



# Comparison of Patterns Shapes and Patterns Texture for Identification of Malaria Parasites in Microscopic Image

Jani Kusanti<sup>1</sup>, Yusuf Zain Santosa<sup>2</sup>

<sup>1</sup> Department of Electrical Engineering and Information Technology, Universitas Surakarta

<sup>2</sup> Department of Computer Science and Electronic, Universitas Gadjah Mada

Email: <sup>1</sup>jani\_kusanti@yahoo.com, <sup>2</sup>yusuf\_zain@yahoo.com

## Abstract

Identification of malaria parasites in red blood cells has been done, with the aim of as tools to identify experts microscopic parasites more quickly. This study aimed to compare the level of accuracy in the results to identify and classify parasites based on the pattern shape and texture patterns. The comparison is based on the characteristics of the pattern used, the steps being taken in this study is the image quality improvement process, the process of segmentation with Otsu method, feature extraction process on the image data to be tested. The process of pattern recognition and pattern shapes texture. The last step is to test the identification and classification of plasmodium falciparum parasite into 12 classes using methods Learning Vector Quantization (LVQ). The results of this study indicate that the pattern forms can provide a higher level of accuracy compared to LVQ texture pattern. LVQ with input shape pattern successfully identified 91% of image data correctly and input texture successfully identified 48% of image data properly.

**Keyword:** Malaria parasites, Classification, LVQ Method, Pattern recognition, Pattern shape texture

## 1. INTRODUCTION

Accurate diagnosis and timely mostly done with the aim that the developed system can be an invaluable tool in resolving problems in the field. Any system that is developed each using a different method. Issues raised in this research is how the accuracy of the resulting image using pattern recognition pattern shape and pattern recognition shape texture. Utilization of image processing feature extraction pattern shape and texture patterns to identify malaria parasites has been carried out [1-4]. The study, titled "*Probabilistic Prediction of Malaria using Morphological and Textural Information*" [1]. In the study to identify malaria parasites use of pattern feature extraction. The study, titled "*Automatic Diagnosis of Malaria Parasites Using Neural Network and Support Vector Machine*" [2]. In the study to identify malaria parasites use also feature extraction, feature extraction pattern texture that is second order. The introduction of characteristic use of pattern and texture patterns each have weaknesses and strengths, in this research study was to compare the accuracy of each of the features. This study uses the classification stage of *P. falciparum* malaria, *P. malaria*, *P. vivax* input is the image of a microscope examination. The image is processed using the method of digital image processing with multiple stages including preprocessing, segmentation, morphology, and feature extraction. In the classification

stage (training and testing) using LVQ. After the classification, it will get the level of accuracy that is generated by the system. As a solution to the above problems, the researchers devised a system to compare the accuracy results with the title "Comparison of Patterns Shapes and Patterns Texture for Identification of Malaria Parasites in Blood".

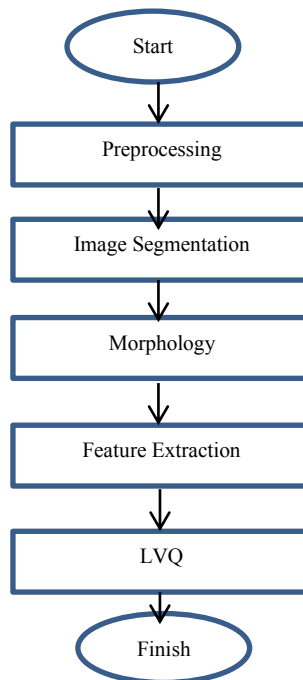
## 2. METHOD

### 2.1. Malaria

Malaria caused by plasmodium parasite. There are four species of the parasite that causes malaria in human plasmodium vivax, plasmodium malaria, plasmodium ovale and plasmodium falciparum. Each plasmodium has four class, differentiated by the shape or particular characteristics. In this study used three types of plasmodium, with each divided into four types of plasmodium. To the researchers grouped into twelve classes (p. falciparum ring (PFR), p. falciparum schizonts (PFS), p. falciparum trophozoites (PFT), p. falciparum gametocytes (PGF), p. malaria gametocytes (PMG), p. malaria schizonts (PMS), p. malaria trophozoites (PMT), p. malaria ring (PMR), P. vivax gametocytes (PVG), P. vivax ring (PVR), P. vivax schizonts (PVS), P. vivax trophozoites (PVT)). Training data is used as the data 320, the test data used as many as 120 data.

