Automated Detection Of Malaria Parasite Based On Thick And Thin Blood Smear Images

Parikshit Shembekar, *Student, PCE,* Niraj Yadav, *Student, PCE*, P. Jayajith , *Student, PCE*, Mohammed Shaikh, *Student, PCE,* Satishkumar verma*, Professor, PCE*

***Abstract*—Malaria is a mosquito borne infectious disease of humans and other animals caused by Parasites(a type of microorganism) of the genus plasmodium. Diagnosis of malaria is done by microscopic examination of blood. But this diagnosis method is time consuming and requires pathologists. The paper aims to introducing fast and accurate method based on image processing for malaria parasite identification. Based on morphological operations total number of cells are counted. Infected cells are analyzed based on intensity profiles within the cells. This result is validated by comparing with manual analysis. This approach is used in rural areas where fewer experts are available and the delayed diagnosis may lead to complication in the patient's health. The performance measure is sensitivity and specificity.**

***IndexTerms*— Binaryimage,FeatureExtraction,Grayscaleimage,K-meansclustering.,Morphological Operation,RBC'S and Parasite Count,Thick and Thin images, Thresholding,**

I. INTRODUCTION

Malaria is a mosquito-borne infectious disease affecting humans and other animals caused by parasitic protozoans (a group of single-celled microorganisms) belonging to the *Plasmodium* type. According to the World Health Organization (WHO) malaria remains one of the most widespread infectious diseases of mankind, with 40% of the world's population at risk and more than 240 million infections each year. Recent estimates indicate a staggering mortality rate of over 1.2 million deaths annually, of which the majorities are children and pregnant women living in sub-Saharan Africa Malaria causes symptoms that typically include fever, fatigue, vomiting, and headaches. In severe cases it can cause yellow skin, seizures, coma, or death. Symptoms usually begin ten to fifteen days after being bitten. If not properly treated, people may have recurrences of the disease months later.

The disease is most commonly transmitted by an infected female Anopheles mosquito. The mosquito bite introduces the parasites from the mosquito's saliva into a person's blood. The parasites travel to the liver where they mature and reproduce. Malaria is typically diagnosed by the microscopic examination of blood using blood films.

Current state of the art for medical diagnosis and research purposes involves drawing a blood sample from a patient or research subject. This blood sample is smeared onto a slide and stained in order to color cell nuclei. Because mature red blood cells do not possess nuclei, the stain only strongly marks malarial parasites. The slide can then be examined under a microscope in order to count the number of infected red blood cells.

A. RBC Cells

Red blood cell is considered infected if at least one parasite can be detected within its interior. White blood cells and free-floating parasites are not considered. The current state of the art involves manual counting by a laboratory technician or other individual, who can distinguish staining artifacts from actual nuclei, white blood cells, and (depending on specific requirements) life cycle and species of malarial parasites . Although manual counting is relatively inexpensive to implement, adequate sensitivity requires proper training and supervision of technicians. This poses problems for both medical care providers in impoverished regions of the world as well as laboratory settings which may benefit from automation of a tedious and time-consuming task. Conceivably, automation of this task could both facilitate laboratory efficiency as well as provide an alternative diagnostic tool in conjunction with mobile phone based microscopy in developing countries. Also experts are needed to do parasite detection, who may not be always available. So having automatic malaria parasite detection system is advantageous.

B. Thick smear

The thick smear of correct thickness is the one through which newsprint is barely visible. It is dried for 30

minutes and not fixed with methanol. This allows the red blood cells to be hemolyzed and leukocytes and any malaria parasites present will be the only detectable elements. However, due to the hemolysis and slow drying, the plasmodia morphology can get distorted, making differentiation of species difficult. Thick smears are therefore used to detect infection, and to estimate parasite concentration.

C. Thin smear

The thin smear images air dry for 10 minutes. After drying, the thin smear should be fixed in methanol. This can be done by either dipping the thin smear into methanol for 5 seconds or by dabbing the thin smear with a methanol-soaked cotton ball. While fixing the thin smear, all care should be taken to avoid exposure of the thick smear to methanol.

D. Architecture of Proposed System

Proposed system uses the techniques like image pre-processing, image segmentation, feature extraction. For the study of microscopic blood cell images are captured from thin smear slide after proper staining. It is observed that different image segmentation algorithms can be applied on the obtained colour images, grey images obtained after grayscale conversion and binary images obtained after image binarization. The proposed method is based on two parameters i.e. number of parasite identified within the image and RBC count of the respective image block. From the figure 1 it can be inferred that the proposed scheme is divided into two distinct phases where in the first phase, the image recognition is done after segmenting the image and in the second phase image classification is achieved based infected RBC and non infected RBC.

