JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

ANALYSIS AND DETECTION OF MALARIA DISEASE BY DEEP LEARNING MODELS

Dev Jignesh Mehta,
High School Student
Reliance Foundation School
Surat, India

Abstract: *Medical diagnosis is growing filed in today's era. One of popular infectious diseases in the world is Malaria which is spread through Plasmodium parasites. Lack of availability of Malaria diagnostic facilities, efficient technology is required for that. Recent deep learning algorithms are used to solve any image classification problems with highest accuracy in less amount of time. The datasets for malaria disease detection are downloaded from kaggle which made up of large number of infected and uninfected images. Two deep learning models namely VGG 19 pre-trained model and Own CNN model are applied to detect malaria disease. Deep convolution neural network Training model will train hundreds of images and if user enter some unseen image then prediction model will detect disease. These two methods classify images into training and testing images, then apply model algorithm and then do analysis to find out accuracy of model. In the analysis part image acquisition means it will take an image from your device and give it to the two different CNN model for the image processing part. Model has large set of data for identification of result. Own CNN model was already trained with the hundreds of images, so it will predict that whether the person is infected with disease or not, Using only input as an image. Outcome of CNN model give better accuracy then VGG 19 pretrained model.*

Keywords: Convolution Neural Network, Image Acquisition, Infected, Uninfected etc.

1. Introduction

As per current World malaria report, 241 million cases in 2020 & 227 million cases in 2019 so huge number of cases are rising day to day [1]. Malaria death report indicates that 69000 deaths were in 2019 which will increase in 2020 with 627000 cases [1]. Spreading of Malaria due to mosquito bite infected person and noninfected person bite by infected mosquito. Malaria disease starts to enter with bloodstream of person and reach to their liver. When this infection becomes mature, it leaves liver and red blood cell infected by it.

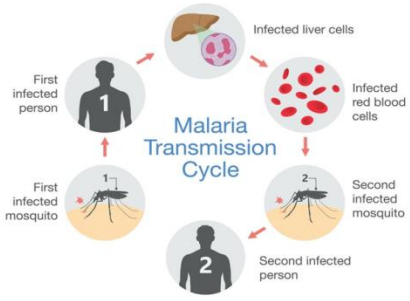


Figure 1: Malaria Transmission Cycle [13]

Researchers mention various conventional diagnostic tools &techniques for examination of Malaria disease like microscopy test [2,3,5,7,9,10,12] ,rapid diagnostic test (RDT)[3,4,5,6,7,8,9,12] , Plasmodium species-specific (PCR)[10] and many more. Clinical diagnosis processes depend on laboratory human specialization so sometimes examinations of disease give inaccurate results [11]. Traditional method to diagnosis for malaria is to take blood samples of patient, put it into a slide and use of microscope to calculate

infected RBC count [12]. Traditional methods apply to diagnosis of disease are very slow and it takes more time. Now a day, researchers have developed various deep learning models to identify and classification of infected malaria images from whole datasets [11].

This research paper highlights datasets of malaria disease infected and non infected images which gather from kaggle online datasets. Then it identify that input image of malaria is infected or not infected using two deep learning model namely VGG pretrained model and own creation of CNN model. Then after analysis of both models have performed.

2. Datasets Description

Malaria disease datasets are gathered from online data resource kaggle which is open-sourced. These datasets are made up of two types of images which are parasitized and uninfected. These images are in PNG format. All datasets images are resized with 64 x 64 pixels to speed up processing and analysis. Figure 2.1 and 2.2 are samples of Malaria disease for uninfected and parasitized. Datasets contain training and testing images. We have taken 13870 parasitized and 13870 uninfected training images & 91 parasitized and 43 uninfected testing images for disease analysis. Table 2.1 shows datasets summary for disease detection.

