

Detection of Noisy Blood Images Indicating Prodromal Diseases

Sohni Muneer

*Institute of Information and Communication Technology
Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan
Email- sohani.muneer@outlook.com*

Tariq Jamil khanzada

*Institute of Information and Communication Technology
Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan
Email- tariq.khanzada@faculty.muet.edu.pk*

Samiullah Jatoi

*Institute of Information and Communication Technology
Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan
Email- Sami_jatoi@hotmail.com*

Abstract- *Malaria is an entirely preventable and treatable disease caused by mosquito, but still it is one of the main causes of increasing illness and instigate fatality rate in both developed and especially in the underdeveloped world. Malaria is a serious infection, disease which needs to be recognized and controlled and a quick diagnosis is required. It is caused by Plasmodium parasites that can harm human red blood cells to a great extent. To improve the accuracy of the detection or diagnosis of the infection, numerous researches have explored diverse remedies. An image pre-processing technique used to enhance the quality of detection of malaria parasite in the infected blood cells. A well trained staff and tools are required for medical diagnosing, which is accomplished using digital image processing to detect and identify the presence of the parasites in RBC's images. By applying image processing techniques the microscopic blood clot images can be analysed more efficiently and accurately. We present a method to detect the infected malaria parasites for four plasmodium types applied to 165 various images of the dataset. Our experiments utilize diverse images and the results achieved the Sensitivity of 90 %, PPV of 90% and the overall accuracy of 88%. Which is better than the reported cases the best of our knowledge.*

Keywords – *Malaria, RBC, image processing, RGB to gray scale, Feature extraction, decision tree classifier.*

I. INTRODUCTION

Malaria is a mosquito borne disease and it's typically spread to the human beings by the bite of female mosquitoes Anopheles. It is widely spread in tropical and subtropical climate of the region. As claim by Center for disease control (CDC) approximate 214 million malaria cases are reported in 2015 [1] and children are majorly affected in the Africa region. This infection can be mild or even severe that can lead to fatality of the victim. Microscopes are utilized to detect the infection by pathologists which is a manual process and the chances of false detection are very high. In case of any false detection, the disease normally converts to a more serious state.

There are four types of Plasmodium parasites that can infect and cause malaria in human beings these are

1. P. falciparum
2. P. ovale
3. P. vivax and
4. P. Malaria

Most of the common symptoms are headache, vomiting, body shivering and fever and which start to manifest 7-10 days after the bite of the infected mosquito [2]. However, we have a large portion of malaria positive cases of falciparum or vivax in the blood test. Falciparum is the most critical part from the plasmodium family and it may lead to death [3]. The collection of infected parasites in human blood samples is known as parasitemia. It is most important

to identify the parasitemia in malaria and medication because by identifying parasitemia one can get the intensity of the disease [4]. Determination of malaria through blood smear examination is difficult and dull and the chances of obtaining an inaccurate result are very high [5]. In this research, we propose a method for diagnosis of infected parasites to expel the human mistake, recognizing the existence of prodromal infected parasites in the blood test through image processing and automation. It is a similar investigation of two techniques adversary distinguishing parasites, the first strategy depends on Segmentation and the second uses features extraction utilizing least separation of a classifier. We construct the parasites identification framework which is an efficient way to the goal that it is unaffected by the remarkable situations and accomplishes high rates of compassion, specificity, positive expectation and negative estimate values. Testing of the blood samples through the microscope is time-consuming and error prone especially when we have a large number of samples and moreover required reliable and efficient analysis. So it is significant to improve the automate image analysis. Various methods have been proposed for detection of malaria and several classification techniques has been used. Some of them are Naive bayes Classifier, Minimum Distance Classifier, Artificial Neural Network (ANN). ANN is used to classify and recognize a pattern of an object and solve pattern recognition problems with a neural network. Minimum Distance Classifier work when the distance between means of different classes is large. but here we use Decision tree Classifier is applied to classify infected blood cell images from normal blood cells. In Training phase we have binary classification process that carried through decision tree.

II. RELATED WORK

Numerous works has been reported on the scope of identifying plasmodium parasite using thin blood smear images. In this section a number of previous work is reviewed and compared.

