Automated Enumeration of Malaria Parasite Using SVM Classifier

[1]Sayali N. Dharpal, [2]Dr.A.V. Malviya

[1]M.E. Student, [2]Assistant Professor
Department of Electronics and Telecommunication Engineering
Sipna College of Engineering and Technology, Amravati, Maharashtra, India.

***Abstract*: The current highest standard level of malaria diagnosis is the manual, microscopy-based analysis of stained blood smears or in other words devouring procedure requiring skilled specialists. This paper introduces an algorithm that recognizes and counts red platelets (RBCs) and additionally stained parasites with the end goal to perform a computation of malaria parasite. This proposed framework help to build up a totally automated system for classification of malaria parasite contaminated erythrocytes are segmented from the pre-processed images. Statistical and color features are extracted and given to the SVM binary classifier which describes Malaria infected erythrocytes on blood smears. The outcomes indicate 98.98% sensitivity and 97.02% accuracy for detecting infected red blood cells.**

Index Terms: **Malaria, Parasite, erythrocyte, Feature Extraction, SVM classifier.**

I. INTRODUCTION

Malaria is one of most broad transmittable diseases and a huge society health emergency. Malaria is serious parasite infection and critical reason for death normal in humid and subtropical atmosphere of territory. As indicated by World Health Association (WHO) likely to happen in around 3.3 billion cases. Intestinal sickness is a basic illness spread through mosquitoes. The significant reason behind malaria is a parasite considered Plasmodium that contaminates female anopheles mosquitoes. Tainted individual shows different therapeutic appearances from mild to genuine ones that may cause to death. Intestinal sickness is a serious worldwide disease; on the off chance that it has been not treated or recognized from the underlying stage it could be more basic or at some point may prompt death. The parasites passing through a complicated life cycle inside the human body, utilizing the red platelets as hosts. There are lots of techniques to identify malaria, among them manual microscopy is viewed as "the gold standard". Anyway in view of the different advances basic in manual estimation, this analytic method takes excessively time.

Intestinal sickness diseases are recognized physically by pathologists who observes the microscopic images of strained blood records on glass slides and ascertain the contaminated platelets. There is a chance to happen human mistake, so computer based characterization utilizing digital image processing techniques gives preferred result over the manual judgments of malaria. Intend of this work is to develop a recognition technique to accurately detect malaria parasites present in images.

This paper discusses a fully automatic system for area of tainted erythrocytes from blood. Statistical and color features are extracted for preparing of SVM classifier. System is supervised. The paper is organized as follows section 2 summarizes literature related to detection of Malaria parasite infected cells, classification of blood cells. Section 3 describes system architecture which includes preprocessing, segmentation, feature extraction and SVM classifier Section 4 and 5 includes results and conclusions of this paper.

II. RELATED WORK

A number of studies on the possibility of detecting plasmodium parasites using images of thin blood smear have been done in the past. In this section a number of these studies are reviewed.

In [1], the authors proposed a system for automatic detection of malaria parasite from blood images. This system employs image segmentation techniques to detect malaria parasites from images acquired from Giemsa stained peripheral blood samples.

Another comparative study, [2], has presented an enhanced technique for Malaria Parasite Detection. This technique employed cell segmentation, where the segmentation process consists of various steps, such as image binarization using Poisson's distribution based minimum error thresholding, followed by morphological opening for the purpose of refinement. Furthermore, in [3], detection of Plasmodium infected erythrocytes, was investigated using a technique called holography.

Ross et al., [4] proposed a technique for automating malaria diagnosis using light microscopy. Here, a light microscope fitted with a digital camera was used to capture image of Giemsa stained blood slides. After images were captured they were loaded to a Personal Computer (PC) for processing. Image processing techniques and neural network classifiers were used. Infected erythrocytes were positively identified with a sensitivity of 81% while the accuracy for species determination was 73%.

