Journal of Applied Intelligent System (e-ISSN : 2502-9401 | p-ISSN : 2503-0493)
Vol. 2 No. , April 2017, pp. 12 – 20

Plasmodium Falciparum Identification in Thick Blood Preparations Using GLCM and Support Vector Machine (SVM)

Farah Zakiyah Rahmanti*[1], Novita Kurnia Ningrum[2]
Universitas Dian Nuswantoro, Imam Bonjol 207 Semarang, (024) 3517261
E-mail : farah_zakiyah@dsn.dinus.ac.id[1], novita.kn@dsn.dinus.ac.id[2]*
**Corresponding author*

Septian Enggar Sukmana[3], Prajanto Wahyu Adi[4]
Universitas Dian Nuswantoro, Imam Bonjol 207 Semarang, (024) 3517261
E-mail : septian.enggar@dsn.dinus.ac.id[3], prajanto@dsn.dinus.ac.id[4]

Abstract - Malaria is one of the serious diseases that require rapid handling, otherwise, it can lead to death. One of the causes of malaria parasites is plasmodium falciparum which can cause severe or fatal malaria. Handling a medical late can increase the risk of death. Therefore, it takes a rapid identification system with a high percentage of accuracy to reduce the risk of death. This research aims to build an identification system of plasmodium falciparum in thick blood film using Gray Level Co-occurrence Matrix (GLCM) and Support Vector Machine (SVM). The GLCM is used to get texture feature values such as contrast, correlations, energy, and homogeneity from images. Those values are processed and as an input of classification using SVM. The research result using SVM for accuracy value of plasmodium falciparum identification can reach 93.33%.

Keywords - malaria parasite, plasmodium falciparum, feature extraction, support vector machine (SVM)

1. INTRODUCTION

Malaria is a serious disease problem that requires immediate attention because this disease can cause death. Plasmodium Falciparum is one of malaria parasite which can cause tropical malaria with fatal brain malaria. It attacks symptoms every two days (48 hours) once [1]. Plasmodium falciparum can be found in Indonesia and with this high mortality rate, malaria needs to be treated as quick as possible. Therefore, the accurate and timely diagnosis of malaria infection is essential to control and cure the disease.

Conventional methods of identification malaria parasite are generally carried out by paramedics when they thoroughly examine blood performed using a microscope [8]. This conventional way as known as a microscopic observation. The diagnosis of malaria can be done if found malaria parasites in the blood of patients. The result obtained in each view overpass later in the analysis, so that would be obtained information on whether or not malaria parasites in the blood of patients. However, conventional methods can make a difference if the diagnosis is made by different experts. The detection of this disease is time-consuming and subjective factors are very high [5].

It has been a lot of research which is done with regard to malaria, the one with the title "LVQ (Learning Vector Quantization) Method for Identification of Plasmodium Vivax In Thick Blood Film", explains that the value of accuracy obtained amounted to 93% for not parasites, 84% for vivax trophozoite, 74% for vivax schizont, and 84.50% for vivax gametocyte. These results are pretty good in identifying plasmodium vivax [3].

The other research is about "Plasmodium Vivax Classification from Digitalization Microscopic Thick Blood Film Using Combination of Second-Order Statistical Feature Extraction and K-Nearest Neighbor (K-NN) Classifier Method". The result of research produces a high value enough accuracy to achieve 95% [4].

Research this time is unlike with previous studies. Because this research is using images of thick blood preparations which contain the malaria parasites, in this case, is plasmodium falciparum. Moreover, this research is using second order statistics approach with Gray Level Co-occurrence Matrix GLCM for features extraction and Support Vector Machine (SVM) for its classification.

2. RESEARCH METHOD

There are four processes in this research, they are taking data, image preprocessing features extraction, and classification.

