INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 7, Issue 3 - V7I3-2221)
Available online at: https://www.ijariit.com

Malaria Parasite Detection System using Deep Learning and Image Processing

Gaddayi Pravallika
pravallikagaddayi1999@gmail.com
Andhra University College of Engineering for Women, Visakhapatnam, Andhra Pradesh

M. Sion Kumari
msmondru@gmail.com
Andhra University College of Engineering for Women, Visakhapatnam, Andhra Pradesh

Dora Aasritha
aasrithadora@gmail.com
Andhra University College of Engineering for Women, Visakhapatnam, Andhra Pradesh

Gandepalli Vandana
gvandana.lucky@gmail.com
Andhra University College of Engineering for Women, Visakhapatnam, Andhra Pradesh

Gandrapu Suswitha
suswitha22@gmail.com
Andhra University College of Engineering for Women, Visakhapatnam, Andhra Pradesh

ABSTRACT

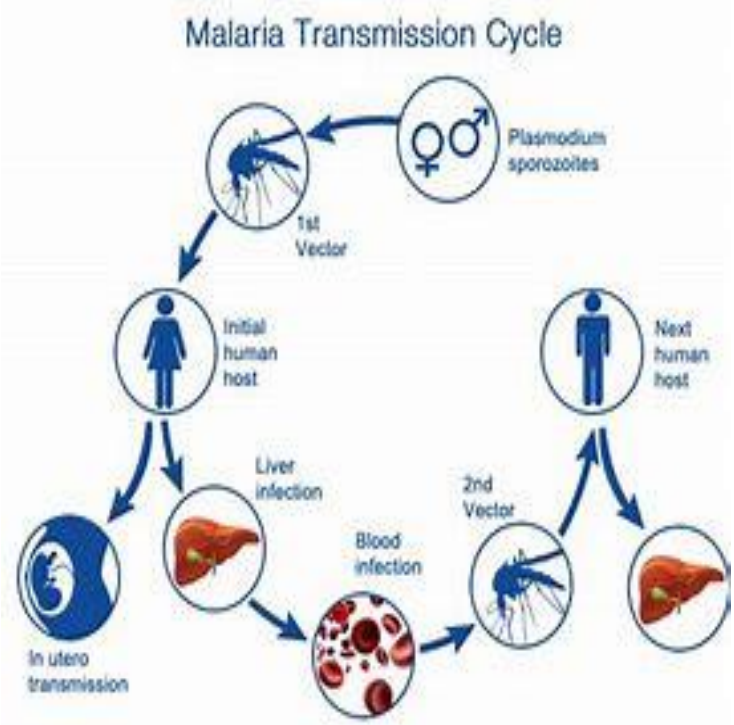
Malaria is a mosquito-borne blood disease caused by Plasmodium parasites which are deadly, infectious, and life-threatening. The conventional and standard way of diagnosing malaria is by visual examination of blood smears via microscope for parasite-infected red blood cells under the microscope by qualified technicians. The given method is inefficient, time-consuming and the diagnosis depends on the experience and the knowledge of the person doing the examination. Image processing based Automatic image recognition technologies has been applied to malaria blood smears for diagnosis before. However, the practical performance has not been up to expectation. With the early prediction results, healthcare professionals can provide better decisions for patient diagnosis and treatments. This motivates us to make malaria detection and diagnosis fast, easy and efficient. To get quick results for the malaria tests, we proposed a model that involves Deep Learning and Image Processing. In this paper, we developed a model using Convolutional Neural Networks (CNNs) classifier that predicts whether the input image is malaria parasitized or not. The CNN model has many convolution blocks that detect even the tiniest possibility of plasmodium parasite present in our input. The proposed model is also evaluated using a large amount of data to increase its accuracy and correctness while detecting the malaria parasite.

