

Diagnosis of Malaria from Peripheral Blood Smear Images using Convolutional Neural Networks



M.Mohana, V. Vani, Shri Dikshanya K.N, Shruthi.B, Vinothini V

Abstract: Malaria is a deadly disease brought about by Plasmodium parasites which affects the general population through the bites of female mosquitoes, called "malaria vectors." There are about five parasites species that cause malaria in human body, and two of the species namely P. falciparum, P.vivax pose the greatest threat. The most prominent technique to detect malaria is by taking blood smear samples to check if the RBC is affected by parasite under the microscope by qualified experts. It is a complex technique and the diagnosis depends on the experience and inside of the person who performs the examination. Malaria blood smear have been diagnosed earlier using image processing methods based on machine learning. This was not effective so far. Convolutional Neural Network (CNN) is use in this system which helps in classifying the cells present in the blood smear images as infected or uninfected.

Keywords : Malaria, Computer aided diagnostics, Convolutional neural networks, feature extraction, classification.

I. INTRODUCTION

Malaria is a mosquito-borne disease which affects millions of people worldwide. Malaria can be caused by several species of Plasmodium parasites including Plasmodium Ovale, Plasmodium Malariae, Plasmodium Vivax and Plasmodium Falciparum, which has a complex life cycle that can make affected individuals turn out to be wiped out with influenza like disease, high fever, and chills.

Manuscript published on November 30, 2019.

* Correspondence Author

Dr.M.Mohana*, Associate Professor, Department of Information Technology, Easwari Engineering College, Ramapuram, Chennai -89.

Dr. V. Vani, Associate Professor, Department of Information Technology, Easwari Engineering College, Ramapuram, Chennai -89.

Shri Dikshanya K.N, Department of Information Technology, Easwari Engineering College, Ramapuram, Chennai -89.

Shruthi.B, Department of Information Technology, Easwari Engineering College, Ramapuram, Chennai-89.

Vinothini V, Department of Information Technology, Easwari Engineering College, Ramapuram, Chennai -89.

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

It might even prompt to demise. The Plasmodium parasites increases in number in the liver and taint Red Blood Corpuscles (RBC) present in the blood. Generous advancement has been made worldwide to control and eradicate Malaria; however it keeps on being a huge general medical issue with roughly 3.2 billion individuals worldwide in danger for the disease. In 2016, there were approximately 216 million instances of malaria in 91 nations [1].

The Malaria Diagnostic tests are classified as microscopic tests and non-microscopic tests. Clinical diagnostics relies on signs of illness of the patient and furthermore dependent upon physical findings amid the examination of patients. The underlying side effects of malaria incorporate migraine, coldness, fever, sweat, muscle agony, retching, nausea and dysentery. In severe cases, the clinical features may include paleness, splenomegaly, hypoglycemia, thrombocytopenia, renal dysfunction, and neurologic changes [2]. The previously mentioned manifestations are also found in some diseases like common viral diseases and influenza. In extreme cases, clinical discoveries are prominent and the dimension of doubt for malaria may increase. The Clinical findings must be affirmed by the laboratory test [1, 2]. The best way of diagnosing malaria is by using blood smear samples and detecting plasmodium parasites present in the blood smear using a microscope. Manually screening malaria is a complex process .A large amount of time is taken to analyze the sample which leads to approximation and deviation of values. The later advancement makes use of high resolution camera that capture images of blood slides to improve consistency in diagnosing malaria .Segmenting of RBC is done using Computer Aided Diagnosis approach. Here the infected and uninfected blood smear images is classified using deep learning algorithm .The computerized version of malaria diagnosis replaces the current laboratory.

II. LITERATURE SURVEY

In [3], a trained model based on a DBN which classifies 4100 peripheral blood smear images into the parasite or non-parasite class is used. Here, Deep belief networks are pre-trained by piling up RBM (Restricted Boltzmann Machines) using the contrastive divergence strategy for pre-training. In order to train the DBN, the features are extracted from the blood smear images and the visible variables of the Deep Belief Networks are initialized. A concatenated feature of texture and color is used as a feature vector.

