iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VIII **Month of publication:** August 2022

DOI: https://doi.org/10.22214/ijraset.2022.46212

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VIII August 2022- Available at www.ijraset.com

Malaria Parasite Detection using Deep Learning

Yash Akruwala[1], Kathan Prajapati[2]
[1]*Lalbhai Dalpatbhai College of Engineering,* [2]*Government Engineering College, Gandhi Nagar, Gujarat Technological University*

Abstract: ***Malaria is a serious disease. Nowadays, People need to go to laboratories or Laboratory technician comes to people`s house to do test for malaria. This process is time-consuming. This disease must be treated fast otherwise people will suffer through some serious complications such as liver failure, jaundice, and many more. For better treatment, reports have to be come out as quickly as possible. This paper makes accessible a review of work that has been done in space of reducing time for taking the test for malaria with help of Deep learning and Image processing.***
Keywords: Malaria, Convolutional Neural Network (CNN), OpenCV, Tensorflow, Keras, Deep learning

I. INTRODUCTION

Malaria is a severe disease transmitted by mosquitos. If not treated in time, then it can cause some serious problems. World Health Organization (WHO) reported in the World Malaria report 2021 that alone in 2020, there was 241 million malaria in the whole world. Also, 627000 people died because of malaria. This represents about 14 million more cases in 2020 compared to 2019, and 69000 more deaths. Approximately two-thirds of these additional deaths were linked to disruption in the provision of malaria prevention, diagnosis, and treatment during pandemic COVID 19. The latest data also shows the worst-case scenario projected by WHO – a doubling of malaria deaths in sub-Saharan Africa – did not come to pass.

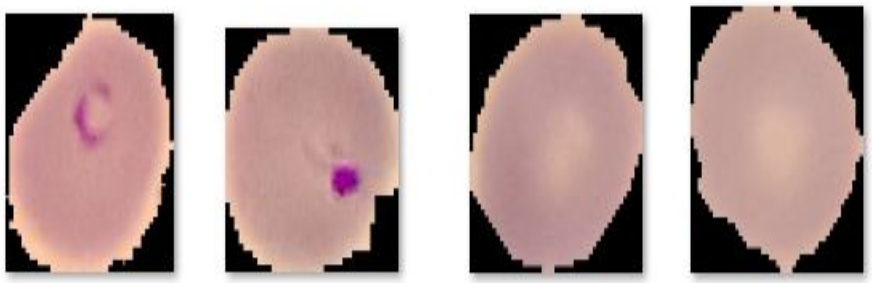
According to the National Center for Biotechnology Information (NCBI) which is part of the State's National Library of Medicine, they suggest PCR technique is the efficient method for malaria diagnosis. PCR has shown higher sensitivity and specificity than conventional microscopic examination of stained peripheral blood smears and now seems the best method for malaria diagnosis.

And advancement in Artificial intelligence allows analyzes samples faster and more d accurately than the n human eye would do. There are many ML methods and Deep learning (DL) is used in medical image classification. DL techniques are now integrated to the identification of COVID 19 cell in lungs. Many of the DL techniques include segmentation tasks and feature extraction or a combination of both.

DL techniques are also used for the classification of malaria parasites in blood cells from microscopy images. Two types of malaria images are widely used for the classification of malaria and that are thin blood smear images and thick blood smear images. Thick blood smear images are more efficient to the find presence of malaria due to thick layer of red blood cells. Thin blood smear images are more useful ad clinical to find stages of malaria. With help of a Convolution neural network, we can extract features ad classify from parasitized and uninfected images.

