

AUTOMATIC DETECTION OF MALARIA PARASITES FOR ESTIMATING PARASITEMIA

TEJASHRI CHAUDHARI¹, Prof.D.G.AGRAWAL²

1(P G Student, Department of Electronics and Communication Engineering S S G B COE & T, Bhusawal, Maharashtra, India)

2(Professor, Department of Electronics and Communication Engineering S S G B COE & T, Bhusawal, Maharashtra, India)

Abstract

Malaria is the leading cause of morbidity and mortality in tropical and subtropical countries. Conventional microscopy used in diagnosis of the disease has occasionally proved inefficient since it is time consuming and results are difficult to reproduce. Alternative diagnosis techniques which yield superior results are quite expensive and hence inaccessible to developing countries where the disease is endemic. In this Project, an accurate, rapid and affordable model of malaria diagnosis using stained thin blood smear images was developed. The method makes use of the morphological, colour and texture features of Plasmodium parasites and erythrocytes. Images of infected and non-infected erythrocytes were acquired, pre-processed, relevant features extracted from them and eventually diagnosis was made based on the features extracted from the images. Diagnosis entailed detection of Plasmodium parasites, differentiation of different Plasmodium parasite stages. Image pre-processing entailed reducing the size of the acquired images to speed up processing and median filtering to remove salt and pepper noise. Neural network classifiers were then trained and used to detect and determine the life stages and species of Plasmodium parasites. Classification accuracy of 100%, 92%, 91%, and 85% for detection of infected erythrocytes, stages determination, respectively was achieved with respect to results obtained by comparing different feature extraction techniques and different classifiers. In this project we have used three classifiers SVM Naïve bayes and neural network classifier and two feature extraction techniques DWT and GLCM feature extraction technique. The study revealed that artificial neural network (ANN) classifiers trained with colour, morphological, and texture features of infected stained thin blood smear images are suitable for detection and classification of Plasmodium parasites into their respective stages and species.

Keyword- *Feature Extraction, SVM Classifier, DWT, GLCM, Neural network, Navie Bayes, MATLAB, Malaria, RBC, Trophozoite, Schizonts, Gametocytes*

I. INTRODUCTION

Malaria is a serious disease caused by a blood parasite named Plasmodium spp. It affects at least 200 to 300 million people every year and causes an estimated 3 million deaths per annum. Diagnosis and medication of it is necessary. In blood sample visual detection and recognition of Plasmodium spp is possible and efficient via a chemical process called (Giemsa) staining. The staining process slightly colorizes the RBCs but highlights Plasmodium spp parasites, white blood cells (WBC), and artifacts. Giemsa stains nuclei, chromatin in blue tone and RBCs in pink color. It has been shown in several field studies that manual microscopy is not a reliable screening method when performed by non-experts. Malaria parasites host in RBCs when it enter in blood stream.

In Malaria parasitemia count it is important step to segment RBCs from blood image and classify it as parasite infected or normal. In thin blood images morphology of cells can be observed clearly. The present paper describes the techniques used in segmenting normal and infected RBCs for purpose of Malaria parasitemia (number of infected blood cells over total red blood cell) count the system architecture which includes pre-processing, cell segmentation, RBCs segmentation, feature extraction and classification.

Malaria is a common but serious protozoan disease caused by peripheral blood, spleen or liver parasites of the genus Plasmodium.

There are four species of Plasmodium parasites that cause malaria in humans. These are Plasmodium

falciparum, Plasmodium ovale, Plasmodium vivax and Plasmodium malariae. These species of Plasmodium attack red blood cells and undergo various life stages namely, early trophozoites, mature trophozoites, gametocytes and schizonts.

It is estimated that approximately 781,000 people of the 225 million people infected worldwide by the disease succumb to this menace annually. Majority of these deaths are children from sub-Saharan Africa In Kenya malaria accounts for 30 – 50% of all outpatient attendance and 20% of all admissions in health facilities. It is also estimated that 20% of all deaths of children under five years is due to malaria The disease causes a heavy economic burden to those affected in terms of the costs incurred to treat the disease and absenteeism from work and school.

