Malaria Detection using Microscopic Image Analysis: A Convolution Neural Network Based Approach

Uzair Hussain, Ahmad Ali, Kashif Sultan[*], Asim Alvi, and Muhammad Waleed Khan

Software Engineering Department, Bahria University, Islamabad, Pakistan
Corresponding author: Kashif Sultan (e-mail: Kashif.buic@bahria.edu.pk)

Abstract—**Malaria is a potentially fatal disease which is caused by Plasmodium parasite. These parasites are transmitted to humans through the bites of female Anopheles mosquitoes which play the role of disease vector. Five types of plasmodium cause malaria named P. Falciparum, P. Vivax, P. Ovale, P. Knowlesi, and P. Malariae, Among the Plasmodium parasites, Falciparum and Vivax are particularly lethal to humans. Therefore, early detection of malaria is mandatory to avoid the loss of human life. Different automatic/semi-automatic malaria detection techniques are available in the literature, which reduces the chance of human errors in the prognosis of malaria. In recent years, deep learning-based methods have proven to be effective for object detection. Therefore, such methods have caught the attention of researchers to use for the detection of malarial parasites in human blood. In this paper, we proposed a Convolutional Neural Network (CNN) model, which detects malarial parasites in microscopic images of human blood samples with high accuracy. The proposed model comprises 15 layers. It has 8 convolution layers with ReLu activation function, 4 max-pooling layers, 1 flattening layer, and 2 fully connected layers. The proposed method has been evaluated using various statistical measures against existing state-of-the-art methods. The quantitative measures show the effectiveness of the proposed model. It has a 97.42% testing accuracy, 97.42% sensitivity, 97.41% specificity, 97.70% precision, 97.42% recall , 97.97% F1-score , 97.41% Area Under Curve (AUC), and 94.82% Mathews correlation coefficient.**

Keywords—**Deep Learning, Malaria Detection, Convolution Neural Network,**

I. INTRODUCTION

According to the world health organization (WHO), 228 million cases of malaria were reported in 2018 across the world[1]. The early symptoms of malaria are similar to the flu-like virus and it starts within a few days after the bite of the mosquito[3]. Early treatment is essential for the people infected with malaria to avoid further spread. To diagnose malaria, pathologists put a drop of blood over the glass plate and measure the infected cells through a microscope[3]. To build a system that minimizes the dependency of skilled pathologists and replicates human performance in malaria detection, there is a demand for feature extraction and mapping of the best suitable algorithm according to features in the existing techniques[4]. A recent study states that Convolutional Neural Network (CNN) can learn features by itself like human perception. This automatic feature extraction provides a high level of results for image classification [3]. Deep learning (DL) models have proven their efficacy on medical images [6]. In recent years, DL models such as CNN, VGGNet have achieved higher accuracy to classify images. Automated detection of malaria includes dataset acquisition, data preprocessing, feature extraction, and machine-learning techniques that need human input while DL models like CNN, ResNet, and VGGNet extract these features automatically. DL models show good results in malaria parasite detection, although this field needs security and privacy of patient data[9]. Our

proposed CNN model achieves better results in accuracy, sensitivity, specificity, precision, recall, F-score, Area Under Curve (AUC), and Matthews Correlation Coefficient (MCC). Moreover, the proposed method shows outstanding performance when compared with the existing state-of-the-art models[2][7]. Figure 1 represents the block diagram of malaria detection, in general. These techniques consist of five stages i.e. (i) data acquisition, (ii) data preprocessing, (iii) segmentation, (iv) feature extraction, and (v) classification of data[5]. The rest of the paper is organized as follows: Section 2elaborates related work carried out for diagnosis and classification of the malaria parasite. Section 3 discusses the proposed method and data collection process. In Section 4, results are presented and Section 4 provides a discussion on experimental results. Last but not the least, Section 5 concludes the paper.
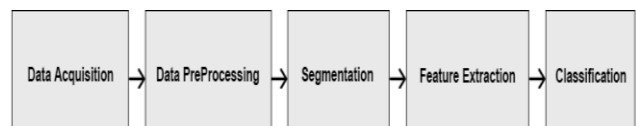

