■ 194

Blood image analysis to detect malaria using filtering image edges and classification

Murk Hassan Memon[1], Tariq Jamil Saifullah Khanzada*[2], Sheeraz Memon[3], Syed Raheel Hassan[4]
[1]Graduate Student, Computer Systems Engineering,
Mehran University of Engineering & Technology, Jamshoro, Pakistan
[2]Professor, Computer Systems Engineering,
Mehran University of Engineering and Technology, Jamshoro, Pakistan
[3]Associate Professor Institute of Information and Communication Technology,
Mehran University of Engineering and Technology, Jamshoro, Pakistan
[4]Assistant Professor, Department of Computer Science,
Faculty of Computing and Information Technology, King Abdulaziz University, Saudi Arabia
*Corresponding author, e-mail: Murk.hassan@outlook.com[1], Tariq.khanzada@faculty.muet.edu.pk[2],
Sheeraz.memon@faculty.muet.edu.pk[3], rhassan1@kau.edu.sa[4]

Abstract

Malaria is a most dangerous mosquito borne disease and its infection spread through the infected mosquito. It especially affects the pregnant females and Children less than 5 years age. Malarial species commonly occur in five different shapes, Therefore, to avoid this crucial disease the contemporary researchers have proposed image analysis based solutions to mitigate this death causing disease. In this work, we propose diagnosis algorithm for malaria which is implemented for testing and evaluation in Matlab. We use Filtering and classification along with median filter and SVM classifier. Our proposed method identifies the infected cells from rest of blood images. The Median filtering smoothing technique is used to remove the noise. The feature vectors have been proposed to find out the abnormalities in blood cells. Feature vectors include (Form factor, measurement of roundness, shape, count total number of red cells and parasites). Primary aim of this research is to diagnose malaria by finding out infected cells. However, many techniques and algorithm have been implemented in this field using image processing but accuracy is not up to the point. Our proposed algorithm got more efficient results along with high accuracy as compared to NCC and Fuzzy classifier used by the researchers recently.

Keywords: feature vectors, median filter, RBC, RGB to gray scale, SVM classifier

1. Introduction

Malaria is one of those kind of dangerous disease that causes a death because it directly affects to liver and then move towards red blood cells from there the life cycle of malaria also starts [1]. The majority of deaths are caused by Plasmodium falciparum disease that spreads globally. The common symptoms of malaria disease leads to headache, fever, vomiting, tiredness which may leads to coma and deaths. According to world health organization (WHO), there are many cases of malaria, estimated 200 million cases of malaria fever yearly [2]. Majority of malarial cases are found in poor countries where the pollution is high. When a person is bitten by an infected mosquito, then its parasites undergo inside the human body and destroy the red blood cells [3].

After that, it is recommended by the contemporary researchers to utilize an image processing technology [4, 5]. Various methods have been proposed for malaria testing and several classification techniques have been applied. Some of them are Minimum distance classifier, Naive baye's classifier and neural networks [6]. Minimum distance classifier works well when the distance between means of different classes is large. The limitation of Naive baye classifier is that the statistical properties of pattern classes are unknown [7]. In Neural network classification the accuracy decreases by the increasing number of features. However, its performance can be better by using minimum features.

In this work, initially the images are loaded. Median filter is applied for remove noise, it is also the best smoothing technique. Feature vectors have been achieved and used to detect

abnormalities in blood cells. The support vector machine (SVM) classifier is then applied in order to classify infected blood cell images from normal blood cells. SVM classifier is not used for detection of malaria infection till now to the best of our knowledge.

In this section we demonstrate the basics of related work in the area of malaria diagnosis. Rose et al., [8], suggested a diagnostic process of malaria using light microscopy. In this technique the images are used for preprocessing and then accuracy of species are determined by the Artificial Neural Network (ANN) classifier. The accuracy is not more than 73%. In 2014, Kareem [9] lodged an application to detect malaria using blood images. This application is based on Annual ring ratio (ARR) method and estimates the infected cells from blood images. Rahman [10], presented a method for detection of malaria parasites from thin blood smears. In first part of a system Morphological operation is used to extract RBC from an image with 95% accuracy and in further part it is able to detect and classify the malaria species along with 100% accuracy. Bhatt and Prabha [11], proposed an approach to detect and count abnormalities in red blood cells efficiently. The main purpose of this application is to overcome the time management. Form factor threshold is applied to find the abnormalities in red cells. Sreekumar [12] proposed an approach for counting total no of red cells and the shape of red cells.