### 2.2. Research Method

A system built on the research carried out in five stages, shown in Figure 1,



**Figure 1.** The process of identification systems malaria-causing parasite

### **Preprocessing**

In this process carried out several stages, the initial stages carried Region of Interest (ROI), which limit only the desired area. The process to get ROI one of which is by way of cropping an image. Step process is then performed using the contrast setting contrast stretching.

### **Segmentation**

Segmentation is a technique for dividing an image into multiple areas or regions where each region has similar attributes. In this research methods used include threshold segmentation. In this process required a limiting value called the threshold value. Image intensity value greater than or equal to the threshold value will be changed to white or 1 and image intensity values less than the threshold value will be changed to black or 0. So the output of the threshold is in the form of a binary image. After this is done the next step threshold edge detection process is carried out by using canny edge detection. Edge detection operation is done through a process of convolution matrix using a kernel matrix.

### **Morphological**

Morphology is the object structure of the image associated with the form. Morphological operations used in this research is closing. Closing operation is useful for smoothing the contours and eliminating small holes. Closing operations carried out by performing dilation operations first and then followed by erosion operations.

### **Feature extraction**

In this study, the decision is based on the characteristics of texture and shape feature extraction. Texture feature extraction using texture features include a mean contrast, the average correlation, mean energy, mean homogeneity.

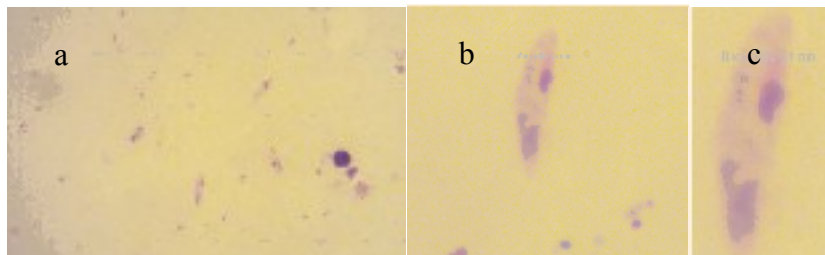
### **LVQ**

Learning Vector Quantization (LVQ) is a method in classification patterns supervised (supervised). The input vectors will be grouped in the same class.

## **3. RESULTS AND DISCUSSION**

### **3.1. Preprocessing**

The imagery used is a color image, after going through the process of cutting the desired area. Figure 2 shows the results of the ROI image.



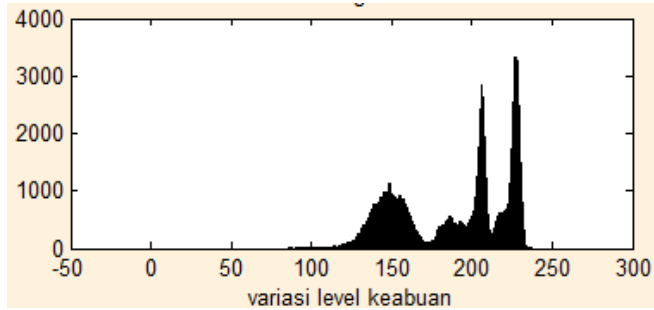
**Figure 2.** ROI image

(a. The original image microscopic examination; b. ROI results from a picture a; c. ROI results from image b)

ROI results using the processed image contrast stretching with the aim of improving the image quality. The values of gray pixels will reach from 0 to 255 (in the image 8bit), in other words all the values of gray pixels used evenly, as indicated by equation 1 [5]

$$h_k = \frac{n_k}{n}, k = 0, 1, \dots, L - 1 \tag{1}$$

Equation 1 states the relative frequency of occurrence of the degrees of gray in the image.



**Figure 3.** Contrast stretching the image histogram results

Figure 3 shows the image histogram shows a lot about the brightness (brightness) and contrast (contrast) of an image. The value of the image pixel count cumulative value calculated based on all pixels of gray ( $n_i$ ), in this process the intensity value of 255, as a result of the calculation of the formula  $2^k-1$  ( $k$  is the value that is 8 bit grayscale image so that the resulting value of 255).

$$\begin{aligned} \text{Level of gray } i &= 0 - 255 \\ n_i &= n_{(0-255)} \\ \sum n_i &= \text{total the results of this multiplication by } n_i \end{aligned}$$

Long  $i$ , divided  $n$ , multiplied by  $\sum n_i$ , the result is rounded, image processing results are shown in Figure 4.