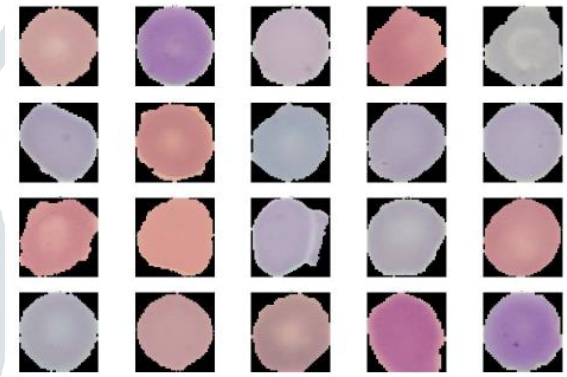


Figure 2.1 Uninfected Malaria RBC Image

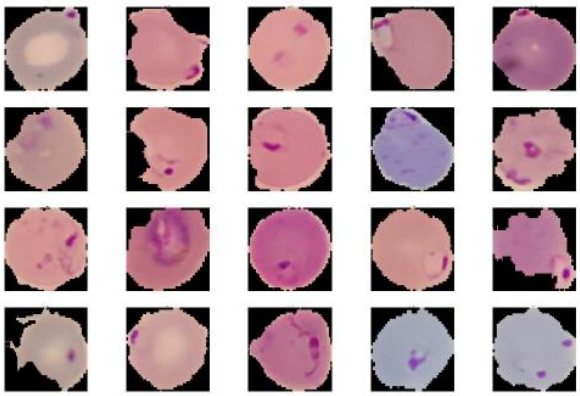


Figure 2.2 Infected Malaria RBC Image

Table 2.1 Malaria disease datasets summary

3. Methodologies to detect Malaria Disease Detection

Convolution Neural Network is model of deep learning which type of Artificial Neural Network is. Application of CNN are used mostly for problems of Image classifications, Image recognition etc [21][22]. Here we used two different CNN models namely pretrained model – VGG 19 and new creation of CNN model to recognition of Malaria disease. For that, publicly available datasets of Malaria diseases is used. Steps to identify disease are described in figure 3.1.

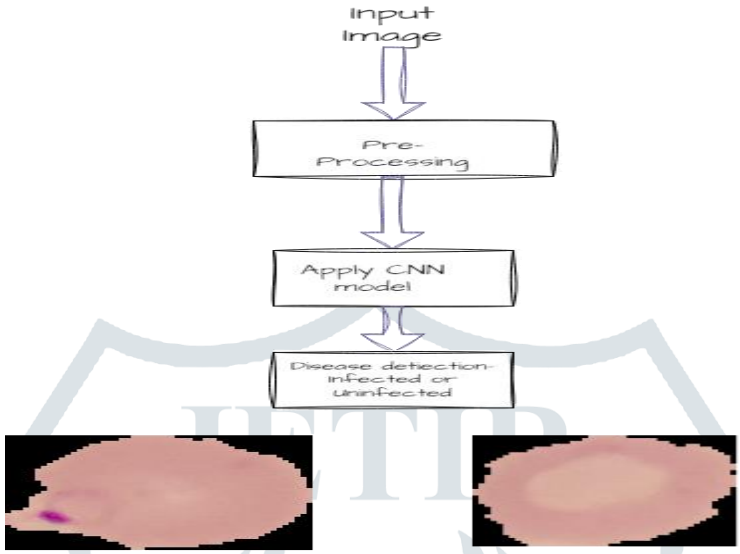


Figure 3.1 Steps for Malaria Disease detection

As per above diagram, The input image is applied to first pre-processing stage which is removed unnecessary noise from RGB cell image and also convert all different input image size into same 64x64 image size. Then we apply to two different convolution neural networks namely VGG 19 and new CNN model to identify malaria disease detection and also do analysis of these two models. Final output of deep learning models is the classified label as infected and non-infected.

3.1 Preprocessing stage

First step based on figure 3.1 is preprocessing which takes malaria images as input. These images sizes are required to resize with same size as all inputs contain different image sizes. [17][18]. so each and every images converted into with 64 x 64 pixels. Then it given to deep learning model.