First, Taking data is in Dinas Kesehatan Propinsi Jawa Timur (East Java Health Department). The second process is preprocessing, it is used for processing of the digitalization microscopic thick blood film before feature extraction stage. It aims to get ROI (Region of Interest) by cropping manually and resizing images. The third process features extraction, it uses Gray Level Co-occurrence Matrix (GLCM) to get texture features values such as contracts, correlations, energy, and homogeneity. Fourth, the classification process is using Support Vector Machine (SVM) with linear kernel function. The steps in this result will be defined in block diagram in Figure 1.

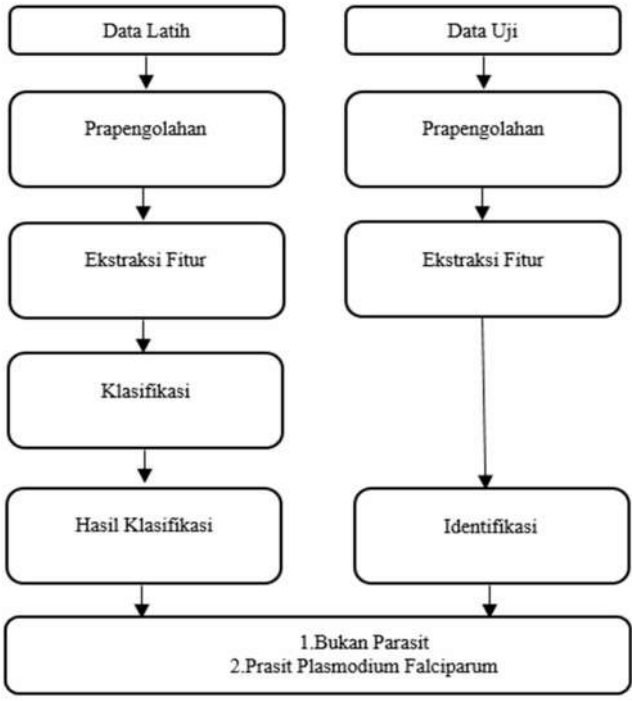


Figure 1. Block Diagram

2.1. Image Data

To acquire the data, a digital single-lens reflex (DSLR) camera with 1000x magnification is connected to the microscope. It takes some redundant processes to capture the object of malaria. During capturing process, we are accompanied by experts and we also confirmed the result of image preprocessing to experts from East Java Health Department.

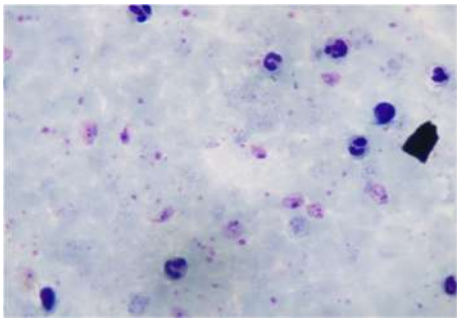


Figure 2. Sample View of Capturing Malaria Parasites in Microscope

2.2. Image Preprocessing

This stage results are getting the region of interest (ROI) by cropping each object area manually, then the resizing images are performed to get an image which has 64 x 64 pixels wide image area such as Figure 3. In this activity, there are 90 images of malaria parasites object which consist of 60 training images and 30 testing images.

