International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 15 (2016) pp 8754-8759

Identifying the Developmental Phase of Plasmodium Falciparum in Malaria-Infected Red Blood Cells Using Adaptive Color Segmentation And Back Propagation Neural Network

Kusworo Adi*[1], Sri Pujiyanto[2], Rahmat Gernowo[1], Adi Pamungkas[1] and Ari Bawono Putranto[1]
[1]*Department of Physics, Faculty of Science and Mathematics, Diponegoro University, Indonesia.*
[2]*Department of Biology, Faculty of Science and Mathematics, Diponegoro University, Indonesia.*
**Corresponding Author:*

Abstract
Malaria is a medical emergency that must be dealt with as it has affected millions of sufferers in 90 countries each year. Malaria is caused by a parasite that infects the red blood cells. It is spread to other people by the Anopheles mosquito. One of the plasmodium that causes malaria is plasmodium falciparum. This plasmodium causes tertiana malaria, which is the most potent malaria type that could even cause death. This research is aimed at designing a system able to identify the developmental phase of plasmodium falciparum in red blood cells using adaptive color segmentation and classify that phase using back propagation neural network. Color segmentation is made possible by converting the color space that is used to be based on the RGB (Red, Green, Blue) components into the HSV (Hue, Saturation, Value) color space. A thresholding is then conducted on the Saturation component. The morphological parameters used to differentiate the developmental phase of plasmodium falciparum are area ratio and eccentricity, whereas the back propagation neural network algorithm is used to classify that phase. The results are 87.80% accuracy during training, and 87.14% accuracy during testing.

Keywords: Malaria, Plasmodium falciparum, Adaptive color segmentation, Back propagation neural network.

INTRODUCTION

Malaria is a global phenomenon that has affected millions of patients in 90 countries around the world every year. Malaria is caused by a parasite infecting the red blood cells injected via the sting of the Anopheles mosquito [1]. Malaria could cause death especially within the high-risk group consisting of babies, toddlers, and pregnant mothers. Malaria can also directly cause anemia and reduces work productivity [2].

There are four species of plasmodium that cause malaria; P.vivax, P.ovale, P.malariae, and P.falciparum. P.falciparum and P.vivax are the most commonly found types. None the less, P.falciparum is the deadliest. P.falciparum may cause organ failure and blood abnormalities in patients. P.falciparum may also cause cerebral malaria, which is can be deadly without prompt handling. In short, malaria is a medical emergency that must be dealt with right away [3].

One of the techniques capable of analyzing red blood cell samples infected with malaria is the digital image processing. The advantages of malaria infected red blood cell samples analysis using digital image processing are ease and swiftness, compared to direct observation. Research on the application of digital image processing to analyze malaria has been conducted by many researchers across the globe [3,4,5,6,7].

The image processing methods employed in this research are characterization and classification of malaria parasites (plasmodium vivax and plasmodium falciparum). Image segmentation is carried out using the watershed method, while image classification is conducted using two algorithms, namely Bayesian learning and Support Vector Machine (SVM). Results show that the Bayesian learning algorithm has higher malaria parasite classification accuracy of 84%, compared to the SVM algorithm that has an accuracy of 83.5% [3].

This research also identifies the developmental phase of plasmodium falciparum in red blood cells using an Artificial Neural Network algorithm. Image segmentation is done by converting RGB images into grayscale images that are then undergoing the thresholding process using the Otsu method. The extracted characteristics to identify plasmodium falciparum developmental phase is the binary pattern. This plasmodium falciparum binary pattern is then used as inputs in the learning process using the back propagation neural network algorithm. The resulting accuracy of the system developed here is 87.5 % [4].