Keywords— *Malaria Detection, Plasmodium, Deep learning , Convolutional Neutral Networks(CNN), Tensorflow, OpenCV, Keras, Flask.*

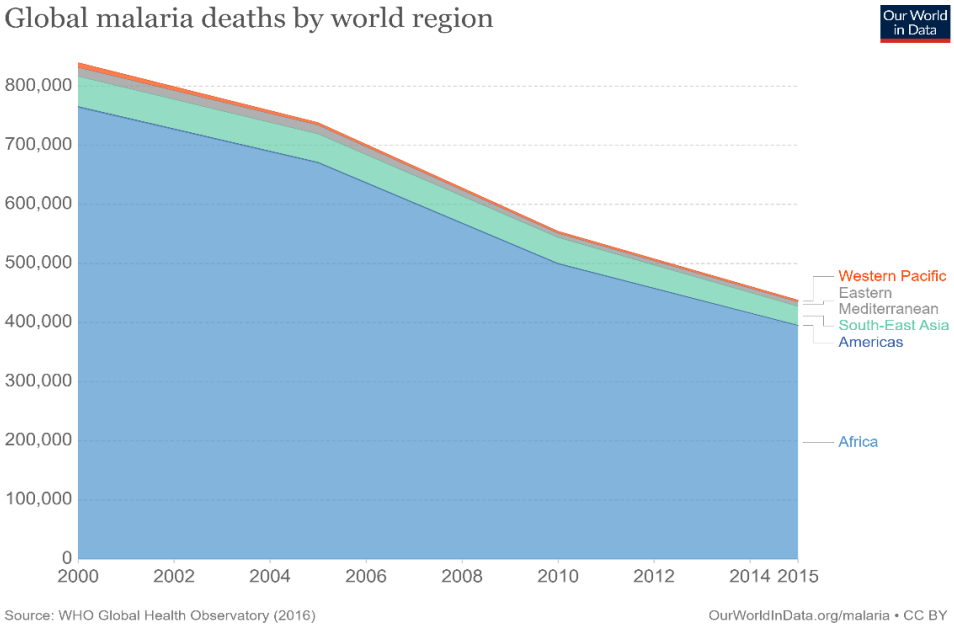
1. INTRODUCTION

Malaria is a deadly infectious disease caused by the Plasmodium parasite which is transmitted by the bites of female Anopheles mosquitoes . This female mosquito is itself a parasite, the females visiting humans for occasional meals of blood. During feeding female infected mosquitoes pass on the malaria parasite from their salivary glands. The mosquito is described as vector. Plasmodium which is of immature form is injected into the blood of humans by the mosquito. This form disappears from the blood stream as it enters various cells of the body, particularly the liver. Here its multiples to produce large numbers of a form which can infect other liver cells. Each parasite in a red blood cells is divided several times further.

So, the RBC‘s is in the form of circular shape. When the RBC‘s is infected by malaria parasite, the shape of RBC‘s will be bursts and the shape of RBC‘s will be changed. Malaria is a mosquito-borne disease caused by Plasmodium parasite which is fact, efficient and life threatening. Throughout, an estimation of 3.2 billion people is at high prospect (greater than 1 in 1000 is the chance of getting malaria in a year). According to the reports till now, there were 212 million new cases of malaria worldwide in 2015 (range 148–304 million). The WHO African Region has been resolved for most global cases of malaria (90%), followed by the South-East.



Microscopy examination is used as one of the prime standards for the detection of malaria to identify existence of parasites in a blood drop from thick smears. However, thin blood smears are used for distinguishing the species of parasite and the development of malaria stages. Examination through a microscopy is commonly used since it is cheap but time-consuming. The examination depends on the quality of blood smear and a skilled person who is expert in the classification and examination of uninfected and parasitized blood cells.



As we can see here, even though there is substantial decrease in the deaths in the previous two decades, the number of deaths is still very high.

Visual detection and recognition of Plasmodium in RBC is possible via chemical process. The staining process somewhat colorizes the RBCs but highlight Plasmodium, WBCs and platelets. The detection of Plasmodium requires detection of the stained objects. However, we need to analyzed stained objects further to determine if they are parasites or not to prevent false diagnosis. Several methods exist for malaria detection.

