International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-9 Issue-1, May 2020

Malaria Parasite Detection using Deep Learning in Thick Blood Smears

M. Sai Tejomaayi, M. Divija, R. Pavani, Nara. Sreekanth

Abstract: ***Detection of malaria disease is done by finding the presence of malaria parasite or plasmodium in blood smear. Here malaria parasites are detected in thick blood smears. This paper proposes a version to detect the presence of malaria parasite(plasmodium) in thick blood smears automatically with the help of deep learning and not using microscopy examinations and chemical tests. This detection will be done using two steps, that is, intensity-based screening which is the preprocessing step, the first step, that extracts candidates for processing, and next is customized convolutional neural network (CNN), the processing step, which takes the preprocessed images and detects whether malaria parasite is present or not. Hold-out(3:1) technique is used for evaluation of the model. The model has achieved an accuracy reaching 91%. The two preprocessing and processing steps improves object detection of the system. Malaria is usually detected using chemical tests and microscopy examinations. This process requires a lot of resources mainly laboratories. Parasitologists who are experienced are sometimes difficult to find, so manually counting the malaria parasites can be prone to major errors. Due to which the cost for testing and even time for malaria diagnosis increases drastically. Since the traditional process of malaria parasite in blood smears detection has many drawbacks it needs a sophisticated, accurate diagnosing equipment or system which has low cost. This system can be used in regions and areas where there are constraints on resources, time of people and cost which they can afford. This system provides many advantages to rural diagnostic centres where the supplies are limited and not easily accessible.***

Keywords: ***Convolutional Neural Network, Deep Learning, Malaria Parasite, Blood Smears.***

I. INTRODUCTION

Motivation:

There has been an observation where many suffer and die due to malaria. Some of the deaths happen due to delay of diagnosis. In rural regions, the facilities provided by the hospitals are poor. The person suffering need to travel to a region where malaria diagnosis is done. This causes people a lot of inconvenience. Hence this application is convenient and easy to use, particularly in resource limited areas.

1.1 Problem Statement

This study reveals up to what extent the methods used in the application would replace the existing methods. The real challenge lies in better performance over a period of time.

1.2 Applications in Real Time

Rural regions can install this application in their health care centres handled by a technician who sends reports to the doctors. The application is portable and can be easily installed and used in resource limited areas, and saves time when compared to chemical tests.

1.3 Objective

Our objective is to develop an application which can be operated easily, fast and accurate. This application should be beneficial in regions where there are limited resource and thus save lives of the rural people.

II. EXISTING SYSTEM

Earlier segmentation and morphological operations were used which produced less accuracy. Later CNN was proposed which gave an accuracy more than 90%.

With similar idea, the work has been done by Feng Yang et al [1]. In their paper they mentioned about using intensity based global minimum screening and customized convolutional neural network.

III. PROPOSED SYSTEM

In this paper we are using intensity-based screening which is a variation of intensity based iterative global minimum screening with modification of training and testing datasets by taking inspiration from the work done by Feng Yang et al [1]. In this paper we are using intensity-based screening for selection of objects for detection which are given to the customized CNN model which gives the result whether the thick blood smear image contains malaria parasite or not.

IV. 4. USE OF THE PROPOSED WORK

We have used intensity-based screening instead of intensity based iterative global minimum screening as proposed by Feng Yang et al and modifying the training and testing datasets for making the process simple[1].

