STM JOURNALS
Journal of Artificial Intelligence Research & Advances
ISSN: 2395-6720
Volume 7, Issue 1
www.stmjournals.com

# Malaria Parasite Detection in Thin Blood Smear Images Using Deep Learning

***Rouf Ahmad Tantray*[1,*]**, ***Zahoor Ahmad Najar*[2]**
[1]Student, Department of Information Technology, Central University of Kashmir, Jammu and Kashmir, India
[2]Assistant Professor, Department of Information Technology, Central University of Kashmir, Jammu and Kashmir, India

## *Abstract*
*Algorithms in deep neural networks are being used to show the presence of malaria and to estimate the depth of infection by automatically counting individual uninfected and infected RBCs in images of thin blood smears. During the training period, the relationship was tried on a set of 13600 images from several thin blood spreads and experiment was conducted. My dataset is divided into 80% for training and 20% for testing. The relationship between the results from the algorithm and expert human readers was r = 0.866. Using image analysis results, the level of parasitemia may be achieved by applying methods to images of thin film smears. In this work, convolutional neural networks like CNN and ResNet are used for classification purposes and then classify them either infected or uninfected. Since my dataset contains only two prevalent species like Plasmodium vivax and P. falciparum that are predominant in their different cell stages and can cause 99% of the total infections. The results indicate varied distribution prevalence of P. vivax and P. falciparum, and also the mixed species infection due to these two species.*

***Keyword:*** *Malaria parasite, deep learning, computer vision, intelligent systems, convolutional neural network.*

***Author for Correspondence:** E-mail: ahmedrouf120@gmail.com

## INTRODUCTION
Transmission of malaria is caused by female anopheles mosquito bites. Then it gets mixed up with the blood stream through blood transfusion. The tropical and subtropical area particularly in the African region makes malaria a dreadful disease and causes several million deaths in those regions. One of the reasons is that the temperature is optimal and the breeding can takes place faster as compared to other regions. The World Health Organization's (WHO) [1] 11th World Malaria Report summarizes global progress in the fight against malaria up to the end of 2017. Every year people in millions die due to malaria disease. Some reports suggest that there were about 214 million cases of malaria globally in 2015 and about 438000 malaria deaths. The sub-Saharan African countries are the worst sufferers. This is because of the environmental conditions of some Africa regions that are favorable for mosquito's breeding.

The World Malaria Report 2017 [1] tells us the fact that the progress against the malaria has been stalled in several countries and that the world was unlikely to achieve the WHO Global technical strategy for malaria 2016–2030. Our main goal must be to have a proper strategy including efforts to intensify the response in the highest burden countries.

This study related to check the intensity of parasitemia inside the infected cell aims to build one such automated system—an intelligent system that can detect the presence of malaria parasites and identify amongst the specific species present in blood samples under observation.

The parasitemia of *Plasmodium vivax* is not too much heavy. The patients' blood samples which are infected erythrocytes contain only one trophozoite. In this trophozoite stage, the structure of the parasite is well-defined, having a prominent central vacuole, whereas *falciparum* is most recognizable having sickle-shaped gametocyte. *P. falciparum* is considered the most severe threat among all the species of *Plasmodium* vector, posing high rate of complications and higher death rates every year including in India.

