Enhanced Detection of Malaria Disease Using Image Processing and Machine Learning

[1]Navya N, [2]Meghashree A C
[1]Student, [2]Asst Professor
[1]Digital Electronics and Communication, MTech,
[1]VTU Post Graduation Center, Mysuru, India

Abstract: The Malaria is a genuine irresistible infection. As indicated by the World Health Organization, it is in charge of about one million passing's every year. There are different procedures to analyze malaria of which manual microscopy is viewed as the best quality level. Malaria is mosquito-borne blood illness brought about by parasites of the class Plasmodium. A specialist examine the microscope image and detects weather the person is infected or not, this is tedious procedure and has various problems. This paper, presents an image processing technique that can be used to detect the malaria disease. The method mainly uses Otsu thresholding, GLCM, later Machine Learning to learn, identify and decide the sorts of infected cells as indicated by its highlights, for this SVM is used.

IndexTerms **- Malaria, plasmodium, RBC, Otsu thresholding, GLCM, SVM.**

I. INTRODUCTION

Malaria is an intense irresistible, infectious disease occurs due to peripheral blood parasite called plasmodium. Which is a major public health issue causing death, mostly in infants and pregnant women and it also causes anemia and reduces productivity. It is a majorly caused in many of the tropical and subtropical regions. World health organization has declared that every year 250million people are affected by malaria and 1 million people are dead due to this disease.

History of malaria stretches from its prehistoric origin as a zoonotic disease in primates of Africa till 21st century. For many years, herbal remedies were been used to treat malaria. Afterword's they prescribed prophylactic quinine and new drugs like chloroquinine and artemisinins, which were used in order to resist the scourge. Nowadays artemisinin drug is present in every redmine which is used to treat the malaria disease. There are few varieties of plasmodium that infect humans they are: Plasmodium falciparum, Plasmodium malariae, Plasmodium ovale, Plasmodium vivax, Plasmodium knowlesi.

II. LITERATURE SURVEY

This section describes about the various methods and some modern methods applied for detection of malaria disease.

A new image processing system for detection and quantification of plasmodium parasites in blood smear slide were constructed and they developed Machine Learning algorithm to learn, detect and determine the types of infected cells according to its features [1].

The research work they have proposed a system for automating the manual work done by a technician in order to cut down the human error and increasing the accuracy of the malaria diagnosis. The System is tested for a dataset of 80 images of a thin blood smear. The infected cells are extracted using HSV segmentation. This approach will be beneficial for the rural areas, with a scarcity of experts [2].

The focus was on segmentation and detection of microscopy images infected with malaria parasites using the color-based cascading method. The method begins with color normalization process, followed by gamma correction , then noise reduction , exposure compensation, edge enhancement, fuzzy c-means cluttering, and lastly, morphological processes [3].

In this paper they are using thin blood smear for detection of malaria disease. The problem they have noticed is clarity and brightness of most images is different. Hence there are errors in detection. To overcome this problem partial contrast stretching technique has been used by malaria the S component in the HIS color space was used. This was also incorporated with a* component in CIEL* a* b* color space in order to obtain complete malaria cell [4].

III. PROPOSED SCHEME

The malaria blood smear is collected using various medical sites and it is resized to 300X400 pixel values. Based on survey, it is found that various segmentation techniques are applied. The proposed method possesses simple and effective detection method for malaria detection. Based on the image obtained for segmentation Otsu thresholding is used. Further the process is has shown in block diagram.

