A Combined Algorithm for Malaria Detection from Thick Smear Blood Slides

Kaushik Chakrabortya, Arnab Chattopadhyayb, Amlan Chakrabarti, Tinku Acharyad and Anjan Kr Dasguptae*

Department of Biochemistry and Center for Excellence in Systems Biology and Biomedical Engineering, University of Calcutta, Kolkata 700019, India

Abstract

Thick blood smear examination is a necessary part for rapid screening of malaria parasite. Primary diagnosis of malaria by thick smear examination is cheap and highly sensitive, advocated by the World Health Organization (WHO). For the examination of thick blood smear, manpower and time can be reduced by using automated computational techniques. These techniques would facilitate such detection that are usually based on morphology and in few cases on analysis of colors. In this work, we propose a combined algorithm consisting of morphological operations and color based pixel discrimination technique to identify malaria parasites from thick smear images of Plasmodium vivax. Using morphological operation segmentation of cells from thick blood smear image is done and color based pixel discriminator distinguished malaria cells from segmented image. Evaluation of percentage of detection and False Positive Rate (FPR) shows that our proposed algorithm has significantly higher predictive rate, and lower FPR as compared to any existing methodology when tested on same input slides. Importantly the reported method does not need any training set and assumes unsupervised methodology. This makes the portability of the approach at various scales as discussed in the paper.

The experiment has been carried out using JSB (Jaswant Singh Battacharya) stained thick blood smear images. It is assuming that our algorithm can identify malaria in less perfect dirty slides. This makes the algorithm more powerful and robust.

Keywords: Thick smear; Malaria; Schizont form; Segmentation; Morphology; Euler number

Introduction

Malaria management is a challenging problem all over the globe particularly in Asian and African continents. Presently, even 110 years after the Nobel Prize of Ronald Ross for his work on malaria, people in the European region are also at risk from diseases carried by vectors both within the region and when traveling abroad. While treatment of malaria itself is a challenging problem its quick detection is also a problem with no less significance. There are mainly four species of malaria parasites infecting human beings [1] namely, *Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale* and *Plasmodium malariae. Plasmodium vivax*, is found mainly in tropical and subtropical areas and has a severe clinical manifestation [2]. Rapid detection of presence of the parasite in human blood and early institution of anti-malaria drugs are the mainstay of management of the disease. WHO recommends that all cases of suspected malaria be confirmed using parasite-based diagnostic testing (either microscopy or rapid diagnostic test) before administering treatment [3]. In the malaria detection test, microscopy based diagnosis has the central importance for species differentiation, parasite quantification, management of severe disease. Additionally, the method may be amenable to larger section of society because of its scalability and low running cost.

Two types of blood smears, thick and thin, are prepared from the blood of patients, who are clinically suspected to be suffering from malaria. The thick smear is more useful for parasite detection whereas the thin smear is particularly used for identification of malaria species. When the parasite load is low, malaria may be detected about 20 times more rapidly in thick smear than in thin smear [3]. The methods based on which image analysis of blood smears have been made, broadly fall into two classes namely, analysis based on morphology and that based on color. Some of the reported methods use supervised training sets and lack of availability of appropriate training sets may delimit the scalability of such methods.

Among the morphology based analysis as presented by Toha et al. [4], we find usage of cluster analysis on thick smear images and Euclidean distance as classifier for malaria cells. Morphology based analysis is also reported by Frean [5], where the authors proposed counting of parasites from individual microscope images. Among other such methods we find the work of Somasekar [6], who presented a linear programming based model and used a graphical approach in image segmentation and morphological operations to detect presence of malaria parasite. Again, Edison *et al.* [7] analyzed the changes of *Plasmodium vivax* in thin smear erythrocytes images by image filtering and edge detection techniques. The work of Anggraini et al. [8], shows segmentation of blood cells components from thin smear slide. There are additional reports on binary parasite detection scheme [9] that is based on a modified K nearest neighbor (KNN) classifier of the infectious species and life cycle. Other work that deserves mention is by Elter et al. [10] who reported a technique of detection from thick blood smear images based on Support Vector Machine.