FIGURE 1. Image Processing technique for malaria detection

II. RELATED WORK

CNN-based deep learning is the recent state-of-the-art technique used for detection, classification, and segmentation[6]. Rajaraman, et al. [2] proposed a 16-layers

CNN model and achieves 95.46% accuracy in malaria detection. Vijayalakshmi et al.[4] proposed a transfer-learning system for malaria-affected cells and used VGG19 network along with a support vector machine (SVM). The VGG-SVM model was compared with LeNet-5, AlexNet, and GoogleNet and was found better in accuracy (i.e., 93.13%) than all these models. Liang et al. [7] proposed a 16-layers CNN architecture and the accuracy was 97.37%. Hung et al. [10] trained the Faster R-CNN model over vivax images and detect the object in malaria images. They applied Region Proposal Network (RPN) based CNN, as a first stage, on malaria images for binary classification of RBC and non-RBC objects, and applied AlexNet, as a second stage, on non-RBC objects for detection of further classes.

Narayanan et al.[11] proposed a CNN model and compared its performance with the transfer learning models, i.e. ResNet, VGG19, VGG16, and Dense Net. Reddy et al.[12] trained ResNet-50 over 27,558 falciparum images for malaria classification; claimed to achieve 95.91% training accuracy and 95.4% validation accuracy. Prasad et al. [15] develop a decision support system to detect malaria from thin blood smear images. This study includes 200-images, which are plasmodium vivax and falciparum. Bibin et al. [16] applied Deep Belief Network, for the first time for malaria detection, and 4100 infected and non-infected images were used.

Timely detection of malaria provides a way to detect malaria parasites in an accurate aspect. Expert technicians require for the effective detection of malaria. The whole process is annoying and faulty due to the subject and might be false medication and even death of the patient [16]. Dong et al. [17] applied automated measurement of malaria diagnosis through DL techniques. The dataset includes malaria images and classifies the three CNN models that are LeNet, Alex Net, and Google Net, and compared with SVM. Deep learning models have an accuracy of 95%, which is more than 92% obtainable by SVM.

Pamungkas et al.[19] used a decision tree algorithm for malaria classification which gives 87.67% accuracy. Razzak et al. [26]used artificial neural network for malaria classification, dataset was falciparum, vivax, malariae and ovale. Suryawanshi et al. [29]used SVM and Euclidean distance classifier, SVM gives 93.33% accuracy while Euclidean distance classifier provides 83.33% accuracy in malaria classification. Sharma et al.[33]predict malaria using SVM and ANN. SVM has high accuracy and early detection of malaria than ANN, 1680 samples are used.

Go et al.[34]used 6 machine learning techniques for malaria classification and proposed that SVM has high results with an accuracy score of 96.78% than other machine learning approaches that are logistic regression, decision tree, k nearest neighbor (KNN), linear discriminant classifier, quadratic discriminant classifier. For this purpose, 280 infected images are used and classification is done with 10-fold validation. Park et al.[35] identify plasmodium falciparum stages that are schizont and trophozoite using Linear Discriminant Classification (LDC), Logistic Regression (LR), and KNN. LR is based on the maximum

likelihood method, while KNN finds adjacent points and classification based on the maximum nearest majority. Infected and uninfected 1237 plasmodium falciparum images were used for classification purposes with equal dataset splitting.

Rosado et al.[36] used SVM classifier and achieved 98.2% sensitivity and 72.1% specificity in the identification of white blood cells (WBC) of Giemsa stained blood cells and 80.5% sensitivity, 93.8% specificity in the identification of P. falciparum trophozoite stages.