In [4], the peripheral blood smear pictures obtained from microscope are used as image source and pre-processing techniques include illumination correction to adjust the lighting of images and noise reduction is performed as noise such as grains might be introduced in image acquisition process, RBC's are segmented from the image, feature extraction, selection is performed later and finally the classification of different stages of malaria have been investigated. Segmentation is done by using marker controlled watershed transformation and approximately ninety six features including shape, size and texture of RBC's have been extracted. Improved version of classification precision is achieved by using Bayesian classification and support vector machine (SVM). In [5], In order to identify RBCs and parasites present on microscopic slides, Morphological and Threshold selection techniques have been used and Image features are extracted based on color, texture and the geometry of the cells . In addition to that features that makes use of expertise of classification and mimics the features of technical experts. To distinguish between true and false positives, Back propagation Feed forward neural network based to state the classifier have been used, which later diagnoses the species of the disease. In [6], RBCs infected with malaria parasites found in the microscopic images is found using an edge-based segmentation that has been developed to improve the diagnostic process. Illuminance correction and color effect reduction is done with color space transformation and Gamma equalization. Clustering is done using Fuzzy C-means, to extract the infected erythrocytes, which is later proceeded for the final segmentation.

III. PROPOSED SYSTEM

A. Image Source

The blood smear images were contributed by Jane Hung of MIT and the Broad Institute of Cambridge, MA. Images are in image file format such as .png or .jpg format. There are about three sets of images which consist of 1364 images (~80,000 cells) with different researchers having prepared each one from Brazil (Stefanie Lopes), Southeast Asia (Benoit Malleret), and from time course (Gabriel Rangel). Blood smears were stained with Giemsa reagent.

B. Pre-Processing

The pre-processing stage is the first stage in image processing. Image pre-processing is done, in order to improve the quality of image and correct the size of the image. The image used is being taken as an input might have poor quality, for example the image might have noises which could have arisen. In addition to that image brightness level can be corrected through pre-processing. Here, pre-processing is performed which consist of two operations on the blood smear image data as it can have huge measure of repetitive data. So as to invariably achieve better classification exactness and accuracy, noise reduction and image segmentation should be done in order to separate blood cells and remove those unwanted image pixels simultaneously. First, each image tile should be converted into a grayscale image from the color space followed by the thresholding

operation. The noise is generated as the by product of thresholding can degrade the quality of the microscopic blood smear images and gives lower classification performance. Therefore, the isolated noise pixels were eliminated by applying bilateral filter. This filter is used in the effective removal of noise while keeping edges sharp. Bilateral filter takes a Gaussian filter in space, but one more Gaussian filter which does the function of pixel difference. Gaussian function of intensity difference makes sure that only those pixels with similar intensity to central pixel are considered for blurring. So it preserves the edges since the pixels at edges will have large intensity variation.

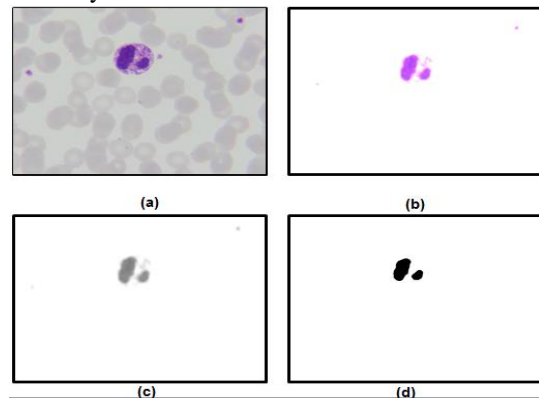


Fig.1: (a) Original Image, (b) After Bilateral Filter (c) After Gray scale conversion (d) Segmented Image

C. Convolutional Neural Networks

The convolutional neural network (CNN or ConvNet) is a class of deep artificial neural networks that found application mostly in image and sound analysis. CNN usually analyze part or region of the sample which called receptive field. What differs ConvNets from other neural networks is that for each receptive field the weights are shared among all neurons in filter of convolutional layer. In that way feature detection is independent from spatial position of the feature in the feature field.