## RELATED WORK AND LITERATURE SURVEY

Detecting malaria parasite images of thin blood smears, it is required to have a proper knowledge of observing the image. Using proper image analysis techniques, one can infer from the images the detailed FoVs (Field of Views) and corresponding to them also exists the various types of features like color, rotation and pixel density values. The learning rate is determined through the use of various neural network techniques. The diagnosis of malaria is done through demonstration of malaria parasites in the blood by preparing blood smears. A thin smear is used in determining the infecting species. Such smears are usually prepared whenever blood is drawn from a fingertip, applied to a slide, and stained by one of the stains such as *Giemsa* to highlight the parasite. A similar research was done by Anggraini et al. [2]. The proposed algorithm does not consider for segmentation the individual pixels of the image. Instead, a global threshold was obtained by varying the contrast among pixels, and was used to classify each pixel as belonging to either foreground or background. To do this, pre-processing of the images to obtain "uniform" images was done by converting all input images to gray scale. The images were then filtered to normalize the pixel intensities around a median value before obtaining the histogram of the individual images and expanding each until the two intensity classes, foreground and background, become distinct given threshold intensity. Some researches employ more morphological techniques than the two previous ones. The work done by Das et al. [3] uses thresholding, marker-controlled watershed algorithm, and Haralick textural features extraction. Another work by Kareem et al. [4] also uses morphological image transformations to infer malaria parasites from Giemsa-stained blood films. First, a grayscale image is dilated and eroded to highlight the parasites and platelets (foreground). The cells and parasites are then identified based on a combination of annular ring ratio method, size, and intensity variation. In conjunction with the techniques explored in the previously mentioned researches, machine learning techniques serve as the core for decision systems for malaria. Some like Das et al. [3], employ generative methods such as multivariate regression models. Anggraini et al. [2] used Bayes decision theory to classify images after segmentation. Other researchers use discriminative models to classify malaria images.

D. Bibin et al. [5]—malaria parasite detection from peripheral blood smear images using DBN studies in the domain of deep learning indicate that there is a profound increase in the use of deep belief networks (DBNs) due to the advancements in the effective layer-by-layer learning method. DBNs have been used in various machine learning areas, such as image analysis, speech recognition, information retrieval, voice activity detection (VAD) and natural language understanding. DBN-related models are also used for 3-D object recognition, hand-written character recognition, motion capture data modeling, information retrieval, and machine transliteration. Compared to the existing studies in the classification of malaria parasites, this study contributes by providing a more efficient and simplified approach. The first article to apply deep learning to malaria diagnosis is by Liang et al. [6], who use a convolutional neural network to discriminate between infected and uninfected cells in thin blood smear images after applying a conventional level-set cell segmentation approach. This is an ideal application for deep learning because images of segmented RBCs are a natural input for a convolutional neural network.

Gopakumar et al. [7] suggests CNN for image analysis that operate directly on focus stack of images is employed to identify malarial infection and compare the results of detection of malaria-infected RBCs in terms of sensitivity and specificity obtained by a support vector machine classifier trained on the statistical and textural features extracted from the suspected parasite locations.

Hung et al. [8] choose to use a deep learning based framework called faster region-based convolutional neural network (Faster R-CNN) because all the previous attempts were particularly made in order to automate the process of identifying and quantifying malaria using complex workflows for image processing and machine learning classification such as features from a predetermined set of measurements (intensity, shape, and texture) have been the basis of the top performing object detection models in recent years. In contrast to previous methods, this one avoids the task of segmentation and does not rely on general features for classification.

## TWO BEST APPROACHES

### Deep Learning

The full-sized images which are used for training the deep learning models, we select a patch around the point of local minimum in range of 50×50 crops of each full image. This is done in order to augment the number of training examples and cut down on training time. Enough crops of each full image were taken such that the density of cells contained in each crop will be at least twice the number of cells contained in the full-sized image up to a maximum of 100 crops [9]. Also, there will be some imbalance in the cell count where the training data is highly imbalanced and underrepresented towards RBCs, in order to create a more balanced training set, we have rotated each crop by 90 degrees, which is a kind of augmentation where the cell count of underrepresented cell roughly four times and removed crops containing only RBCs. The deep neural network like CNN does not depend on hand crafted features, it works on custom build models where it learns while training data and predicts accordingly. It directly operates on RGB candidate patch which is obtained from the best focused image having higher variance measure (Figure 1).

Dealing with patches around the suspected parasite locations, the input to CNN is decided as 32*32 which is decent enough to hold the neighborhood in making the final decision. We have chosen standard kernel size 7*7 for the feature extraction. Being a binary classification, the output neurons were set to 2 that can be either infected or uninfected. Once the input and output are set, a reasonably deep CNN must give a fair classification.