Diaz et al., [6] Proposed an approach to identify and for the assessment the life stages of the parasite. For the perception and deviation of parasite life stages, train classifiers are used. The study described a sensitivity of 74% for stage identification and 94% of the infected blood cells. The claim that the technique is completely automatic was rejected because of human interposition during classifier's training and the diagnosis had to be performed every time.

Wongsakorn Preedanana., 2016 [7] Detection of infected parasites are implemented on classification, automatic segmentation, feature extraction, thresholding and morphological opening remove small objects. Watershed method is used to remove overlapping cell and the overall accuracy is 97 %.

Monlica wattana, 2017 [8] presented the procedure in which the red blood cell images are adjusted to the brightness level by enhancing the partial contract stretching techniques. This method is appropriate for infected cell image segmentation and as a result the complete and clear cell images are attained.

According to J.Somasekar [9] two stage thresholding method for segmentation of infected parasite cells in microscopic RBC images for helping diagnosis were used. The infected parasites are identified by the proposed thresholding techniques that contains maximizing the difference of an original image and thresholding selection from one image for segmentation of malaria parasites.

Mihar Bhatt, Shashi prabha in 2015 [10] presented the procedure to count RBC abnormalities efficiently in a minimum time. The main approach of this method is to overcome the time limitation. In this case the medical fields take high time that may lead to a death to very dangerous disease such as malaria. To find out the abnormalities in RBC's form factor threshold is applied.

Yonathan Ferry Hendrawan [11], presents a paper based on segmentation and detection of infected malaria parasites using with color-based cascading method. The Method starts with color normalization procedure, it is followed by gamma correction, after that exposure compensation, noise reduction, and edge enhancement, fussy c-mean clustering, and last morphological operation is performed.

III. PROPOSED METHODS

The aim of this research is to detect the infected malaria parasite cells from non-infected one. Automatic diagnosis of microscopic blood cell can be an effective analytical tool and enhances the accuracy as well. The basic steps are to implement to detect the infected malaria parasites which as the following the repository. The First images are loaded into the system from CDC (Laboratory Identification of parasites of Public Health Concern) and then pre-processing and resizing is performed. After that they are converted into gray scale images, furthermore, we have segmentation, feature extraction and dilation is performed. Finally, detection of parasites is performed by using the classifier. **Figure 1** defines the steps involved in detecting plasmodium parasites.

3.1 Data Acquisition

The images of thin Giesma Stained blood smears have been acquired from online dataset library “CDC (Laboratory Identification of parasites of Public Health Concern)”. 165 images size of 300*300 is assumed from the CDC website [12], and another dataset from Parasitology Laboratory, Faculty of medicine, Indonesia is considered. There is a lot of variation between images due to heterogeneous structures of the dataset.

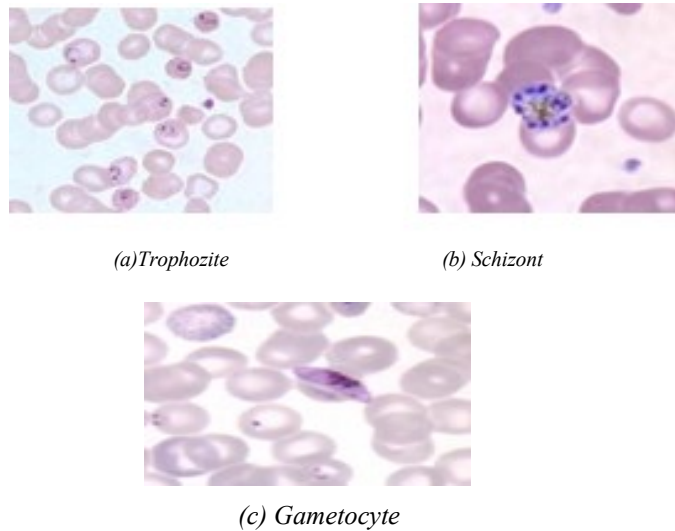


Figure 1. Stages Plasmodium species

3.2 Image Pre-processing

Where an image is loaded the main step is image pre-processing to enhance the visual appearance of images and to improve the manipulation of the dataset. The enhancement technique can emphasize image artifacts or even leads to a loss of information if not correctly performed [13]. Pre-processing is exerted to achieve these objectives. Firstly, the image is resized for the purpose of digital zooming and to speed up the processing speed. Secondly, the acquired image is further enhanced by reducing the noise and then intensified the image contrast is performed for better visual analysis. Image processing is performed on low resolution images. Wiener method is used to remove noise by adapting thresholding to remove slight unwanted areas from the image to achieve the clear view of the region of interest. The obtained image is forwarded for further processing.