3.2 VGG 19 Model

VGG model takes input images with size 224*224 RGB image. It applies to preprocessing layers convert all images into 224*224 pixel size. VGG19 consists of 19 weight layers which have 16 convolution layers, 3 fully connected layers, 5 MaxPool layers and 1 SoftMax layer. Figure 3.2 consist of VGG 19 architecture.

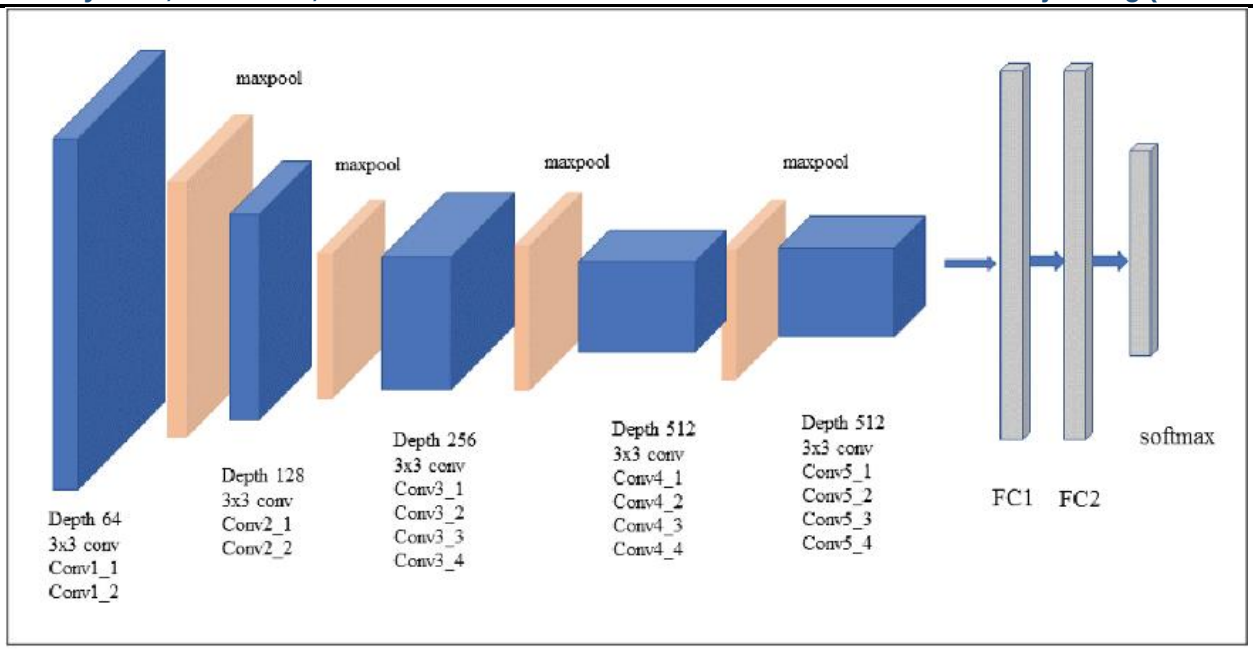


Fig. 3. VGG-19 network architecture

As per above network architecture, 16 convolution layers are used. In which , first two convolution layer CON1_1 and CON1_2 apply with 64 filters with 3*3 filter size with stride of 1 are used. Next, two convolution layer CON2_1 and CON2_2 apply with 128 filters with 3*3 filter size with stride of 1 are used. Then 4 convolution layers CON3_1, CON3_2, CON3_3 and CON3_4 are applied with 256 filters with 3* 3 and stride 1. Then again 4 convolution layers CON4_1, CON4_2 and CON4_3 and CON4_4 with 512 filters with 3*3 and stride 1 and finally 4 convolution layers CON5_1,CON5_2, CON5_3 and CON5_4 with 512 filters with 3*3 and stride 1. As per above diagram, five maxpooling layers with size 2*2 and 3 fully connected layers with size 4096 are used.

3.2 Convolution Neural Network (CNN) Model

CNNs contain a combination of layers which transform an image into output the model can understand. In neural networks, Convolution neural network is one of the main categories to do images recognition, images classifications. Object's detections, recognition faces etc., are some of the areas where CNNs are widely used [14-18].