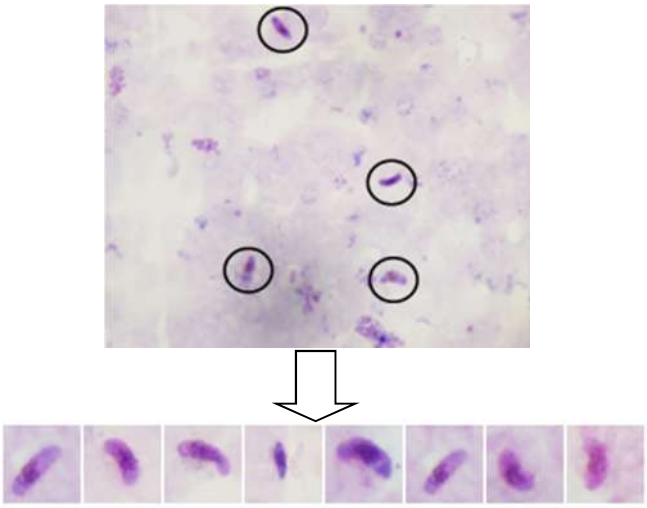


Figure 3. Result of Preprocessing Step

2.3. Feature Extraction

Feature extraction is related to texture analysis. For example, an early step in image classification is image characteristic quantization into a specific group with a suitable value. Getting the feature, it can be acquired by using second orde statistical.

Gray Level Co-occurrence Matrix (GLCM) is a technique to get the second rode statistical with calculating the probability of its adjacency relationship between two pixels in specific distance d and degree orientation θ. In this research, we use four-degree directions with degrees orientation interval 45◦. Therefore we use 0◦, 45◦, 90◦, and 135◦.

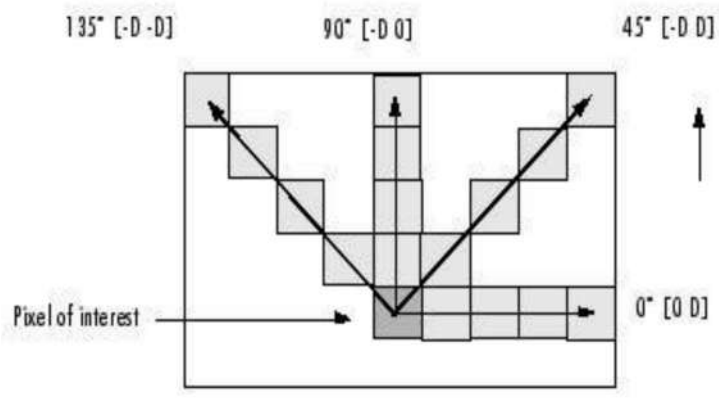


Figure 4. GLCM and Four Angle Directions with 45 Degree Interval [6]

The feature extraction stage uses Gray Level Co-occurrence Matrix (GLCM) to get texture feature values such as contracts, correlations, energy, and homogeneity.

1. Contrast (CON): this feature refers to the amount of color or grayscale differentiation that appears inside the malaria parasite image. Contrast value will be 0 if the pixel adjacency has the same value [6].

$$CON = \sum_i \sum_j |(i,j)|^2 \, p(i,j) \quad (1)$$

2. Correlations (COR): this feature shows the linear dependence of the degree of a gray image, which can indicate the presence of image linear structure [6].

$$COR = \sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma i \, \sigma j} \quad (2)$$

where
μi is the average value of line elements matrix p(i,j) [7].
The equation of μi can be viewed below:

$$\mu i = \sum_i \sum_j i \, p(i,j) \quad (3)$$

μj is the average value of column elements matrix p(i,j) [7].
The equation of μj can be viewed below:

$$\mu j = \sum_i \sum_j j \, p(i,j) \quad (4)$$

σi is standard deviation line elements in matrix p(i,j) [7].
The equation of σi can be viewed below:

$$\sigma i = \sqrt{\sum_{i,j} (i-\mu i)^2 \, p(i,j)} \quad (5)$$

σj is standard deviation column elements in matrix p(i,j) [7].
The equation of σj can be viewed below:

$$\sigma j = \sqrt{\sum_{i,j}(j-\mu i)^2}\, p(i,j) \quad (6)$$

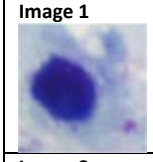
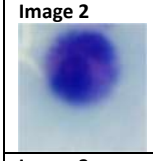
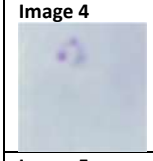
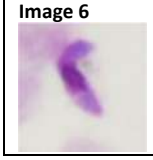
3. Energy (ENG): this feature indicates the size of uniformity properties image. This feature will get a high value when the pixel values are similar to each other.