**Figure 4.** Image results contrast stretching

### 3.2. Segmentation

Thresholding techniques or a threshold value is the focal point in the application segment that aims to separate certain parts with other parts [4]. The process begins by calculating the area of the image from the contrast stretching, the process of calculating the average value of the pixel. The pixel values are calculated per line to a

height of the image, starting from the initial pixel until the last pixel, the values obtained are accumulated. Furthermore, the pixel value is calculated for each column as wide as the image, the values obtained will be accumulated. This process obtained from the threshold value or threshold. Threshold value =  $j$ ,  $j$  is within range  $0 \leq j \leq m-1$ .  $m$  = degrees of gray of the image,  $\mu$  = the number of pixels with the average value of the degrees of gray.

Its function refers to the equation (2) - (8). [6]

$$\eta(k) = \sigma_B^2(k) / \sigma_r^2 \quad (2)$$

$$\sigma_B^2(k) = \frac{[\mu_\tau \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \quad (3)$$

$$\sigma_r^2 = \sum_{i=1}^L (i - \mu_\tau)^2 P_i \quad (4)$$

$$\mu_\tau = \mu(L) = \sum_{i=1}^k iP_i \quad (5)$$

$$\mu(k) = \sum_{i=1}^k iP_i \quad (6)$$

$$\omega(k) = \sum_{i=1}^k P_i \quad (7)$$

$$P_i = n_i / N, P_i \geq 0, \sum_{i=1}^L P_i = 1 \quad (8)$$

The pixels in the image represented in degrees of gray  $L [1, 2, \dots, L]$ . The number of pixels with the degree of gray  $i$  denoted by  $n_i$  and the total number of pixels with  $N = n_1 + n_2 + \dots + n_i$ .  $P_i$  is a representation of a histogram,  $k$  is the threshold value. Distance variant is denoted by  $\sigma_B^2(k)$ . Range for  $x$  from 0 to  $L-1$ , count each  $\sigma_B^2(k)$  and the value of  $k$  refers to the value of the largest distance variant which is the result of the threshold  $T$ . The result of the segmentation is shown in Figure 5.

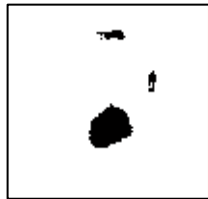


Figure 5. Segmentation result

### 3.3. Morphological

Morphological processes carried out to improve the shape of the object in order to produce features more clearly, so that when carried out measurements on the image will look the area being measured. Morphological processes used by many researchers to impose limits on the image to be measured. Closing process is the process of dilation by adding pixels to the boundary edges of objects, followed by a process of erosion that is part of an object that has been restricted by the pixels transferred on the object boundary. So we get the image that has been separated between the foreground and background. Example of images with a closing process shown in Figure 6.



Figure 6. Closing result

### 3.4. Feature Extraction

Feature extraction using texture patterns co-occurrence matrix, the results obtained by calculating the probability of adjacency relationship between two pixels within a certain angle and orientation. [6] and [7] using co-occurrence matrix with the characteristics of entropy. In this study the results of the characteristics using the average value of the four characteristics of image texture that is GLCM  $0^{\circ}$ , GLCM  $45^{\circ}$ ,  $90^{\circ}$  and GLCM  $135^{\circ}$ . GLCM co-occurrence Matrices that have been obtained transpose and get a symmetrical angle is the angle  $180^{\circ}$ . The sum of the matrix and the matrix transpose summed co-occurrence this makes the phase angle between the symmetric angle  $0^{\circ}$  and angle  $180^{\circ}$ . Further normalization where normalization is obtained by sum of all symmetric, then used as dividers for all pixels contained in the symmetric. The last stage calculation consists of contrasting texture characteristics, correlation, energy and homogeneity and form features comprising area, perimeter, metric and eccentricity. Figure 7 shows the results of the characteristic patterns of shapes and Figure 8 shows the results of the characteristics of texture patterns.