Manuscript received on April 02, 2020.
Revised Manuscript received on April 15, 2020.
Manuscript published on May 30, 2020.
* Correspondence Author
M. Sai. Tejomaayi *, department of CSE, BVRIT HYDERABAD College of Engineering, affiliated to JNTUH, Hyderabad, India. Email: 16wh1a0527@bvrithyderabad.edu.in
M. Divija, department of CSE, BVRIT HYDERABAD College of Engineering, affiliated to JNTUH, Hyderabad, India. Email: 16wh1a0534@bvrithyderabad.edu.in
R. Pavani, department of CSE, BVRIT HYDERABAD College of Engineering, affiliated to JNTUH, Hyderabad, India. Email: 17wh5a0501@bvrithyderabad.edu.in
Dr. N. Sreekanth, Associate Professor, department of CSE, BVRIT HYDERABAD College of Engineering, affiliated to JNTUH, Hyderabad, India. Email: sreekanth.n@bvrithyderabad.edu.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

V. METHODOLOGY

The dataset consists of 909 images including positive images and negative images, which have been divided into training set, the validation set and testing set after preprocessing using intensity-based screening. We have taken the images from the link provided, in the work, by Feng Yang et al [1]. The training and testing dataset is split such that the count of images is in the ratio 3:1.

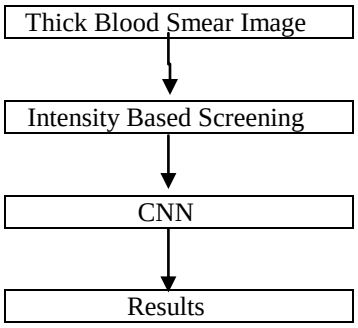

Fig 5.1 : Methodology

5.1 Intensity-based Screening

In a dyed microscopic image of thick blood smear nuclei of malaria parasite and WBCs have the same dye intensity absorbed. This includes a series of methods which help in extracting these objects for processing through the next step, that is, customized CNN model.

5.1.1 Median filter

This filter is used for noise removal from the image which makes it more ready for the processing step. It is not a linear filtering method.

5.1.2 Laplacian of gaussian

Laplacian of gaussian blur is used as blob detection [2]. It removes the noise that is detected the blobs. Blob contains almost constant values when pixelated. This is the bright region on dark background and vice versa. All the blobs from image are extracted using this method.

5.1.3 Otsu thresholding

Otsu thresholding generated binary images by identifying the pixels that fall in the threshold range [4]. It classifies them as foreground or background. It identifies all the objects that fall in the threshold range which are WBCs and malaria parasite. Fig 5.2. shows the detected objects after Otsu thresholding.

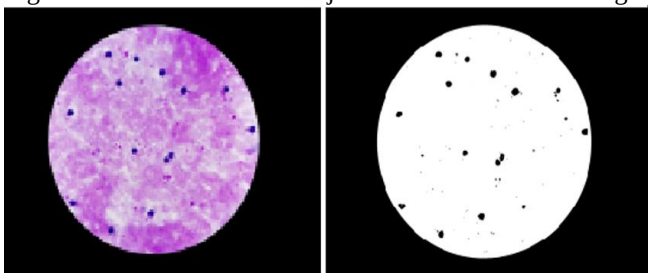

Fig 5.2. Candidate objects detected after applying Otsu thresholding

5.2 Customized CNN model

Customized convolutional neural network with 7 Layers is used for processing the images and giving the result, based on the work done by Feng Yang et al [1]. It classifies the candidate objects of the image as parasite or not. It contains three fully connected layers, three max pooling layers, and a softmax layer.

Fully connected layer is used so that there is connection to the next layer's activation layer.
Max-pooling layer is used to extract the most prominent feature map.
Softmax layer is the final layer which helps in detection and prediction of the objects finally.

The architecture of the model is shown in Fig 5.2, the input being the images in training set for building the model and test set for checking the accuracy of it.

K. Simonyan [3], have used a model, which selected six convolutional layers initially and three comparing max-pooling layers, there design VGG19 to stop feature maps at 64 @ 5 x 5, followed straight forwardly by the combination of dropout layers and fully connected layers. For training this model hold-out technique is used with ration on 3:1 of training set and testing set.