There are few other reports consisting of color based analysis. Mandal et al. [11] presented an algorithm to segment RBC and malaria parasite from infected RBC using an optimized normalized cut method using in RGB, YCbCr and HSV color space. Makkapati et al. [12], proposed a scheme based on HSV color space that segments RBC and parasites

| Sl. No.	Author(s)	Method	Objective
1	Toha et al [4]	Cluster analysis on thick smear images	Classification of malaria cells
2	Frean [5]	Morphology based analysis	Count parasites from individual microscopic images
3	Somasekar [6]	Linear programming based Image segmentation and morphological operations	Detection of malarial parasite
4	Edison et al. [7]	Image filtering and edge detection	Analyze the changes of *Plasmodium vivax* in thin smear erythrocytes images
5	Anggraini et al. [8]	Filtering and Multiple Thresholding	Detection of Erythrocyte from thin blood smear Image
6	Tek, F B et al. [9]	Area granulometry, Illumination, correction, Scale correction, Foreground and background separation,Color normalization, Feature extraction	Identification of malaria parasite from thin blood smear image and study the life cycle stage of species
7	Elter et al. [10]	Image Filtering and SVM (Support Vector Machine)	Detection of malaria parasite from thick smear image
8	Mandal et al. [11]	Optimized normalized cut method using in RGB, YcbCr and HSV (Hue Saturation Value) color space	Detection of malaria parasite from infected RBC (Red Blood Corpuscle)
9	Makkapati et al. [12]	Optimal saturation threshold on HSV (Hue Saturation Value) color space	Segmentation of RBC (Red Blood Corpuscle) and Malaria parasite
10	Prasad et al. [13]	Android based image analysis and decision support system	Identification of species from Giemsa stained thin smear images.
11	Hirimutugoda et. al. [16]	Color normalization, morphological operation, ANN (Artificial Neural Network) based classifier	Detection of erythrocytes from thin blood smear image.
12	Prescott et. al. [14]	Color and Shape pattern of parasite and RBC (Red Blood Cells)	Detection of malaria parasite from thick and thin smear image and find parasitaemia
13	Linder et. al. [15]	Local Binary Pattern (LBP), Local Contrast, Scale Invarient Feature Transform (SIFT) and SVM (Support Vector Machine) classifier	Detection of ring-stage trophozoites from Giemsa stained thin smear image

Table 1: Tabular View of Past Work.
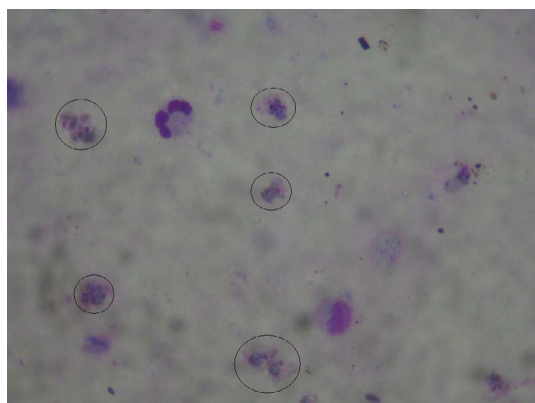
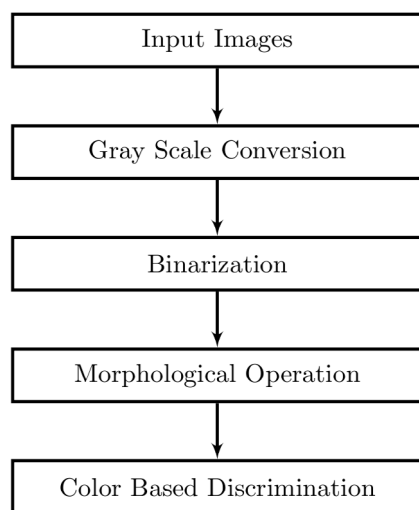**Figure 1:** Thick smear image of malaria infected patient (Encircled cells are malaria).

Figure 2: Steps of our proposed methodology.

by detecting dominant hue range and by calculating optimal saturation thresholds. Prasad et al. [13], proposed a detection technique based on HSV color-plane and implemented an android application for the final confirmation on the decision for treatment of malaria. In the report of Prescott et al. [14], we find detection of malaria parasite from both thick and thin smear images using color and shape pattern. Linder et al. [15] identify ring-stage trophozoites from Giemsa stained thin smear image using Local Binary Pattern (LPB), Local Contrast, Scale Invarient Feature Transform (SIFT) operator. Their approach is purely supervised because they used SVM (Support Vector Machine). The only example known to us where color based analysis is combined with morphology based methods is the work by Hirimutugoda et al. [16].The work is however based on a supervised method using normalized color and cell size features that was further classified using Artificial Neural Network.