TABLE I. LITERATURE OF THE EXISTING TECHNIQUES

Author	Deep Learning Model	Dataset
Rajaraman, et al.[2]	Pre-trained 16-layers-CNN	27,558 falciparum dataset
Bana *et al.[3]*	CNN-Model	Thin Blood Smears
Vijayalakshmi et al. [4]	*VGG-SVM*	*falciparum images*
Das *et al.*[5]	Multi-Layer Perceptron	Falciparum and vivax images
Liang et al.*[7]*	*16-Layers CNN*	*27,558 falciparum Images*
Hung et al.[10]	Faster R-CNN	Vivax images
Reddy et al.[12]	ResNet-50	27,558 falciparum Images
Bibin et al.[16]	6-Layers deep CNN	Thin Blood images
Dong et al. [17]	LeNet, AlexNet, GoogleNet	1034 infected and 1531 non-infected cells
Mehanian et al. [23]	Computer Vision System	Falciparum dataset
Horstmeyer et al.[24]	CNN	falciparum
Linder et al.[25]	Computer Vision Algorithm	Mamic Database
Machine Learning Model		
Nugroho et al.[27]	Histogram	Thin Blood Images
Pamungkas et al.[19]	Decision Tree	Falciparum dataset
Suryawanshi et al.[29]	Euclidean Distance Classifier and SVM	Infected and non-infected RBC cells
Das et al.[30]	Machine Learning Model	falciparum and vivax images

Jagtap et al. [37]proposed a machine learning model with high accuracy results when compared to SVM and bayesian models in the study. Savkare et al.[38] used SVM for malaria detection and achieved 96.42% accuracy. SVM classifier is used to classify blood cells as infected or non-infected. Linear, polynomial, and RBF kernels are used for the detection of malaria, and RBF kernels give high accuracy on 71 images in the identification of stages than other SVM kernels. Pandit et al. [39]used ANN for malaria detection and gives 90% accuracy. Table I represents the literature on the existing machine learning and deep learning techniques that were used for malaria detection.

Our proposed CNN architecture has achieved greater accuracy, sensitivity, precision, accuracy, specificity, recall, f-score, AUC, and MCC than the model of transfer learning.

III. PROPOSED METHOD

The proposed method is comprised of different steps, i.e., (1) data collection, (2) pre-processing, (3) proposed model, (4) splitting the dataset, (5) training the dataset, (6) testing the proposed model and the last section is about CNN-architecture. Each block of the diagram is explained as follows: Figure 2 describes the malaria parasite types and their species, Figure 3 explains the CNN architecture for malaria detection and Figure 4 describes the training and validation accuracy/loss.

A. DATA COLLECTION

The Dataset was collected from the GitHub repository [21] and Kaggle [22], it contains 34,625 microscopic images. Figure 2 represents different types of the malaria parasite and their species. Deep learning models for malaria detection[2][7][12] were evaluated on a 27,558 image dataset, our proposed 15-layers CNN model uses 34,625 microscopic images and shows better results.
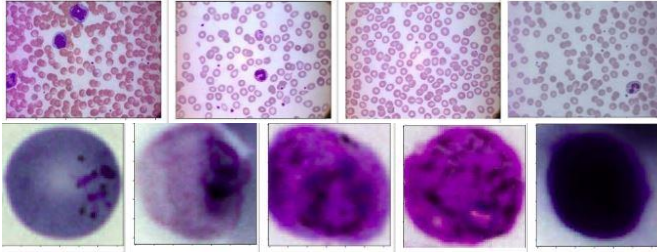

FIGURE2. In this figure Malaria parasites types are in first row i.e. falciparum, vivax, ovale, malariae and its species are in the second row i.e. ring, trophozoite, schizont, gametocyte and leukocyte

Github repository[21]contains malaria Plasmodium types, i.e., falciparum, vivax, and ovale parasites, while Kaggle[22]contains all species of malaria that are ring, Leukocyte, Gametocyte, Trophozoite, Schizont.

B. PRE-PROCESSING

In order to preprocess the dataset, we resize the image into 64x64 pixels for the training and testing.

C. PROPOSED MODEL

The proposed model is comprised of eight convolutional layers, four pooling layers, one flatten and two fully connected layers. The sequence of layers is such that each two convolutional layers are followed by a pooling layer. Then, one flatten and two fully connected layers, are applied in the network. There are 848,018 parameters in the network and all parameters are trained. The batch size is 86, and total number of steps per epoch is founded by dividing 86 to total trainable examples, all filters have 3 x 3 size. Filter starts from 16 and ends on 128, two 16, two 32, two 64, and after that two 128. Relu [13]is used as an activation function and Batch normalization[14]is employed in the proposed model to avoid vanishing gradient problem.