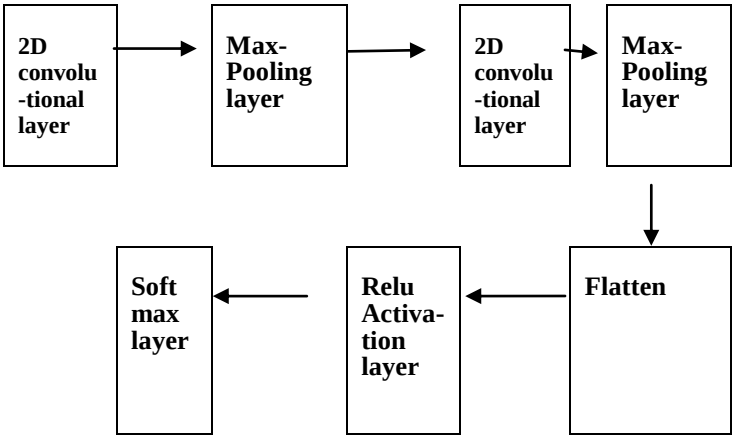

Fig 5.3. Customized CNN Model

VI. RESULT AND EVALUATION

After training the model the accuracy achieved is 0.9169 using the hold-out technique. The resultant image consists of red boxes resembles malaria parasite in it, if it's present and the result is displayed as positive (Fig 6.2) else negative (Fig 6.4). This is reliable system as it is less time consuming when compared to traditional models and the maintenance cost is also less. The total trainable parameters generated are 3696930. It takes ten seconds for producing the result (8 seconds for processing and 2 seconds of obtaining the result).

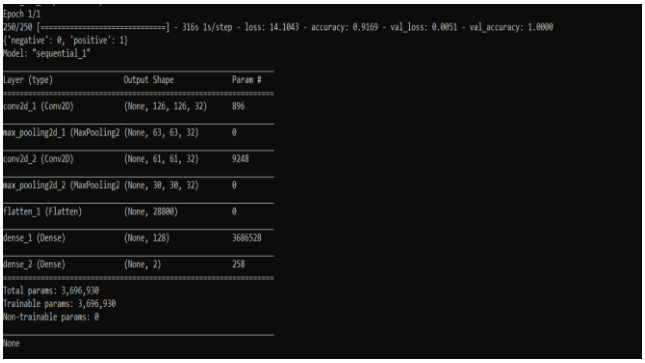

Fig 6.1 Summary customized CNN model

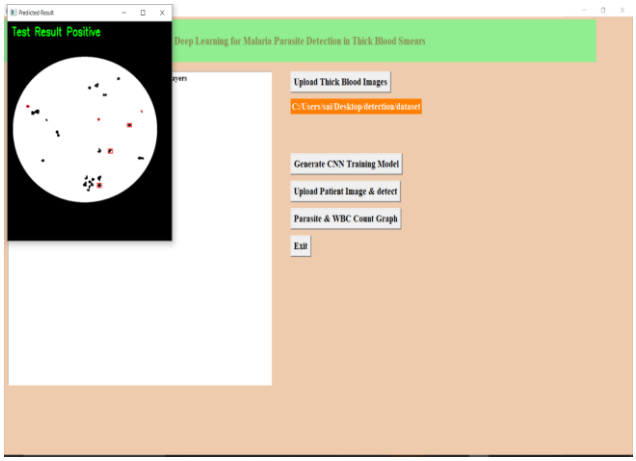

Fig 6.2: Result when malaria parasaite is present

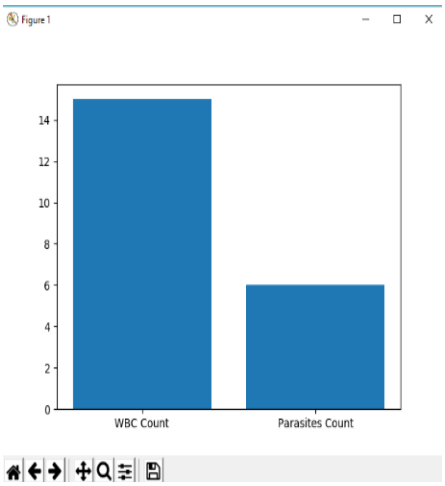

Fig 6.3: Graph of WBC vs malaria parasite count graph when the result is positive

This graph is obtained using the subset of candidate objects generated. When the output result is positve all the objects generated after blob detection and minimum are enclosed in red boxes indicates malaria parasite. This tells the intensity of the disease. The parameters of this graph is taken per blood smear radius.

