Detection of diseases via blood analysis using Image processing Techniques

Prof.(Ms).Kajal Jewani

Professor, Computer Engineering Dept.,V.E.S.I.T.

Mumbai University.

Mumbai, India

kajal.jewani@ves.ac.in.

Krishna Boddu, Pratik Gurnani, Kiran Solapure

Computer Engineering, V.E.S.I.T.

Mumbai University.

Mumbai, India

krishna.boddu@ves.ac.in, pratik.gurnani@ves.ac.in, kiran.solapure@ves.ac.in

Abstract—Blood related diseases like Malaria, Leishmaniasis and Acute Leukemia are responsible for the deaths of millions of people each year. Early diagnosis of the disease is necessary for their correct identification and treatment. Malaria, Leishmaniasis and Acute Leukemia are diagnosed by drawing blood sample from the patient's body and observing the thin blood smear under the microscope to check for irregularities. This requires skill and expertise and is prone to human error. The proposed method constitutes an android application which acts as a portable and inexpensive means of diagnosis via image processing and analysis.

***Keywords—Malaria, Leishmaniasis, Acute Leukemia, Sobel edge detector, Harris corner detector, k-means clustering algorithm, Otsu Thresholding*.**

I. INTRODUCTION

Modern medicine has advanced to the point that once deadly diseases like malaria are now easily treatable, and its cure is widely available and affordable. Despite this, an estimated 429,000[3] deaths were caused last year, in part due to delayed diagnoses. Diagnosis requires drawing blood from the patient and analyzing the thin blood smear under microscopic observation. It is dependent on the availability and expertise of pathologists, and is still prone to human error. Moving towards an automated diagnosis provides efficiency, accuracy and higher availability at low expense.

Essential information regarding diseases in question:

1. Malaria is a deadly parasitic disease caused by protozoan parasites of Plasmodium family and is transmitted through the female Anopheles mosquito. In the human body, the parasites go through a complicated life cycle in which they grow and reproduce.

2. Leishmaniasis is a parasitic disease spread by the bite of phlebotomine sand flies. It is caused by parasites of leishmania type. Leishmaniasis can be in three main ways: cutaneous, mucocutaneous, or visceral leishmaniasis.[4]

3. Leukemia is a cancer of blood that usually begins in bone marrow and results in high numbers of abnormal white blood cells.

II. LITERATURE SURVEY

The paper[1] presents the usage of Sobel Edge Detector for highlighting the regions with high spatial frequency that are related to edges. The authors also mentioned the usage of a mathematical operator named Harris Corner Detector for detecting the corner points. The pixel positions of the corner points computed are used for detecting the stages of malaria.

The paper[2] presents highly efficient techniques and procedures over traditional method which requires highly trained pathologist for detection of malaria parasite species

and its life cycle stages. This paper has implemented segmentation of malarial parasite by k- means clustering applied to the a*b component of the L*a*b color space helps in improving the efficiency of the algorithm as features are extracted for only the parasites rather than the entire group of erythrocytes present. This method works with an accuracy of 90.17% and sensitivity of 90.23%.

III. COMPARISON

For both the papers, initial image acquisition and pre-processing techniques used are similar.

In the paper[1], after the preprocessing step, area of the infected cell is found, and then using sobel edge detector the boundary of the infected cell is obtained. The size of the malarial parasite growing from one stage to another stage (Plasmodium Species) increases. For example when a ring trophozoite splits itself to form schizont stage, the size of the latter is bigger than the former. So in this paper, metric, calculated as Harris corner points per area is used to determine the stage of malarial parasite in RBCs. The observation obtained from this paper is that the value of metric obtained in early stage is initially low and is increased in further stages. Detecting the stages of malaria is one of the benefit of using Sobel edge detector and Harris Corner detector. One of the advantage of this method and technique implemented over segmentation is that we can give statistical analysis whether it is true positive, false positive, false negative, true negative which increases overall accuracy of the system.

In the paper[2] ,detection of infected cell is done by applying k-means clustering on the a*b of L*a*b color space. The authors highlight the extraction of feature of the infected cells and avoid rest of the erythrocytes present in the segmented image. 300 images of infected and uninfected sample is given as a dataset to the classifier consisting of 13 classes (12 classes for 3 life cycle stage for each of the 4 malaria species and 1 class for non-infected blood samples).Two classifiers are used KNN and SVM for training the system. The KNN provides more accuracy and sensitivity as compared to SVM.