Morphological image processing techniques used for erythrocyte segmentation could not produce satisfactory results for erythrocytes which are heavily clustered [5]. The sizes of erythrocytes were determined using granulometry with circular structuring element (SE). The assumption was erythrocyte shapes are circular. This is not always the case. Sometimes erythrocytes shapes are deformed especially if they are infected with diseases such as sickle cell or if they appear in clusters.

Diaz et al., [6] developed a technique for detection, quantification of parasitemia and parasite life stages. Pixels' color features were extracted and used to train classifiers for detection and determination of parasite life stages. Clustered erythrocytes were resolved by use of template matching before parasitemia was estimated. The study reported a sensitivity of 94% for detection of infected erythrocytes and 79% for stages identification. The technique was not fully automatic as it called for human intervention during training of the classifier every time diagnosis had to be made.

Di Ruberto et al., [7] proposed a technique of automatically detecting and quantifying malaria parasites infection in blood images of patients. The method employed a modified watershed algorithm to segment erythrocytes. There were two alternatives proposed for classifying parasite stages. One was the use of morphological thinning, where skeletons of parasites images were used to categorize parasites into their respective stages of infection. The second option was use of colour histograms similarity. The efficiency of the segmentation algorithm proposed reduces with the degree of clustering of erythrocytes. Similarly the accuracy of colour histogram similarity for classification of parasites would depend on the imaging parameters and illumination conditions under which the image being probed is taken. The detection accuracy of parasitemia reported was relatively low, 50%.

Most of the techniques proposed in the previous works didn't address the distinction of plasmodium parasites from the rest of stained objects (artefacts) in the blood sample. In this work, this limitations of the previous works have been addressed. A novel method of segmenting erythrocytes and plasmodium parasites using Support Vector Machine (SVM) has been developed. This technique has solved the problem of distinguishing between the plasmodium parasites and other stained objects (artefacts) in images of thin blood smears.

III. SYSTEM ARCHITECTURE

The system model is implemented is shown in Figure 1.Initially an image is obtained and pre-processed, it is then divided into various locales and feature extracted. Next, an appropriate classifier is utilized to categorize the features into their diverse classes. At last, a decision is made about the data passed on by the image based on the classes of features found by the classifier. Figure below gives a block diagram of the recognition of plasmodium parasites in thin blood smears.

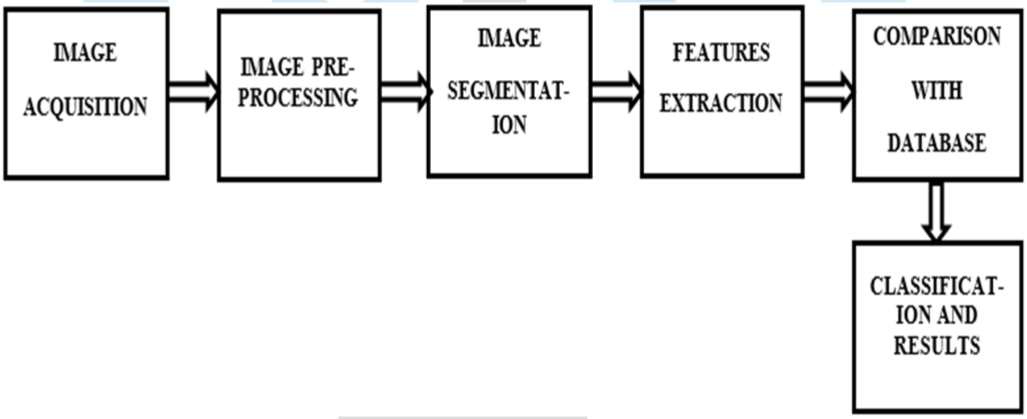

Fig -1: Block Diagram of the recognition of plasmodium parasites

A. Image acquisition

Thin blood smear images were acquired from the Centre for Disease Control (CDC) website [8] and captured from the Reference Laboratory of Malaria, in Sudan Ministry of Health.