However, in this paper the we propose the filtering mechanism followed by the feature vector extraction which is helpful to find the area, shape and roundness of infected blood cell stage and finally to apply the SVM classifier to separate the infected images from the normal ones, which is curial step being applied and finds the optimal and efficient way of classifying infections and detection of malaria at various stages.

2. Proposed Technique

The aim of this research work is to use SVM classifier to classify infected cells from non-infected one. First images are pre-processed and resized. Next it is converted into gray scale image, then median filtering operation is applied on RGB image. Furthermore feature vectors have been used and the last step is plasmodium detection that has been performed by using suitable SVM classifier. Figure 1 shown below defines the flow of the proposed work of detecting plasmodium parasites. Each of these steps are elaborated in the following subsections.
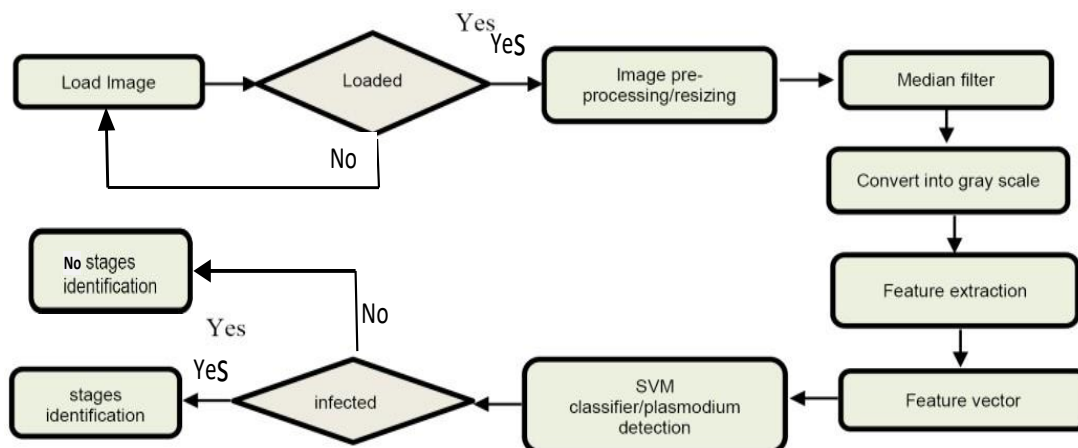
Figure 1. Steps of proposed technique

2.1. Pre-processing

The first step is to load the image. The image preprocessing involve the operation that has many basic features. It helps to resize the image in order to maintain standard size of all images, to speed up processing. In this work blood image samples have been taken from two image resources Centers for Disease Control (CDC) that contains images with 300*300 magnification [13]. Other resource is Clinical and Medical (CE LAB) that contains 251*201

magnification [14]. It is necessary to reshape these samples with the same size 300*300. RGB images are processed from both resources. Figure 2 (a) shows CDC image and Figure 2 (b) shows RGB image processed from CDC image library at CE LAB, similarly Figure 3 (a) and Figure 3 (b) show the images before and after application of Median filter respectively, and discussed in next subsection.
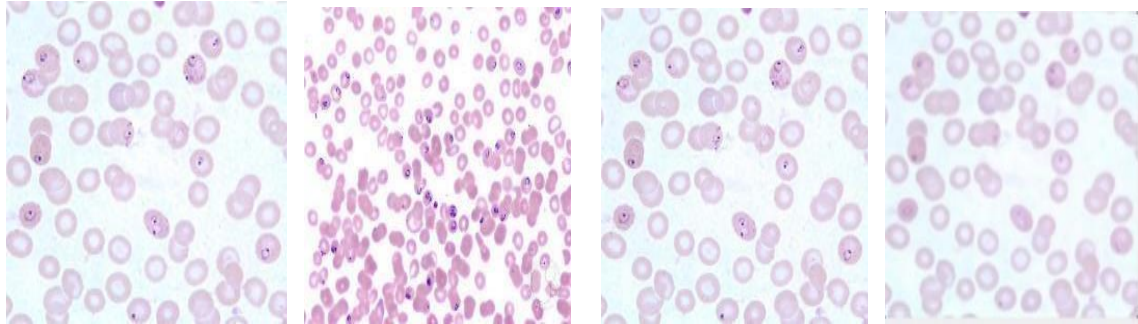
(a)   (b)   (a)   (b)

Figure 2. (a) RGB from CDC   Figure 3. (a) RGB before median filter
(b) RGB from CE lab   (b) RGB after median filter

2.2. Median Filter
Median filter is a nonlinear digital filtering technique. It helps to remove the noise. In this work median filter operation has been applied on RGB images in order to preserve the edges, retaining useful information. Figure 3 (a) and Figure (b) shows the results of median filter operation is applied before and after RGB image. After applying median filter to RGB image some sort of noise is being removed. The next step is conversion to grayscale which is brief in next subsection.