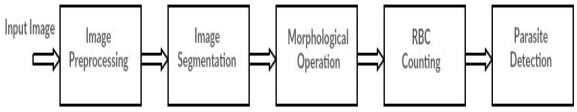


Fig 1. Block diagram of Proposed System

II. METHODOLOGY

The design is essentially an image processing based on thick and thin smear images. The Flowchart for thick and thin blood smear images as shown in figure 2.

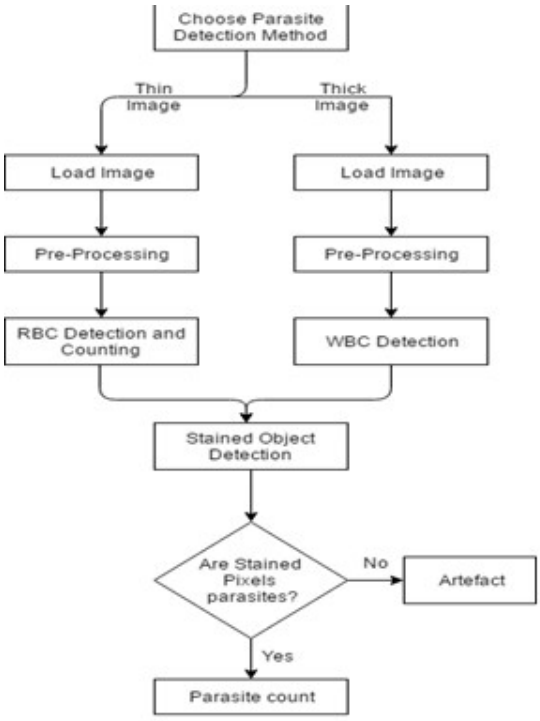


Fig 2. Flowchart on System Response Based on Input

The flowchart is used to depict the flow of work of the given software. It shows the various steps involved from obtaining the input, classification a of the input and the processing it accordingly. It gives a basic gist how the application will work Given below in Figure 2 is the flowchart of the proposed system.

The following steps are performed for thick smear blood images.1) Image Acquisition, 2)Preprocessing, 3)WBC Detection, 4)Segmentation, 5)Parasite Count.

1. Image Acquisition

Each slides are been examined by experts, who have verified and given a species-specific and parasitemia diagnosis. Images were captured using a built-in digital camera microscope. Images were then analysed manually. Images were saved in the JPEG format in 720 x 480 pixels size. These images are transmitted for pre-processing. Figure 3 Show thick blood image with multiple parasite and contain one wbc cell.

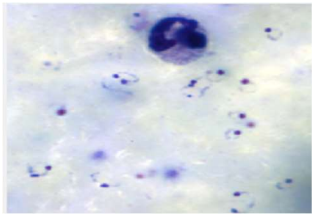


Fig 3.Input image

2. Preprocessing

The aim of preprocessing is to obtain images with low noise, high contrast than original images for the further processing. Due to camera calibration and staining variability of blood smear, changes may occur in illumination and color distribution of blood images. This particular problem poses difficulties for classification of blood cells since it is hard to deal with proper segmentations of objects with quite similar colors. This process contains two operations image enhancement and noise reduction. We have used spatial filtering (median filter) for noise reduction. The median filter replaces pixel value with the median of its neighboring value.

3. WBC Detection

There are several types of WBC, with different morphological characteristics, but every WBC has a nucleous surrounded by cytoplasm. When a thick smear is prepared, the cytoplasm membrane is destroyed, thus the WBC nucleus being only visible. The WBC detection is divided into 2 main tasks: 1) WBC Nuclei Candidates Segmentation: segment all the elements on the image that are possible candidates of WBC nucleus, which stain blue to almost black with Giemsa stain. Figure 4 show the detected wbc cell.

Fig 4. Detected WBC cell

4. Segmentation

The Segmentation of thick smear blood images are done by using k-mean clustering. *k*-means clustering is a partitioning method. The function kmean partitions data into *k* mutually exclusive clusters, and returns the index of the cluster to which it has assigned each observation. Unlike hierarchical clustering, *k*-means clustering operates on actual observations (rather than the larger set of dissimilarity measures), and creates a single level of clusters.