Technically, deep learning CNN models to train and test, each input image will pass it through a series of convolution layers with filters (Kernals), Pooling, fully connected layers (FC) and apply Softmax function to classify an object with probabilistic values between 0 and 1.

Convolution layer is explained in figure 3.2.1. In the below figure two matrixes are there one matrix 5*5 is the input matrix and another 3*3 matrix is filter which is applied in every convolution neural network. In one convolution neural network filters are used for recognize or identify every single activity in the image.

Fig 3.2.1 Input image 5 x5 convolve Operation with 3x3 filter matrix

In the malaria tissue image, there is also lots of different identification so using that different identity the decision can be making. So basically, to classify one image into multiple images to identify the accurate image the filters are used.

So basically, different filters are used to identify the different activity, like identify the horizontal line, vertical line, circle, triangle and many more things. So how many filters are used in convolution neural network is defined within the VGG19 model. We have used pre-trained model so filters are defined within a model.

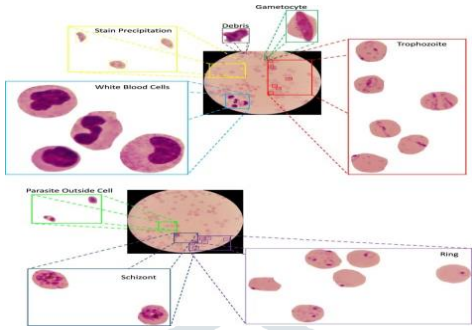


Figure 3.2.2. Disease image after apply filters

You can see in above image if we apply multiple filters to disease one image then how it will identify multiple things and convert it into multiple pixels.

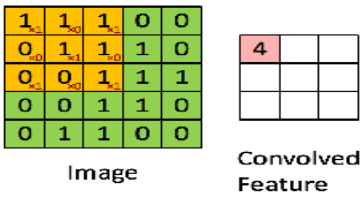


Figure 3.2.3 Convolved operation output

Image will convert in multiple images like we can see in above image one image will convert in stain precipitation, Debris, Gametocyte, Trophozoite, White Blood Cells, Parasite Outside Cell, Schizont, Ring. So all this will come from applying filters on the one image.

Max pooling operation functionality is shown in figure 3.2.4. This operation show both types of pooling operations namely max pooling and average pooling which apply to convolved feature matrix [21][22] .

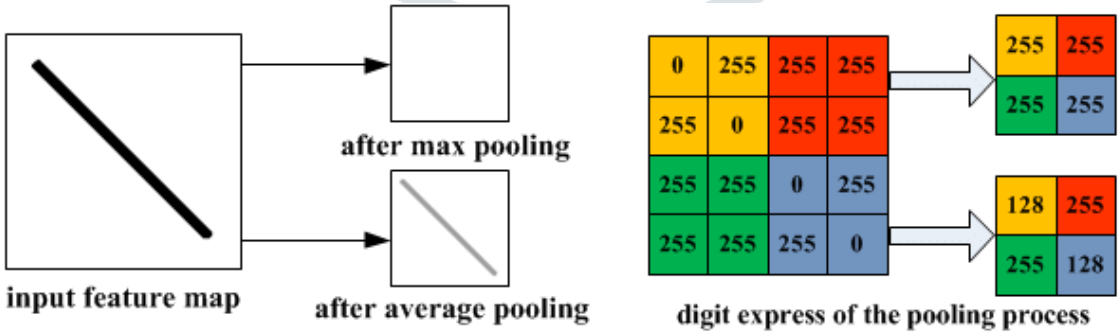


Figure 3.2.4 MaxPooling Operation

We can see in above figure that there is an input feature image we are going to max pooling that image. So, there is a 4*4 convolved matrix of image and we are using 2*2 max pooling matrix.