2.3. Conversion in Grayscale
Conversion of gray scale is performed on the resultant image discussed in the last sub section. Grayscaled image is then converted to binary image along with 0.9 intensity value as shown in Figure 4 (a) and Figure 4 (b). Background value is converted into foreground pixels by filling holes, shown in Figure 4 (c). The noise is removed by using bwareaopen shown by Figure 4 (d). This operation is use to remove the small objects whose values are less than 300 pixels. After median filter the RGB image is converted into grayscale image and some operations have been applied on grayscale image in order to further proceed to remove the noise. The results of these operations are shown in Figure 4.
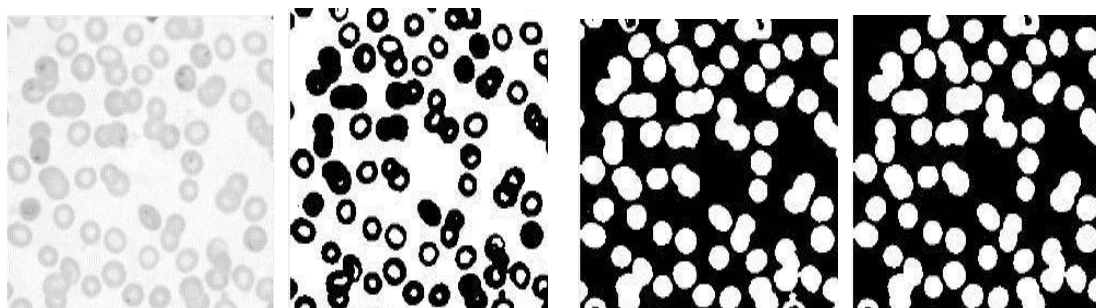
(a)   (b)   (c)   (d)

Figure 4. Gray scale, binary conversions, holes filing and noise reduction
(a) gray image (b) binary image (c) holes filing (d) noise reduction

2.4. Feature Extraction

Feature extraction is applied to segment erythrocyte green component images, shown in Figure 5 (a). It produces better results of finding out parasites. Malaria infected cells are shown by purple dots in it, shown in Figure 5 (b). Figure 6 (a) and Figure 6 (b) show the red number of cells and parasites respectively. The next step is to prepare feature vector, which are useful in detection of disease
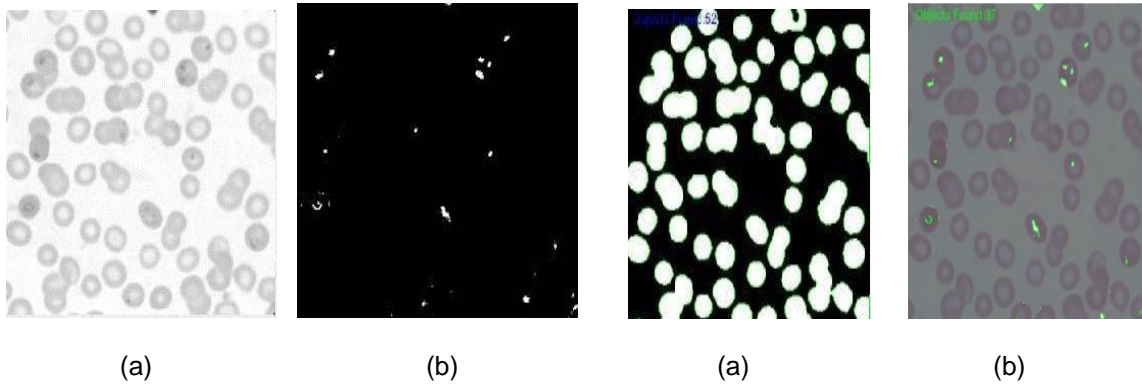

(a)   (b)   (a)   (b)

Figure 5. Green and purple components
(a) green component (b) purple plane

Figure 6. Total red cells and parasites
(a) total red cells (b) parasites

2.5. Feature Vector

Feature vectors have been used for detection of infected cells, which are Form factor, Roundness and Area, those are discussed in Table 1.