5. Parasite count

The image after segmentation and corresponding contour plot are used for detecting parasites from image. The intensity change in cell dimensions of all the cells is located by scanning its contour plot. Thus we get the total number of infected blood cells in an image. Figure 5 show the number of parasite in given image.

Fig 5. Parasite count

The following steps are performed for thin smear blood images.1)Image Acquisition, 2)Preprocessing, 3)Segmentation, 4)Morphological Operation, 5)Cell Detection .

1. Image Acquisition

In any image processing project, data collection plays an important role. Finding the required datasets is a prime task. Once available data sources are identified, they need to be selected, cleaned, constructed and formatted into the desired form. The database is created by capturing the microscopic images of blood films. As all the cells in blood are very transparent it is stained by using Giemsa stain. This enables to recognize and observe all types of cells. Images are taken from different blood samples. These images are used as raw data for malarial parasite count as shown in figure 6 and shows the number of cells, in which two cell are infected.

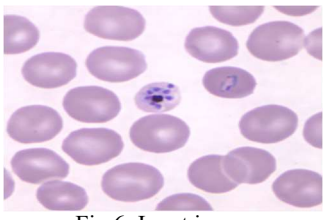
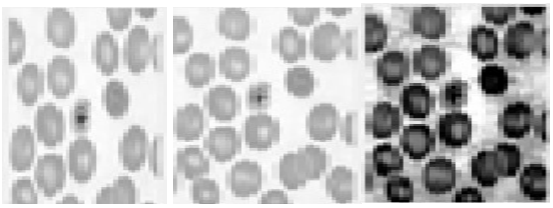


Fig 6. Input image

2.. Preprocessing

The segment of input image pixels is selected for further processing. The input image may have low brightness and contrast. Real-world images are highly susceptible to noisy, missing, and inconsistent data. Low-quality data will lead to low-quality mining results. Hence it is essential to pre-process the data. Pre-processing methods uses small neighborhood of pixel in input image to get the new value of brightness in output image. The different pre-processing methods like normalization, filtering, image plane separation

etc. are used. Here normalization of input image is done as preprocessing method. It is the process that changes the range of pixel intensity. It expands the dynamic range of pixel values in an image into the range in which the image appears more normal. There are a number of pre-processing techniques. In our work, we mainly aim at median filter and histogram equalization. The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing like edge detection on an image.



(a) (b) (c)

Fig 7. Preprocessing Operation performed on thin image

Fig 7. (a) RGB to Grayscale

Fig 7. (b) Filtered Image

Fig 7. (c) Histogram image

3.. Segmentation

Segmentation divides the image into its constituent regions or objects. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries in images. In our study we first remove noise, adjust intensity of the image, perform gray threshold and convert the image to binary form. Segmentation techniques used are sobel, prewitt.

i. Sobel Operator

Sobel operator is used within edge detection algorithms where it creates an image emphasising edges. it is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function. At each point in the image, the result of the Sobel–Feldman operator is either the corresponding gradient vector or the norm of this vector. The Sobel–Feldman operator is based on convolving the image with a small, separable, and integer-valued filter in the horizontal and vertical directions and is therefore relatively inexpensive in terms of computations. On the other hand, the gradient approximation that it produces is relatively crude, in particular for high-frequency variations in the image.

ii. Prewitt Operator

Prewitt operator calculates the *gradient* of the image intensity at each point, giving the direction of the largest possible increase from light to dark and the rate of change in that direction. The result therefore shows how "abruptly" or "smoothly" the image changes at that point, and therefore how likely it is that part of the image represents an *edge*, as well as how that edge is likely to be oriented as shown in figure 8.