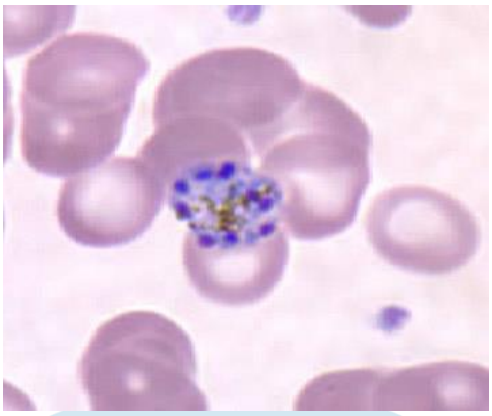

Fig -2: Input Image sample

B. Image pre-processing

The motivation behind pre-processing is to expel unwanted objects and noise from the image to facilitate image segmentation into significant areas, the histogram equalization is shown in fig-3. CLAHE of an input image.

In this paper, we used bilateral filter for smoothening of color picture and In range filter is used for edge sharpening. This result is subtracted from unique to upgrade the image. The bilateral filter is a non-linear digital filtering technique, used to remove noise from images. In bilateral filtering each pixel is supplanted by a weighted normal of its neighbors. This point of view is basic since it makes it easy to secure instinct about its conduct, to modify it to application-specific prerequisites. It depends just on two parameters that demonstrate the size and complexity of the features to protect as appeared in fig-4. Bilateral Filtering of input image. In-range filtering is a methodology that cleans up appearances and takes into consideration particular featuring of particular data shown in fig-5. After pre-processing, image is send to cell segmentation block to segment cells.

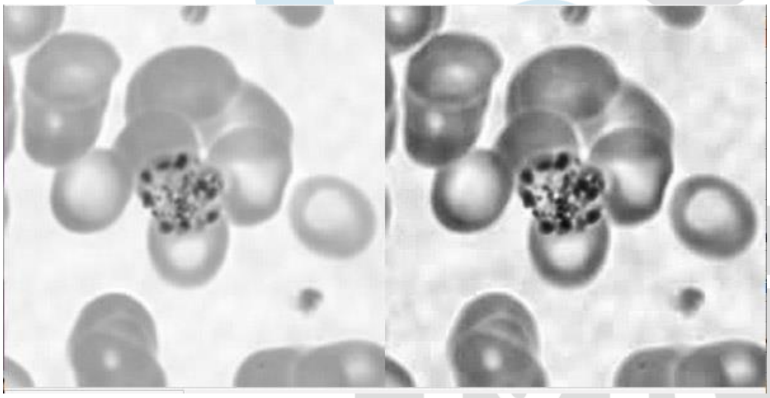

Fig-3: CLAHE of an input image

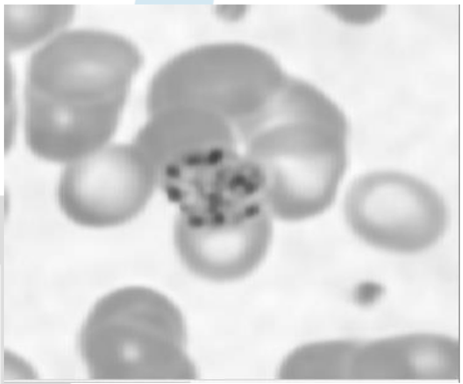

Fig-4: Bilateral Filtering of input image

C. Image Segmentation

Image segmentation is the major advance to analyze images and extract information from the preprocessed image appeared in fig-5. Image segmentation is a mid-level preparing method used to break down images and can be characterized as a handling strategy used to classify or cluster a image into a few disjoint parts by gathering the pixels to form a district of homogeneity dependent on the pixel qualities like gray level, color, texture, intensity and other features [9], [10]. The reason for the segmentation process is to get more data about the locales of interest in an image, which helps in explanation of the object scene. The primary objective of segmentation is to clearly separate between the object and the background in an image.