The diagnosis of the disease requires powerful and expensive tools unavailable for the poorest countries of the world, where often the disease is endemic. Microscopic malaria diagnosis is, by far, considered to be the most effective diagnostic method, but it is highly time-consuming and labour intensive. The accuracy of the system solely depends on the expertise of the microscopist . Other techniques widely involved in Malaria diagnosis are Rapid Diagnostic Tests (RDTs) and Polymerase Chain Reaction (PCR) tests . However, the accuracy of these tests depends on the extent of infection with sensitivity directly proportional to the level of infection. Various automated malaria related diagnostic studies are described in.

II. LITERATURE SURVEY

Zoueu et al. (2008) [1], proposed a method of diagnosing malaria without labelling the parasites using a light microscope with LEDs emitting in the range of UV to IR replacing the classic white light. The microscope was fitted with a digital camera to capture images formed at the eye piece. It was reported that parasite images having the best contrast were recorded for blue light. This study, however, failed to address the effects of chromatic aberration which are common in multispectral imaging using classical optics. Chromatic aberration causes images of specimen in a light microscope to be formed in different focal points for different wavelengths of light used to illuminate the specimen. This would therefore pose a challenge in automating the detection and classification of Plasmodium parasites because adjustment of focal point would be required

for every wavelength used for illumination. The technique was also dependent of a human operator for switching between LEDs and making the diagnosis

Brydegaard et. al.(2011) [2], proposed an improved version of multispectral microscope based on light emitting diodes (LEDs). The LEDs emitted lights in 13 spectral bands ranging from ultra-violet (UV) to near infra-red (IR). The dispersive optical components of the instrument were made of quartz to reduce achromatic aberration and lens fluorescence in illumination profiles towards the ultraviolet region. The device was also fitted with an imager for capturing images in the 13 spectral range of the LEDs. The instrument was interfaced to a computer and switching of LEDs and image capturing was done using LAB-VIEW software. It was reported that the instrument could detect Plasmodium parasites in non-stained thin blood smear images. However, species and stages differentiation of Plasmodium parasites was not addressed.

Minh-Tam Le et. al. [3] ,proposed a comparison-based analysis, which differentiates solid components in blood smears. The semiautomatic method uses statistical measures and cross referencing validations yields a reliable detection scheme. The nucleated components are identified using adaptable spectral information. Cells and parasites are isolated from the background, by comparing the input image with an image of an empty field of view. The range of erythrocyte sizes is determined by input of isolated RBC.

Jesus Angulo et. al. [4], presents a technique to automatically detect the working area of peripheral blood smears stained with Giemsa. The approach consists of two stages. First, an image analysis procedure using mathematical morphology is applied for extracting the erythrocytes, the centers of erythrocytes and the erythrocytes with center. Second, the number of connected components from the three kinds of particles is counted.

D. Ruberto et. al. [5] follow morphological method for detection of parasites in Giemsa stained blood slides. Different objects in blood are identified using their dimensions and color. The parasites are detected by means of an automatic thresholding based on morphological approach, using Granulometrics to evaluate size of RBCs and nuclei of parasite. A

segmentation method using morphological operators combined with the watershed algorithm.

Silvia et. al. [6], proposed a technique for estimating parasitemia. Template matching is used for detection of RBCs. Parasites are detected using variance-based technique from grayscale images and second approach is based on color co-occurrence matrix. Support Vector Machine (SVM) as the classifier which exploits the texture, geometry and statistical features of the image.

III. OBJECTIVES:

The aim of this research was to develop a system that would offer speedy and accurate malaria diagnosis in human blood media based on the colour and morphological features of Plasmodium parasites and infected erythrocytes.

The specific objectives were as follows:

1. To develop suitable algorithms for:

- a) Detecting Plasmodium parasites
- b) Classifying the parasites into their life stages

2. To Compare the accuracy :

- a) Performance Accuracy using different classifier
- b) Stage classification accuracy with different feature extraction technique.

IV. SYSTEM ARCHITECTURE

1. Image Size Reduction and Noise Reduction:

In order to develop a standard algorithm to process images from both sources, all images were rescaled to have the same size. Since no enlargement was needed for images, the size of their spatial resolution (256 by 256 pixels) which corresponded to the minimum size of images was adopted to be the standard image size of the pre-processed images. Filtering operation using a median filter was performed to the rescaled images from both sources to reduce noise in the images.