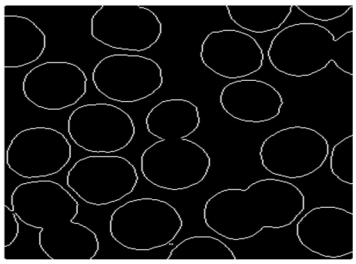


Fig 8. Prewitt segmentation Performed

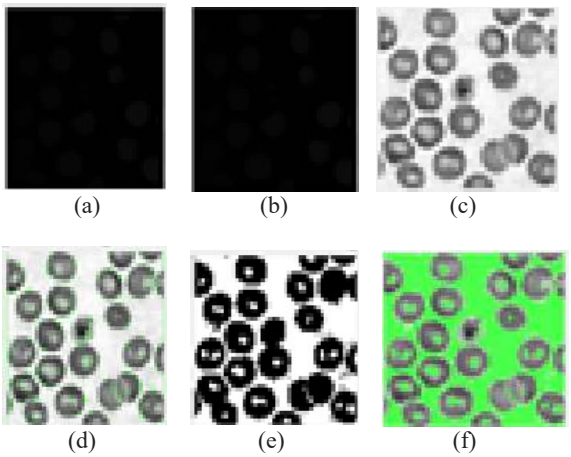
4. Morphological Operation

Morphological operations are image processing operations which processes images based on shapes. It applies a structuring element of specific shape and size on input image. The output image is created by comparing the value of each pixel with its neighbours. These operations are sensitive to the shape of the structuring element. Morphological operations used are dilation and erosion. Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries. The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. In the morphological dilation and erosion operations, the state of any given pixel in the output image is determined by applying a rule to the corresponding pixel and its neighbors in the input image. The rule used to process the pixels defines the operation as a dilation or an erosion. Further operations such as holes filling, overlaying is carried out which helps in detection of infected cells. The

i. Dilation

Dilation is one of the two basic operators in the area of mathematical morphology, the other being erosion. It is typically applied to binary images, but there are versions that work on grayscale images. The basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels (*i.e.* white pixels, typically). Thus areas of foreground pixels grow in size while holes

within those regions become smaller.



(a) (b) (c)

(d) (e) (f)

Fig 9. Morphological Operation performed on thin image

Fig 9. (a) Dialation 9. (b) Holes Filing

Fig 9. (c) Adaptive Histogram

Fig 9. (d) Overlay of Original image

Fig 9. (e) Masked Image

Fig 9. (f) Overlay of Masked image

5. Cell Detection

After all the required image processing techniques are performed, the total number of cells are detected. Here we count the cells having low density and connected using bwlabel() and find maximum number of components that are connected. High density cells are those which are not affected so find such cells having holes. Fill such image region and holes .Repeat the same procedure to get the count of infected cells. Finally superimpose the infected cells onto original image. Figure 10 show two cells are infected.

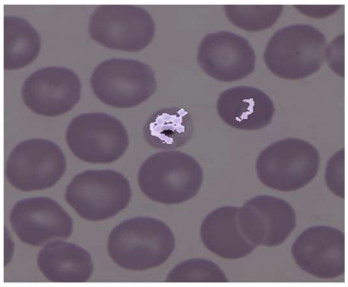


Fig 10. Infected Cell of thin image

Feature extraction starts from an initial set of measured data and builds derived values intended to be informative and nonredundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. In this the RBC counting, Contour formation and Parasite counting is done.

i. RBC Counting

Red Blood Cell (RBC) extraction is a very important and vital step in RBC counting. As there is a possibility of other elements to be present on the smear, only RBCs need to be extracted. RBCs are extracted based on their specific color. The area of one blood cell is calculated in terms of pixels. By calculating the area of all the cells in image and dividing it by area of one cell we get the total number of blood cells.

ii. Contour Formation

If the blood cell is infected, specific intensity changes for corresponding red, green and blue planes is observed. Hence by locating the intensity changes in a cell dimension, infected cells can be found out. Contour plot joins the pixels having same intensity. The contour plots of images are formed from the selected color segment (250 × 250 pixels) of original microscopic image.

iii. Parasite Counting

The image after morphological operation and corresponding contour plot are used for detecting parasites from image. The dimension of each cell is obtained from RBC counting step. The intensity change in cell dimensions of all the cells is located by scanning its contour plot. Thus we get the total number of infected blood cells in an image. As we have the total RBC count and malarial parasite count.

i. Performance measure

Sensitivity is defined as the probability (percentage) that patients with the infection will have a positive result using the test under evaluation. Specificity is defined as the probability (percentage) that patients without the infection will have a negative result using the test under evaluation. The values for sensitivity and specificity are expressed in terms of true positives (TP),

$$\text{Sensitivity} = TP/TP+FN \quad (1)$$

$$\text{Specificity} = TN/TN+FP \quad (2)$$

We have plotted the ROC graph of our results. We have used 10 samples of thick smear and 10 samples of thin smear.