Based on that background and results of some related earlier researches, this research is aimed at developing an identification system algorithm for the developmental phase of plasmodium falciparum in red blood cells. The identified developmental phases of plasmodium falciparum are trophozoite, which is the phase at which a parasite develops, schizont, which is the phase when a parasite multiplies, and gametocyte, which is the phase when a parasite forms its gender [8]. Color segmentation is made possible by converting the color spectrum that is used to be based on the RGB (Red, Green, Blue) components into the HSV (Hue, Saturation, Value) color spectrum. A thresholding is then conducted on the Saturation component. The morphological parameters used to differentiate the developmental phase of plasmodium falciparum are area ratio and eccentricity.

HSV (HUE, SATURATION, VALUE) COLOR SPACE

HSV is a space representing colors observed by the human eye. Hue is a color representation from different wavelengths. Saturation represents the scale of color purity, and Value or sometimes referred as brightness, represents the intensity of reflected object image taken by eyes [9]. An illustration of

color space transformation from RGB to HSV is shown in Fig. 1.

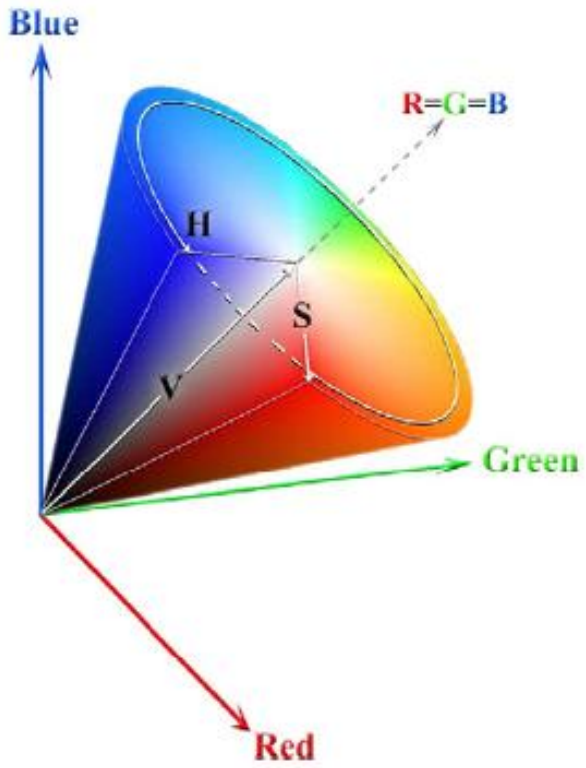


Figure 1: Transformation of color space from RGB to HSV [9].

The following equations are used to obtain the H, S, V values of the corresponding R, G, and B, scales [10]:

$$r=\frac{R}{(R+G+B)}, g=\frac{G}{(R+G+B)}, b=\frac{B}{(R+G+B)} \quad (1)$$

$$V=\max(r,g,b) \quad (2)$$

$$S=\begin{cases} 0, \text{if } V=0 \\ 1-\dfrac{\min(r,g,b)}{V}, \text{if } V>0 \end{cases} \quad (3)$$

$$H=\begin{cases} 0, if\ S=0 \\ \dfrac{60*(g-b)}{S*V}, if\ V=r \\ 60*\left[2+\dfrac{b-r}{S*V}\right], if\ V=g \\ 60*\left[4+\dfrac{r-g}{S*V}\right], if\ V=b \end{cases} \quad (4)$$

$$H=H+360, \text{if } H<0 \quad (5)$$

THRESHOLDING

Image thresholding becomes the focal point in image segmentation application due to its intuitive and simplistic properties in its implementation. Fig. 2 shows a histogram of image intensity f(x,y) consisting of a bright object on a dark background, which has an intensity level categorized in the dominant mode. In order to extract the object from its dark background, a threshold T that divides these modes is chosen. Any point (x,y) for which f(x,y) ≥ T is called the object point, whereas the other points are named the background point. In other words, an image given a threshold g(x,y) is defined as[11]:

$$g(x,y)=\begin{cases} 1, if\ f(x,y)\geq T \\ 0, if\ f(x,y)<T \end{cases} \quad (6)$$

The pixel given value 1 relates to the object, while the pixel given value 0 relates to the background. Threshold picking using bi-modal histogram visual analysis is depicted in Fig. 2.