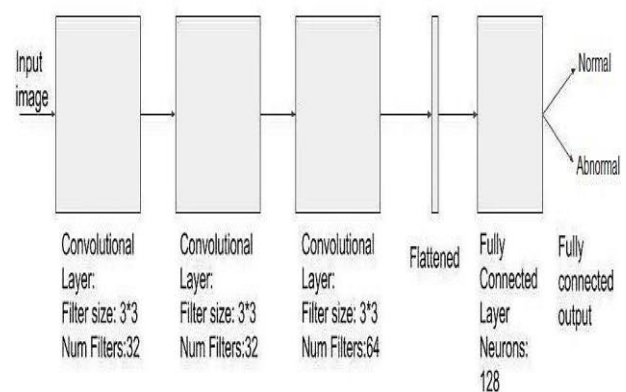
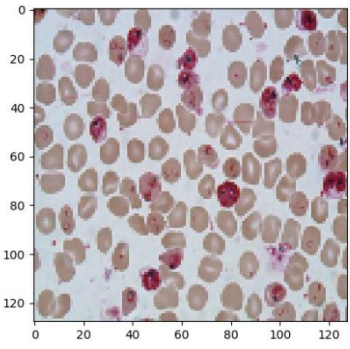
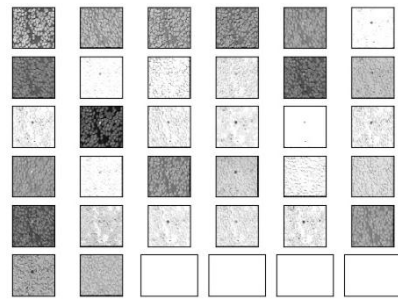
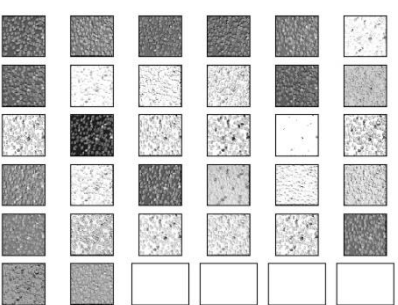
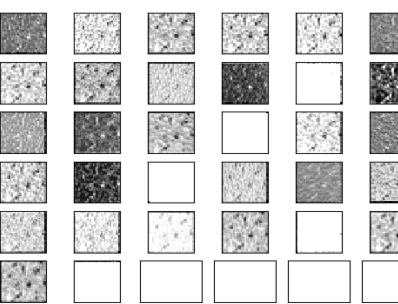


Figure 2. Block diagram for malaria parasite detection Based on CNN

The block diagram of Malaria Parasite detection is based on Convolutional Neural Network which is shown in Fig.2. The CNN based Malaria parasite detection is classified into two phases such as training and testing phases.

The number of images is divided into different categories by using label name such as infected and uninfected. In the training phase, preprocessing, feature extraction and classification loss function is performed to create a prediction model. First, label the training image set. In the preprocessing stage, image segmentation and noise-reduction is done in order to extract the blood cells and to remove the redundant image pixels simultaneously.

TABLE I. The visualized output of Convolutional layers

Layer	Output
Input	
Convolutional Layer 1	
Convolutional Layer 2	
Convolutional Layer 3	

Convolutional layer: This layer applies convolution operation to the input of the receptive field and passes the result to the next layer. Usually comprises of several filters that detect

different features. First convolutional layer usually detects simple features. Next convolutional layer will detect more complex elements and so on.

Filter size: 32, 64 Kernel size:3x3

Pooling layer: This layer collects the output of a cluster of neurons and combines it into single neuron of the next layer. Pooling can use different aggregation functions (max, average, etc.) to choose the value that represents the cluster. Pooling layer reduces size of the spatial representation (therefore the number of parameters) but retains translational invariance. It helps to control over-fitting.

Filter size: 2x2, Type: max

Fully connected/Dense layer: This layer has properties of traditional perceptron neural network in which every neuron from one layer is connected to every neuron in the next layer. Dense layers are the last layers that perform the final classification.

Filter size: 4 neurons with softmax activation.

Activation layer: This layer introduces non-linearity to the representation through resetting to zero negative input from the convolution layer. Among different activation functions ReLU proved to be the most reliable.

IV. RESULT AND DISCUSSION

In the proposed system, the dataset contains affected and unaffected blood smear images collected from the Broad Institute of Cambridge, MA. Images are in image file format such as .png or .jpg format. There are about three sets of images which consist of 1364 images (~80,000 cells) with different researchers having prepared from each one.

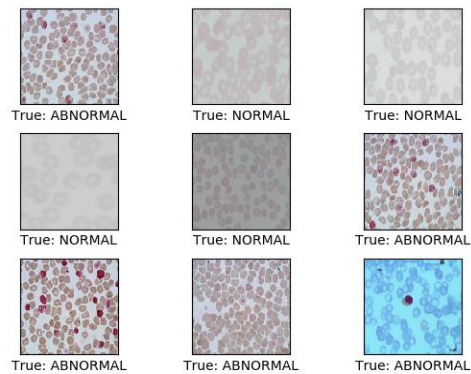


Figure 3. Experimental Results

In this work, malaria parasite detection is performed by using Convolution Neural Networks (CNN). Simulation is performed using python language. The accuracy of our work is calculated and compared with the system which makes classification based on Deep Belief Networks (DBN). The training, validation and validation loss accuracy are calculated to find the efficiency of malaria detection classification scheme. In existing technique, DBN based classification is performed for Malaria detection, but the parameters of optimization is not possible for large dataset and initialization makes the training process computationally more expensive. In the proposed CNN based classification, complexity and computation time is low and accuracy rate is high.