In this paper, superpixel calculation is used in which superpixels total pixels into perceptually critical atomic locales that can be used to supplant the inflexible structure of the pixel grid in images as appeared in fig-6. In this way, image primitives and redundancy can be decreased incredibly. It is more advantageous and powerful to compute image features based on regions than pixels.

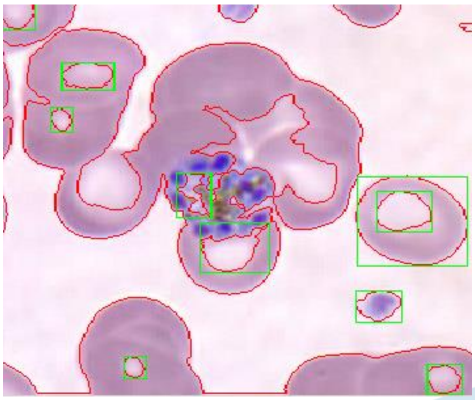

Fig -5: Post-processed image

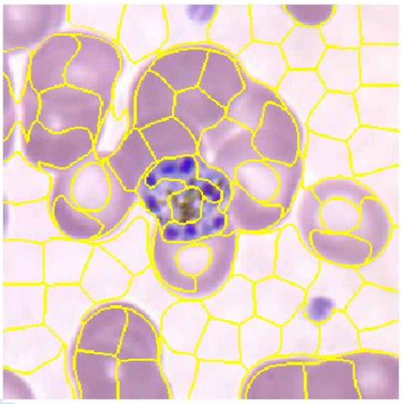

Fig -6: Superpixel segmentation of input image

D. Feature Extraction

The essential destinations of feature extraction are lessening the computational complexity of the consequent procedure and facilitating a reliable and precise recognition for unknown novel information, being the last objective especially important for computer vision and pattern recognition systems. Also, the in-depth understanding of the domain-specific knowledge.

In this work, the input image is partitioned into small regions, from which LBP histograms are extracted, and the neighborhood histograms are additionally connected into a spatially upgraded feature vector. Likewise, a few varieties even augmentation the feature vector length radically, for instance, Expanded LBP, VLBP and Gabor Wavelets based LBP. The determined LBP-based element vector outfits an over-total portrayal with repetitive data [10], which could be diminished to be more limited and discriminative. Additionally, when assembling real-time systems, it is similarly wanted to have LBP-based representation with diminished feature length. For every one of the reasons, the issue of LBP include choice has as of late been addressed in many literatures.

Run based Methodology Uniform example is a suitable control to pick LBP features, and it has been received in existing work. There are moreover unique guidelines which could be used. For instance, Lahdenoja et al. [11] proposed a symmetry level arrangement for uniform examples to furthermore reduce the length of LBP include vectors. The symmetry level is turn invariant according to the definition. The most symmetric example contains a comparative number of zeros, demonstrating a symmetric edge, while the precedents with the slightest symmetry level are the ones involving just ones or zeros. It is affirmed that the examples with high symmetry level happen simply more a significant part of the time in the pictures with more discriminative power [11].

E. SVM Classifier

A Support Vector Machine (SVM) performs classification by constructing an N-dimensional hyper plane that optimally separates the data into two categories. SVM models are closely related to neural networks.

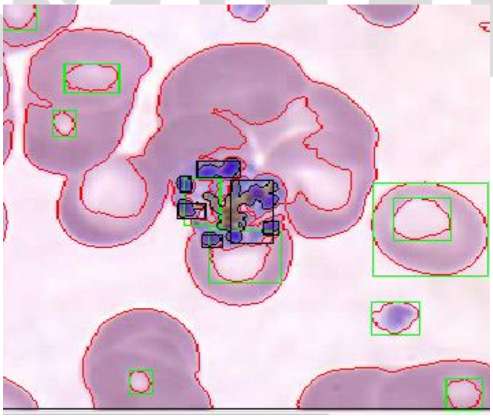

Fig -7: Detected parasite infected cells

Support Vector Machine (SVM) models are a close cousin to classical multilayer perceptron neural networks. Utilizing a kernel work, SVM's are an training strategy for polynomial, radial basis function and multi-layer perceptron classifiers in which the weights of the system are found by tackling a quadratic programming issue with direct requirements, instead of by comprehending a non-convex, unconstrained minimization issue as in standard neural system preparing.