Malaria parasite (MP) in blood sample can be identified by using image segmentation and feature extraction using minimum distance classifier. Based on Image Acquisition, Image Pre-processing, Image Smoothing, Thresholding and Dilation image segmentation is done. Feature extraction uses two phases in architectural model: 1) Training Phase and 2) Recognition Phase which helps to recognize the MP. In this project, we focus 1)automated detection and quantification of malaria detection, 2) strategy to determine infected image using machine learning 3) discuss to improve the predictive value for detection of infected cells.

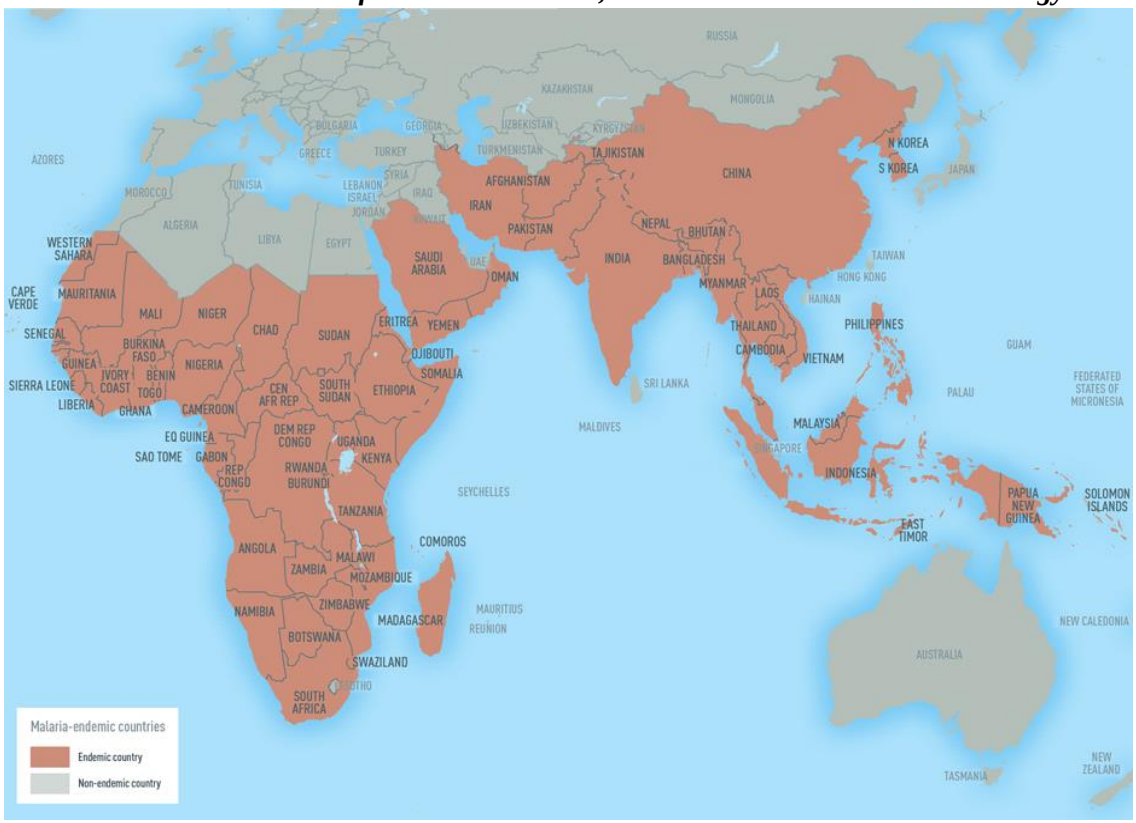


Fig 1: Malaria-endemic countries in the Eastern Hemisphere

Latest data on malaria trends in 87 countries. Each year, quite 400,000 people die of malaria – a preventable and treatable disease. An estimated two thirds of deaths are among children under the age of 5 . The 2020 edition of the planet malaria report takes a glance back at key events and milestones that helped shape the worldwide response to the disease over the last 2 decades – a period of unprecedented success in malaria control that saw 1.5 billion cases and seven .6 million deaths averted. The disease is curable but early detection holds the key. The world Health Organization (WHO) African Region, with an estimated 215 million cases in 2019, accounted for about 94% of cases. Traditional approaches for malaria detection are very time consuming, may produce inaccurate reports due to human errors and are laborious for extensive diagnosis. This motivates us to propose an automatic detection of malaria applying deep learning techniques which is fast, easy and effective. Our deep learning–based model can detect malarial parasites with an accuracy of 96.18%.