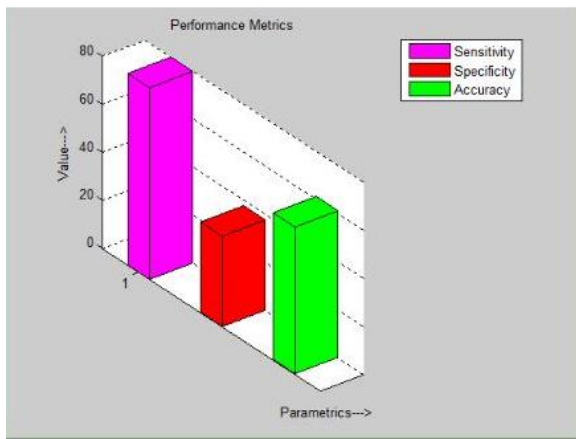


Figure 4. Performance metrics

V. CONCLUSION

The main goals of this research work is to design automatic malaria parasite classification with high accuracy, high performance and low complexity. The conventional malarial classification is performed using level method-based segmentation, textural features are extracted based on Gray level run length matrix (GLRLM) and DBN based classification are carried out. The conventional method has low complexity and high computation time, meanwhile accuracy is low. In order to reduce computation time and to improve accuracy, CNN based image classification is introduced in the proposed scheme. Also, the classification results are given as Infected or Uninfected Malaria cell images. Python language is used for simulation of the results and cell image dataset was contributed by Broad Institute in Cambridge, MA for classification purpose. The training accuracy, validation accuracy and validation loss are calculated for the proposed system. In rural areas where experienced pathologists are not available, this work paves way for accurate and cost effective diagnosis of malaria. Since this study uses deep learning techniques, it could be easily used for diagnosis of variety of other hematology-related abnormalities.

REFERENCES

1. Chan, M. (2015). World malaria report. Geneva: World Health Organization.
2. Kettelhut, M. M., Chiodini, P. L., Edwards, H., & Moody, A. (2003). External quality assessment schemes raise standards: evidence from the UKNEQAS parasitology subschemes. *Journal of clinical pathology*, 56(12), 927-932.
3. Bibin, D., Nair, M. S., & Punitha, P. (2017). Malaria parasite detection from peripheral blood smear images using deep belief networks. *IEEE Access*, 5, 9099-9108.
4. Das, D. K., Ghosh, M., Pal, M., Maiti, A. K., & Chakraborty, C. (2013). Machine learning approach for automated screening of malaria parasite using light microscopic images. *Micron*, 45, 97-106.
5. Ross, N. E., Pritchard, C. J., Rubin, D. M., & Duse, A. G. (2006). Automated image processing method for the diagnosis and classification of malaria on thin blood smears. *Medical and Biological Engineering and Computing*, 44(5), 427-436.
6. Somasekar, J., & Reddy, B. E. (2015). Segmentation of erythrocytes infected with malaria parasites for the diagnosis using microscopy imaging. *Computers & Electrical Engineering*, 45, 336-351.
7. Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), 786-804.

8. Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), 610-621.
9. Kumar, P. S., & Dharun, V. S. (2016). Hybrid brain MRI segmentation algorithm based on K-means clustering and texture pattern matrix. *International Journal of Applied Engineering Research*, 11(6), 4343-4348.
10. Elter, M., Haßlmeyer, E., & Zerfaß, T. (2011, August). Detection of malaria parasites in thick blood films. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 5140-5144). IEEE.

AUTHORS PROFILE



M. Mohana obtained her Bachelor's Degree in Computer Science and Engineering from University of Madras. Then she obtained her Postgraduate degree ME in Computer Science and Engineering from Anna University and she has completed her Doctorate in Computer Science Engineering from Anna University in 2017. She has presented papers in various National, IEEE & IET International Conferences and in International Journal. Her current research interests are Mobile Database Systems, Data Mining, Security in Distributed Systems, Data and Information Security. She has more than 18 years of teaching experience and she is currently working as Associate Professor in the Department of Information Technology, Easwari Engineering College, Chennai, Tamilnadu, India.



V. Vani is working as Associate Professor in Department of Information Technology at Easwari Engineering College, Anna University, Chennai, Tamil Nadu. She completed her Doctorate in Computer Science Engineering from Anna University in 2016 and Masters of Engineering in Computer Science and Engineering from Anna University in 2009 in Computer Science and Engineering and Bachelors from Bharathiyar University, Coimbatore, TamilNadu in 2002 in Information Technology. She has experience of 17 years in the field of teaching. She has guided many UG/PG projects and has published papers in National/International Journal and Conferences. Her primary research interest is Natural Language Processing, Text mining, Machine Learning and Software Engineering