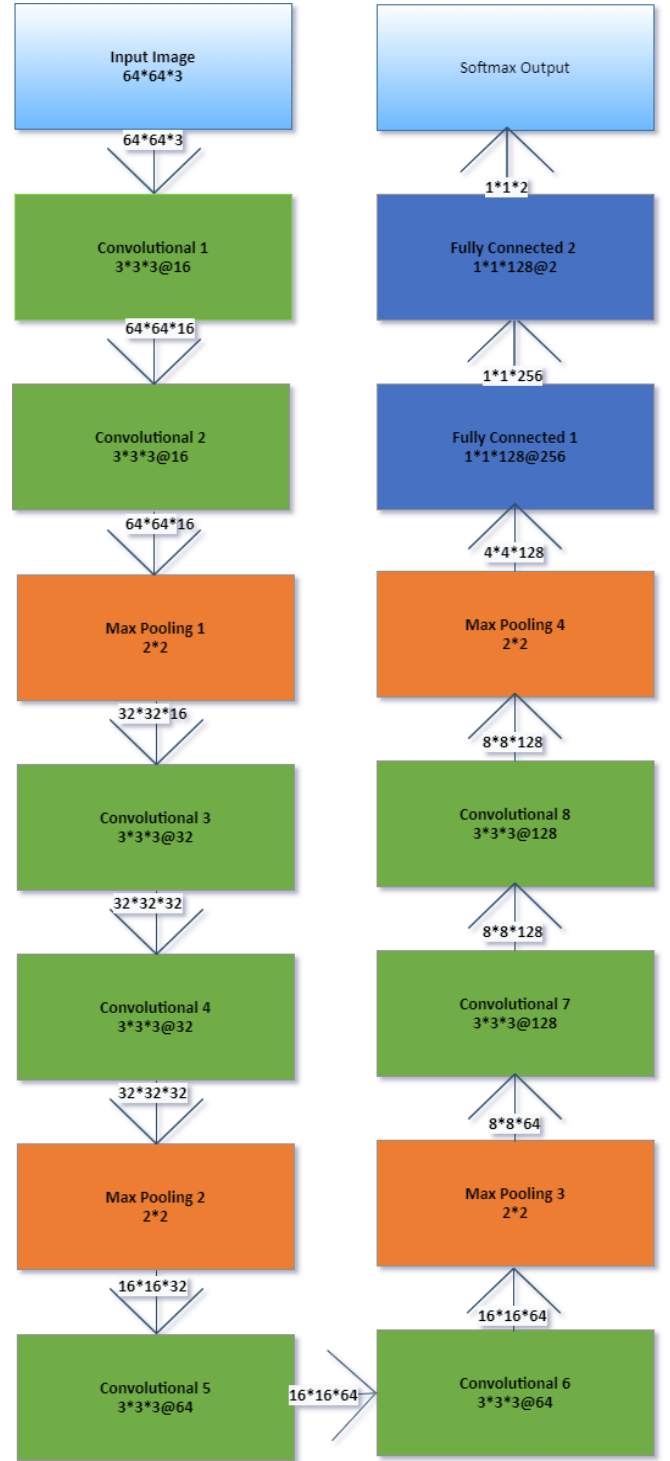Figure 3 represents the 15-layers CNN architecture for malaria detection.

FIGURE 3.CNN Architecture for Malaria Detection

D. IMPLEMENTATION AND TRAINING

In order to implement the proposed model, we employed Python version 3.6.9, Keras 1.14.0 and Tensor Flow 2.3.1, Ubuntu 18.04, Jupyter Notebook, 8GB RAM. We used 70% of the dataset for training,, 15% dataset for validation and 15% dataset for testing. Softmax is used as the activation function on the fully connected layer.

Categorical_crossentropy is used as loss parameter, and this parameter applied for the two label classes. ADAM is used as an optimizer, and it is applied to speed up the training process. The number of epochs is 30.