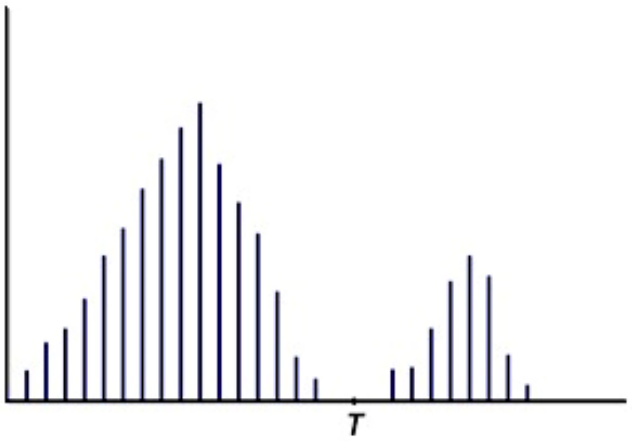


Figure 2: Threshold picking using bi-modal histogram visual analysis [11].

EXTRACTION OF MORPHOLOGICAL FEATURES

Two of the morphological features extracted from an object on an image are area ratio and eccentricity. Area ratio is the ratio between the pixels forming an object and the overall pixels of an image, while eccentricity is length comparison between the minor and major axes [12,13]. An illustration of eccentricity calculation is given in Fig. 3.

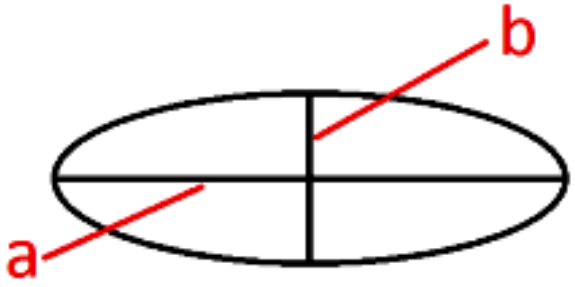


Figure 3: Illustration of eccentricity calculation.

The value of eccentricity is known using this formula [12]:

$$e=\sqrt{1-\left(\frac{b}{a}\right)^2} \quad (7)$$

where e is eccentricity value; a is the length of the major axis, and b is the length of the minor axis.

BACK PROPAGATION NEURAL NETWORK

The learning algorithm for back propagation neural network comprises of two phases. First, the input vector/pattern is given to an input layer. The network then propagates this input pattern from the input layer into the first hidden layer. This is further inputted into the following hidden layers until an output value is generated by the output layer. Second, if the output vector/pattern differs from the expected output value, errors will be calculated and re-propagated back from the output layer all the way back to the input layer. Weight is modified during this back propagation [14]. An architecture of the back propagation neural network is shown in Fig. 4.

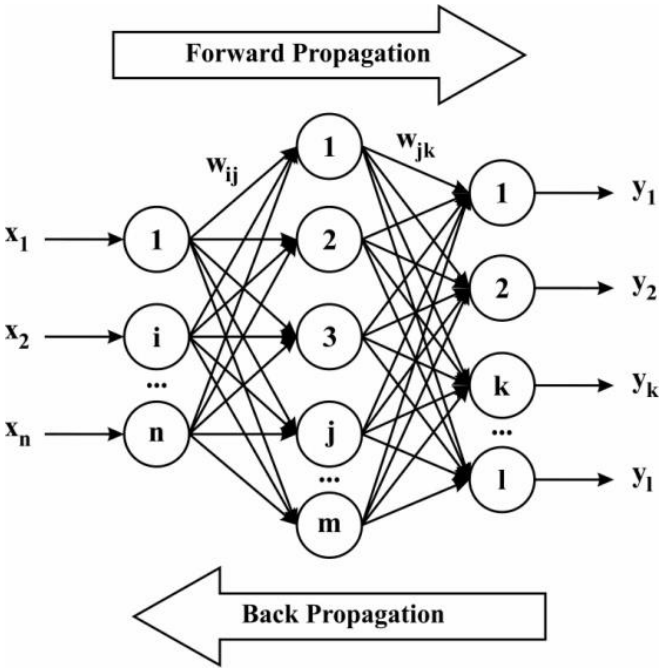


Figure 4: Back propagation neural network architecture [14].