All the methods are summarize below in Table 1

In this paper, we for the first time, describe an unsupervised approach in which color and morphology based algorithms are put together to formulate an algorithm for *Plasmodium vivax* detection from thick smear slide. Our approach has a comparative higher predictability and lower false positive rate.

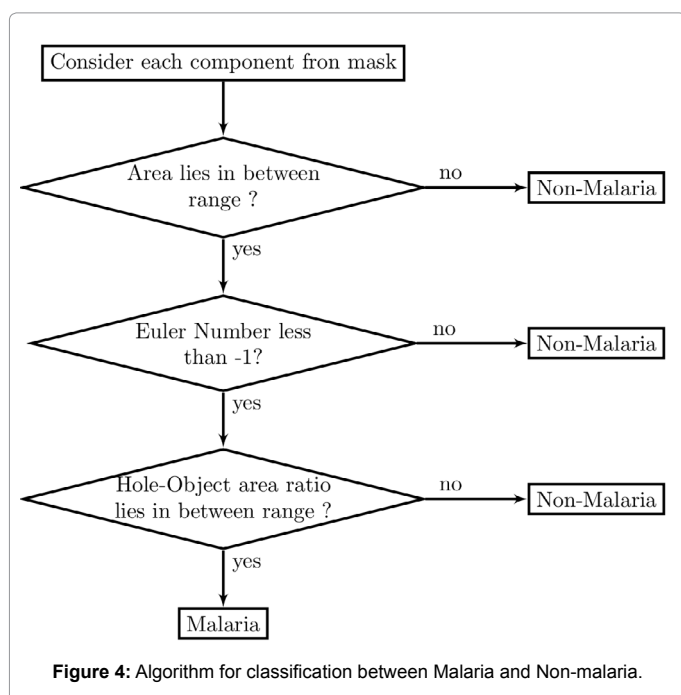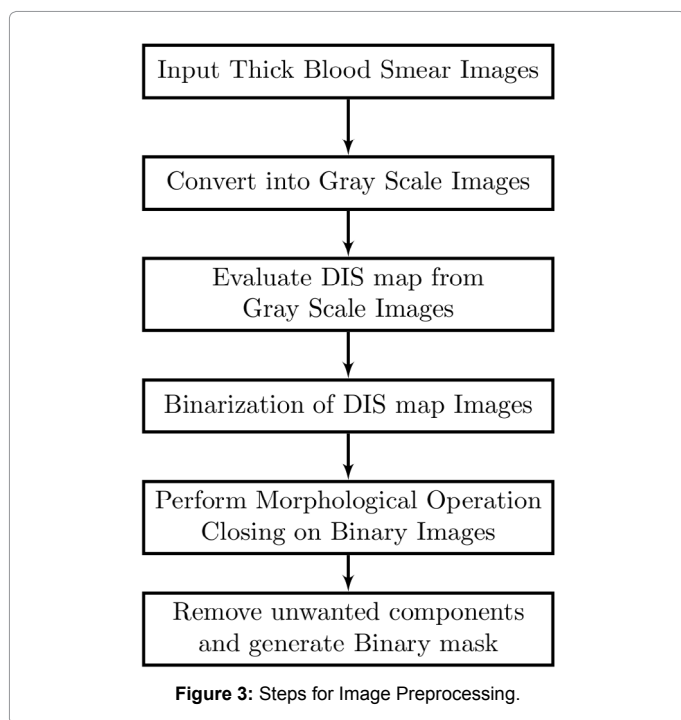## Problem Definition and Motivation

A thick smear image of *Plasmodium vivax* is shown in Figure 1. In practice, the following features of *Plasmodium vivax* malaria affected cells distinguish them from other species of malaria:

- Ill-defined circumference

- Blue colored parasite cytoplasm area with a violet chromatin grid

- Brown dots of Schuffner's stipplings

Using the above mentioned features we have developed an automated malaria detection technique to offer as a tool to the investigator to accelerate the investigation process of the presence of malaria and save their valuable time. It is particularly useful in a thickly populated country where doctor patient ratio is very low.