TABLE II. MACHINE LEARNING, DEEP LEARNING AND IMAGE PROCESSING ARCHITECTURES FOR MALARIA DETECTION

Author	Accuracy	Model
Proposed Model	**97.42%**	**15-layers-CNN**
Rajaraman et al. [2]	*95.46%*	*16-layers-CNN*
Vijayalakshmi et al.[4]	*93.13%*	*VGG19*
Liang et al.[7]	*97.37%*	*16-layers-CNN*
Prasad et al. [15]	96%	Image-Processing
Daset al.[5]	96.73%	Multi-Layer Perceptron
Reddy et al.[12]	95.4%	ResNet-50
Bibinet al.[16]	96.21%	6-Layers deep CNN
Dong et al. [17]	95%	LeNet, AlexNet, GoogleNet
Mehrjouet al.[9]	87.5%	Image-Processing
Pamungkaset al.[19]	87.67%	Machine Learning
Imran Razzak et al.[20]	95.10%	Machine Learning and Deep Learning
Narayananet al.[11]	96.7%	Machine Learning and Deep Learning

In the whole process, 15% dataset is used for testing that contained 5194 examples of images and this achieved a testing accuracy of 97.42%. Implementation of the proposed CNN model and results are available at https://github.com/uzairhussain777/Malaria-Detection-Using-CNN-97.42-Accuracy.

IV. RESULTS & DISCUSSION

The results of our model demonstrate its accuracy for malaria detection using microscopic images. The proposed CNN model achieves 97.42% accuracy. Table II represents that the proposed model has high accuracy in malaria detection than existing image processing, machine learning, and deep learning techniques. Previously, a 16-layer model [2] was proposed in which the dataset of 27,558 falciparum images was used. Whereas, in our research study we have used the dataset with 34,625 parasite images that contain all four plasmodium and their species as shown in Fig. 2. Computational cost is lower than in previous work[2], and the accuracy of detecting malaria is high as shown in Fig. 4. Table III represents the results of the CNN model and transfer learning model on the same dataset and proposed CNN model achieves better results in accuracy, specificity, sensitivity, recall, precision, f-score, MCC, and AUC.

TABLE III. REPRESENTS THE RESULTS OF CNN MODEL AND TRANSFER LEARNING MODEL ON THE SAME DATASET

Measures	Transfer Learning	CNN-Model
Accuracy	94.35%	97.42%
Specificity	91.75%	97.41%
Sensitivity	91.76%	97.42%
Recall	95.39%	97.42%
Precision	95.12%	97.70%
F1-Score	95.16%	97.97%
MCC	88.89%	94.82%
AUC	94.58%	97.41%

Figure 4 represents the training accuracy, validation accuracy, training loss, and validation loss. Table IV shows the training, validation accuracy and loss. Figure 4 shows that our proposed model achieves 97.42% accuracy for malaria detection. It is obvious from the Figure 4 that the proposed model has a loss lower than 1%.

TABLE IV. TRAINING, VALIDATION ACCURACY AND LOSS

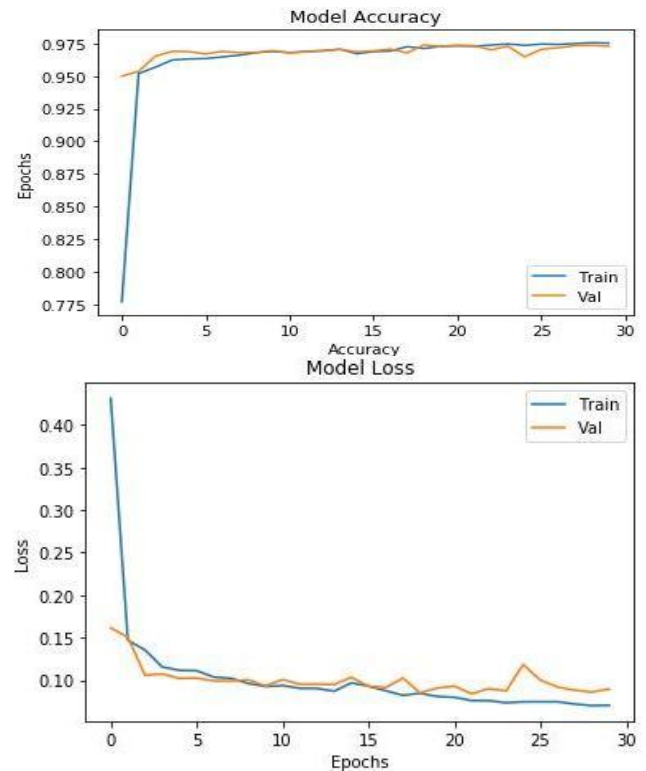| Metric	Values
Training Accuracy	97.56
Training Loss	.0703
Validation Accuracy	97.40
Validation Loss	.0842