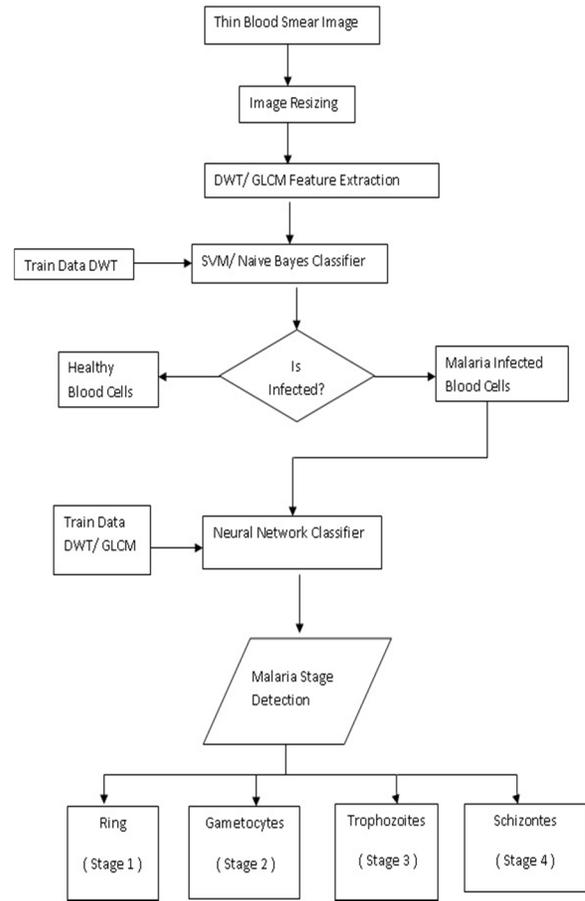


Figure 1: Proposed Model of Malaria Diagnosis System

2. Feature Extraction

After an image has been segmented, the resulting aggregate of segmented pixels usually is represented and described in a form suitable for further computer processing. Representing and describing regions involves two choices. An object can be represented and described in terms of its external characteristics (its boundary) or in terms of its internal characteristics (the pixels comprising the region).

A. DWT Feature extraction:

The low frequency components (smooth variations) constitute the base of an image, and the high frequency components (the edges which give the detail) add upon them to refine the image, thereby giving a detailed image. Hence, the smooth variations are demanding more importance than the details. Separating the smooth variations and details of the image can be done in many ways. One such way is

the decomposition of the image using a Discrete Wavelet Transform (DWT).

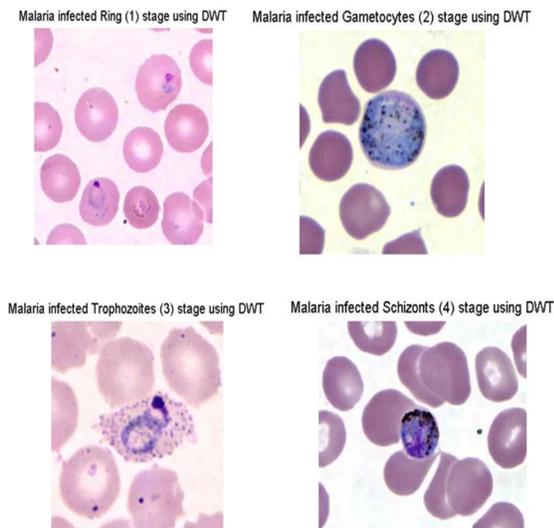


Figure 2: Stage classification using DWT

B. GLCM Feature Extraction:

In statistical texture analysis, texture features are computed from the statistical distribution of observed combinations of intensities at specified positions relative to each other in the image. According to the number of intensity points (pixels) in each combination, statistics are classified into first-order, second-order and higher-order statistics. The Gray Level Cooccurrence Matrix (GLCM) method is a way of extracting second order statistical texture features.