3.3 Gray Scale Conversion

After the pre-processing the conversion of grayscale is performed on the resultant image. Gray scale image is then converted to binary image. All grayscale algorithms utilize the basic three-step process:

Step 1. Get the red, green, and blue value of pixels

Step 2. Use a fancy math to turn those numbers into a single gray value.

Step 3. Replace the original red, green, blue values within the new gray value.

The next step is about morphological opening in which background is converted into foreground pixels by filling holes and then the noise is removed by using `bwareaopen`. This operation is used to remove small objects whose value is less than 300 pixels.

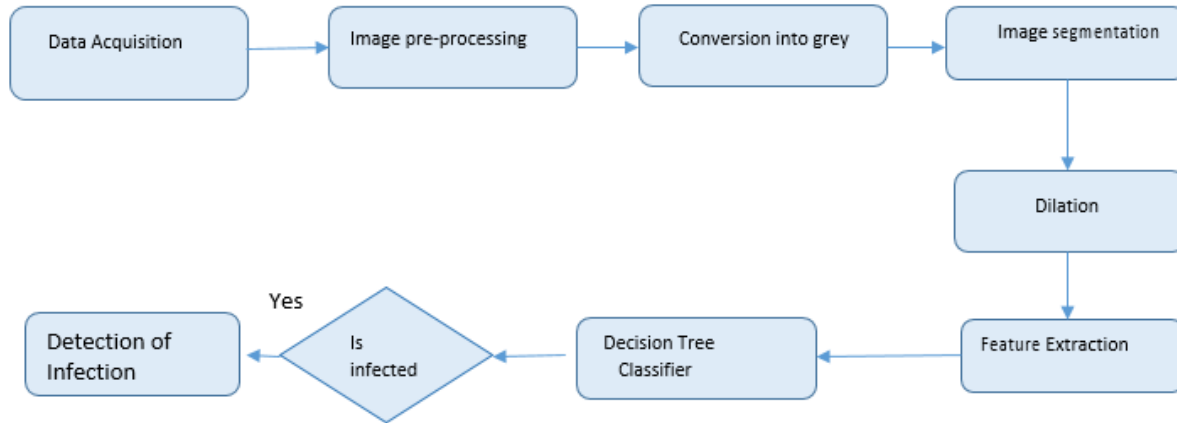


Fig. II. Elaborated Steps

The standard size of all binary images is about 300 pixels. The result of these images is:

The next step is described below is segmentation.

3.3 Image Segmentation

Image Segmentation is a process of partitioning a digital image into meaningful regions with the respect to particular a application. The segmentation is based on the measurement taken from the image which may be Gray level, texture, color, motion or depth. For cell segmentation two steps should be performed: (1) Removing artifacts (2) Filling the holes. Artifacts are the objects that remain as foreground pixels in the resulted photo for thresholding. Thresholding is the simplest method of image segmentation. From a grayscale image thresholding can be used to create binary images. Assumptions for thresholding the intensity values are different in different regions. Within each region it represents the corresponding object in a scene, where the intensity values are similar. Local adaptive thresholds select an individual's thresholding for each pixel based on the range of intensity values in its local neighborhood, then closing operator is applied in order to fill all holes in the cells that occur because the parasites have different grayscale levels from normal cells.

The next step is to perform dilation which is useful in detection of disease.

3.4 Dilation

The basic operation in mathematical morphology is dilation. It is also known as erosion for binary images it is develops gradually and the foreground pixels of boundary region are enlarged .The size of the foreground pixel area grows while region within holes become smaller [14].