Proposed Methodology

Current approaches for malaria parasite detection needs image acquisition from thick smear malaria slide. All images were generated from thick blood smear slides of human. Each slide is stained using JSB (Jaswant Singh Battacharya) staining agent and examined under light microscope with 10X100 magnification by expert. There is no loss of parasite during staining in thick blood smear; artifacts and parasites are observed in their natural location.In brief, proposed methodology is illustrated in Figure 2.

Figure 3: Steps for Image Preprocessing.

Figure 4: Algorithm for classification between Malaria and Non-malaria.

Malaria detection using binarization and morphological operation

Image Pre-processing: In this stage, following techniques were used to convert the source image to binary image. At first thick smear images were converted into gray level image. From the gray level image, Difference in Strength map (DIS map) [17] for each pixel is calculated. After processing all the input pixels, the DIS image is obtained. In DIS image, the larger the DIS value, the more the pixel is likely to be located at the edge. In the next stage, we applied binarization technique. In this stage, we applied Zack thresholding [18] method. Morphological operation closing [19] is performed on the resultant binary image. In closing at first dilation and then erosion is performed with a 3x3 structural element, in which all elements of the matrix is 1. Next, from the binary image, we removed most of unwanted components based on area of the components. It has been observed that on 10×100 light microscopy the component of interest can hold more than 0.0018% of total area and less than 0.014% of total area. After removing the unwanted components, mask is produced. We elaborate the segmentation process in the following sub-section. The steps are shown in Figure 3.

Image segmentation: To segment out malaria cells from blood smear images two features were used, Euler number [19] and ratio of black and white pixel of an individual object. Euler Number [19] is a feature related with a single connected component. It returns whether a single object has any hole or not.

The selection of Euler number as a feature helps us to discriminate malaria and non-malaria cells. The mask generated in this step contains both malaria and non-malaria cells and the porosity of malaria affected cell mask was higher than others. From the observation, we found that malaria affected cells have Euler number less than -1. It has eliminated unwanted objects.

It is observed that most of the malaria affected cell's whole area and object area ratio lies between 0 to 0.5 and Euler number is always less than -1. The classification method is summarized in Figure 4.

Malaria detection using color based discrimination

Image pre-processing: HSV [19] (Hue, Saturation, Value) is a color space, which describes the colors perceived by human eyes. Hue is an attribute by which humans perceive the colors like Red, Green, Blue, Yellow, Magenta, and Cyan like primary chromatic components of colors or Purple etc. like intermediate chromatic components color. Saturation describes the colorfulness of a color perceived by human. Value signifies the intensity level of color i.e. how much light is reflected or emitted by an object. The HSV color system is represented by a cone. In HSV cone value varies along the vertical axis from 0 to 1. Saturation is the measure of radius of base circle of the cone from center to circumference and various from 0 to 1. Hue varies along the periphery of the base circle of the cone, varies from angle 0° to 360°. In hue, six colors- red, yellow, green, cyan, blue, magenta are spaced 60° apart from each other with starting color red at 0° .From empirical observation of our experiment it is found that for JSB stained image, a range in Hue and Saturation parameter make a differentiation between our target and background, where target initially includes malaria cells, leukocytes and platelets. In this experiment the range of hue exists in between 66% and 75% of the total range and saturation varies from 0.45 to above. This gives us a very encouraging result.

Discrimination of malaria cell: In this stage also Euler Number [19] is chosen as a discriminating feature by utilizing the porosity of objects. In the preprocessed image it is shown that each malarial cell