Table1. Comparison between KNN and SVM

Classifier (%)	Accuracy (%)	Sensitivity (%)
KNN	90.17	90.23
SVM	84.2	89.5

The overall accuracy of the of the system is more which is 97.73% when implemented using Sobel Edge detector and Harris corner detector whereas when implemented segmentation as mentioned in paper 2 it works with an accuracy of 90.17% and sensitivity of 90.23%.

Although the overall accuracy obtained by using Sobel edge detector is more, but the dataset required for analyzing is also more as compared to k mean.

IV. PROPOSED METHODOLOGY

Our process uses k-means clustering on pre-processed image and uses Otsu's method of local thresholding to verify the segmentation. Segmentation performs the goal of removal of background and plasma and separation of infected cells from healthy cells. After thresholding, the image is cleaned up with repeated erosion and dilation with round, disk-shaped RBC like element and further morphological image processing is performed. Hole filling operation is performed to fill the gaps.

K-means Clustering Algorithm: It is an unsupervised learning algorithm that solves the clustering problem. The idea is to classify the given dataset[6][7] through clusters .For this we need to define k centers. The location of the k centers could be arbitrary, but it is good if we keep the centers as far as possible from each other. Next we have to take each point from the dataset and associate it to the nearest center. When there is no point left, it means the step 1 clustering is done. In the next iteration we need to select k new centers as a barycenter of the clusters obtained from previous step, and the same process is continued and we can notice that the k centers change their location until the centers end up at a fixed location. The algorithm aims at minimizing the below equation[5],

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} \left(\left\| x_i - v_j \right\| \right)^2$$

||xi-vi||=Euclidean distance between xi and yi

'ci'= number of data points in 'i'th cluster

'c'=number of cluster centers

Otsu Thresholding: It is use to perform clustering based thresholding. It reduces the gray level image to binary image. It is based on very simple idea of finding the threshold to minimize the inter class variance[8] of thresholded black and white pixels. It operates directly on gray level histogram so it's fast once the histogram is computed. The histogram is divided in two classes and the inter-class variance is minimized. This method involves iterating through all the possible threshold values and calculating the measure of spread of pixel levels on each side of the threshold that is the pixels that either fall on foreground or background then it finds the threshold value where the sum of foreground and background spreads is at its minimum.

V. RESULT ANALYSIS

The flow of detection is implemented as follows for the detection of Ring Trophozoite Stage of Malaria-:

1. Ring Trophozoite Stage of Malaria (Stage 1)

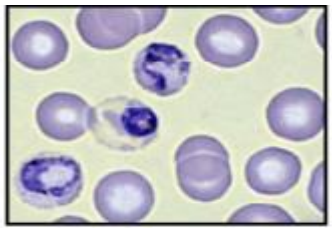

Fig1.Original malarial blood smear

2. Original image is binarized to get the next image

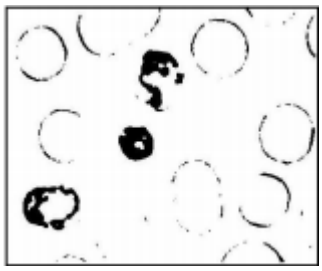

Fig2. Binarized Blood smear

3. Inclusion bodies detected inside RBCs are shown in Pink

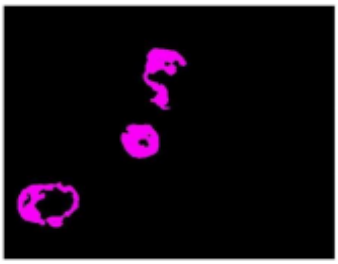

Fig3. Blood smear of malaria after transforming it to RGB

4. After this the inclusion bodies are shown along with the background RBCs

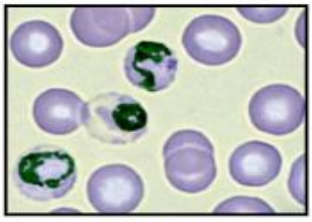

Fig4. Infected corpuscles are shown along with the background of smaller ones

5. Inclusion bodies are edge detected by SOBEL method

Fig5.Infestation of RBCs detected by SOBEL Operator

6. Harris corner point is applied on those bodies

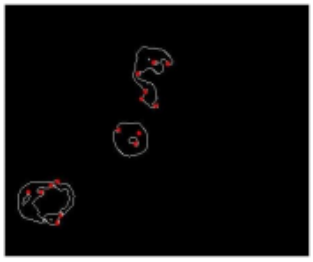

Fig6.Harris corner points detected on infestation

Statistical Analysis-:

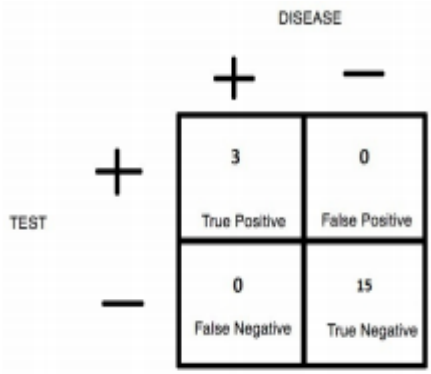

Table2: Results obtained for different blood samples based on parasite count obtained-:

sample	RBC count in the sample	True +ve(a)	True -ve(b)	False -ve(c)	False +ve(d)	Accuracy (a+b)/ (a+b+ c+ d)* 100	Overall accuracy of the system
1	18	3	15	0	0	100%	
2	13	1	12	0	0	100%	
3	11	2	9	0	0	100%	97.73%
4	13	2	10	0	1	92.30%	
5	12	0	12	0	0	100%	
6	17	2	14	1	0	94.117%	

In this paper[2] segmentation algorithm is tested on 120 images of malaria infected thin blood smear parasite. Features such as entropy, Hu's moments, GLCM features, GRLM features are obtained from the segmented image. A total of 90 features are extracted for all the processed images. Extracted features are ranked according to their F-static values obtained by applying one-way ANOVA. Feature set to train the classifier is reduced by only considering the most significant features as ranked by their F-static value.

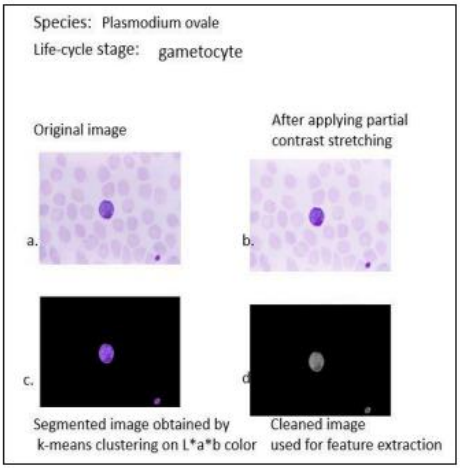

Fig7. Result obtained from k means clustering segmentation

VI. CONCLUSION

Sobel edge detector method provides superior results, however it requires an extremely large dataset for accurate solutions. K-means clustering is more complex than Sobel method in implementation, but requires a relatively smaller dataset. We will be implementing segmentation via k-means clustering and verifying using an alternative method to ensure accuracy.

VII. REFERENCES

[1] Detection of presence of Parasites in Human RBC In Case of Diagnosing Malaria Using Image Processing
Published in: Inventive Computation Technologies (ICICT), International Conference on
Date of Conference: 26-27 Aug. 2016
Date Added to IEEE Xplore: 19 January 2017

[2] Detection of presence of parasites in human RBC in case of diagnosing malaria using image processing
Published in: Image Information Processing (ICIIP), 2013 IEEE Second International Conference on
Date of Conference: 9-11 Dec. 2013
Date Added to IEEE Xplore: 09 January 2014

[3] Fact Sheet: World Malaria Report 2016
World Health Organization
ISBN 978-92-4-151171-1
Publication date: 2016

[4] Leishmaniasis Fact sheet No. 375
World Health Organization
Publication date: January 2014

[5] An Efficient k-means Clustering Algorithm: Analysis and Implementation
Tapas Kanungo, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman and Angela Y. Wu

[6] R. Donida Labati, V. Piuri, F. Scotti, "ALL-IDB: the acute lymphoblastic leukemia image database for image processing"
2011 IEEE Int. Conf. on Image Processing (ICIP 2011) pp. 2045-2048, September 11-14, 2011
ISBN: 978-1-4577-1302-6.

[7] Malaria Parasite Images – Isfahan MISP Datasets
Medical Image and Signal Processing (MISP) Research Center

[8] Otsu Thresholding – The Lab Book Pages
Dr. Andrew Greensted, June 17, 2010