First, we take a 2*2 matrix from the convolved 4*4 matrix values are 0, 255, 255, 0 shown in yellow color from this value we select maximum value, so here 255 is maximum so we take it and put it in our max polling 2*2 matrix. Again, take another 2*2 matrix from

convolved 4*4 matrixes shown in red color and find maximum value and put it in the max pooling 2*2 matrix. So, repeat this process until max pooling matrix will fill completely. Outcome of pooling layer is to generate number of features which reduction in size

Fully Connected layer, Flatten and Softmax layer (Classification):

Flatten take input from pooling layers and convert feature vector into single vector which will become input for fully connected layer [14-18]. Flatten operation shown in following figure 3.2.4 which converts 4*4 matrixes into single vector. Fully Connected Layer (also known as Hidden Layer) is the last layer in the convolution neural network. This layer is a combination of Affine function and Non-Linear function. Fully connected layer is also known as Hidden layer.

Multiple such layers can added based on depth for classification purpose. Then final output receive from fully connected layer is given to softmax or sigmoid function for final classification. The combination of Flatten Layer with Fully Connected Layer and Softmax Layer used for Classification of Deep Neural Network[21][22].

Figure 3.2.5 Flatten Operation

CNN model [21][22] is developed to detection of malaria disease. This model is created with 3 convolution layers with 32 filters with input size 48 by 48, Max pooling layers with size of 2*2, four Batch normalization and four drop out layers. CNN model architecture is shown in following table 3.3.1.

4. **Result analysis of Malaria disease detection**

Malaria disease detection analysis is performed with VGG 19 and CNN model with various epochs which described in 4.1 and 4.2.

4.1 Our Own Created CNN (Convolution Neural Network) Model:

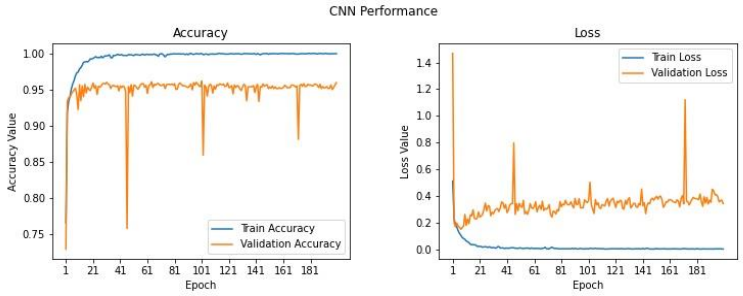


Figure 4.1 Our Own CNN models Accuracy and Loss

In above figure we can see that training accuracy is almost 99%, testing accuracy is increasing and reached up to almost 90-95%. Training loss is nearest to the 0.01 and testing loss is fluctuating between 0.2-0.4 which was good for now.

4.2 VGG19 (Pretrained Model)

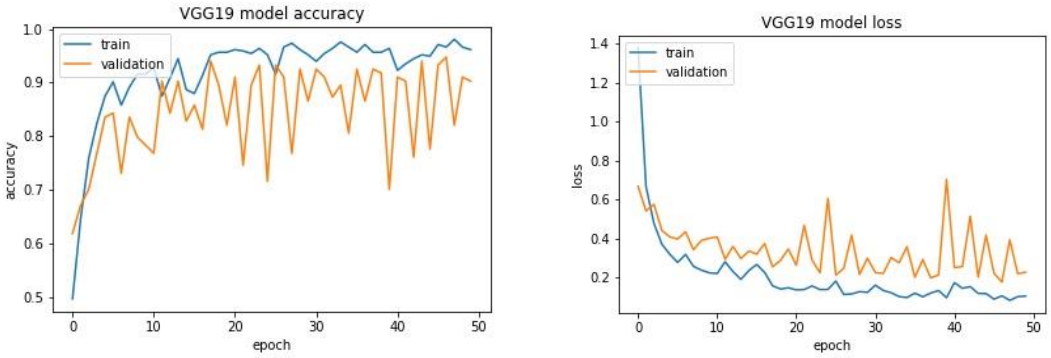


Figure 4.2. VGG 19 Model Accuracy and loss

VGG 19 is pretrained model, which is already available on google. Just we have to import and apply some required parameter and train on dataset. We got 85-90% accuracy and 0.2-0.5 loss with this model.