$$ENG = \sum_{i,j} p(i,j)^2 \quad (7)$$

4. Homogeneity (HOM): this feature is measuring homogeneity of the image. Homogeneity will get high value when all the pixels have a uniform value.

$$HOM = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \quad (8)$$

Table 1. Features Extraction Result

Data	Name File	CON	COR	ENG	HOM
Image 1 	104.jpg	0.1305	0.9715	0.2474	0.9348
Image 2 	116.jpg	0.1034	0.9775	0.2938	0.9438
Image 3 	502.jpg	0.0365	0.8373	0.7696	0.9818
Image 4 	501.jpg	0.0734	0.8362	0.4936	0.9633
Image 5 	605.jpg	0.1086	0.9170	0.3666	0.9457
Image 6 	619.jpg	0.0667	0.9532	0.5091	0.9676

2.4. Classification

There are many machine learning techniques to solve linear separation problem such as Support Vector Machine (SVM). This method computes and finds optimum function separation (the best hyperplane) that separates two classes in space input, this description can be seen in Figure 5. In this research, we use SVM with linear kernel function which features extraction value is described as input in the classification process.

Generally, kernel functions that commonly used are linear, polynomial, Radial Basis Function (RBF), and tangent hyperbolic sigmoid. This kernel function is used to determine feature surface, so the linear function can be known. Next, SVM can be performed although classification map still not determine to what used for.

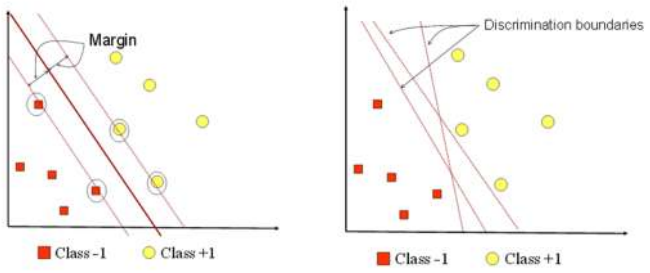


Figure 5. SVM Finds The Best Hyperplane to Separate Class -1 And Class +1 [2]