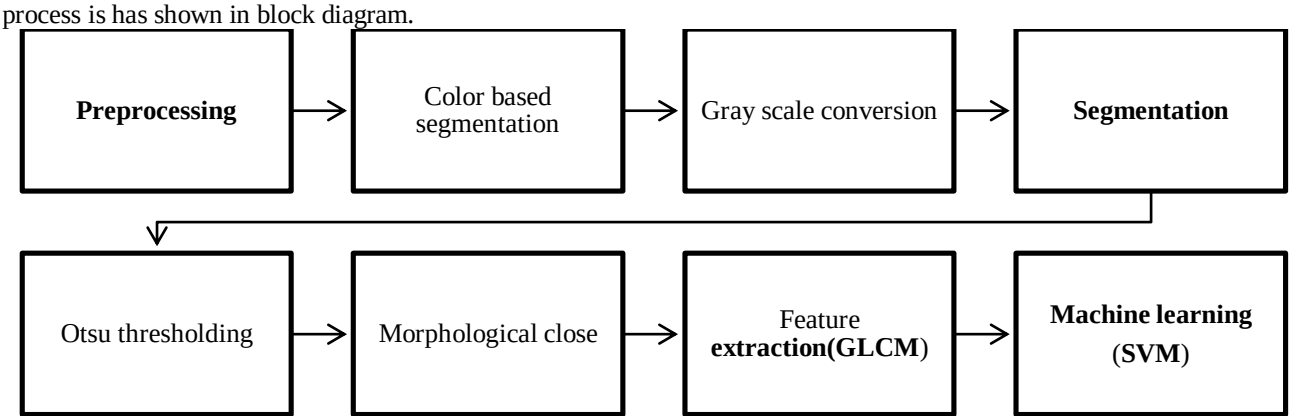

Fig1: Block diagram of the overall process

3.1 Preprocessing

The preprocessing includes color based conversion and gray scale conversion. The images were taken from websites which is of 1400X1600 and 256X256 sizes. The image is resized to 300X400. Initially the image is of RGB color space where RGB means Red, Green, and Blue respectively. It is converted to HSV color space, where hue represents the color, saturation means the amount of white color mixed and value for amount of color mixed with black color. The HSV conversion is used because we can spot the region of interest. After obtaining the color space converted image as in fig 3, the image is converted to gray scale. The luminance of a pixel estimation of a gray scale picture ranges from 0 to 255. The change of a RGB image into a gray scale picture is changing over the RGB esteems (24 bit) into gray scale esteem (8 bit).

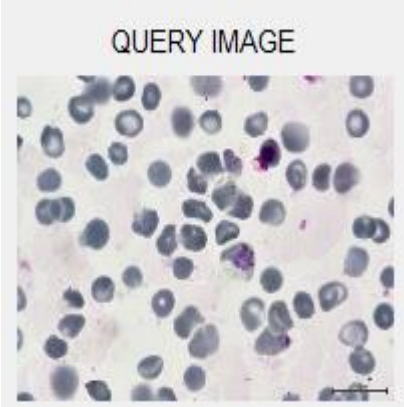

Fig 2: Input image

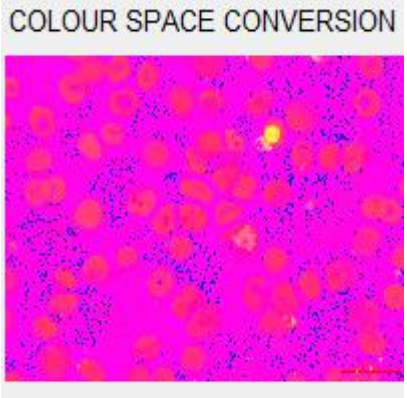

Fig 3: Color space conversion

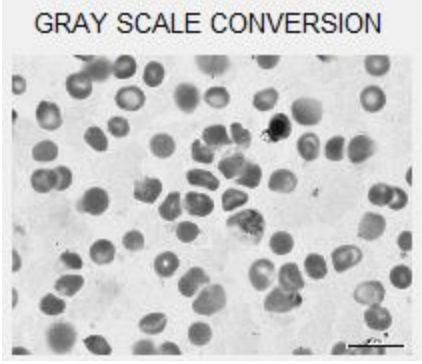

Fig 4: Gray scale conversion

3.2 Segmentation

Segmentation is a process in which the digital image is partitioned to multiple images. In the proposed method for segmentation Otsu thresholding is used. The Otsu method selects a threshold value as the values between two peaks where one is foreground and other is background. The threshold value is obtained by trial and error method and it is applied to the image. Thresholding is done in order to delete the unwanted background and to obtain only the region of interest. After thresholding the image is as shown in Fig5.

3.2.1 Morphological close

After thresholding, morphological closing operation is performed, so that any unwanted region in the images which is still present after thresholding is eliminated. It performs erosion and dilation operation, where erosion removes a layer from

outskirt of blob and similarly dilation adds a layer of pixel to it and image is binary converted so that all the leftover minute unwanted data will be removed. Finally a boundary is drawn around the region of interest which is as shown in Fig6.