The back propagation neural network algorithm, shown in Table 1.

Table 1: Algorithm of back propagation neural network.

Step 1: Initialization

Initialize al weights on both hidden and output layers; determine activation functions used for each layer.

Set learning rate.

Initialize all biased weights using random numbers in the range of [-0.5,0.5] or use uniform distribution for small range (Haykin, 1999):

$$\left(-\frac{2.4}{F_i}, +\frac{2.4}{F_i}\right)$$

F_i is the number of input on neuron I in the ANN.

Step 2: Activation

Activate the network by setting the inputs, $x_1(p), x_2(p), \dots, x_n(p)$, and the expected outputs, $y_{d1}(p), y_{d2}(p), \dots, y_{dn}(p)$.

a. Calculate gained output from the neurons in the hidden layer::

$$v_j(p) = \sum_{i=1}^{n} x_i(p).w_{ij}(p)$$

$$y_j(p) = \frac{1}{1 + e^{-v_j(p)}}$$

n is the number of input on neuron j in the hidden layer.

b. calculate gained output from the neurons in the hidden layer:

$$v_k(p) = \sum_{j=1}^{m} x_j(p).w_{jk}(p)$$

$$y_k(p) = \frac{1}{1 + e^{-v_k(p)}}$$

Step 3: Weight Renewal

Weight is renewed when error is reversely propagated in the ANN. Error is reversed in line with the direction of output signal.

a. Calculate error gradient for neurons in the output layer:

$e_k(p) = y_{dk}(p) - y_k(p)$

$\delta_k(p) = y_k(p) \times [1 - y_k(p)] \times e_k(p)$

Calculate weight correction:

$\Delta w_{jk}(p) = \eta \times y_j(p) \times \delta_k(p)$

Renew weight of neurons in the output layer:

$w_{jk}(p+1) = w_{jk}(p) + \Delta w_{jk}(p)$

b. Calculate error gradient for neurons in the hidden layer:

$$\delta_j(p) = y_j(p) \times \left[1 - y_j(p)\right] + \sum_{k=1}^{l} \delta_k(p).w_{jk}(p)$$

Calculate weight correction:

$\Delta w_{ij}(p) = \eta \times x_i(p) \times \delta_j(p)$

Renew weight of neurons in the hidden layer:

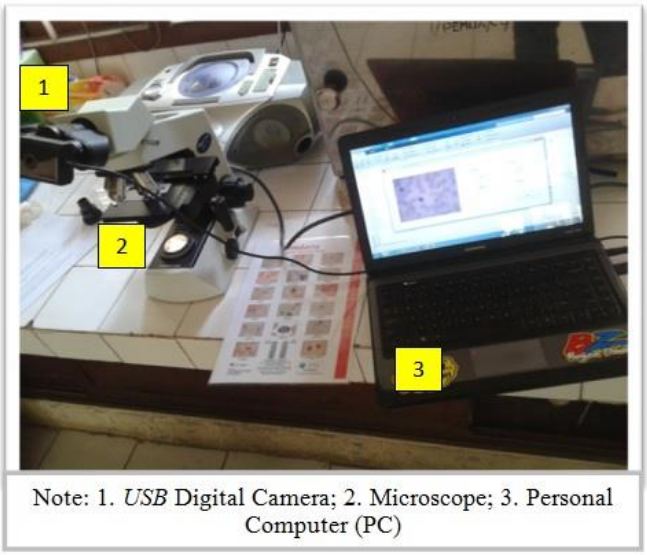
$w_{ij}(p+1) = w_{ij}(p) + \Delta w_{ij}(p)$

Step 4: Iteration

Raise one for p iteration, back to step 2, and repeat the process until the error criterion is reached.