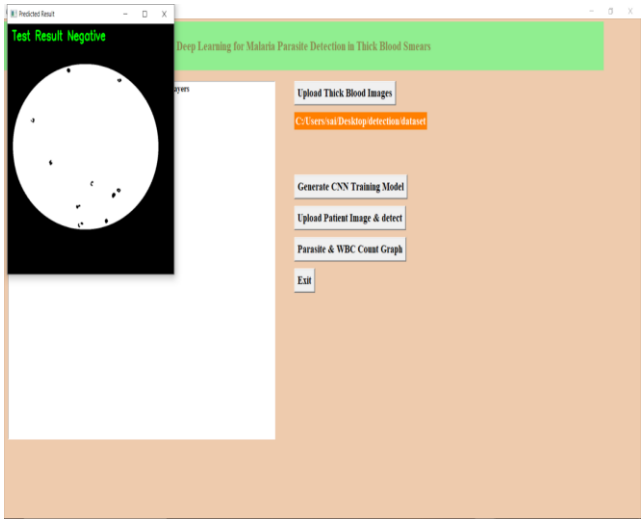

Fig 6.4 Result when the malaria parasite is absent

Total images		Training set		Testing set	
Posistive 562	Negative 347	Positive 375	Negative 232	Positive 187	Negative 115

Accuracy: 91.69%

Fig 6.5 Accuracy and division analysis

VII. CONCLUSION

From the drawn results, we have achieved 91.69% accuracy using hold-out technique, which took just 10 seconds. The usage of two methods(preprocessing step and processing step): intensity based screening and customized convolutional neural network enhanced the efficiency of the system which will be helpful for healthcare sector. Therefore the proposed work is highly beneficial in resource limited regions and saves a lot of time. The results obtained by the system makes it reliable when compared to the traditional systems by overcoming their drawbacks in view of supplies,time and cost. This system also avoids the manual errors while counting the malaria parasite during malaria diagnosis.

REFERENCES

1. Feng Yang, Poostchi M, Yu H, Zhou Z, Silamut K, Yu J, Maude RJ, Jaeger S, Antoni SK," Deep learning for smartphone-based malaria parasite detection in thick blood smears", IEEE, Booked Health Inform, 2019, Sep 23.
2. Julisa bana Abraham, "Plasmodium detection using simple CNN and clustered GLCM features", Jan 2018
3. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", *arXiv Prepr. arXiv1409.1556*, p. [Online] Available: https://arxiv.org/abs/1409.155, 2014.
4. N. Otsu, "A threshold selection method from grey level histograms", Page No: 62-66, Jan 1979
5. K. S. Makhija, S. Maloney, and R. Norton, "The utility of serial blood film testing for the diagnosis of malaria", Pathology, vol. 47, no. 1, pp. 68–70, 2015.
6. Z. Liang, A. Powell, I. Ersoy, M. Poostchi, K. Silamut, K. Palaniappan, P. Guo, M. A. Hossain, A. Sameer, R. J. Maude, J. X. Huang, S. Jaeger, and G. Thoma, "CNN-based image analysis for malaria diagnosis", in Proc. BIBM, ShenZhen, China, 2017, pp. 493–496.