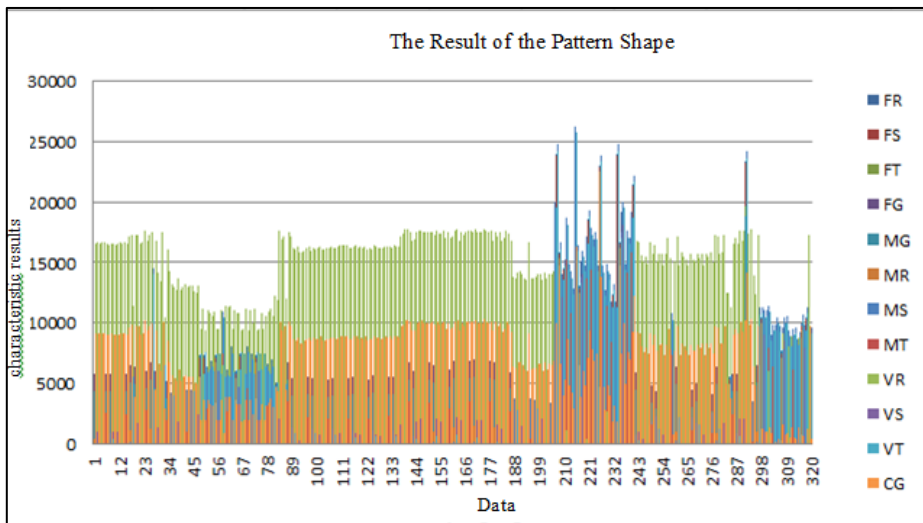


Figure 7. The results characterize the image of the pattern shape

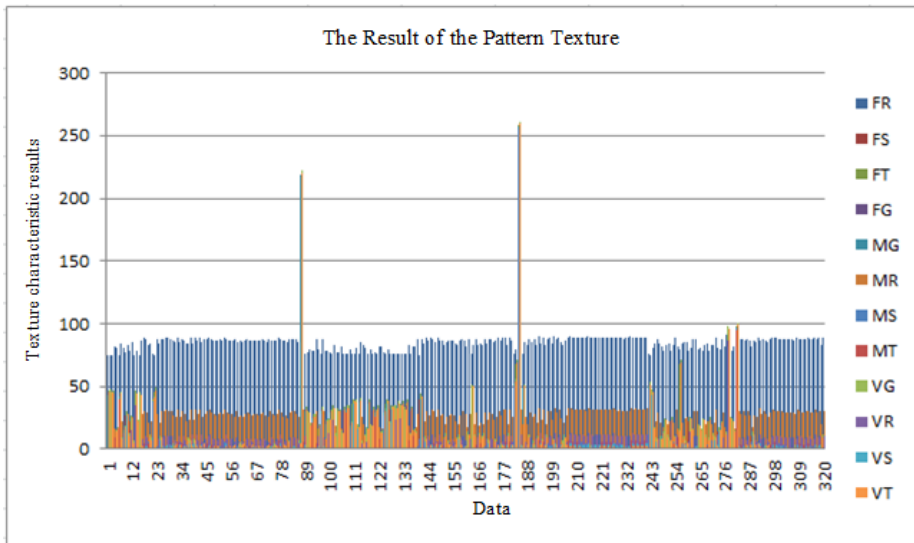


Figure 8. The Result of The Pattern Texture

### 3.5. LVQ

LVQ (Learning Vector Quantization) is a method in the Neural Network to conduct supervised learning on the layer. Pattern classification methods with each unit-output represents a particular class or a particular category. Used reference vector (Vector Reference / codebook). Weight vector of an output unit that became the benchmark for the class / category is represented by the output. The approach taken is to classify the input vector based on the proximity of the weighted input vector (minimum Euclidean distance squared method). Figure 9 shows the architecture LVQ

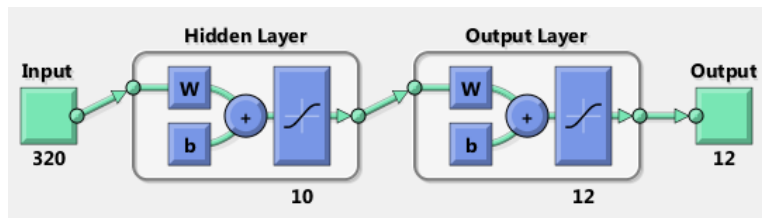
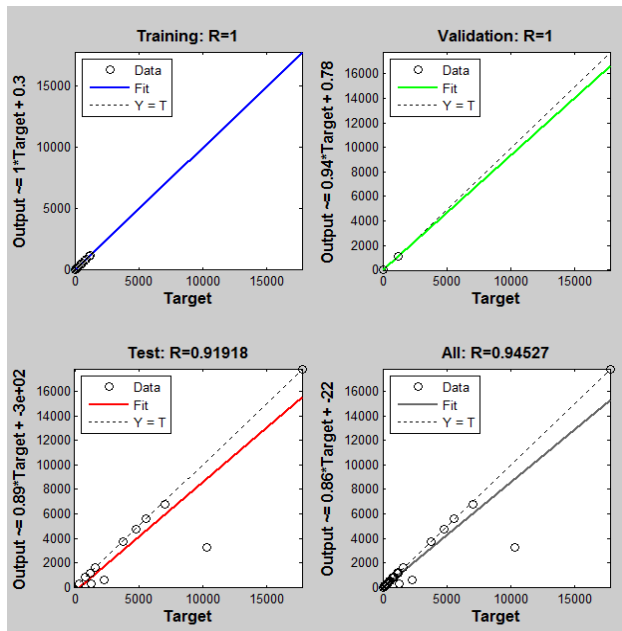


