakraborty, C., Computational microscopic imaging for malaria parasite detection : a systematic review, *Journal of Microscopy*, Vol. 260, Issue 1, February 2015

[7] Dewi, R., "Penyakit Tropis Mengancam di Indonesia", http://kesehatan.kompasiana.com/medis/2010, accessed date July 2nd, 2014

[8] Fausett, L., 1994, Fundamentals of Neural Networks, Prentice Hall, Inc

[9] Gitonga, L., Memeu, D.M., Kaduki, K.A., Mjomba, A.C.K. and Muriuki, N.S., "Determination of Plasmodium Parasite Life Stages and Species in Images of Thin Blood Smears Using Artificial Neural Network", Open Journal of Clinic Diagnostics, Scientific Research Publishing Inc., 2014.

[10] Goswami, B., 2010, Texture Based Image Segmentation using GLCM, Dissertation, Faculty Council For UG and PG Studies In Engineering and Technology, Jadavpur University

[11] http://www.voaindonesia.com, "Upaya Pengendalian Malaria Selamatkan 3.3 juta Nyawa", accessed date November 16th, 2015

[12] Kareem, S., Kale, I., Morling, R.C.S., "Automated Malaria Parasite Detection in Thin Blood Films: A Hybrid Illumination and Color Constancy Insensitive, Morphological Approach", IEEE Asia Pasific Conference On, Kaohsiung, 2nd - 5th of Desember 2012

[13] Komagal, E., Kumar, K.S., Vigneswaran, A., Recognition and Classification Of Malaria Plasmodium Diagnosis, International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 1, 2013

[14] Liwan, A.S., " Diagnosis dan Penatalaksanaan Malaria Tanpa Komplikasi pada Anak", Cermin Dunia Kedokteran, Volume 42, No.1, 2015.

[15] Makkapati, V.V. and Rao, R.M., "Ontology-based Malaria Parasite Stage and Species Identification from Peripheral Blood Smear Images", International Conference of the IEEE, Boston, 30th of August – 3rd of September 2011

[16] May, Z., Aziz, SSAM, Salamat, R., "Automated Quantification and Classification of Malaria Parasites in Thin Blood Smears", International Conference on Signal and Image Processing *Application (ICSIPA)*, 2013

[17] Memeu, D.M., Kaduki, K.A., Mjomba, A.C.K., Muriuki, N.S. and Gitonga, L., "Detection of Plasmodium Parasites from Images of Thin Blood Smears", *Open Journal of Clinic Diagnostics*, 2013

[18] Otsu, N., 1979, A threshold selection method from gray-level histogram, *IEEE Trans. Syst. Man Cybernet.*, Vol. 9, No. 1, pp. 62–66, Jan. 1979

[19] Refaeilzadeh, P., Tang, L. and Liu, H., "Cross Validation", *Encyclopedia of Biometric*, pp 206-206, 2008

[20] Ritter, N. and Copper, J., "Segmentation and border identification of cells in image of peripheral blood smear slides", Proceeding of Thirtieth Autralasian Computer Science Conference, pp. 161-169, 2007.

[21] Savkare, S.S. and Narote, S.P., Automated System for Malaria Parasite Identification, International Conference on Communication, Information & Computing Technology (ICCICT), India, 2015.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[22] Soni, J., "Advanced Image Analysis based system for automatic detection of malarial Parasite in Blood Images Using Susan Approach", International Journal of Engineering Science and Technology (IJEST), Volume 3 No. 6, June 2011.

[23] Suryawanshi, S. and Dixit V.V., "Comparative Study of Malaria Parasite Detection using Euclidean Distance Classifier & SVM", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 2 Issue 11, November 2013.

[24] Suwalka, I., Sanadhya, A., Mathur, A. and Chouhan, M.S., "Identify Malaria Parasite Using Pattern Recognition Technique", ICCCA, Dindigul, 22nd - 24th of February 2012.

[25] Thung, F. and Suwardi, I.S., "Blood Parasite Identification using Feature Based Recognition", International Conference on Electrical Engineering and Informatics, Bandung, 17th - 19th of July 2011.

[26] Widodo, S. and Wijayanto, Texture Analysis To Detect Malaria Tropica In Blood Smears Image Using Support Vector Machine, International Journal of Innovative Research in Advanced Engineering (IJIRAE), Volume 1 Issue 8, September 2014.

[27] Witoelar, A., Biehl, M. and Hammer, B., "Learning Vector Quantization: Generalization Ability and Dynamics of Competing Prototypes", Dagstuhl Proceedings, Schloss Dagstuhl - Leibniz - Zentrumfür Informatik, 2007.

ORIGINALITY REPORT

5%	5%	1%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.jatit.org Internet Source	4%
2	publikasi.dinus.ac.id Internet Source	1%
3	S.S. Savkare, S.P. Narote. "Automated system for malaria parasite identification", 2015 International Conference on Communication, Information & Computing Technology (ICCICT), 2015 Publication	1%

Exclude quotes	On	Exclude matches	< 25 words
Exclude bibliography	On		