A faster, low cost, and reliable alternative to microscopic detection of Malaria is proposed. The study proposes a image processing model for detection of malaria infected cells. We use image processing techniques to detect parasite-infected red blood cells in thin smears on standard microscope slides. The most widely used present day method is examining thin blood smears under a microscope, and visually searching for infected cells. A clinician manually counts the number of parasitic red blood cells - sometimes up to 5,000 cells (according to WHO protocol) .Malaria could be forestalled, controlled, and relieved all the more adequately if an increasingly precise and effective symptomatic techniques were accessible.

2. BACKGROUND AND RELATED WORK

Traditional method of detecting malaria disease is using microscope which is time consuming and is difficult, which needs considerable expertise of laboratory technician. People who are bitten by female anopheles mosquito infected with P. Falciparum are most in danger of dying from malaria. Most of the research has found that a person with little or no immunity to malaria such as young children, pregnant women, or travelers coming from areas with no malaria is most likely to become sick or die. Poor people living in rural areas who lack access to health care are at greater risk for the disease.

Most of the methodologies for detecting malaria disease are based on two criteria: (i) images acquired under well controlled conditions; (ii) the need of proper microscope equipment. Both criteria are different to accomplish in endemic area of malaria, where this type of equipment is scare or non-existent in health care facilities. So, L. Rosado and his team proposed different methodology approach for image processing of malaria-infected thick blood smears by using images exclusively acquired with low cost and accessible tools such as Smartphone. The methodology was divided into three main block; Optical Circle Detection, WBC Detection and Trophozoites Detection. It used two different Smart phones, HTC One S and LG Nexus 5, with image resolution ranging from 1456 × 2592 to 1944 × 2592 pixels. L. Rosado and his team proposed the method that only represents a component of mobile-based framework for malaria parasite detection. They do not identify and count all possible species- stages combinations of MP that potentially infect humans.

There has been a significant amount of research during the last decades for cost-effective solutions to support interoperable healthcare in reducing diseases. For instance, Neto et al. proposed a simulator for simulating events of epidemiology in real time. Kaewkamnerd et al. proposed an image analysis system consisting of five phases for malaria detection and classification. Anggararini et al. developed an application applying image segmentation techniques for separating blood cells' background.

N. A. Khan and his team proposed a computer vision - based approach to identify the MP from light microscopy images. The research deals with the challenges involved in the automatic detection of malaria parasite tissues. It is based on pixel - based approach. They used K-means clustering (unsupervised approach) for the segmentation to identify malaria parasite tissue. The purpose of K-means clustering is that the clusters of things with an equivalent target category are identified. The predictions for new data items are made by assuming that they are of the same type and nearest to the cluster center. They used chemical change that has permeabilization , fixation, mounting and marking which may be a little more harder to do in rural areas due to insufficient materials and expertise.

To summarize, the related work mentioned above largely used different pretrained CNN variants such as AlexNet, VGG-16, ResNet-50, Xception, DenseNet-121. However, the downside is that these results obtained through the feature extraction and subsequent training that required long time in some cases a little over 24hours. A faster, low cost, and reliable alternative to microscopic detection of Malaria is proposed. The study proposes a image processing model for detection of malaria infected cells. We use image processing techniques to detect parasite-infected red blood cells in thin smears on standard microscope slides. In contrast, we build a simpler and computationally efficient CNN model with considerably less trainable parameters, yet producing comparable or better results.