Table 1. Feature Vector Parameter with Definition and Equation

Name	Definition	Equation
Form factor	The form factor is used to measure shape metric. Form factor threshold is fixed, its value is equal to 1 for perfect circle. For all other non circular cells its value varies and is less than 1.	$Form\ Factor = \dfrac{(4*pi*area)}{(Perimeter^2)}$
Roundness	Normal red blood cells are round in shape and abnormal cells are having different variations in size. For a perfect circle the value must be equal to 1.	$Roundness = \dfrac{4*area}{(MajorAxisLength^2*pi)}$
Area	Shape of blood cells becomes rigid due to presence of infection. Normally the size of infected cells are larger than normal cells [15]. Then obviously the surface area increases because of larger cells producing larger surface area and volume.	Area = regionprops(BW2,'area')

2.6. Plasmodium Detection

Detection of plasmodium parasites has been done by SVM. This classifier is trained with some feature vectors, which are discussed in section 3.5, those feature vectors found to be the most appropriate method for detection of plasmodium parasites. This method is able to count total number of cells shown in Figure 6 (a) are 52 cells and parasites shown in Figure 6 (b) are 37.

2.6.1. Classification of Malaria Parasites and Their Life Stage

Four species of plasmodium have been recognized in infected blood cells, which are P.falciparum (p.f), P.vivax (p.v), P.ovale (p.o) and P.malarie (p.m). These four species have 4 life stages, which are Ring (R), Trophozite (T), Gametocyte (G) and Schizont (S).

Life stages of species can be detected by using different feature values. However, in this work Plasmodium falciparum Life stages has been detected along with some feature values are shown in Table 2.

Table 2. Four Life Stages of Plasmodium Falciparum

Name	Stage	Appearance of erythrocyte	Form factor	Roundness	No. of parasites	Shape of erythrocyte
Ring	First	Small chromatin dots	0.2-0.6	0.3-0.7	<50	
Trophozite	Second	Compact, dark pigment	0.1-0.8	0.2-0.5	>50	
Gametocyte	Third	Crescent shape, brown pigments	0.1-0.5	0.3-0.7	>200	
Schizont	Last/Fourth	Development of merozoites	0.1-0.7	0.5-0.8	>400	

2.6.2. Expected Values of Uninfected Cells

Number of red cells in normal/uninfected human are high in comparison of infected ones, number of RBC are significantly lower in malaria patients [16]. Uninfected cells having 0 parasites. In some cases minimum number of parasites may also be acceptable [17]. If the value of form factor is equal to 1, then it is said to be normal cell having 0.9 value is also adjustable [18]. Normal cells are round in shape, for this purpose roundness is measured, having 0.9 or 0.8 roundness value is said to be normal cells [19]. Surface area of normal cells are smooth, when cells having even a small amount of infection the area of those cells becomes rigid [20].

3. Results Analysis

These feature vector values have been tested on 100 images. Table 3 shows the summarization of results, achieved through our detection mechanism.

Table 3. Sample Acquisition Values for Feature Vectors of Four Life Stages of Fours Species

Image	Total Cells	Parasites Found	Form Factor	Roundness	Mean Area	Shape
P.F-R	8	15	0.2	0.3	5669	Square
P.F-T	11	72	0.4	0.4	3309	Square
P.F-G	9	287	0.3	0.5	4214	something
P.F-S	7	417	0.1	0.5	6564	Square
P.V-R	12	99	0.4	0.4	3366	Square
P.V-T	1	223	0.1	0.2	79094	Something
P.V-G	16	197	0.4	0.4	3283	Something
P.V-S	9	60	0.5	0.3	5307	Square
P.O-R	7	67	0.2	0.1	5161	Square
P.O-T	9	150	0.3	0.4	5809	Square
P.O-G	9	12	0.2	0.2	4757	something
P.O-S	9	36	0.4	0.5	5604	Square
P.M-R	8	31	0.3	0.5	5551	Square
P.M-T	16	309	0.4	0.3	3527	Square
P.M-G	13	63	0.3	0.4	3501	Square
P.M-S	7	90	0.3	0.4	6241	something

The total number of parasites amongst the total number of cells, along with form factor, roundness, mean area and shape are shown, which detect the infected cells and its stage, already detailed in Table 2. Four species along with four life stages that are 4*4=16 images. The performance of the proposed method is evaluated using statistical properties shown in Table 4, where TP denotes true positive, TN denotes true negative, FP denotes false positive and FN denotes false negative respectively. Summary of the results based on these statistical properties are shown in the shape of confusion matrix in Table 5.