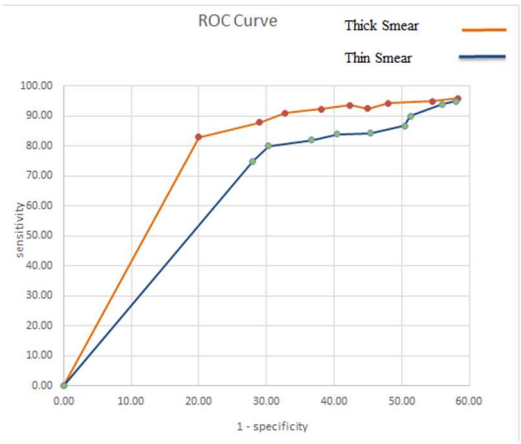


Fig 11. ROC Graph

III. CONCLUSION AND DISCUSSION

From the graph we can see that we have obtained 95% sensitivity for thick smear and 80% sensitivity for thin smear.

The proposed system provides a robust automated system for detection of malaria parasites in thin and thick blood films. The detection of Malaria parasites is done by pathologists manually using microscopes. So, the chances of false detection due to human error are high, which in turn can result into fatal condition. This system curbs the human error while detecting the presence of malaria parasites in the blood sample by using image processing techniques. We achieved this goal by using Image Segmentation, Morphological operations, edge detection technique to detect malaria parasites in images acquired from digitized blood samples. The system acts in a robust manner so that it is unaffected by the exceptional conditions and achieved high percentages of sensitivity, specificity and prediction values It generally validated by comparing with manual analysis. This approach is used in rural areas where less experts are available and the delayed diagnosis may lead to complication in patient's health.

As further enhancement, Support Vector Machine (SVM) techniques can be used to analyze and classify the parasite species based on their shapes.

IV. ACKNOWLEDGEMENT

We have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and organizations. We would like to extend our sincere thanks to all of them.

We are highly indebted to Dr. Satishkumar Varma for his guidance and constant supervision as well as for providing necessary information regarding the project and also for his support in completing the project.

We wish to express our deep thanks towards Head of Department, Dr. Madhumita Chatterjee of department of Computer Engineering and Principal, Dr. R. I. K. Moorthy for making all the facilities available in college required for completion of the project.

We would like to express our special gratitude and thanks to pathology persons for giving us such attention and time.

My thanks and appreciations also go to our colleagues in developing the project and people who have willingly helped us out with their abilities.

V. REFERENCES

[1] F. B. Tek, A. G. Dempster, and I. Kale, "Computer vision for microscopy diagnosis of Malaria," Malaria Journal 2009.

[2] Andrew G Dempster, Izzet kale and F.Boray Tek, "Computer vision for microscopy diagnosis of malaria," 2009.

[3]S. S. Savkare, S. P. Narote "Automatic Detection of Malaria Parasites for Estimating Parasitemia," ICASSP 2009.

[4] IshaSuwalka, AnkitSanadhya, AnirudhMathur and Mahendra S Chouhan, ―Identify Malaria Parasite Using Pattern Recognition Technique‖ GeetanjaliInstitute of Technical Studies Dabok,Udaipur, India,2011

[5]Saowaluck, ApichartIntarapanich, MontriPannarat, Sastrachaotheing, ‖Detectionand Classification Device for Malaria Parasites in Thick-Blood Films‖.In Proceedings of 6th IEEE International conference on Intelligent Data Acquisition and Advanced Computing systems, September 2011. Vol 1 pp. 435-438.

[6] WHO: World Malaria Report 2012 World Health Organization.

[7] K M Khatri, V R Ratnaparkhe, S SBalchandra : Image Processing Approach For Malarial Parasite Identification, International Journal of Computer Applications(IJCA),2013.

[8] Aimi Salihah Abdul-Nasir, Mohd Yusoff Mashor, Zeehaida Mohamed, "Colour Image Segmentation Approach for Detection of Malaria Parasites Using Various Colour Models and kMeans Clustering," 2013.

[9] Pallavi T. Suradkar,"Detection of Malarial Parasite

in Blood Using Image Processing," 2013.

[10] Sanjay Nag, Samir Kumar Bandyopadhyay, "Identification of malaria infection using hsv colour model and dynamic thresholding with image binarization," 2016

[11] Ronald L. Lipsman, Jonathan M: A Guide to MATLAB: For Beginners and Experienced Users : 2nd Edition

[12] Raman Maini Dr. HimanshuAggarwal Study and Comparison of Various Image Edge Detection Techniques.