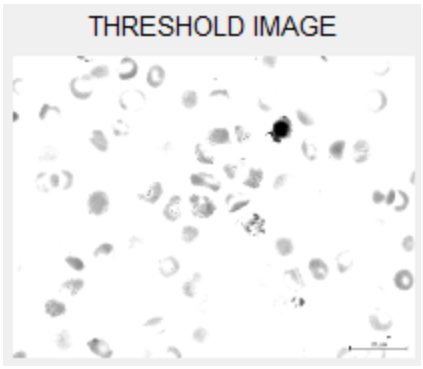

Fig 5: After applying threshold value to the image

Fig 6: ROI marked image

3.3 Feature extraction

The GLCMs are put away in an I x j x n framework, where n is the quantity of GLCMs determined normally because of the distinctive introduction and relocations utilized in the calculation. Generally the qualities I and j are equivalent to 'NumLevels' parameter of the GLCM figuring capacity graycomatrix(). Note that matlab quantization esteems have a place with the set {1,..., NumLevels} and not from {0,...,(NumLevels-1)} as gave in certain references

In spite of the fact that there is a capacity graycoprops() in Matlab Image Processing Toolbox that registers four parameters Contrast, Correlation, Energy, and Homogeneity. Along with that still more features are considered, totally making 13 features. They are mean, entropy, variance, kurtosis, SD, RMS, smoothness, skewness, energy, homogeneity, contrast, correlation, IDM

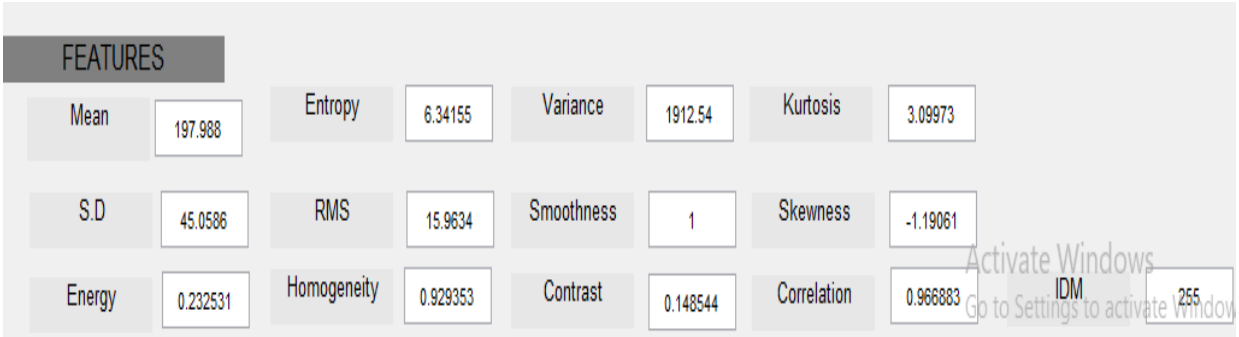

Fig 7: GLCM feature extracted values

3.4 Machine learning (SVM)

SVM is a supervised machine learning algorithm. When the information has precisely two classes SVM (support Vector machine) can be used. It is a model that learns from past input and makes future predictions as output. Having set of points of two types in N dimensional place SVM generates a (N-1) dimensional hyperplane which separates 2 classes. A SVM orders information by finding the best hyperplane that isolates all information purposes of one class from those of the different class. The best hyperplane for a SVM implies the one with the biggest edge between the two classes. Edge implies the maximal width of the section parallel to the hyperplane that has no inside information focuses. The help vectors are the information indicates that are nearest the isolating hyperplane, these focuses are on the limit of the section. The accompanying figure delineates these definitions, with + demonstrating information purposes of sort 1 and – showing information purposes of sort –1.

IV. RESULTS

After obtaining the features the SVM is used to train and test the data, for a new image the system checks whether the blood smear is infected or not, which is based on SVM training data set. Finally the output will be displayed as malaria infected or uninfected. To check whether the system is feasible or not we will calculate the specificity, sensitivity, accuracy. Where,

Sensitivity, Ss, of a test is characterized as the likelihood of positive test consequences of the nearness of disease.