Table 4. Statistical Properties

Name	Definition	Equation
Sensitivity	The sensitivity (Se) of a test is defined as probability of positive test result, when disease is present and test is positive.	$Se = \dfrac{(TP)}{(TP+FP)} *100$
Specificity	The specificity (Sp) of a test is defined as probability of negative test result, when disease is absent and test is negative.	$Sp = \dfrac{(TN)}{(TN+FN)} *100$
Accuracy	Accuracy (A) is measured by adding all those instances whose predicated output values match up the ground truth. Accuracy is all about that, how much efficient the tests are.	$A = \dfrac{(TN+TP)}{(TP+TN+FP+FN)} *100$

Table 5. Confusion Matrix

		Target Class		
Output Class	0	5 5.0%	1 1.0%	83.30% 16.7%
	1	2 2.0%	92 92.0%	97.9% 2.1%
		78.40% 28.6%	88.90% 1.1%	97.0% 3.0%
		0	1	

From confusion matrix of Table 5, we can see that, TP=5, TN=92, FP=1, FN=2. Total images on which the statistical calculations are performed are 100. Calculating the sensitivity in (1), (2) and (3), specificity in (2) and accuracy in (3) we achieve the performance results.

$$Se = \frac{5}{(5+1)} * 100 = 83.3\% \tag{1}$$

$$Sp = \frac{92}{(92+2)} * 100 = 97.8\% \tag{2}$$

$$A = \frac{(5+92)}{(5+92+1+2)} * 100 = 97\% \tag{3}$$

It is clear from the Efficiency analysis results compared in Figure 7 that SVM classifier achieves a better performance which is having 97% accuracy compared with previous methods of NCC and Fuzzy.
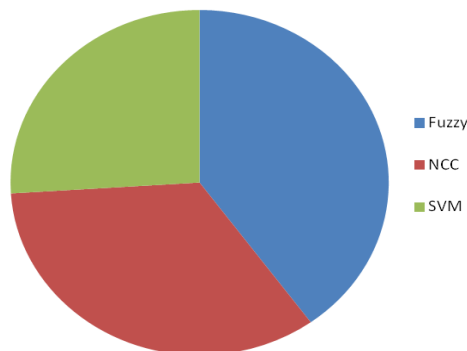

Figure 7. Comparison of proposed algorithm for Efficiency with NCC and Fuzzy

From Figure 7 where 30, 50 and 100 are number of images. The accuracy is comparatively high on increasing number of image samples. Fuzzy and NCC having good accuracy on using minimum no. of images, its accuracy decreases by increasing no. of image

samples. Whereas from Figure 8, various techniques have been deployed but time is again a non-existent service for such dangerous diseases. Efficiency of this proposed system have been compared with other previous methods by estimating the time. This Figure 8 clearly shows that SVM required less time for implementing 100 images. Our proposed method supersede all the concurrent methods applied to detect infected blood cells.
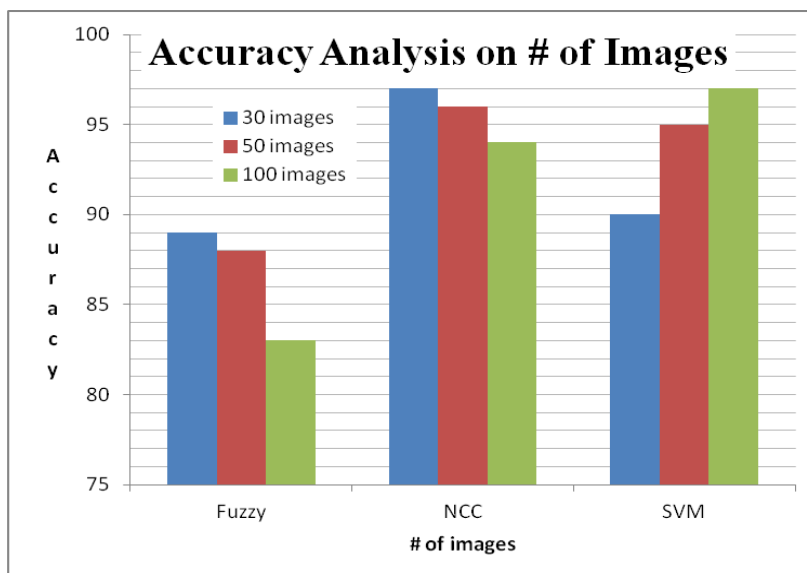

Figure 8. Accuracy analysis comparison with the counterpart methods

4. Conclusion

In this paper an attempt has been made to detect malaria through SVM classifier by using microscopic blood images. The aim of this work is to detect malaria by finding out abnormalities in red cells and to determine the life stages of malaria. Infected images have been processed along with suitable feature vectors that include, total number of red cells and number of parasites present. Form factor threshold value is fixed, for all circular objects the value must be equal to 1 otherwise it varies for all non-circular objects. Normal cells are round in shape, however, for a perfect circle the value of roundness is 1 or 0.9 is also considered as a circle. SVM classifier is being trained with some data along with these feature vectors. Implementation has been performed on almost 100 images after that, the sensitivity of this proposed system is 83.3%. Specificity is 97.8%. Overall accuracy of the system is 97%.