Figure 5: Steps for Combined Algorithm.
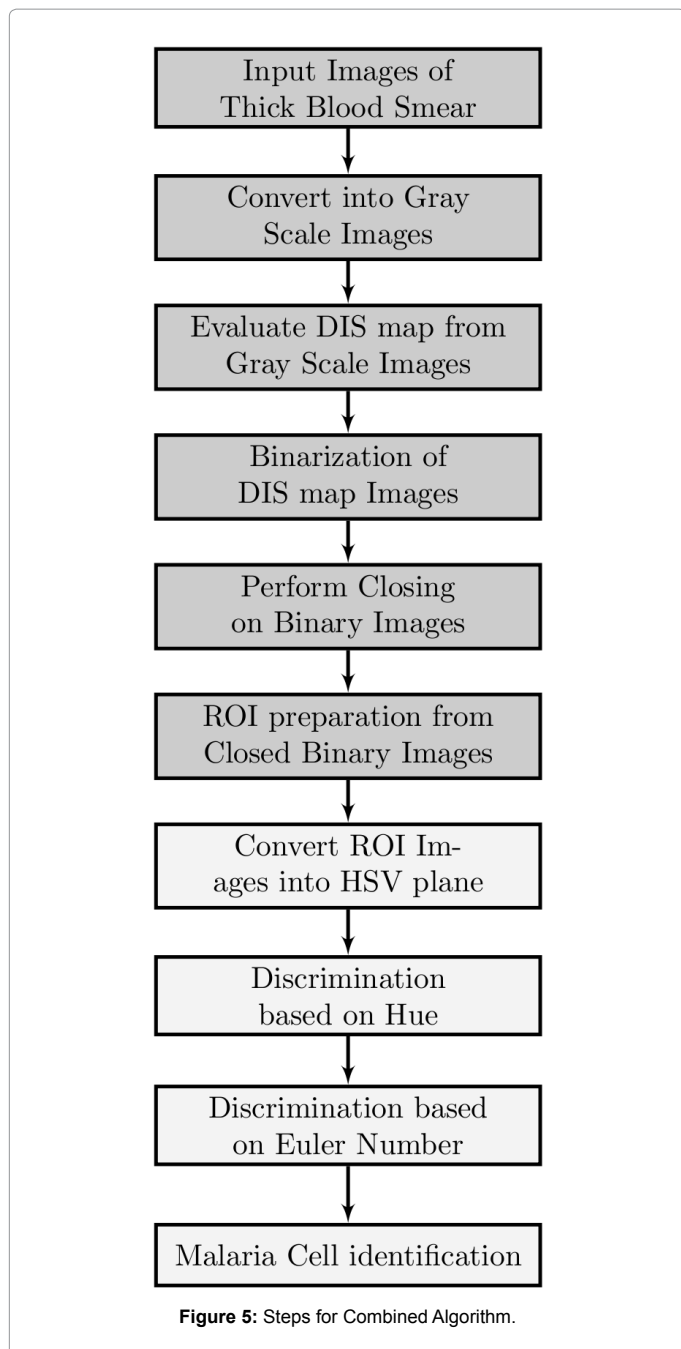
holds Euler number =-1 but some non-malaria cell also has Euler number =-1. So, to discriminate these two sets, we apply significantly small area screening because each non-malaria cells holds a significantly large area. Discrimination based on area is possible when there are a large number of particles are present with a large variety of area.

After the screening based on area, it is observed that resultant set contains both malarial and non-malarial objects whose area is like malaria cells. To eliminate those objects we consider each small object and measure their center. Then from the center, search its neighbor if there is any other objects within a circle (from the center of object of interest and radius is 30 pixels) then that two objects (both object of interest and neighbor object) have a high possibility of malaria.

Combined algorithm of malaria detection

In this approach we have combined the above two algorithm. At first binarization and morphological operations are performed and then color based segmentation is used.

The images are converted into gray scale image and then DIS map [17] is produced from gray scale images. Then the DIS map images are birarized with Zack thresholding [18] method and morphological operation closing [19] is done with a 3x3 square element. From the closed image, using Euler number [19] feature is used for first screening. These steps washed the background and generate an image where only our ROI (Region of Interest) is present. The Figure5 shows the steps where purple colored boxes represent the processes from Binarization and Morphological Algorithm and pink colored boxes represents the processes from Color Based Discrimination. A sample source image and output image images are shown in Figures 6a & 6b. On the back ground removed image, we applied a color based screening technique [11,12]. For this, the images are converted into HSV plane. The possible malaria pixels are screened out based on Hue value to produce the intermediate image. Once more we extract the Euler number feature from the intermediate image.