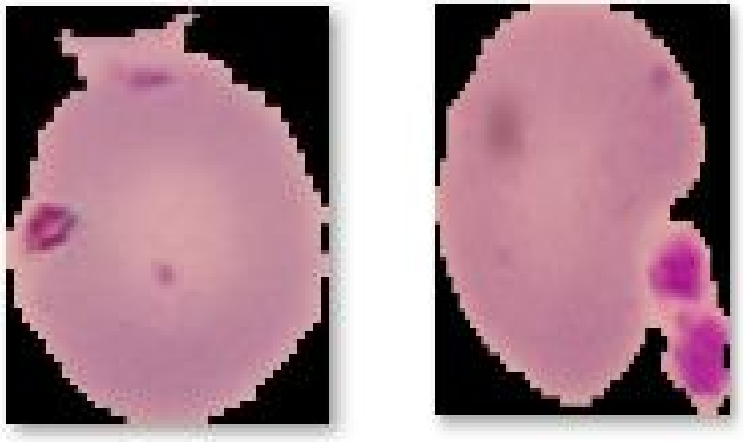
II. DATA-SET

The dataset is uploaded in Kaggle which is more similar data that is collected by Chittagong Medical College Hospital, Bangladesh of 150 malaria infected and 50 healthy patients thin blood smears. In this dataset there are main three types of images are present and those are parasitised cells, uninfected cells and uninfected cells containing impurities. There are 27,558 cell images.



(a) Parasitized cell (b) Uninfected cell

©IJRASET: All Rights are Reserved



(c) Uninfected cell, impurities and artefects

III. METHODS

A. Pre Processing

Distributing parasitized and uninfected images into different folder. There are 22046 blood smear images are for train dataset and 5512 blood smear images for test dataset. Using OpenCv find infected and uninfected area with help of findContours and contourArea() method and after that convert them into csv file for both train dataset and test dataset.

In model, these CSV are read by using the read_csv method of pandas. Then images are reshaped in (50,50,1) converting color of images into gray and type is float32. With using that reshaped images, we are transforming them into pixels.

By using LabelEncoder from scikitlearn library, assigning labels to the pixels. Assigning 0 to infected and 1 for uninfected. Label encoded data is a numpy array which is transformed into binary matrix the where class axis is placed last, to_catgorical () method of keras library is used for generating binary matrix.

B. Model

It is a convolutional neural network with three convolutional layers. On the first layer, there are 16 filter shapes (50,50,1) and used "relu" as a layer activation function. Set padding as a 'same' to preserve spatial dimensions of volume such that output volume size matches the input volume size. After that we applied maxpooling in model. After all, three convolutional layers, applied maxpooling for extraction. In all three convolutional layers used 'relu' as a activation function and 'same' as a padding.

After that layers there are two dense layer and before each dense layer used dropout layer from preventing data from overfitting. I first dense layer used 'relu' as an activation function and in second dense layer used 'softmax' as an activation function.

ReLU is defied by:

$$f(x) = \max(0, x)$$

here ReLU function returns value if function receives positive value x. if function receives negative value than it returns 0. Output range is zero to infinite.

Softmax is defined by:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

Adaptive learning rate optimization algorithm used was Adam. Adam is a combination of RMS prop and stochastic Gradient Descent with momentum. Also, categorical cross-entropy is used as a loss function. This loss function is used in multi-class classification related tasks.

There are tasks in which an is belongs to one of many possible categories and the model has to decide which Our model image has two categories such as parasitized and uninfected and our model has to decide one. So categorical crossentropy is the most suitable loss function for calculating loss for our model.

Images are divided into batch size of 50 and 441 for epochs. And used 35 epochs to the model.