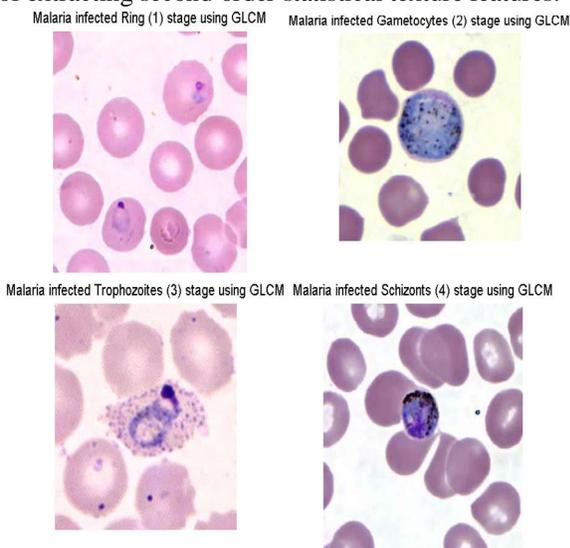


Figure 3: Stage classification using GLCM

3. Support vector machine (SVM) Classifier:

Support Vector Machine (SVM) is used in this paper for classification of the affected or not affected images by malaria. The statistical learning theory provides a framework for studying the problem of gaining knowledge, making predictions, making decisions from a set of data. In simple terms, it enables the choosing of the hyper plane space such a way that it closely represents the underlying function in the target space.

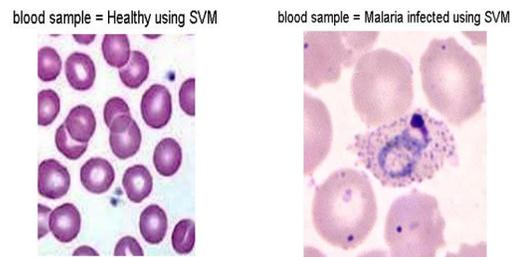


Figure 4: Detection of malaria using SVM Classifier

4. Artificial Neural Networks (ANN):

Artificial Neural Network (ANN) has been employed together with image processing techniques to automate the diagnosis of malaria using the morphological features of erythrocytes in blood images. Artificial Neural Networks (ANN's) have proven to be a promising paradigm for Intelligent Systems. Neural networks have been trained to perform complex functions in various fields of application including pattern recognition, identification, classification, speech, vision and control systems. ANN's have the advantage of learning by example and the ability to generalize from their training data to other data. They are fault tolerant in the sense; they can produce correct outputs from noisy and incomplete data. ANN's are relatively inexpensive to build and train.

5. Naive Bayes classifier:

A Naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem (from Bayesian statistics) with strong (naive) independence assumptions. A more descriptive term for the underlying probability model would be "independent feature model".

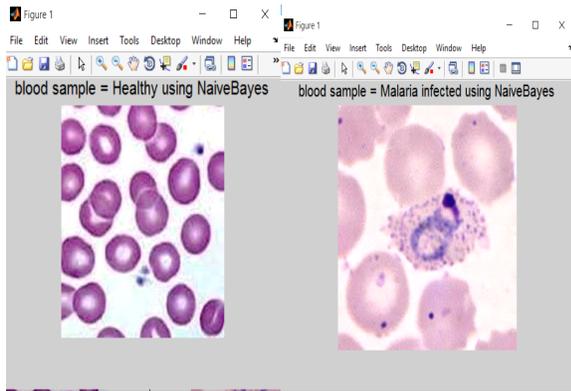


Figure 5: Detection of malaria using Naive Bayes Classifier

The naive Bayes classifier is designed for use when predictors are independent of one another within each class, but it appears to work well in practice even when that independence assumption is not valid. It classifies data in two steps:

1. **Training step:** Using the training data, the method estimates the parameters of a probability distribution, assuming predictors are conditionally independent given the class.
2. **Prediction step:** For any unseen test data, the method computes the posterior probability of that sample belonging to each class. The method then classifies the test data according the largest posterior probability.

V. RESULT AND PERFORMANCE ANALYSIS

The performance of the malaria diagnosis system was evaluated in four main stages: detection of *Plasmodium* parasites, classification of *Plasmodium* parasites life stages,. In each of these stages, the predictions made by different classifiers and different feature extraction techniques were compared and from them the accuracy was obtained.

Detection of *Plasmodium* parasites was done by using two types of classifiers

- 1 SVM Classifier
- 2 Naïve Bayes Classifier

The network was trained with RGB color features as this was found to be the most accurate method of detecting *Plasmodium* parasites. The network searched through the images and identified images infected by *Plasmodium* parasites. DWT feature extraction technique was used for these classifier.