The sensitivity is given by

Ss = (100TP) / (TP+FP)

Specificity Sp is characterized as the likelihood of a negative test outcome given in the absence of disease. The specificity is given by

Sp = (100TN) / (TN+FN)

The precision is the way close a measure esteem is real it is given by

A = (100(TP+TN)) / (TP+FP+TN+FN)

All things considered, from the acquired esteem plotted in the chart as appeared as follows as in fig 8. In the event that the esteem lies between 0.45-0.55, at that point the proposed technique is working properly and desired output is obtained.

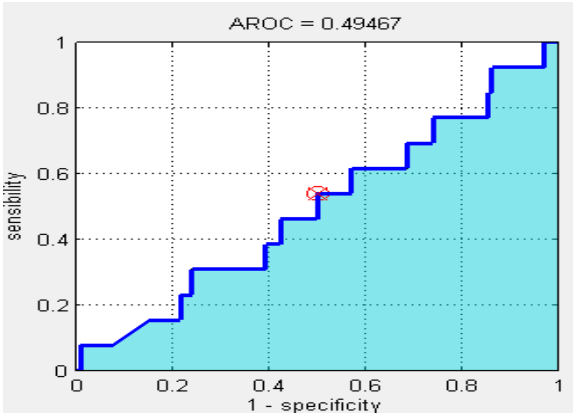

Fig 8: AROC curve

V. CONCLUSION

The detection of Malaria parasites is completed by pathologists by physically utilizing Microscopes. Along these lines, the likelihood of false discovery because of human blunder is high and thus can result into lethal condition. This framework will limit the human mistake while identifying the nearness of malaria parasites in the blood test by utilizing image preparing and limit human blunder. In this proposed framework, Feature Extraction is utilized to distinguish malaria parasites in blood smear. The framework is unaffected by the uncommon conditions and can benefit from outside assistance to accomplished high rates of affectability, explicitness, positive forecast esteem. Furthermore, by this method one can discover the seriousness of the disease and can be treated from the beginning time. Additionally, we built up a methodology which utilizes machine learning algorithm.

VI. ACKNOWLEDGMENT

This project is partly supported by the assistant professors of Visvesvaraya Technological University, Post-Graduation Center, Satagalli, Mysuru.

References

[1] Malaria detection using Image processing and machine learning by Suman Kunwar, Manchana Shrestha, Rojesh Man Shikhrakar

[2] Malarial Parasites Detection in RBC Using Image Processing Shipra Saraswat1, Utkarsh Awasthi2, Neetu Faujdar3

[3] Colour image segmentation for malaria parasites detection using cascading method by Yonathan Ferry Hendrawan, Cucun Very Angkoso, Rima Tri Wahyuningrum

[4] Improvement of complete malaria cell image segmentation by Monlica Wattana,TIpwalee Boonsri

[5] Detection of malaria parasites using digital image processing by ahmedelmubarak Bashir, Zeinab A.Mustafa, Islah Abdelhameid

[6] Detection of malaria parasite in giemsa blood sample using image processing by Kishor Roy, Shayla Sharmin, Rahma Bintey Mufiz Mukta, Anik Sen

[7] An Image Analysis Algorithm for Malaria Parasite Stage Classification and Viability Quantification by Seunghyun Moon1,4, Sukjun Lee2 , Heechang Kim1 , Lucio H. Freitas-Junior3 , Myungjoo Kang4 , Lawrence Ayong2 , Michael A. E. Hansen1 *

[8] A Combined Algorithm for Malaria Detection from Thick Smear Blood Slides by Kaushik Chakrabortya, Arnab Chattopadhyayb, Amlan Chakrabarti, Tinku Acharyad

[9] An Image Analysis Algorithm for Malaria Parasite Stage Classification and Viability Quantification by Seunghyun Moon , Sukjun Lee, Heechang Kim, Lucio H. Freitas-Junior, Myungjoo Kang, Lawrence Ayong, Michael A. E. Hansen

[10] Adaptive Threshold determination based on entropy in active contour without edge method for malaria parasite candidate detection by Faza Maula Azif, Hanung Adi Nugrohororgan Kaufmann Publishers Inc, 2008.