### Feature Extraction and Classification

For each 300×300 pixel input image location, 50×50 pixel patches are extracted with each individual cell image. The pixel region displays pixel information: (x, y) pixel intensity value from set of blood images. The distance measure of the object with respect to the axis x data and y data is determined with angular distance. With the help of this angular distance, the number of pixels is calculated using image analysis classifying the malaria as either non-parasite RBC and parasite. The third stage of three-stage object detection procedure involves Softmax classifier which generalizes the logistic regression classifier. Softmax classifier (SMC) is added to effectively differentiating malaria-infected and noninfected blood smear images. As the model accepts high-level data as input and reconstructs an output of better clarity and accuracy than the input [10]. This reconstructed output maps between the input and the class labels fed to the supervised model of SMC produces a practical solution of value 0 and 1, that can be interpreted as malaria-infected (abnormal) or noninfected (normal) thin blood smear image (Figure 2).

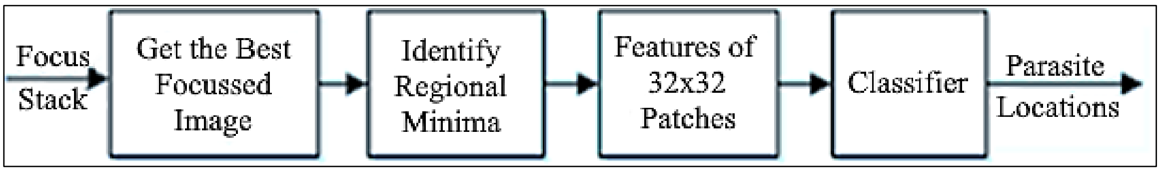


***Fig. 1:*** *Flowchart Showing Detection of Malaria Parasite Locations.*

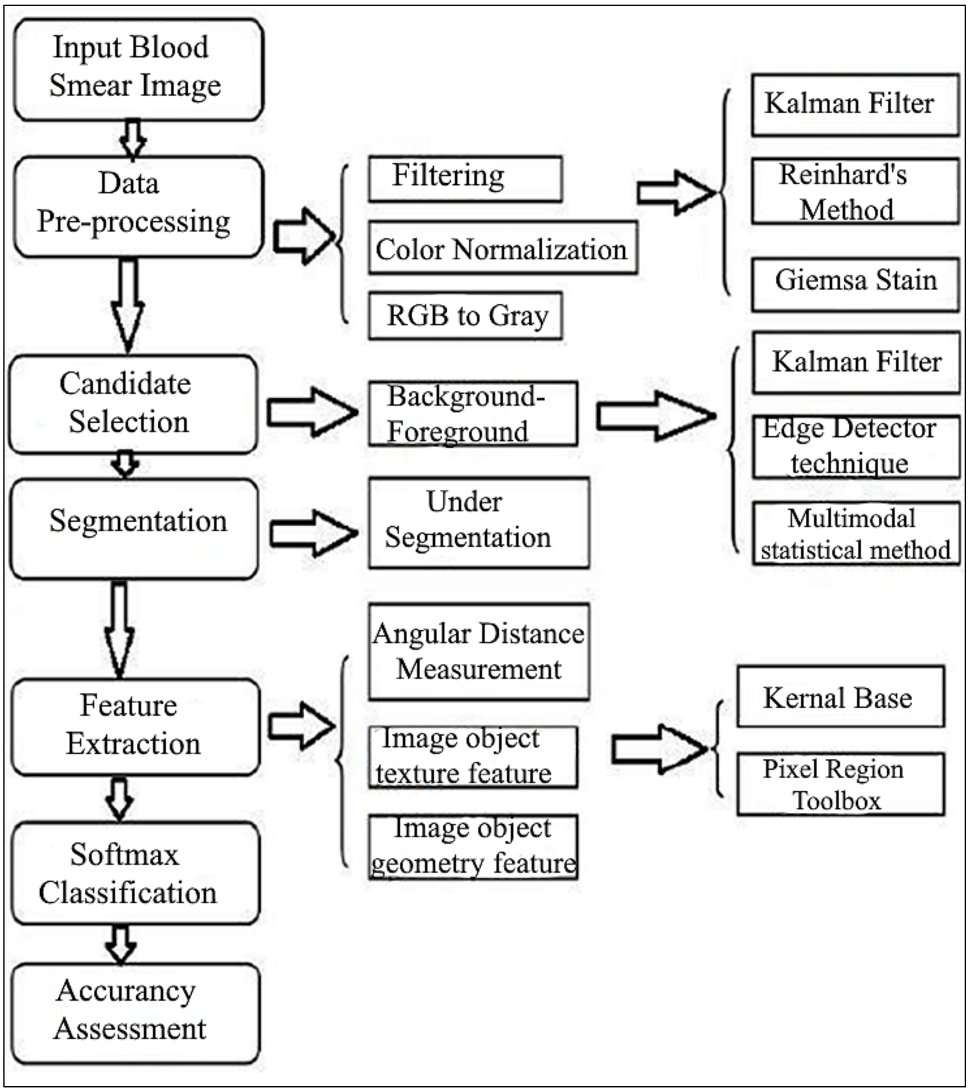


***Fig. 2:*** *Three-stage Image Detection Workflow Object Detection Method.*

## CONCLUSIONS

Conventionally, the accuracy of malaria diagnosis depends on the human expert. We will propose an automated diagnosis system using deep learning principles that will exclude the human expert from the process or serve as an aid for an expert to lower workload and improve accuracy. To propose an algorithm based on CNN that can deal with the problems inherent in low quality images.

To conclude my contributions, we have investigated and proposed a solution to the diagnosis of malaria in thin blood smear images. Also splitting the problem into two parts, firstly localization of erythrocytes, and secondly, the classification of these as uninfected or infected. The system must be able to provide quantification and exact locations of uninfected and infected blood images. Such a system would either enable a fully automated diagnosis or act as an aid to improve accuracy and lower workload of human experts. In this paper, we have proposed an automated state of the art machine which works on custom-built features, portable, cost-effective prototype system with necessary instrumentation. Image analysis/classification algorithms that can be used to quantify malaria parasitemia. The results produced in this paper suggests that implementing CNN over traditional classifiers like SVM gives considerable improvement in terms of higher specificity and higher sensitivity and also suggest operating on focus stack of images for malaria diagnosis.