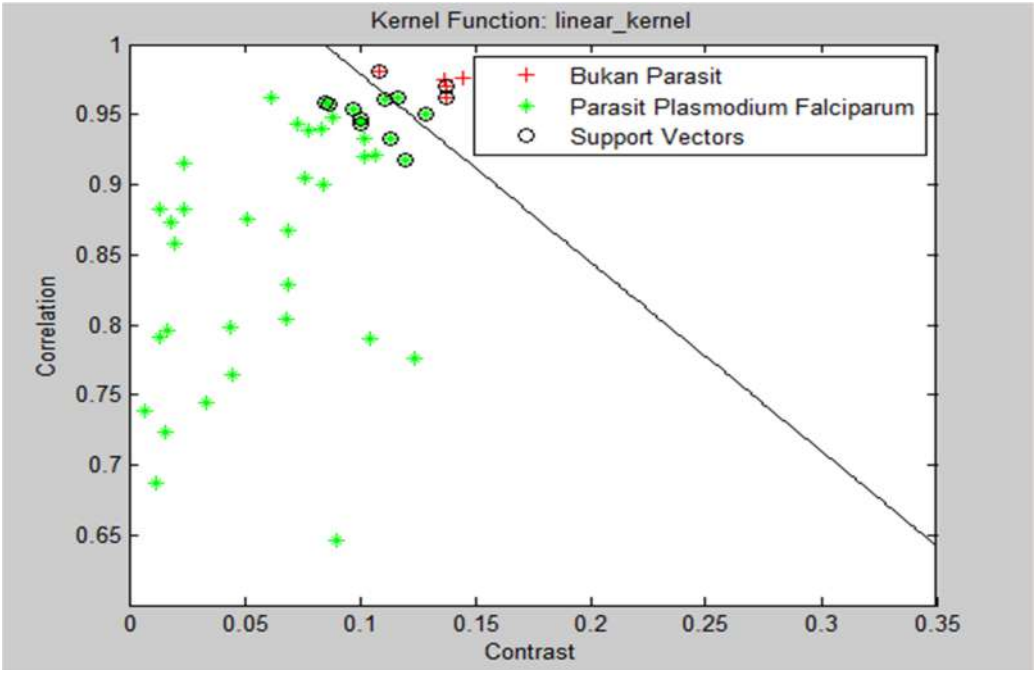


Figure 6. SVM Training Result With Contrast And Correlation Value

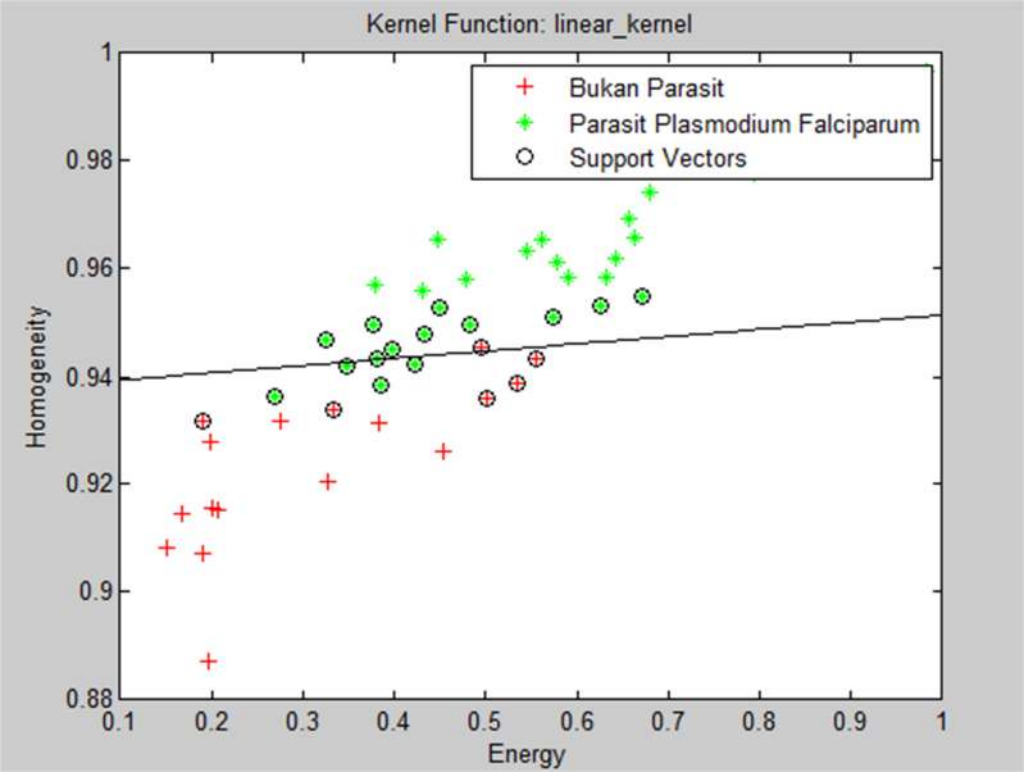


Figure 7. SVM Training Result With Energy And Homogeneity Value

SVM training result uses linear kernel function with contrast and the correlation value in range 0 – 1 as in Figure 6. SVM training result uses linear kernel function with energy and homogeneity value also has range 0-1 as in Figure 7.

3. RESULTS AND DISCUSSION

The experiment and testing are performed using 60 training data and 30 testing data. Based on Figure 6, training process is run using SVM linear kernel function to contrast and correlation values. The result shows in Tabel 2 with accuracy value 93.33% as in Tabel 4.