©IJRASET: All Rights are Reserved

TABLE I
MODEL STRUCTURE

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 50, 50, 16)        160
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 25, 25, 16)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 25, 25, 32)        4640
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 12, 12, 32)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 12, 12, 64)        18496
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 6, 6, 64)          0
_________________________________________________________________
dropout (Dropout)            (None, 6, 6, 64)          0
_________________________________________________________________
flatten (Flatten)            (None, 2304)              0
_________________________________________________________________
dense (Dense)                (None, 200)               461000
_________________________________________________________________
dropout_1 (Dropout)          (None, 200)               0
_________________________________________________________________
dense_1 (Dense)              (None, 2)                 402
=================================================================
Total params: 484,698
Trainable params: 484,698
Non-trainable params: 0
```

IV. RESULTS

The final result of the model is based on basis of accuracy matric. The final accuracy of the train model after all epochs. Accuracy matric is calculated according to

Accuracy = (TP + TN)/(TP+TN+FP+FN)

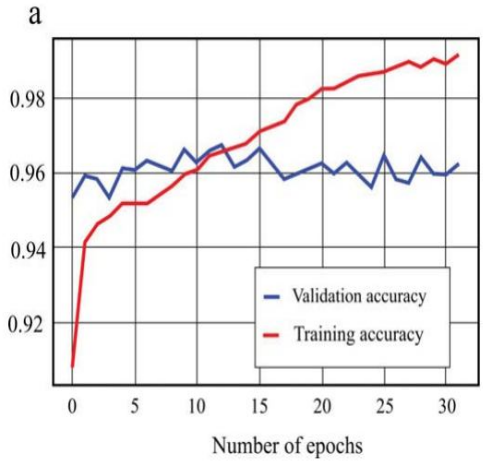
Here,

TP = number of cells that are parasitized classified correctly.

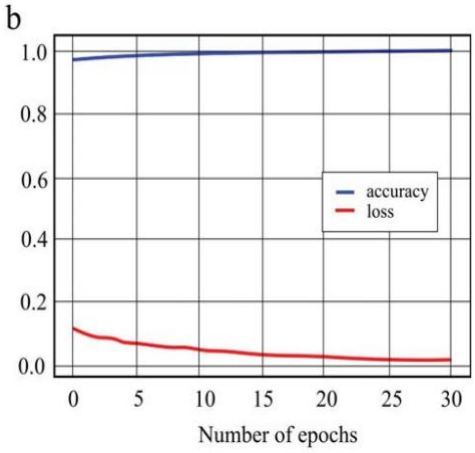
TN = number of cells that are uninfected classified correctly.

FP = number of uninfected cells that are wrongly classified as parasitized.

FN = number parasitized cells that are wrongly classified as uninfected.



(a) Accuracy vs epochs



(b) loss and accuracy vs epochs

Test model accuracy after epochs is 0.9501 and loss is 0.2554.

Below showed example in which randomly used image test dataset is uploaded in model and it gave result accurately.

V. DISCUSSIONS

Our model used smaller convolutional neural network gave accuracy of 0.9501 in test model. Also, smaller image is derived from larger images using interpolation. Very deep neural networks are used to extract small features from the images. But in our database images can be classified by singularity point like infected or parasitized. So we can say that shallow convolutional neural networks are more suitable for this type of classification.