Results and Discussion

In this experiment, we have used our own database which contains 75 images. The size of the images are 500×500 and 1024×768. The images are collected from JSB (Jaswant Singh Battacharya) stained thick blood smear slides. The staining of slides have been performed by a heterogeneous group of technician. The database was created by the capturing images from the slides by a 4.4 mega pixel digital camera in different lighting conditions. For microscopy, Nikon E200 microscope with 10×100x magnification was used. At first the images have been checked and validated by experts and then we executed our algorithm on those images. As our goal is to detect weather a patient is malaria

Figure 6: (a) Original Image (b) Background washed image (Region of interest).
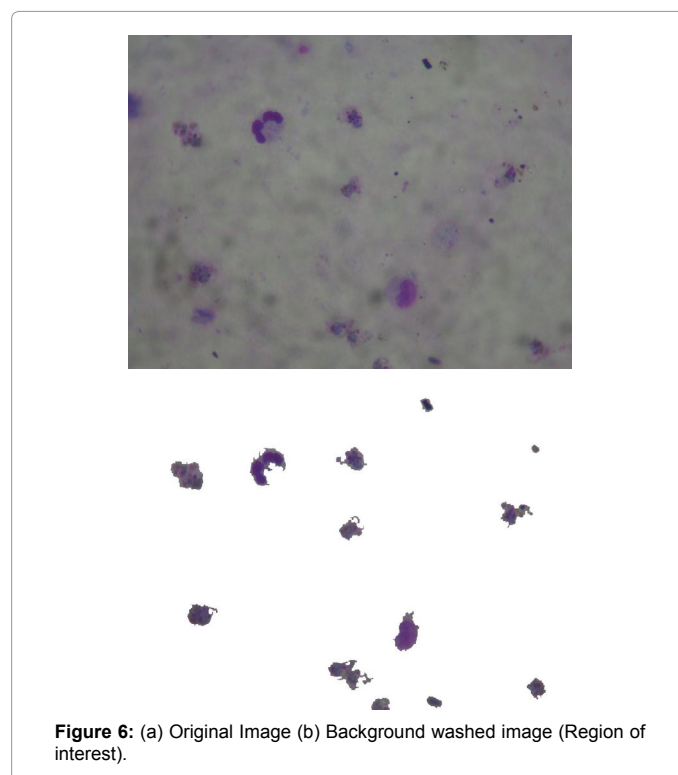
Sl. No.	Method	% of Detection	False Positive Rate
1	Morphological Operation (M) (our method)	79 ± 2.3	0.10 ± 0.02
2	Color Discrimination(C) (our method)	90 ± 1.5	0.04 ± 0.01
3	M+C (our method)	94.5 ± 1.5	0.10 ± 0.02
4	Method Proposed by [13]	46.5	0.303