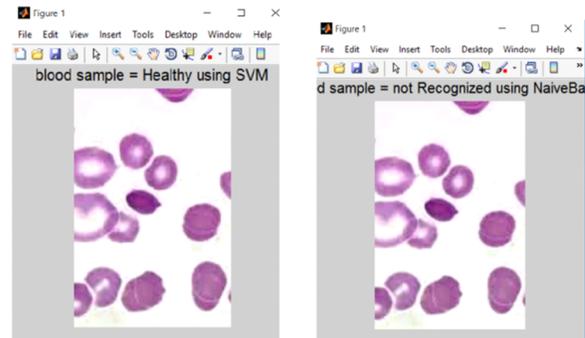


Figure 6: Comparison of same blood sample with Different Classifier to detect the accuracy

Classifier Type	Accuracy (%)
Naive Bayes	92.85
SVM	100

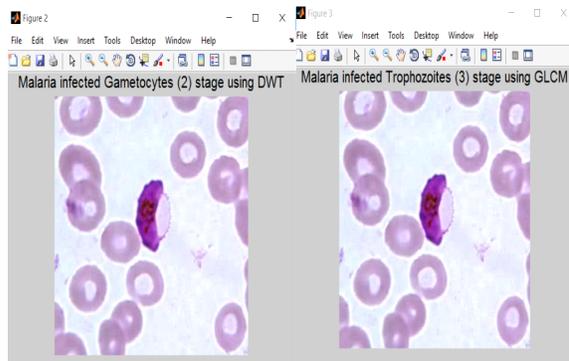


Figure 7: Comparison of Performance Accuracy

Feature Extraction	Accuracy (%)
DWT	91.667
GLCM	85.4166

V. APPLICATION AND FUTURE SCOPE

Android Application to Diagnose Malaria

The application enables the user to capture the image using the mobile phone camera and saves it into the gallery of the phone. Because of the larger size and memory requirements, the image has to be resized. Usually the image capture using the state of the art smart phones has around 20 Mpixel of size. Hence it will undergo down sampling in order to speed up the process. The user interface developed will display the

results of the diagnosis along with statistical data of total number of RBCs, WBCs, infected RBCs, life stages along with the location of pixels and the Parasitemia.

Separate display option for closer analysis of detected RBCs, WBCs and infected cells are also provided along with display of closed image and ratio transformed images. For precise information, the results will be stored in a folder, with the same name of the input RGB picture, in the mobile phone SD card to be available at any time. The format of the stored pictures is PNG to avoid loss of compression and the file is stored in TEXT format in order to be used in all the platforms.

V. CONCLUSION

- A system for detecting Plasmodium parasites, identifying their life stages and comparison of their accuracy using different types of classifier and different feature extraction techniques using images of thin blood smears stained with Giemsa was developed.
- A total of 96 images were used to train and test the performance of the system. The system recorded 100% accuracy in detecting the presence of Plasmodium parasites using SVM Classifier and 92.85% accuracy in detecting the presence of Plasmodium parasites using Naïve Bayes Classifier.
- 92% accuracy was obtained while recognizing stages using DWT Feature Extraction using Neural Network classifier and 85.41% accuracy in recognizing stages using GLCM Feature Extraction using Neural Network classifier.
- It was therefore concluded that detection of Plasmodium parasites can be adequately performed by a SVM classifier. On the other hand, stages recognition to be adequately classified using Neural Network classifiers. Using DWT feature extraction technique.

VI. REFERENCES

- 1) Coarentin Dallet *, Saumya Kareem*, Izzet Kale*“Real Time Blood Image Processing Application for Malaria Diagnosis Using Mobile Phones” 2014 IEEE