Figure 9. Architecture LVQ

Train results obtained are shown in Figure 10.



**Figure 10.** Test results with the pattern shape

From the test results by using the characteristic pattern generated form accuracy rate of 91% for the Test and 94% for Target. As for the test using a texture pattern characteristic of the results shown in Figure 11 with an accuracy rate of 49% for the Test and 55% for Target.



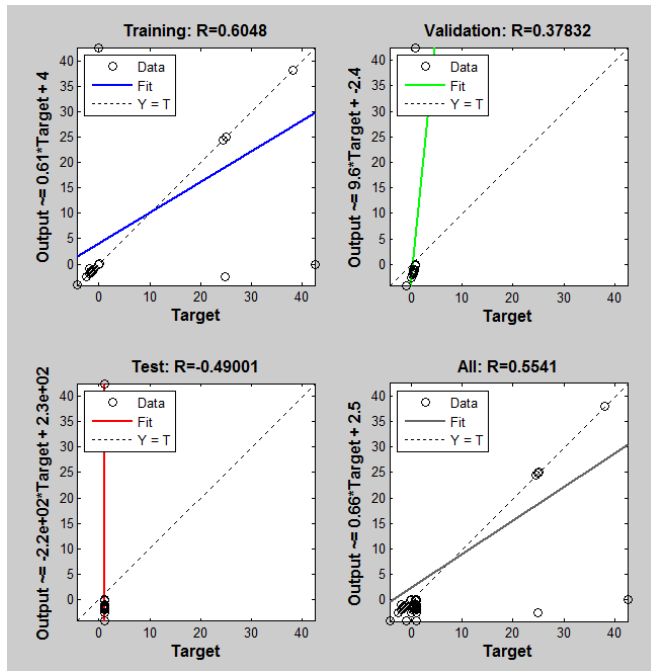


Figure 11. Test results with texture pattern

#### 4. CONCLUSION

The results of testing the use of pattern with the characteristic area and perimeter as well as the use of texture pattern with the characteristic energy influential to identify image data identified according to the classification of the malaria parasite malaria parasite class. Patterns in the form of image data identifying the malaria parasite is more dominant than the texture patterns. Selection of characteristics to identify image data of malaria parasites is important to help improve time efficiency in the identification process image data, but in this study remains to be developed further by adding another pattern characteristics in pattern recognition, such as geometric patterns and color patterns.

#### 5. REFERENCE

- [1] Das D. and Maiti, A., K., 2011. Probabilistic Prediction of Malaria using Morphological and Textural Information, *International Conference on Image Information Processing (ICIIP 2011)*
- [2] Annaldas, S., Shirgan, S., S., 2015., Automatic Diagnosis of Malaria Parasites Using Neural Network and Support Vector Machine, *International Journal of Advance Foundation and Research in Computer (IJAFRC) Volume 2, Special Issue (NCRTIT 2015), ISSN 2348 - 4853*

- [3] Widodo, S. and Wijiyanto, 2014, Texture Analysis to Detect Malaria Tropica in Blood Smears Image using Support Vector Machine, *International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163 Volume 1 Issue 8 (September 2014)*
- [4] Soleman, M., Ruliah., 2011. Identifikasi Parasit Malaria Plasmodium Falciparum Pada Sediaan Darah dengan Pendekatan Support Vector Machine. *Jurnal VISIKES. Vol.10/No2/September 2011: 89-99.*
- [5] Gonzalez, R.C. and Woods, R.E., 2008, Digital Image Processing Third Edition, *New Jersey 07458: Prentice Hall.*
- [6] Prasetyo, E., 2011. *Pengolahan Citra Digital dan Aplikasinya menggunakan Matlab.* Penerbit ANDI
- [7] Kusumadewi, S., 2004. *Membangun Jaringan Syaraf Tiruan (Menggunakan Matlab dan Excel Link).* Yogyakarta: Graha Ilmu