In the SVM literature, an indicator variable is called a attribute, and a changed attribute that is utilized to characterize the hyper plane is known as a feature. The task of picking the most appropriate representation is known as feature selection. An arrangement of features that depicts one case (i.e., a row of predictor values) is known as a vector. So the objective of SVM modeling is to locate the ideal hyper plane that isolates clusters of vector so that cases with one classification of the objective variable are on one side of the plane and cases with the other classification are on the other size of the plane. The vectors close to the hyper plane are the support vectors.

In this paper, it has been utilized for the acknowledgment and characterization of cells. The fundamental favorable position of the SVM network utilized as a classifier is its great generalization capacity and to a great degree intense learning technique, prompting the worldwide least of the characterized error function. Direct SVM is a straight discriminant classifier working on the principle of most extreme edge between two classes. Results are appeared in Figure-7.

IV. RESULT ANALYSIS

The performance of the employed classification method, was evaluated using the following statistical measures: (a) the accuracy (b) the sensitivity and (c) the specificity. The sensitivity of a test is defined as the probability of a positive test result (or presence of the symptom) given the presence of the disease.

The sensitivity is given by , sensitivity = $\frac{TP}{TP+FN}$ x100 (1)

The specificity of a test, , is defined as the probability of a negative test result (or absence of the symptom) given the absence of the disease.

The specificity is given by , specificity = $\frac{TN}{TN+FP}$ x100 (2)

Finally, the accuracy is how close a measured value is to the actual (true) value.

The accuracy is given by , accuracy = $\frac{TP}{TP+TN}$ x100 (3)

where TP, TN, FP and FN denote the true positive, the true negative, the false positive and the false negative, respectively. The calculation process is shown in Table I.

	Abnormal	Normal	Total
Positive	98(TP)	1(FN)	99
Negative	2(FP)	3(TN)	5
Total	100	4	104

Table I- PERFORMANCE EVALUATION OF CLASSIFICATION

From Table I, It is clear that from the performance evaluation, the accuracy is 97.02%, which means that the SVM classifier gives more accurate result for the data used in this study.

The performance of classifier is defined by the feature used to train the classifier. The results for the experiment are given in table. For the malaria images database the result obtained are as follows.

Method	1.Ross, N.; Morphological and novel threshold selection	2. Arco, J.; Morphological operations,Adaptive histogram equalization	3. Das et al; Mathematical morphology (MM)	4. Khan, M.; Two-stage tree classifier using back propogation feed forward neural networks	5.Kareem, S.; Dilation, erosion	6.Proposed Method
Accuracy	73%	96.46%	88.77%	81%	88%	97.02%
Sensitivity	85%	_	99.72%	85.5%	90%	98.98%

Table 2: Shows accuracy and sensitivity of algorithm for different methods of detection