- 2) Vishnu V. Makkapati and Raghuveer M. Rao “SEGMENTATION OF MALARIA PARASITES IN PERIPHERAL BLOOD SMEAR IMAGES”, „IEEE Transaction 2009 IEEE
- 3) S. S. Savkare and S. P. Narote “Automatic Detection of Malaria Parasites for Estimating Parasitemia”India International Journal of Computer Science and Security (IJCSS), Volume (5) : Issue (3) : 2011
- 4) Pallavi T. Suradkar “Detection of Malaria Parasite in Blood Using Image Processing” ‘International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 10, April 2013
- 5) Matthias Elter, Erik Haßlmeyer and Thorsten Zerfaß “Detection of malaria parasites in thick blood films”,Annual International Conference of the IEEE EMBS Boston, Massachusetts USA, August 30 - September 3, 2011
- 6) M. Tam Le, T. Bretschneider,“A novel semi-automatic image processing approach to determine Plasmodium falciparum parasitemia in Giemsa-stained thin blood smears”, Research article, BMC Cell Biology, 28 March 2008.
- 7) J. Angulo, G. Flandrin, “Automated detection of working area of peripheral blood smears using mathematical morphology”, U. S. National Library of Medicine, Analytical Cellular Pathology 25(1), pp 39-47, 2003.
- 8) C.D. Ruberto, A.G. Dempster, S. Khan, and B. Jarra, "Automatic Thresholding of Infected Blood Images Using Granulometry and Regional Extrema", in Proceedings of International Conference on Pattern Recognition, pp 3445-3448, 2000.
- 9) S. Halim et al., “Estimating Malaria Parasitaemia from Blood Smear Images”, in Proceedings of IEEE international conference on control, automation, robotics and vision, pp 1-6, 2006.

- 10) S. Osowski et al., "Application of Support Vector Machine and Genetic Algorithm for Improved Blood Cell Recognition", in proceedings of IEEE transaction on Instrumentation and Measurement, Vol. 58, No. 7, pp 2159-2168, July 2009.
- 11) WHO: Global report on antimalarial efficacy and drug resistance: 2000-2010
- 12) Anna Rosanas-Urgell, Dania Mueller , Inoni Betuela, Céline Barnadas, Jonah Iga, Peter A Zimmerman, Hernando A del Portillo , Peter Siba, Ivo Mueller and Ingrid Felger , " Comparison of diagnostic methods for the detection and quantification of the four sympatric Plasmodium species in field samples from Papua New Guinea" Malaria Journal 2010, Volume 9, www.malariajournal.com.
- 13) S.Kareem,R.C.S Morling, I Kale "A Novel Method to Count the Red Blood Cells in Thin Blood Films"IEEE International Symposium Circuits and Systems (ISCAS), 2011, Pages 1021-1024
- 14) S.Kareem, I Kale, R.C.S Morling "Automated Malaria Parasite Detection in Thin Blood:- A Hybrid Illumination and Color Constancy Insensitive, Morphological Approach"
- 15) S.Kareem , I .Kale, R.C.S Morling ,"A Novel Fully Automated Malaria Diagnostic Tool Using Thin Blood Films" Pan American Health Care Exchanges (PAHCE2013), Medellin, Columbia, May 2013
- 16) S.Kareem, I Kale, R.C.S Morling "Automated P.falciparum Detection system for Post-treatment Malaria Diagnosis using Modified Annular Ring Ratio Method" IEEE 14th International Conference on Computer Modelling and Simulation, March 2012
- 17) World Health Organization, (2011), "World Malaria Report 2011", World Health Organization Geneva. Available <http://www.who.int/malaria>
- 18) Perandin, F., Manca, N., Calderaro, A., Piccolo, G., Galati, L., Ricci, L., and Chezzi, C. (2004). Development of a real-time PCR assay for detection of Plasmodium falciparum, Plasmodium vivax, and Plasmodium ovale for routine clinical diagnosis. Journal of Clinical Microbiology, 42(3), 1214-1219.
- 19) <http://www.kemri.org>, 1/10/2011.
- 20) http://www.dpd.cdc.gov/dpdx/HTML/ImageLibrary/Malaria_il.htm, 1/10/2011.
- 21) Bansal, R.K., Goel, A., and Sharma, M.K., (2009). MATLAB and its Applications in Engineering, Pearson Education South Asia, New Delhi, India.
- 22) Storey, J., (2010), Basic Malaria Microscopy: Part 1. Learner's Guide, World Health Organization, Geneva.
- 23) Hagan, M.T., Demuth, H. B., and Beale, M., (2010), Neural Network Toolbox™ 6, The Mathworks, Inc. 3 Apple Hill Drive Natick, MA 01760-2098