We used only 3 convolutional layers but we can use more. K.Simonyan and A.Zisserman used 13 convolutional layer in their research of very deep convolutional networks for large-scale image recognition. Using VGG8 or VGG16 or using more convolutional layers that gives you the better accuracy.

VI. CONCLUSION

Our work shows that deep learning can be more efficient for finding malaria parasites on blood smears. We have presented a DL model that gives high accuracy. More work need to done on model so that it can able to identify parasites, impurities and artefacts accurately. However, our model allows identification of more cases accurately. Our model is faster and more accurate to detect parasite in blood cells than manual testing.

REFERENCES

[1] https://www.researchgate.net/publication/349644405_Malaria_Parasite_Detection_Using_Deep_Learning_Methods
[2] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7277980/
[3] https://opensource.com/article/19/4/detecting-malaria-deep-learning
[4] https://www.hindawi.com/journals/wcmc/2020/8895429/
[5] https://www.researchgate.net/publication/322416654_Deep_Convolutional_Neural_Networks_Structure_Feature_Extraction_and_Training
[6] https://arxiv.org/abs/1412.6980v9

Undertaking for Originality of the Work

I Kathan Prajapati, give undertaking that the Project/Research entitled

"Malaria Parasite Detection using Deep Learning"

submitted by me, towards the fulfilment of the requirements for the degree of Bachelor of Engineering in Information Technology of Gujarat Technological University, Ahmedabad, is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. I understand that in the event of any similarity found subsequently with any other published work or any project report elsewhere; it will result in severe disciplinary action.

Signature of Author:



Date: 06/08/2022

Place: Ahmedabad

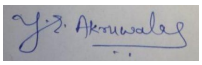
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VIII August 2022- Available at www.ijraset.com

Undertaking for Originality of the Work

I Yash Akruwala, give undertaking that the Project/Research entitled

"Malaria Parasite Detection using Deep Learning"

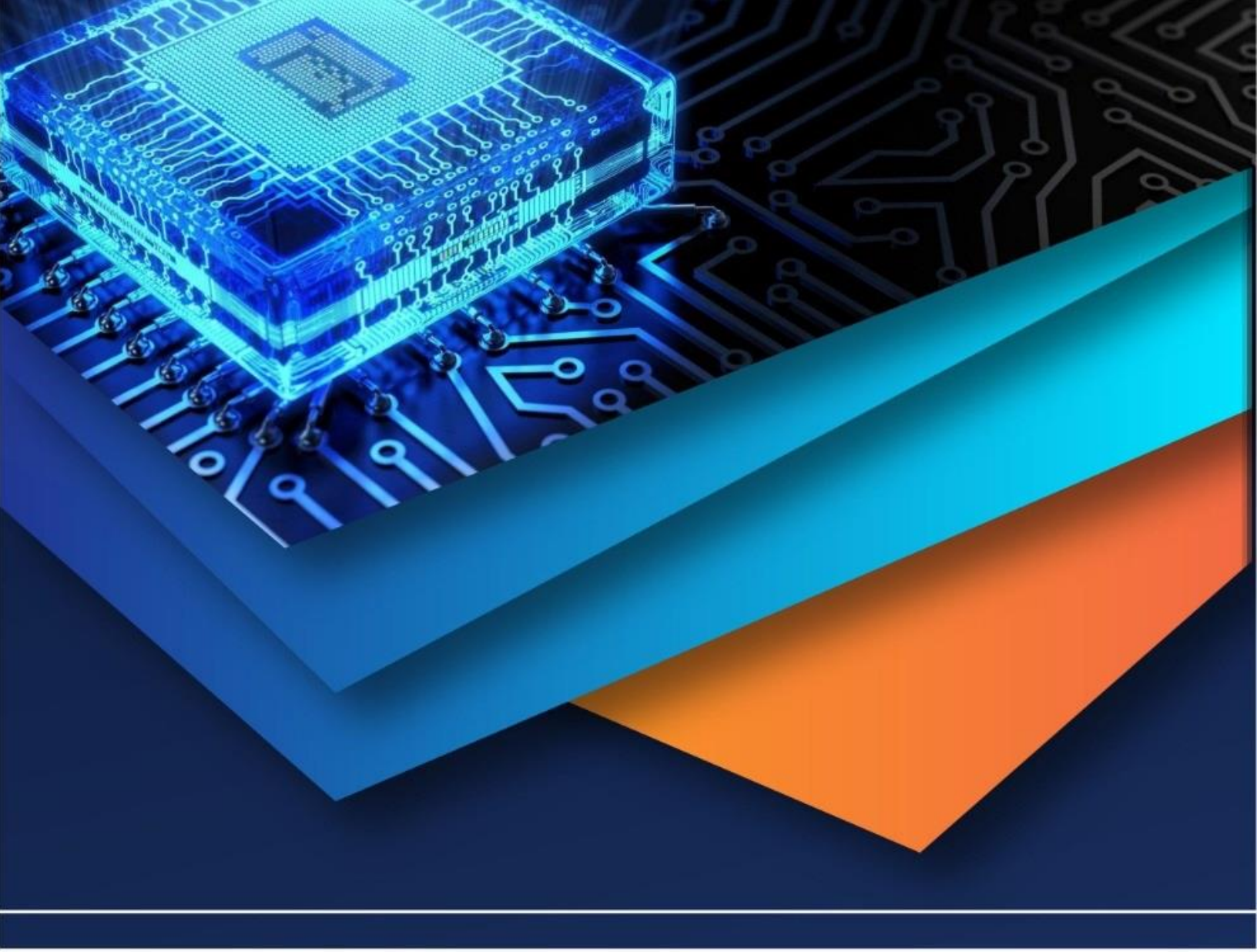
submitted by me, towards the fulfilment of the requirements for the degree of Bachelor of Engineering in Information Technology of Gujarat Technological University, Ahmedabad, is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. I understand that in the event of any similarity found subsequently with any other published work or any project report elsewhere; it will result in severe disciplinary action.

Signature of Author: 

Date: 06/08/2022

Place: Ahmedabad

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)