3.5 Rbc/Feature Extraction

Feature extraction is an important stage to represent the identity of different staging of blood cells. We identify the feature set which can give prominent difference between normal and infected cells and use it for training data. Feature extraction shows different stages so that one can easily identify various stages of red blood cells. We recognize the list of feature set and it gives prominent difference between infected and normal cells and it is utilized for preparing the data set [15]. The first order has statistical analysis which includes mean, standard deviation, entropy, minimum value and smoothness calculated from an image. These operations are defined in subsequent sections.

1) Mean

Mean is the normal value of all pixel intensities in an image. It is characterized by eq (1).

$$\text{Mean} = \sum_{i=0}^{L-1} i \cdot x(i) \quad (1)$$

Where $x(i)$ is the probability function.

The median of the image is identified with the mean and well defined as the central value of all pixels. Hence if all pixel values of image are arranged, the central value can be the median. The mode of an image is the pixel value occurring with greater frequency.

2) Standard Deviation

The standard deviation is called as a variation of an image. The variety of the intensity of the image pixels is measured to acquire the variance of the image. The variance of the image is defined by eq (2):

$$\text{S.D} = \sqrt{\sum_{i=0}^{L-1} (i - b)^2 \cdot x(i)} \quad (2)$$

Where (b) is the mean value and $x(i)$ is the probability function.

3) Entropy

The entropy value shows the complexity of the acquired image. The greater the value, the greater the complexity. It is important to realize that entropy and energy tend to be inverse. Moreover, the entropy value denotes the amount of confined information in the data distribution it is defined by eq (3).

$$\text{Entropy} = - \sum_{i=0}^{L-1} x(i) \log_2(x(i)) \quad (3)$$

Where $x(i)$ is the probability function.

4) Smoothness

Smoothness is incorporated to measure the amount level of smoothness/unpleasantness of the image intensity. It is calculated as shown in eq (4)

$$\text{Smoothness} = 1 - \frac{1}{1 + \sigma^2} \quad (4)$$

Confusion Matrix:

		Abnormal	Normal	Total
Positive	Yes	100	10	PPV 90 %
Negative	No	10	5	NPV 66 %
Total		Sensitivity: 90 %	Accuracy: 88%	

Table1 : Texture Feature Extraction

Data	Total No: cell	Total No: parasites	Mean	Dev	Entrop	Smths
T1	38	15	0.5258	0.1625	6.4826	0.3438
T2	630	25	0.4092	0.2279	7.6672	0.0338
T3	35	14	0.5933	0.1420	6.5322	0.3870
T4	22	8	0.6207	0.0991	6.3437	0.0048
T5	28	36	0.4305	0.2505	7.3020	0.0806
T60	42	5	0.4966	0.1819	7.0116	0.2297
S1	27	5	0.6348	0.1372	4.8815	0.0055
S2	65	24	0.4097	0.2050	7.5916	0.0163
S3	17	11	0.6274	0.1067	5.7330	0.0048
S4	13	6	0.5657	0.1381	5.5097	0.0055
S65	32	3	0.6279	0.1145	5.7443	0.0129
G1	20	36	0.5740	0.1016	4.7473	0.0102
G2	186	12	0.6507	0.1006	6.3924	0.100
G3	8	25	0.6399	0.1649	7.1783	0.0265
G4	72	47	0.5624	0.1249	6.7134	0.3047
G5	32	3	0.6279	0.1145	5.7443	0.0055
G70	18	27	0.6928	0.1084	6.3823	0.5009
G71	6	184	0.5993	0.1543	7.1153	0.233

Mean: calculated mean Dev:standard deviation

Entropy: entropy Smths: smoothness

T1: Trophozite stage of date number 1

S1: Schizont stage of date number 2 and G1: Gametocyte stage of data number3

Detection of plasmodium parasite is done by using classified cells applying classifiers. From these plasmodium parasites classifiers is accountable for analysis that either blood smear images are infected or non-infected. The data will be trained with the features extracted from the infected red blood cells. The algorithm searches and identifies the infected regions of the images and detect the plasmodium parasites. Detection of Plasmodium parasites is totally related to image taxonomy. RBC's are now existing a ring shape malaria parasite on red blood cells. This method counts the total number of cells as shown in figure 4 and finds 48 red cells and parasites are 15.