IMAGE ACQUISITION

The procedures employed in this research include image acquisition using a microscope of 1000x magnification and a USB digital camera of 400 x 320 pixels resolution. The device set-up for image acquisition of red blood cells infected with plasmodium falciparum is described in Fig. 5.



Note: 1. *USB* Digital Camera; 2. Microscope; 3. Personal Computer (PC)

Figure 5: Set-up of the image acquisition instrument.

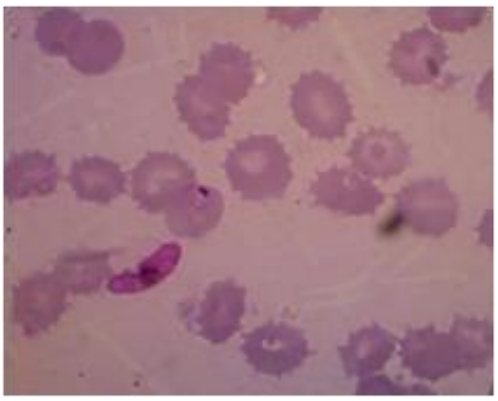
The number of images used in this research is 152 images; 82 of them for training and the other 70 for testing. Color segmentation starts with converting the color spectrum that is used to be based on the RGB (Red, Green, Blue) components into the HSV (Hue, Saturation, Value) color spectrum. A thresholding is then conducted on the Saturation component. The morphological parameters used to differentiate the developmental phase of plasmodium falciparum area ratio and eccentricity.

SYSTEM TRAINING

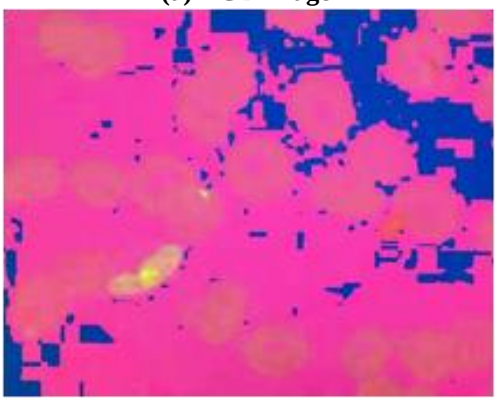
The system training process in this research comprises of three stages; adaptive color segmentation, morphological features extraction, and classification training using back propagation neural network algorithm.

Adaptive Color Segmentation

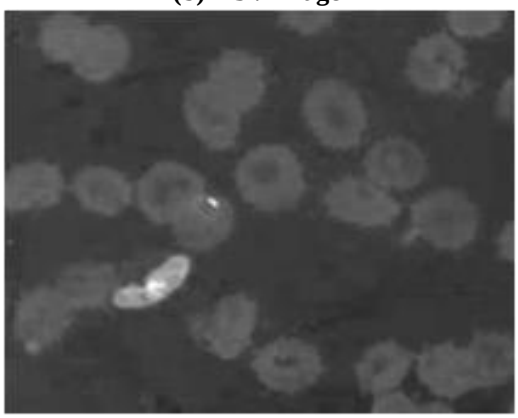
Color segmentation is carried out to separate the object (plasmodium falciparum) from the background (red blood cells). It starts with converting the components of the RGB (Red, Green, Blue) color space into the HSV (Hue, Saturation, Value) color space. Thresholding of the saturation component then ensues and results in a binary image of plasmodium falciparum. The training process of this system uses 82 image data of plasmodium falciparum consisting of 5 gametocyte, 5 schizont, 72 trophozoite phases. An example of the adaptive color segmentation system training process is depicted in Fig. 6.