References

[1]   H Jia. Malaria hijacks liver cells to avoid immune system. Science and Development Analysis. 2016.
[2]   World Health Organization (WHO). 2017.
[3]   P Rakshit, K Bhowmik. *Detection of Presence of Parasites in Human RBC In Case of Diagnosing Malaria Using Image Processing*. International Conference on Image Information Processing (ICIIP). 2013: 329-334.
[4]   A Yasrib, MA Suhaimi. Image processing in Medical Applications. *Journal of Information Technology Impact.* 2003; 3(2): 63-68.
[5]   S Naseera, GK Rajini, B Venkateswarlu, JP Priyadarisini M. A Review on Image Processing Applications in Medical Field. *Research Journal of Phramacyy and Technology.* 2017; 10(10): 3556-3560

[6]   G Vinodhini, RM Chandrasekaran. A Comparative Performance evaluation of neural network based approach. *Journal of King Saud University-Computer and Information Sciences.* 2016: 281: 2-12.

[7]   DM Memeu. *A Rapid Malaria Diagnostic Method Based on Automatic Detection and Classification of Plasmodium Parasites in Stained Thin Blood Smear Images*. M.Sc. Thesis. Department of physics. University of Nairobi. 2014.

[8]   NE Ross, CJ Pritchard, DM Rubin, AG Duse. Automated image processing method for the diagnosis and classification of malaria on thin blood smears. *Medical & Biological Engineering & Computing.* 2006; 44: 427-436.

[9]   C Dallet, S Kareem, I Kale. *Real time Blood Image Processing Application for Malaria Diagnosis Using Mobile Phones*. IEEE International Symposium on Circuits and Systems (ISCAS). 2014: 2405-2408.

[10] HA Rahman, IA Maaly. *Detection and classification of malaria in thin blood slide smears*. International Conference on Communication, Control, Computing and Electronic Engineering (ICCCCEE). 2017.

[11] M Bhatt, S Prabha. Detection of Abnormal Blood Cells Using Image Processing Technique. *International Journal of Electrical and Electronics Engineers (IJEEE).* 2015; 7(01): 89-94, 2015.

[12] A Sreekumar, PA Bhattcharya. Identification of Sickle Cells from Microscopic Blood Smear Image Using Image Processing. *International journal of Emerging Trends in Science and Technology (IJETST).* 2014; 1: 783-787.

[13] Centers for Disease Control and Prevention. https://www.cdc.gov/dpdx/malaria/index.html. Accessed date March 8. 2018.

[14] Online Laboratory Continuing Education for Clinical Laboratories. https://www.labce.com/normal-peripheral-blood-cells.aspx. Accessed date April 27. 2018.

[15] KC Charpe. *Automated Malaria Parasite and there Stage Detection in Microscopic Blood Images*. International Conference on Intelligent Systems and Control (ISCO). 2015; 9.

[16] Kotepui M, Piwkham D. Effects of Malaria Parasite Density on Blood Cell Parameters. *National Center for Biotechnology Information (NCBI).* 2015; 10(3): 1-11.

[17] Adu-Gyasi D, Admass M, Amoako S. Estimating Malaria Parasite Density. *National Center for Biotechnology Information (NCBI).* 2012; 11: 238.

[18] A Sreekumar, PA Bhattcharya. Identification of Sickle Cells from Microscopic Blood Smear Image Using Image Processing. *International journal of Emerging Trends in Science and Technology (IJETST).* 2014; 1: 783-787.

[19] Rahmat, Romi, FS Wulandari, S Faza, MA Muchtar, I Siregar. *The morphological classification of normal and abnormal red blood cell using Self Organizing Map*. IOP Conference Series: Materials Science and Engineering. 2015.

[20] T Ye, NP Thien. Stretching and Relaxation of Malaria Infected Red Blood Cells. National Center for Biotechnology Information. 2013.