FIGURE **4.** Training, validation accuracy, and loss

The state-of-the-art model [2] uses 10 convolutional layers, 5 pooling layers, and 1-flatten, and two fully connected layers, in other words, the 16-layer CNN model was used. To further evaluate the performance of the proposed model, the different performance metric of the proposed model is compared with the transfer learning model. The proposed CNN architecture achieves better results in malaria detection when discussing accuracy, sensitivity, specificity, precision, recall, f-score, AUC, MCC than existing models ResNet50[12], VGG19[4], CNN-Model [2][7].

V. CONCLUSIONS

As malaria is among dangerous epidemics, therefore, there is a need for an intelligent system that detects the malaria disease accurately and efficiently. In this research study, a 15-layers CNN model was presented for malaria detection. The study illustrates that the proposed deep learning-based model achieved better results in diagnosing malaria. Furthermore, to evaluate the performance of the proposed model, the transfer-learning model is also trained on the same dataset and it is observed that CNN model achieves better results in terms of accuracy, sensitivity, specificity, precision, recall, f-score, AUC and MCC being above 97%. This also proves to be better when compared with existing state-of-the-art models such as VGG19, ResNet-50, LeNet, AlexNet, and GoogleNet. In this study, more than 34,000 images are used in the training process, and with lesser complexity, the computation cost is used.

REFERENCES

[1] World Health Organization 2018, [Online] Available: https://www.who.int/malaria/en/ (Accessed on: 7 May, 2020).

[2] Rajaraman et al., Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. PeerJ, 2018. vol. 6, p. e4568.

[3] Bana et al., Plasmodium Detection Using Simple CNN and Clustered GLCM Features. arXiv preprint arXiv:1909.13101, 2019.

[4] Vijayalakshmiet al., Deep learning approach to detect malaria from microscopic images. Multimedia Tools and Applications, 2019: p. 1-21.

[5] Das et al., Automated system for characterization and classification of malaria-infected stages using light microscopic images of thin blood smears. Journal of microscopy, 2015. vol. 257, no.3, p. 238-252.

[6] Poostchi et al., Image analysis and machine learning for detecting malaria. Translational Research, 2018. vol. 194: p. 36-55.

[7] Liang et al., CNN-based image analysis for malaria diagnosis. in 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2016. IEEE.

[8] Pan et al., Classification of malaria-infected cells using deep convolutional neural networks. Machine Learning: Advanced Techniques and Emerging Applications, 2018. vol. 159.

[9] Mehrjou al., Automatic malaria diagnosis system. in 2013 First RSI/ISM International Conference on Robotics and Mechatronics (ICRoM). 2013. IEEE.

[10] Hung et al., Carpenter. Applying faster R-CNN for object detection on malaria images. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017. p. 56-61.

[11] Narayanan et al., Performance analysis of machine learning and deep learning architectures for malaria detection on cell images. in Applications of Machine Learning. 2019. International Society for Optics and Photonics. vol. 11139: p. 111-390

[12] Reddy et al., Transfer Learning with ResNet-50 for Malaria Cell-Image Classification. in 2019 International Conference on Communication and Signal Processing (ICCSP). 2019. IEEE. p. 0945-0949

[13] Mishkin et al., Systematic evaluation of convolution neural network advances on the imagenet. Computer Vision and Image Understanding, 2017. vol. 161: p. 11-19.

[14] Ide et al., Improvement of learning for CNN with ReLU activation by sparse regularization. in 2017 International Joint Conference on Neural Networks (IJCNN). 2017. IEEE. p. 2684-2691.

[15] Prasadet al., Image analysis approach for development of a decision support system for detection of malaria parasites in thin blood smear images. Journal of digital imaging, 2012. vol. 25, no. 4, p. 542-549.