4.3 Model Analysis table

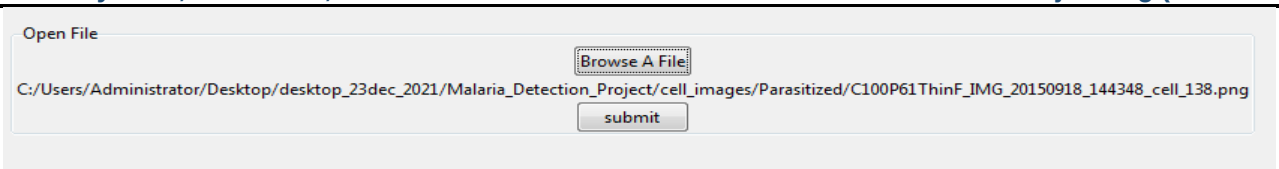
From the below table 4.3 we conclude that we used our own created CNN model for training our dataset. Because that model always gives us around 99% accuracy and 0.2-0.3 loss, which was good to train on our dataset.

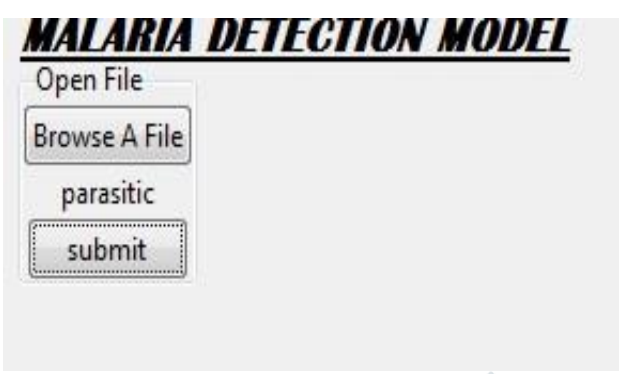
Table 4.3 Models Analysis Table

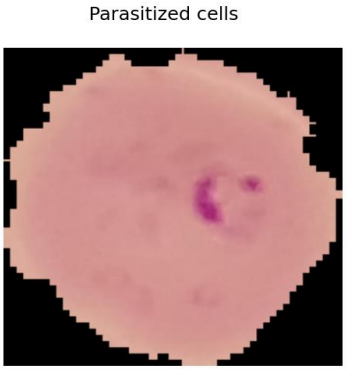
Sr.No	Model	Epochs	Train Accuracy	Validation Accuracy	Train Loss	Validation Loss
1	Created CNN	5	99%	94%	0.013	0.30
		25	98%	91%	0.012	0.41
		50	99%	94.8%	0.015	0.31
		200	99%	95.6%	0.010	0.29
2	VGG19	50	95%	89%	0.14	0.23
		200	965	92%	0.09	0.17

4.4 Model Result

In results we are detecting the image whether the person has a disease or not. So here in result input image and model will tell us that person is Parasitized or Uninfected







CONCLUSION

This RESEARCH will tell us about the one of the techniques of image processing using the Convolution Neural Network (CNN). In this image is taken as an input and it will process using VGG19 model and CNN model. In the model image will process using multiple filters and create a multiple identity from the image. After that max polling layer will make an accurate image of every identity and use it for the future input purpose for identification of same image. After completing the all-convolution network and max pooling layer its output will take as an input in fully connected layer and flatten the images. And before the output activation function softmax layer is there it will just generate the possibilities of the image of same image, and the possibility was not as much efficient then it will again update the filters value and again process everything.

This automated system is analysis with two different model namely VGG19 and CNN model. Created CNN model is tested with various epochs and it gives 99% accuracy while VGG 19 model gives 95% accuracy. This system is useful to classify that how strongly the malaria is infected by doing applying some advance technique in image processing. It will be better if person knows how much they affected by disease and need a good treatment.

REFERENCES

[1] https://www.who.int/news-room/fact-sheets/detail/malaria accessed on dated 9th march 2021 at 2:38 pm)

[2] Oboh, M. A., Oriero, E. C., Ndiaye, T., Badiane, A. S., Ndiaye, D., & Amambua-Ngwa, A. (2021). Comparative analysis of four malaria diagnostic tools and implications for malaria treatment in southwestern Nigeria. International Journal of Infectious Diseases, 108, 377-381.