Based on Figure 7, training process is run using SVM linear kernel function to energy and homogeneity values. The result shows in Tabel 3 with accuracy value 80% as in Tabel 4.

Table 2. Testing Result 1

No	Name File	Class	Result
1	1.jpg	Not Parasite	Not Parasite
2	2.jpg	Not Parasite	Not Parasite
3	3.jpg	Not Parasite	Not Parasite
4	4.jpg	Not Parasite	Not Parasite
5	5.jpg	Not Parasite	Not Parasite
6	6.jpg	Not Parasite	Not Parasite
7	7.jpg	Not Parasite	Not Parasite
8	8.jpg	Not Parasite	Not Parasite
9	9.jpg	Not Parasite	Not Parasite
10	10.jpg	Not Parasite	Not Parasite
11	11.jpg	Plasmodium Falciparum	Plasmodium Falciparum
12	12.jpg	Plasmodium Falciparum	Plasmodium Falciparum
13	13.jpg	Plasmodium Falciparum	Plasmodium Falciparum

14	14.jpg	Plasmodium Falciparum	Plasmodium Falciparum
15	15.jpg	Plasmodium Falciparum	Plasmodium Falciparum
16	16.jpg	Plasmodium Falciparum	Plasmodium Falciparum
17	17.jpg	Plasmodium Falciparum	Plasmodium Falciparum
18	18.jpg	Plasmodium Falciparum	Plasmodium Falciparum
19	19.jpg	Plasmodium Falciparum	Plasmodium Falciparum
20	20.jpg	Plasmodium Falciparum	Plasmodium Falciparum
21	21.jpg	Plasmodium Falciparum	Plasmodium Falciparum
22	22.jpg	Plasmodium Falciparum	Plasmodium Falciparum
23	23.jpg	Plasmodium Falciparum	Plasmodium Falciparum
24	24.jpg	Plasmodium Falciparum	Not Parasite
25	25.jpg	Plasmodium Falciparum	Plasmodium Falciparum
26	26.jpg	Plasmodium Falciparum	Plasmodium Falciparum
27	27.jpg	Plasmodium Falciparum	Plasmodium Falciparum
28	28.jpg	Plasmodium Falciparum	Plasmodium Falciparum
29	29.jpg	Plasmodium Falciparum	Not Parasite
30	30.jpg	Plasmodium Falciparum	Plasmodium Falciparum

Table 3. Testing Result 2

No	Name File	Class	Result
1	1.jpg	Not Parasite	Not Parasite
2	2.jpg	Not Parasite	Not Parasite
3	3.jpg	Not Parasite	Not Parasite
4	4.jpg	Not Parasite	Not Parasite
5	5.jpg	Not Parasite	Not Parasite
6	6.jpg	Not Parasite	Not Parasite
7	7.jpg	Not Parasite	Not Parasite
8	8.jpg	Not Parasite	Plasmodium Falciparum
9	9.jpg	Not Parasite	Not Parasite
10	10.jpg	Not Parasite	Not Parasite
11	11.jpg	Plasmodium Falciparum	Plasmodium Falciparum
12	12.jpg	Plasmodium Falciparum	Plasmodium Falciparum
13	13.jpg	Plasmodium Falciparum	Plasmodium Falciparum
14	14.jpg	Plasmodium Falciparum	Plasmodium Falciparum
15	15.jpg	Plasmodium Falciparum	Plasmodium Falciparum
16	16.jpg	Plasmodium Falciparum	Plasmodium Falciparum
17	17.jpg	Plasmodium Falciparum	Plasmodium Falciparum
18	18.jpg	Plasmodium Falciparum	Plasmodium Falciparum
19	19.jpg	Plasmodium Falciparum	Plasmodium Falciparum
20	20.jpg	Plasmodium Falciparum	Plasmodium Falciparum
21	21.jpg	Plasmodium Falciparum	Plasmodium Falciparum
22	22.jpg	Plasmodium Falciparum	Plasmodium Falciparum
23	23.jpg	Plasmodium Falciparum	Not Parasite
24	24.jpg	Plasmodium Falciparum	Not Parasite
25	25.jpg	Plasmodium Falciparum	Plasmodium Falciparum
26	26.jpg	Plasmodium Falciparum	Not Parasite
27	27.jpg	Plasmodium Falciparum	Plasmodium Falciparum
28	28.jpg	Plasmodium Falciparum	Plasmodium Falciparum
29	29.jpg	Plasmodium Falciparum	Not Parasite
30	30.jpg	Plasmodium Falciparum	Not Parasite