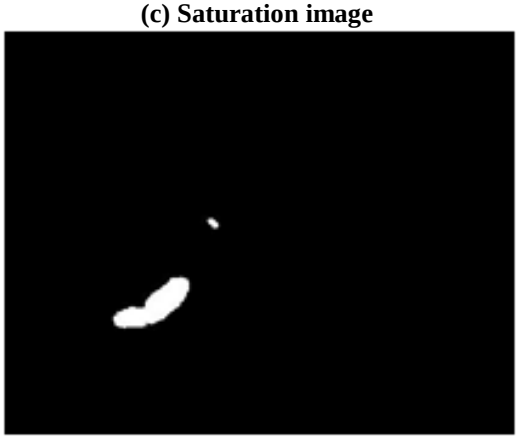
(a) RGB image



(b) HSV image



(c) Saturation image



(d) Binary image

Figure 6: Adaptive color segmentation in the system training process.

Morphological Features Extraction

Features extractions in this research are conducted based on the morphological characteristics of plasmodium falciparum in terms of area and eccentricity. Results of characteristics extraction as given in described in Fig. 4.2. (d) is given in Table 2.

Table 2: Extraction results of plasmodium falciparum morphological characteristics

No.	*Area Ratio*	*Eccentricity*
1	*2.9318e-2*	*9.4691e-1*
2	*1.1248e-4*	*8.3694e-1*

Artificial Neural Network Training

During the training process, data of extraction results are classified using the back propagation neural network algorithm. The network architecture this algorithm is given in Fig. 7.

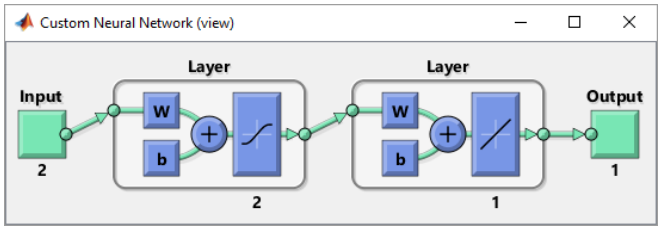


Figure 7: Architecture of back propagation neural network.

The resulting accuracy from classification training process of back propagation neural network using both area ratio and eccentricity parameters is:

$$accuracy = \frac{number\ of\ correctly\ identified\ data}{total\ number\ of\ identified\ data} \times 100\%$$

$$= \frac{72}{82} \times 100\%$$

$$= 87.80\%$$

SYSTEM TESTING

The system testing in this research employs 70 image data of plasmodium falciparum (12 gametocyte, 1schizont, and 57 trophozoite phases). These data were then extracted for their characteristics based on the parameters of area ratio and eccentricity. Both parameters serve as the input in the artificial neural network built by the system training process. The accuracy gained from this classification testing is:

$$accuracy = \frac{number\ of\ correctly\ identified\ data}{total\ number\ of\ identified\ data} \times 100\%$$

$$= \frac{61}{70} \times 100\%$$

$$= 87.14\%$$

Results of adaptive segmentation developed in this research show that the system is able to segment objects (plasmodium falciparum) and background (red blood cells) properly, even though sometimes dust, dirt, and or paint leftovers-that make objects resembling the color of plasmodium falciparum identified as plasmodium falciparum-are present. The morphological parameters of area ratio and eccentricity are essential for the identification of plasmodium falciparum in this research. Area ratio is used to differentiate the size of trophozoite that is smaller compared to the other phases of schizont and gametocyte, whereas eccentricity is utilized to differentiate the round shape of schizont and that of gametocyte, which is more oval.

CONCLUSION

This research has successfully created a system to identify the developmental phase of plasmodium falciparum in blood smears of malaria-infected red blood cells. Image acquisition process is microscopically conducted using a microscope and an USB digital camera. Color segmentation employs the adaptive color segmentation method, extraction process is based on parameters of morphological characteristics, and classification process makes use of the back propagation neural network algorithm. Accuracy of the system in identifying the developmental phase of plasmodium falciparum is 87.80%at the learning process, and 87.14% at the testing process.