[3] Wongsrichanalai, C., Barcus, M. J., Muth, S., Sutamihardja, A., & Wernsdorfer, W. H. (2007). A review of malaria diagnostic tools: microscopy and rapid diagnostic test (RDT). Defining and Defeating the Intolerable Burden of Malaria III: Progress and Perspectives: Supplement to Volume 77 (6) of American Journal of Tropical Medicine and Hygiene.

[4] Dev, V., & Manguin, S. (2021). Defeating malaria in the North-East region: the forerunner for malaria elimination in India. *Acta Tropica*, *222*, 106040.

[5] Jahan, F., Khan, N. H., Wahid, S., Ullah, Z., Kausar, A., & Ali, N. (2019). Malaria epidemiology and comparative reliability of diagnostic tools in Bannu; an endemic malaria focus in south of Khyber Pakhtunkhwa, Pakistan. Pathogens and Global health, 113(2), 75-85..

[6] Odaga, J., Sinclair, D., Lokong, J. A., Donegan, S., Hopkins, H., & Garner, P. (2014). Rapid diagnostic tests versus clinical diagnosis for managing people with fever in malaria endemic settings. Cochrane Database of Systematic Reviews, (4).

[7] Zebaze Temgoua Kemleu, S. G., Ngando, L., Nguekeng, E., Fogang, B., Mafo Kapen, M., Fopa, S. I., ... & Ayong, L. (2021). Diagnostic performance of a rapid whole blood-based RT-LAMP method for malaria diagnosis among apparently healthy blood donors and febrile neonates in Cameroon. *Plos one*, *16*(1), e0246205.

[8] Orish, V. N., De-Gaulle, V. F., & Sanyaolu, A. O. (2018). Interpreting rapid diagnostic test (RDT) for Plasmodium falciparum. *BMC Research Notes*, *11*, 1-6.

[9] Kamaliddin, C., Sutherland, C. J., Houze, S., Cottrell, G., Briand, V., Castaneda Mogollon, D., & Pillai, D. R. (2021). The role of ultrasensitive molecular methods for detecting malaria—the broader perspective. Clinical Infectious Diseases, 73(6), e1387-e1390.

[10] Loomans, L., Conesa Botella, A., D'hondt, A., Verschueren, J., Van den Bossche, D., Van Esbroeck, M., & Jacobs, J. (2019). Accuracy of malaria diagnosis by clinical laboratories in Belgium. Malaria journal, 18(1), 1-7.

[11] Fuhad, K. M., Tuba, J. F., Sarker, M., Ali, R., Momen, S., Mohammed, N., & Rahman, T. (2020). Deep learning based automatic malaria parasite detection from blood smear and its smartphone based application. Diagnostics, 10(5), 329.

[12] Wongsrichanalai, C., Barcus, M. J., Muth, S., Sutamihardja, A., & Wernsdorfer, W. H. (2007). A review of malaria diagnostic tools: microscopy and rapid diagnostic test (RDT). Defining and Defeating the Intolerable Burden of Malaria III: Progress and Perspectives: Supplement to Volume 77 (6) of American Journal of Tropical Medicine and Hygiene.

[13]https://www.mayoclinic.org/diseases-conditions/malaria/multimedia/malaria-transmission-cycle

[14] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90.

[15] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).

[16]. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[17] Dangarwala, K. J., & Hiran, D. (2020). Alphabet Classification of Indian Sign Language with Deep Learning. In *Innovative Data Communication Technologies and Application: ICIDCA 2019* (pp. 569-576). Springer International Publishing.

[18] Dangarwala, K. J., & Hiran, D. (2020). Deep Convolution Neural Network Model for Indian Sign Language Classification. In International Conference on Communication, Computing and Electronics Systems: Proceedings of ICCCES 2019 (pp. 35-44). Springer Singapore.