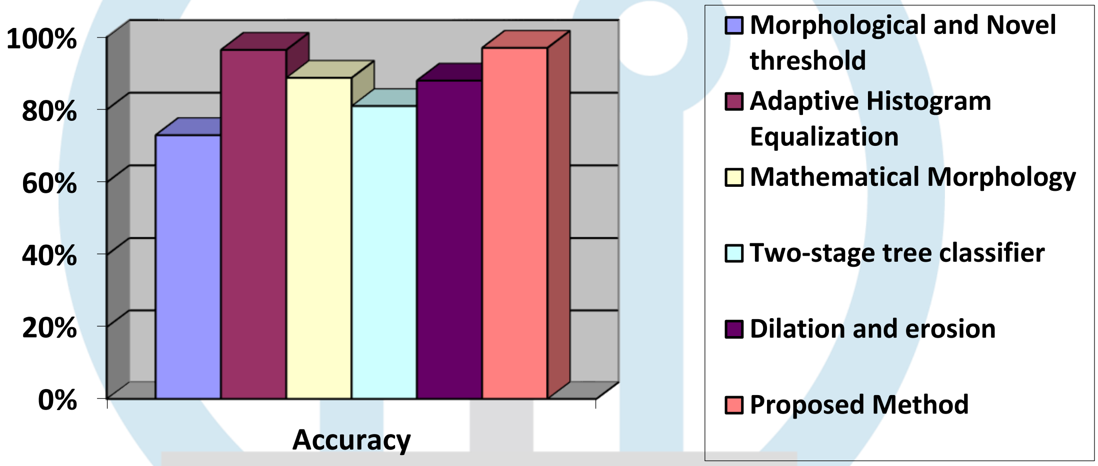

Fig -8: Graphical comparision of Accuracy for different methods of detection

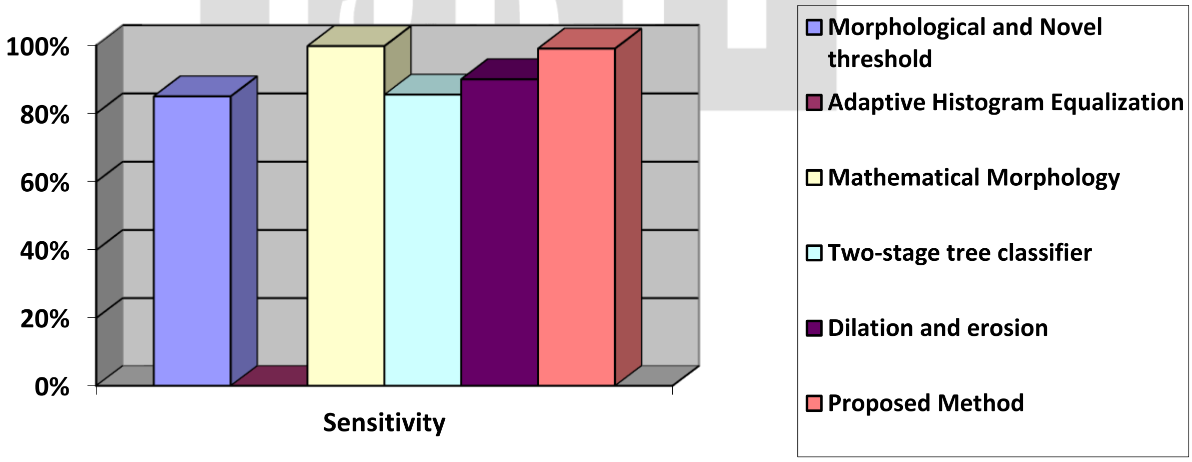

Fig -9: Graphical comparision of Sensitivity for different methods of detection

The performance of system is defined by a classifier used with existing set of a database. The result for the experiment is given in table which shows the accuracy and sensitivity of algorithm for different methods in terms of percentage as shown in Table 2.

The use of parameter extracted and other method work good with ANFIS, Morphological operations, Mathematical morphology, two-stage tree classifier better with dilation and finally best by using SVM as classifier which is shown in fig-8 and fig-9 respectively.

V. CONCLUSION

A system for detecting Plasmodium parasites was implemented. The images utilized in this work were gathered from various sources, at that point the images were handled and certain features were extracted. These features were then used to distinguish the presence of the malaria parasite. In addition, a graphical user interface has been designed to facilitate the use of the system. A total of 104 erythrocytes sub-images were utilized to prepare and test the execution of the system. The outputs of the system were compared to the results of expert microscopists. The results were promising and the sensitivity of the proposed method outperforms most of the other reported methods. The system recorded 97.02 % accuracy and 98.98% sensitivity in detecting the presence of Plasmodium parasites. The SVM classifier improves the accuracy and performance of the system. Moreover, the automated image processing based method introduced in this project is interactive; hence, it is faster and more accurate than manual process.