[16] Bibin et al., Malaria parasite detection from peripheral blood smear images using deep belief networks. IEEE Access, 2017. vol. 5: p. 9099-9108.

[17] Dong et al. Classification accuracies of malaria infected cells using deep convolutional neural networks based on decompressed images. in SoutheastCon 2017. IEEE. p.1-6

[18] Mohanty et al., Automatic Detection of Malaria Parasites Using Unsupervised Techniques. in International Conference on ISMAC in Computational Vision and Bio-Engineering. 2018. Springer. p.41-49.

[19] Pamungkas et al., Identification of plasmodium falciparum development phase in malaria infected red blood cells using adaptive color segmentation and decision tree based classification. International Journal of Applied Engineering Research, 2015. vol. 10, no. 2, p. 4043-4056.

[20] Imran Razzak et al., Microscopic blood smear segmentation and classification using deep contour aware CNN and extreme machine learning. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017. p.49-55.

[21] Awesome_Malaria_Parasite_Imaging_Datasets, [Online]Available: https://github.com/tobsecret/Awesome_Malaria_Parasite_Imaging_D atasets (Accessed on: 7 May, 2020).

[22] Malaria Dataset, [Online] Available:https://www.kaggle.com/dharun/malaria-bbbc (Accessed on: 7 May, 2020).

[23] Mehanian et al. Computer-automated malaria diagnosis and quantitation using convolutional neural networks. in Proceedings of the IEEE International Conference on Computer Vision. 2017. p.116-125

[24] Horstmeyer et al., Convolutional neural networks that teach microscopes how to image. arXiv preprint arXiv:1709.07223, 2017.

[25] Linder et al., A malaria diagnostic tool based on computer vision screening and visualization of Plasmodium falciparum candidate areas in digitized blood smears. PLoS One, 2014. vol. **9**, no. 8, p. e104855.

[26] Razzak et al. Automatic detection and classification of malarial parasite. International Journal of Biometrics and Bioinformatics (IJBB), 2015. vol. 9, no. 1, p. 1-12.

[27] Nugroho et al., Feature extraction and classification for detection malaria parasites in thin blood smear. in 2015 2nd international conference on information technology, computer, and electrical engineering (ICITACEE). 2015. IEEE. p. 197-201.

[28] Quinn et al. Deep convolutional neural networks for microscopy-based point of care diagnostics. in Machine Learning for Healthcare Conference. 2016. p. 271-281.

[29] Suryawanshi et al., Comparative study of Malaria parasite detection using euclidean distance classifier & SVM. International Journal of AdvancedResearch in Computer Engineering & Technology (IJARCET), 2013. vol. **2**, no. 11, p. 2994-2997.

[30] Das et al., Machine learning approach for automated screening of malaria parasite using light microscopic images. Micron, 2013. vol. 45: p. 97-106.

[31] Bashir et al. Detection of malaria parasites using digital image processing. in 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE). 2017. IEEE. p. 1-5.

[32] Ross et al., Automated image processing method for the diagnosis and classification of malaria on thin blood smears. Medical and Biological Engineering and Computing, 2006. vol. 44, no. 5, p. 427-436.

[33] Sharma et al., Malaria outbreak prediction model using machine learning. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 2015. vol. 4, no. 12.

[34] Go et al., Machine learning based in-line holographic sensing of unstained malaria-infected red blood cells. Journal of biophotonics, 2018. vol.11, no.9, p. e201800101.

[35] Park et al., Automated detection of P. falciparum using machine learning algorithms with quantitative phase images of unstained cells. Plos one, 2016. vol. 11, no.9.

[36] Rosado et al. Automated detection of malaria parasites on thick blood smears via mobile devices. 2016.

[37] Jagtap et al. Heuristic scale to estimate premature malaria parasites: scope in microscopic blood smear images. Indian Journal of Science and Technology, 2017. vol. 10, no.8, p. 1-13.

[38] Savkare et al. Automatic system for classification of erythrocytes infected with malaria and identification of parasite's life stage. Procedia Technology, 2012. vol. 6, p. 405-410.

[39] Pandit et al. Artificial neural networks for detection of malaria in RBCs. arXiv preprint arXiv:2016.