Table 2: Comparative Results of our image data. Digits after ± denotes the standard error of the mean.
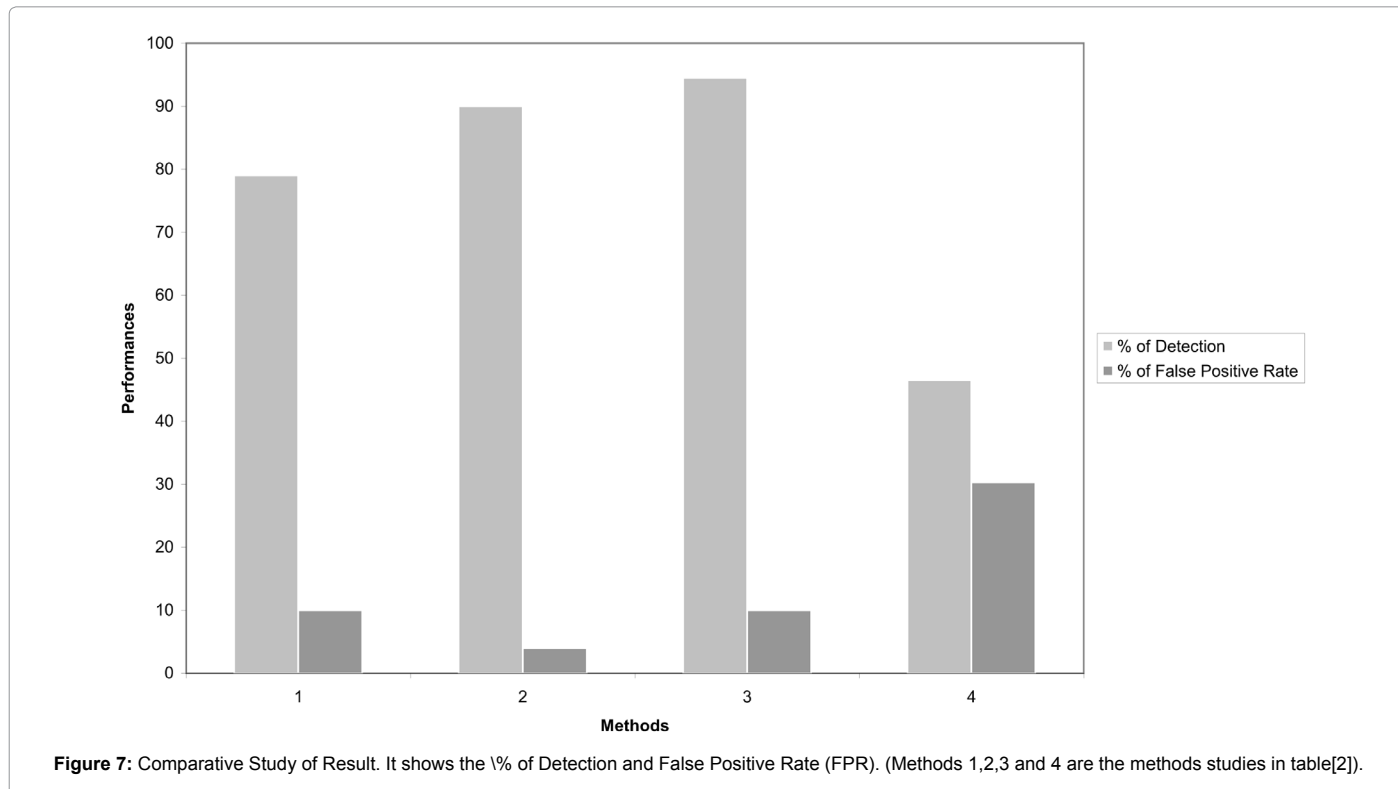

Figure 7: Comparative Study of Result. It shows the \% of Detection and False Positive Rate (FPR). (Methods 1,2,3 and 4 are the methods studies in table[2]).

affected or not, so we classified the result into two groups, whether the element is malaria affected cell or healthy. For this classification we used traditional measures; True Positive (TP),False Positive (FP), True Negative (TN),False Negative (FN).From the above four quality parameters the average sensitivity or True Positive Rate (TPR) and False Positive Rate (FPR) can be calculated using following equations 1 and 2:

$$TPR = TP / (TP + FN) \quad (1)$$

$$FPR = FP / (FP + TN) \quad (2)$$

In this study it is very encouraging that percentage of detection is improved when Binarization and Morphological algorithm and Color Based Discrimination algorithm are cascaded one after another and forms a combined algorithm. In case of False Positive rate, Combined Algorithm returns the maximum FPR value. From these point of view combined algorithm gives us satisfactory result than the individual algorithm.

It may be noted that there is some dependence of the algorithm on the input sample. We chose to use 'dirty slides' assuming that if the algorithm passes test on such slides it would be perhaps more effective in detecting presence of malaria in better prepared slides. Thus when we test our dataset using algorithm described by Prasad et al. [13] the detection by their algorithm is much less accurate than the results given by the authors using their own test dataset. To avoid any confusion we provide comparative results of different algorithms based on the set of test data set chosen by us (Table 2 and Figure 7).

Some extensive studies on ring formed in thin smear images have been reported in literature [20]. In earlier studies in thick smear [4-13] only select images were shortlisted. Our motivation is to make our algorithm powerful and robust so that we can detect malaria affected cells in less perfect slides (where presence of dirt or stain crystal or other slide related spots) for early detection of Schizont form of malaria from thick smear image of *Plasmodium vivax*.

Conclusions

We found that for images with lesser quality as it may happen for low cost and low resolution cameras, our proposed methodology can tackle the problem algorithmically by intelligent combination of color and light based operations. We have used unsupervised method, however possible interface with supervised method leaves the scope for further improvement of the prediction rate and decrease of the false positive rate. The scalability of the method to smart phone based microscopy platform [21] may be another value based addition to this new diagnostic paradigm.