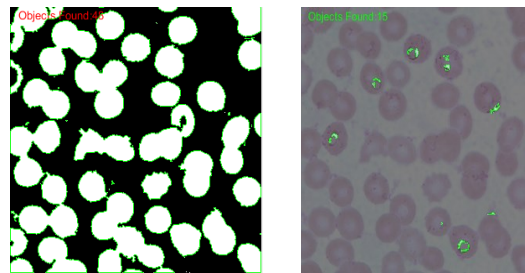


Fig III. (a) Total red cells

(b) Total parasite

IV. Classification

This is the last and essential step where the features of the input image are compared with the decision tree classifier and the infected RBC image. For this, the classifier is trained provides the features extracted in the last step. Then the classifier would be trained to enable it to classify the input images into normal and infected RBC image. To classify the further images into different types of parasites. This is the way how the presence of malarial parasite would be detected in the human blood. The persistence of the life cycle step of the malarial parasite will be useful to provide proper anti-malarial medication. As definite drug amount varies according to the appropriate step. This work will be done using MATLAB16 for the coding and complete processing.

V. RESULTS

In this study the algorithm is tested for 165 images of infected malaria parasites of thin blood smear images. These images then proceed using Matlab 2016 so according to the proposed algorithm and then for each of the image infected parasite images are obtained. Here we used in our system design can be explained with some parameters to test and compare the results such as Sensitivity, PPV (Positive Predictive value), NPV (Negative Predictive value), Accuracy. The Formula given below.

$$\text{Sensitivity} = (TP) / (TP + FP) \quad (5)$$

$$\text{PPV} = (TP) / (TP + FP) \quad (6)$$

$$\text{NPV} = (TN) / (TN + FN) \quad (7)$$

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (8)$$

Where, True Positive is the number of infected parasite images correctly identify as a parasite. False Positive is the number of non-infected parasite images correctly identify as a parasite. True Negative is the number of infected parasite images identify as non-infected parasite. While False Negative is the number of infected parasite images correctly identify as non- infected parasite.

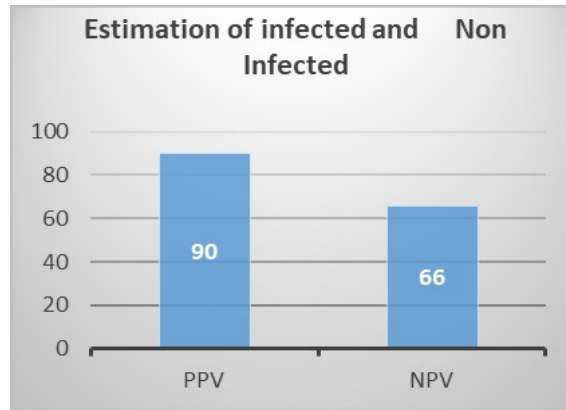


Fig. IV. The Total number of parasites detected

The Figure given above shows the average of infected and non-infected cells being estimated. Whereas, PPV shows positive predictive value as estimated 90%, NPV shows negative predictive value being calculated 66%. These results are taken into consideration of about 165 images.

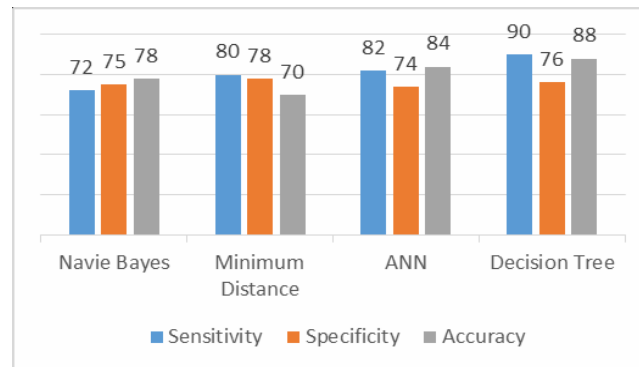


Fig V. Comparison of proposed method's accuracy with other methods based on no. of images.

Clearly shows that our proposed method to succeed in order to compare all the other current methods are applied to detect the infected parasite detection of red blood cells. The accuracy is comparatively high on increasing no: of images. Minimum distance classifier, Naive Baye's and ANN have lower accuracy in order to compare methods [16].