## REFERENCES

1. WHO. Malaria microscopy quality assurance manual-version 2. World Health Organization; 2016. World Malaria Report 2016. World Health Organization; 2016. 3. Tek FB, Dempste.
2. Anggraini D, Nugroho A, Pratama C, Rozi I, Iskandar A, Hartono R. Automated status identification of microscopic images obtained from malaria thin blood smears. *Proceedings of the 2011 International Conference on Electrical Engineering and Informatics*, Bandung, Indonesia. 17–19 July 2011.
3. Das D, Ghosh M, Chakraborty C, Maiti A, Pal M. Probabilistic prediction of malaria using morphological and textural information. *Proceedings of the 2011 International Conference on Image Information Processing (ICIIP)*, Shimla, India. 3–5 November 2011; pp. 1–6.
4. Kareem S, Kale I, Morling R. Automated Malaria Parasite Detection in Thin Blood Films: A Hybrid Illumination and Color Constancy Insensitive, Morphological Approach; Proceedings of the 2012 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS); Kaohsiung, Taiwan. 2–5 December 2012.
5. Bibin D, Nair MS, Punitha P. Malaria parasite detection from peripheral blood smear images using deep belief networks. *IEEE Access*. 2017;5:9099–9108p.
6. Liang Z, A Powell, Ersoy I, Poostchi M, Silamut K, Palaniappan K, Guo P, Hossain MA, Sameer A, Maude RJ, Huang JX, Jaeger S, Thoma G. CNN-based image analysis for malaria diagnosis. 2016 IEEE International Conference on Bioinformatics and Biomedicine 1(BIBM), 15–18 Dec. 2016, Shenzhen, China, pp. 493–496.
7. Gopakumar G, Jagannadh VK, Gorthi SS, Subrahmanyam GRKS. Framework for morphometric classification of cells in imaging flow cytometry. *Journal of Microscopy*. 2016;261(3):307–319p.
8. Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017;39(6):1137–1149p.
9. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*. 2015;115:211–252p.
10. Alex Krizhevsky, Ilya Sutskever, Hinton GE. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*. 2017;60(6):84–90p.

**Cite this Article**
Rouf Ahmad Tantray, Zahoor Ahmad Najar. Malaria Parasite Detection in Thin Blood Smear Images Using Deep Learning. *Journal of Artificial Intelligence Research & Advances*. 2020; 7(1): 28–32p.