REFERENCES

[1] D. A. Ghate and P. C. Jadhav, “Automatic Detection of Malaria Parasite from Blood Images,” Int. J. Adv. Comput. Technol., vol. 1, no. 3, pp. 66– 71, 2012.

[2] M. S. Suryawanshi and P. V. V Dixit, “Comparative Study of Malaria Parasite Detection using Euclidean Distance Classifier & SVM,” vol. 2, no. 11, pp. 2994–2997, 2013.

[3] A. Anand, V. K. Chhaniwal, N. R. Patel, and B. Javidi, “Automatic identification of malaria-infected RBC with digital holographic microscopy using correlation algorithms,” IEEE Photonics J., vol. 4, no. 5, pp. 1456–1464, 2012.

[4] Ross, N.E., Pritchard, C.J., Rubin, D.M. and Duse, A.G. (2006) Automated image processing method for the diagnosis and classification of malaria on thin blood smears. Medical & Biological Engineering & Computing, 44, 427-436.

[5] Tek, F.B., Dempster, A.G. and Kale, I. (2009) Computer vision for microscopy diagnosis of malaria. Malaria Journal, 8, 153.

[6] Diaz, G., Gonzalez, F.A. and Eduardo, R. (2009) A semiautomatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images. Journal of Biomedical Informatics, 42, 296-307.

[7] Di Ruberto, R.C., Dempster, A., Khan, S. and Jarra, B. (2002) Analysis of infectedblood cell images using morphological operators. Image and Vision Computing, 20, 133-146.

[8] CDC, “CDC - Malaria - About Malaria - Biology - Malaria Parasites,” USA Government, 2012. [Online].

[9] W. Khan, “Image Segmentation Techniques: A Survey,” J. Image Graph., vol. 2, no. 1, pp. 6–9, 2013.

[10] A. Verm, M. T. Scholar, C. Lal, and S. Kumar, “Image segmentation: Review paper,” Int. J. Educ. Sci. Res. Rev., vol. 3, no. 2, 2016.

[11] O. Lahdenoja, M. Laiho, and A. Paasio, “Reducing the feature vector length in local binary pattern based face recognition,” in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, 2005, pp. II: 914-917.

[12] Ross, N.; Pritchard, C.; Rubin, D.; Dusé, A. Automated image processing method for the diagnosis and classification of malaria on thin blood smears. Med. Biol. Eng. Comput. 2006, 44, 427–436. Arco, J.; Gorriz, J.; Ramirez, J.; Alvarez, I.; Puntonet, C. Digital image analysis for automatic enumeration of malaria parasites using morphological operations. Expert Syst. Appl. 2014, 42, 3041–3047.

[13] Arco, J.; Gorriz, J.; Ramirez, J.; Alvarez, I.; Puntonet, C. Digital image analysis for automatic enumeration of malaria parasites using morphological operations. Expert Syst. Appl. 2014, 42, 3041–3047.

[14] Das, D.; Ghosh, M.; Chakraborty, C.; Maiti, A.; Pal, M. Probabilistic prediction of malaria using morphological and textural information. In Proceedings of the 2011 International Conference on Image Information Processing (ICIIP), Shimla, India, 3–5 November 2011; pp. 1–6.

[15] Khan, M.; Acharya, B.; Singh, B.; Soni, J. Content based image retrieval approaches for detection of malarial parasite in blood images. Int. J. Biometr. Bioinform. (IJBB) 2011, 5, 97–110.

[16] Kareem, S.; Morling, R.; Kale, I. A novel method to count the red blood cells in thin blood films. In Proceedings of the 2011 IEEE International Symposium on Circuits and Systems (ISCAS), Rio de Janeiro, Brazil, 15–18 May 2011; pp. 1021–1024.