7. S. Rajaraman K. Silamut, M. A. Hossain, I. Ersoy, R. J. Maude, S. Jaeger, G. R. Thoma, and S. K. Antani, "Understanding the learned behavior of customized convolutional neural networks toward malaria parasite detection in thin blood smear images", J. Med. Imaging, vol. 5, no. 3, p. 034501, July 2018.
8. P. A. Pattanaik and T. Swarnkar, "Comparative analysis of morphological techniques for malaria detection," Int. J. Healthc. Inf. Syst. Informatics, vol.
9. WHO, Guidelines For The Treatment of Malaria, Third edition World Health Organization, 2015.
10. K.S. Makhija, S. Maloney, and R. Norton, "The utility of serial blood film testing for the diagnosis of malaria", Pathology, vol. 47, no. 1, pp. 68–70, 2015.
11. WHO, Malaria micropscopy quality assurance manual, Version 2. World Health Organization, 2016.

AUTHORS PROFILE

Miss M. Sai. Tejomaayi, is an under graduate student with Reg No: 16WH1A0527, is pursuing her B.Tech degree in Computer Science and Engineering from JNT University, studying at BVRIT HYDERABAD College of Engineering for Women, Hyderabad, Telengana State., India. She is currently in final year Second Semester. Her area of interest includes Data Mining, Image Processing, Data Science. She has done online certification courses in "data science for engineers" conducted by IITM, "Joy of computing using python" conducted by IIT, through NPTEL and has been awarded merit and gold certificates. She has been certified from IIT Mumbai for course in JAVA. She has completed her internship in the organization called Verzeo in the field of cloud computing and also got certified by Microsoft after training. Her active participation in various workshops and online courses helped in doing this project.

Miss. M. Divija, is an under graduate student with Reg No: 16WH1A0534, is pursuing her B.Tech degree in Computer Science and Engineering from JNT University, studying at BVRIT HYDERABAD College of Engineering for Women, Hyderabad, Telengana State., India. She is currently in final year Second Semester. Her area of interest includes Data Mining, Image Processing. She is certified in RPA by Automation anywhere. She has received elite certification for the course 'Joy of Computing using Python 'conducted by IIT through NPTEL. She has completed her internship in the organization called 'Verzeo', in the field of Cloud computing. She has also been certified from IIT Mumbai for the course in Java. she attended various related workshops and pre-conference, which helped in doing this project.

Miss. R. Pavani, is an under graduate student with Reg No: 17WH5A0501, is pursuing her B.Tech degree in Computer Science and Engineering from JNT University, studying at BVRIT HYDERABAD College of Engineering for Women, Hyderabad, Telengana State., India. She is currently in final year Second Semester. Her area of interest includes Data Mining, Image Processing. She is certified in Robotic Process Automation by Automation Anywhere University. She has done various online certification courses through NPTEL like "Software Project Management" through IITM, "Joy of computing using python" through IIT-B and received merit and silver certificates. She attended her area of interest related workshops and pre-conference, which helped in doing this project.

Dr. N. Sreekanth received his PhD from Sri Satya Sai /university of Technology & Medical Sciences, Sehore, Madhya Pradesh, India in 2018. He received his B.E. degree in Electronics and Communication Engineering from Gulbarga University, studied at Rural Engineering College in Bhalki, Karnataka., India in 1997. He received his M.Tech. degree in Computer Science from JNT University, studied at School of Information Technology in Masab Tank, Hyderabad, India in 2002. He has served PIRMEC, HITSCOE and VGNT; as Assistant Professor and Associate Professor and taught various courses for UG and PG students, during in his 15 years of teaching experience. He is currently working as Associate Professor, Dept of C.S.E, BVRIT Hyderabad college of Engineering for women, Bachupally, Telengana, India. He is pursuing his Ph.D (Part-Time) in Computer Science from Rayalaseema University, Kurnool, A.P., India. His area of research interests includes Digital Image Processing, Multimedia Security, Software Engineering, Cloud Computing and Data Mining. He is a Professional Member of ACM and also, he is a Life Member for IEAE,InSc and IAENG. He has published more than 13 research publications in various National, Inter National conferences, proceedings and Journals.