VI.CONCLUSION

In this paper the detection of infected malaria parasites through decision tree classifier by using microscopic blood images. Matlab tool gives a better platform for algorithm verification the identification of the malarial parasites in the red blood cells need more certainty and for this the features would be considered an important part in it. It will give more certainty when we have more number of features, this method will differentiate the stages schizonts, trophozoite and gametocytes on the basis of the features[17]. It can be further used for the diagnosis of hematological problems and make changes based on the features of parasites. Classifier is being used with some data along with these feature implementation has been performed on almost 165 images after that the sensitivity of this proposed system is 90%. PPV is 80%. The Overall accuracy of the system is 88%. The system would be online to diagnose malaria future work could be done will develop an android app in the future.

REFERENCES

- [1]. Saowaluck Kaewkamnerd, Apichart Intarapanich, "Detection and Classification Device for Malaria Parasites in Thick-Blood Films" in IEEE International Conference in 15-17 Sept 2011, Prague, Czech Republic
- [2]. Miss. Kshipra and C.Charpe, "Automated Malaria Parasite and there Stage Detection in Microscopic Blood Images", International Conference on Intelligent Systems and Control (ISCO) vol.9, 2015 India.
- [3]. Hanung Adi Nugroho, Son Ali Akbar,"Feature Extraction and Classification for Detection Malaria Parasites in thin blood smear" Proc. Of 2015 2nd Int. Conference on information technology, Computer and Electrical Engineering, Indonesia Oct 16-18th.
- [4]. Dr. Arora, " Medical Parasitology", Second edition, CBS publisher and distributors, chapter 5-sporozoa , malaria parasite, pp 67-80.
- [5]. Adedeji Olugboja, Zenghui wang." Malaria Parasite Detection using Different Machine Learning Classifier", Proceeding of the 2017 International Conference on Machine learning and Cybernetics, Ningbo China, 9-12 July, 2017.
- [6]. Pargorn Puttapirat 2016,"Automatic Segmentation of Erythrocytes from Giemsa-Stained thin Blood Films", IEEE, International College, King's Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand.
- [7]. Wongsakorn Preedanana , Montri Phothisonothai,"Automatic Segmentation of Plasmodium Falciparum from Giemsa-stained thin blood smear", Institute of Technology Ladkrabang, Bangkok, Thailand 2016.
- [8]. Monlica Wattana," Improvement of complete malaria cell image segmentation", The twelfth International Conference on Digital Information Management (IDDIM 2017),September 12-14,2017, Kyushu University, Fukuoka
- [9]. Abnormal blood cells using image processing technique. Electronic and telecommunication engineering .IJEET, Volume 07, Issue 01, Jan-June.
- [10]. Yonathan Ferry Hendrawan, "Color Image Segmentation of malaria parasites Detecting using cascading method", 2017 International Conference on Sustainable Information and Technology (SIET), Indonesia.
- [11]. Ary Kusuma Ningsih,"An overview of malaria Identification Techniques for Microscopic blood images", 2016 International Conference on informatics and computing ICIC, Yogyakarta, Indonesia.
- [12]. A. Ms. Deepali Ghatge, "Automatic Detection Of malaria parasite from blood images", International Journal of advanced computer Technology,IJACTI,Volume4,Number1.
- [13]. Hanung Adi Nugroho 2015," Feature Extraction and Classification for Detection Malaria Parasites in Thin Blood Smear", (ICITACEE), Indonesia, Oct 16-18th.
- [14]. W.H organization (2016,August) Fact sheet World Malaria Report 2016.
- [15]. Kotepui M and Piwkharn D, "Effects of Malaria Parasite Density on Blood Cell Parameters", National Center for Biotechnology Information (NCBI) 10.1371/journal.pone.0121057, March 2015.
- [16]. Rahmat Romi S Wulandari, F & Faza, S & A Muchtar, M & Siregar and Ikhsan, "The Morphological classification of normal and abnormal red blood cell using self-Organizing Map", IOP Conference Series: Materials Science and Engineering 2015.
- [17]. Yogang Dai and Haosheng sun, "the Naïve BayesText classification algorithm based on rough set in the cloud platform", Journal of chemical and Pharmaceutical Research, 2014.