Table 4. Accuracy Level of Plasmodium Falciparum Identification Using SVM

Testing Activity	Success Identified	Unsuccess Identified	Accuracy
1	28	2	93.33%
2	24	6	80%

4. CONCLUSION

Based on these activities, there is some conclusion of this research:

1. The result of testing in SVM with linear kernel function in contrast and correlation value take 93.33% accuracy, while in energy and homogeneity take 80%.
2. Better features value of SVM with linear kernel function in contrast and correlation are selected absolutely and they can be used to plasmodium falciparum identification in thick blood image.
3. Next activity would exploit segmentation technique to identify parasites automatically.

ACKNOWLEDGMENT

We are strongly thankful to Dinas Kesehatan Propinsi Jawa Timur (East Java Health Department) for their own digitalization microscopic thick blood films. We also thankful to Mr. Tjuk Agus, Mr. Kartaji, Mr. Achmad Hasan Huda, and Mr. Agung Nugroho from East Java Health Department that already accompanied and guided us.

REFERENCES

[1] Departemen Kesehatan R.I, Modul Parasitologi Malaria, Direktorat Jenderal PPM & PL Direktorat Pemberantasan Penyakit Bersumber Binatang, Diperbanyak oleh Dinas Kesehatan Propinsi Jawa Timur, Jakarta, 2011.
[2] E. Permata, I.K.E. Purnama, M.H. Purnomo, "Klasifikasi Jenis dan Fase Parasit Malaria Plasmodium Falciparum dan Plasmodium Vivax Dalam Sel Darah Merah Menggunakan Support Vector Machine", SETRUM vol.1 no.2 ISSN 2301-4652, Desember 2012.
[3] F.Z. Rahmanti, H.A. Santoso, P.W. Adi, I.K.E. Purnama, M.H. Purnomo, "LVQ (Learning Vector Quantization) Method for Identification of Plasmodium Vivax In Thick Blood Film," ICBETA, Yogyakarta, 2014.
[4] F.Z. Rahmanti, Sutojo, N.K Ningrum, N.K. Imania, M.H. Purnomo, "Plasmodium Vivax Classification from Digitalization Microscopic Thick Blood Film Using Combination of Second Order Statistical Feature Extraction and K-Nearest Neighbor (K-NN) Classifier Method," ICICI-BME, Bandung, 2015.
[5] I.K.E. Purnama, F.Z. Rahmanti, M.H. Purnomo, "Malaria Parasite Identification On Thick Blood Film Using Genetic Programming," ICICI-BME, Bandung, 2013.
[6] M. Darmawan, "Image Processing," 2011, [Online], https://martendarmawan.wordpress.com/category/imageprocessing/, tanggal akses: 17 Desember 2014.
[7] Softscients, "Algoritma GLCM – Analisis Tekstur Citra Digital," 2015, [Online], http://www.softscience.web.id/2014/04/pengolahan-citra-digital-denganc.html?m=1. tanggal akses: 10 Januari 2015.
[8] World Health Organization, "Basic Malaria Microscopy Part I Learner's Guide", Second Edition, WHO Library Cataloguing in Publication Data, Switzerland, 2010.