ACKNOWLEDGMENT

The authors would like to thank helpful people at the Parasitology laboratory of the Faculty of Medicine, Diponegoro University, Semarang, that include among others; dr. Sri Hendratno, DAP&E, Sp.ParK; Prof. dr. Edi Dharmana, M.Sc., Ph.D., Sp.ParK; dr. Dian Puspita Dewi; and Rahmah Afrianti, A.Md. AK for many illuminating discussions on the red blood cell count aspects of this work. They also read the manuscript through and made useful suggestions and corrections.

REFERENCES

[1] Hirimutugoda, Y. M.; and Wijayarathna, G. (2010). Image Analysis System for Detection of Red Cell Disorders using Artificial Neural Networks. *Sri Lanka Journal of Bio-Medical Informatics*, 1(1), 35-42.

[2] Sio, S. W. S.; Sun, W.; Kumar, S.; Bin, W. Z.; Tan, S. S.; Ong, S. H.; Kikuchi, H.; Oshima, Y.; and Tan, K. S. W. (2007). Malaria Count: An Image Analysis-based Program for the Accurate Determination of Parasitemia. *Journal of Microbiological Methods*, 68, 11-18.

[3] Das, D. K.; Ghosh, M.; Pal, M.; Maiti, A. K.; and Chakraborty, C. (2013). Machine Learning Approach for Automated Screening of Malaria Parasite using Light Microscopic Images. *Micron*, 45, 97-106.

[4] Adi, K.; Pujiyanto, S.; Gernowo, R.; Pamungkas, A.; and Putranto, A. B. (2014). Identification of Plasmodium Falciparum Phase in Red Blood Cells using Artificial Neural Networks. *International Journal of Applied Engineering Research*, ISSN

0973-4562, 9(23), 3917-13924.

[5] Pamungkas, A.; Adi, K.; and Gernowo, R. (2015). Identification of Plasmodium Falciparum Development Phase in Malaria Infected Red Blood Cells using Adaptive Color Segmentation and Decision Tree based Classification. *International Journal of Applied Engineering Research*, ISSN 0973-4562, 10(2), 4043-4055.

[6] Savkare, S. S.; and Narote, S. P. (2011). Automatic Classification of Normal and Infected Blood Cells for Parasitemia Detection. *International Journal of Computer Science and Network Security*, 11(2).

[7] Savkare, S. S.; and Narote, S. P. (2012). Automatic System for Classification of Erythrocytes Infected with Malaria and Identification of parasite's life Stage. *Procedia Technology*, 6, 405-410.

[8] Tilley, L.; Dixon, M. W. A.; and Kirk, K. (2011). The Plasmodium Falciparum-Infected Red Blood Cell. *The International Journal of Biochemistry & Cell Biology*, 43, 839-842.

[9] Ledley, R. S.; Buas, M.; and Golab, T. J. (1990). Fundamentals of true-color image processing. *Pattern recognition: Proceedings of the 10th International Conference*, 1, 791-795.

[10] Acharya, T.; and Ray, A. K. (2005). *Image Processing Principles and Applications*. New Jersey: John Wiley & Sons, Inc.

[11] Gonzalez, R. C.; and Woods, R. E. (2004). *Digital Image Processing (2nd ed)*. New Jersey: Pearson Prentice Hall.

[12] Adi, K.; Gernowo. R.; Sugiharto, A.; Firdausi, K. S.; Pamungkas, A.; and Putrato, A. B. (2013). Tuberculosis (TB) Identification in The Ziehl-Neelsen Sputum Sample in NTSC Channel and Support Vector Machine (SVM) Classification. *International Journal of Innovative Research in Science, Engineering and Technology*, 2(9).

[13] Adi, K.; Gernowo, R.; Sugiharto, A.; Pamungkas, A.; Putranto, A. B.; and Mirnasari, N. (2013). Autothresholding Segmentation for Tuberculosis Bacteria Identification in the Ziehl-Neelsen Sputum Sample. *The 7th International Conference on Information & Communication Technology and Systems (ICTS)*, Bali.

[14] Tan, P.; Steinbach, M.; and Kumar, V. (2006). *Introduction to Data Mining*. New York: Pearson Education.