References

1. Rougemont M, Van Saanen M, Sahli R, Hinrikson HP, Bille J et al. (2004) Detection of four plasmodium species in blood from humans by 18s rrna gene subunit-based and species-specific real-time pcr assays. J Clin Microbiol 42: 5636-5643.

2. Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI (2005) The global distribution of clinical episodes of *plasmodium falciparum* malaria, Nature 434: 214-217.

3. H. Reyburn, (2010) New who guidelines for the treatment of malaria. BMJ 28: 340.

4. Toha S, Ngah U (2007) Computer aided medical diagnosis for the identification of malaria parasites, in: Signal Processing, Communications and Networking, ICSCN'07. International Conference on, IEEE. 521-522.

5. Frean J (2010) Microscopic determination of malaria parasite load: role of image analysis. Micrsocopy: Science, Technology, Applications, and Eductaion 862-866.

6. Somasekar J, Reddy B, Reddy E, Lai C (2011) Computer vision for malaria parasite classification in erythrocytes, International Journal on Computer Science and Engineering 3: 2251-2256.

7. Edison M, Jeeva J, Singh M (2011) Digital analysis of changes by *Plasmodium vivax* malaria in erythrocytes. Indian Journal of Experimental Biology 49: 11-15.

8. Anggraini D, Nugroho AS, Pratama C, Rozi IE, Iskandar A A, et al. (2011) Automated status identification of microscopic images obtained from malaria thin blood smears. In: Electrical Engineering and Informatics (ICEEI), 2011 International Conference on, IEEE. 1-6.

9. Tek FB, Dempster AG, Kale I (2010) Parasite detection and identification for automated thin blood film malaria diagnosis, Computer Vision and Image Understanding 114: 21-32.

10. Elter M, Hasslmeyer E, Zerfass T (2011) Detection of malaria parasites in thick blood films. Conf Proc IEEE Eng Med Biol Soc. 2011. 5140-5144.

11. Mandal S, Kumar A, Chatterjee J, Manjunatha M, Ray A (2010) Segmentation of blood smear images using normalized cuts for detection of malarial parasites. in: India Conference (INDICON), 2010 Annual IEEE. 1-4.

12. Makkapati VV, Rao RM (2011) Ontology-based malaria parasite stage and species identification from peripheral blood smear images. Conf. Proc. IEEE Eng Med Biol Soc 2011 2011:6138-41.

13. Prasad K, Winter J, Bhat UM, Acharya RV, Prabhu GK (2012) Image analysis approach for development of a decision support system for detection of malaria parasites in thin blood smear images. J Digit Imaging 25: 542-549.

14. Prescott WR, Jordan RG, Grobusch MP, Chinchilli VM, Kleinschmidt I, et al. (2012) Performance of a malaria microscopy image analysis slide reading device. Malar J 11: 155.

15. Linder N, Turkki R, Walliander M, Mårtensson A, Diwan V, et al. (2014) A malaria diagnostic tool based on computer vision screening and visualization of *plasmodium falciparum* candidate areas in digitized blood smears. PloS one 9: e104855.

16. Hirimutugoda Y, Wijayarathna G (2009) Artificial intelligence-based approach for determination of haematologic diseases. in: Image and Signal Processing, 2009. CISP'09. 2nd International Congress on, IEEE, 2009. 1-5.

17. Yu YW, Wang JH (1999) Image segmentation based on region growing and edge detection. in: Systems, Man, and Cybernetics, 1999 IEEE International Conference on. 6: 798-803.

18. Zack GW, Rogers WE, Latt SA (1977) Automatic measurement of sister chromatid exchange frequency. J Histochem Cytochem 25: 741-753.

19. Acharya T, Ray AK (2005) Image processing: principles and applications, John Wiley & Sons 2005.

20. Anyona SB, Schrier SL, Gichuki CW, Waitumbi JN (2006) Pitting of malaria parasites and spherocyte formation. Malar J. 5: 64.

21. Meankaew P, Kaewkungwal J, Khamsiriwatchara A, Khunthong P, Singhasivanon P, et al. (2010) Application of mobile-technology for disease and treatment monitoring of malaria in the "Better Border Healthcare Programme". Malaria journal 9: 237.