Detection of Malarial Parasite in Blood Images by two classification Methods: Support Vector Machine (SVM) and Artificial Neural Network (ANN) The study [16] presents that there are many systems which describe the computerized methods of image analysis that commonly involves three main phases. In the first phase of pre-processing, luminance of the image is corrected and transformed to a continuing color space. At the second step, a histogram-based image segmentation process is employed which helps in avoiding maximum artifacts and over stained objects. Later, a back propagation neural network was used for classification. This paper set us a attend reference for the steps at granular level yet the examination trusted on singular RBCs as against an entire blood sample. Blood Cells Counting using Python OpenCV A more accurate method of counting blood cells using Python OpenCV is explored in [17]. It uses images of blood obtained by keeping blood samples under a microscope to compute number of cells. Image processing is a method that involves signal processing and mathematical procedure. In this study, the given input images were processed and a blob detection algorithm was used to detect and differentiate RBCs from WBCs. A cell counting method was also used to provide an actual count of the RBCs and WBCs detected. The automation comes with a GUI backed up with a database.

3. PROPOSED MODEL

The architecture of our proposed model is depicted in Fig 3 below.

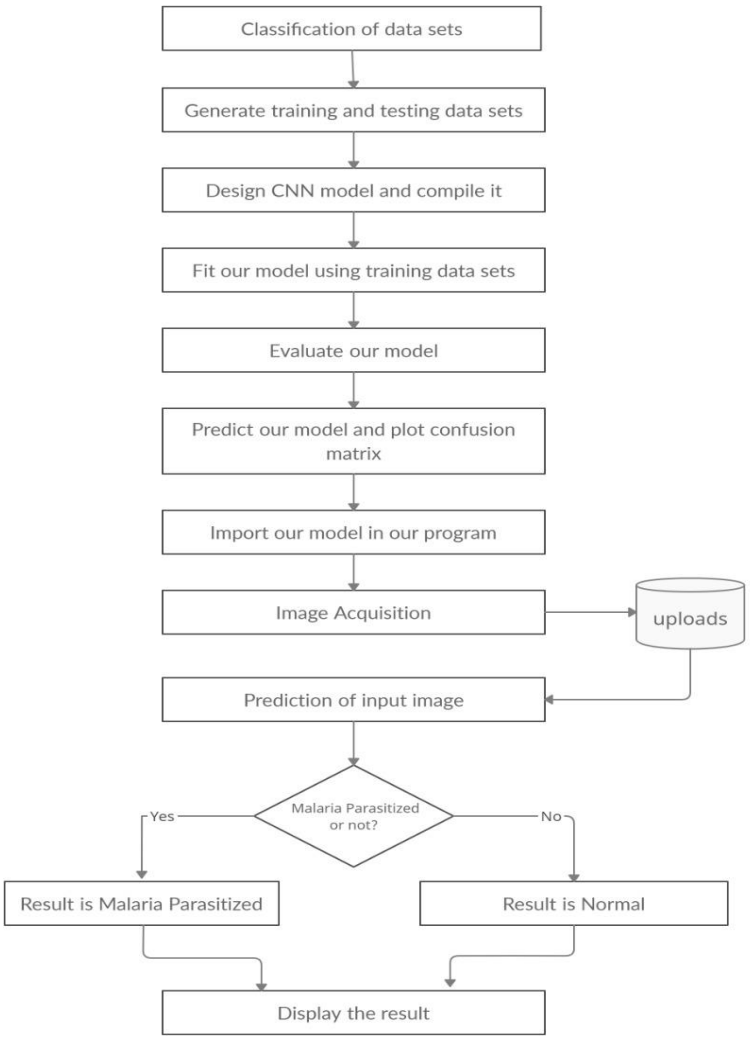


Fig 2: Architecture of proposed model

From the above figure, the entire procedure that is followed in this paper is displayed. This architecture includes the process of acquiring the image input from the camera/webcam using OpenCV library. The input image is temporarily stored in uploads. From uploads the image is given as input to our model. The acquired image is then given to the model to identify whether the cell is malaria parasitized or not. Finally, the output is displayed whether it is parasitized or not.

3.1 Convolutional Neural Networks

In this project, we are going to apply a Deep Learning algorithm, specifically a Convolutional Neural Network (CNN) algorithm that is able to train the algorithm to the given images from a certain cell whether infected or not easily. Since this is a heavy load project we have used Jupyter Notebook.

In Deep learning, to analyse visual imagery we most commonly use a Convolutional Neural Network(CNN) that is a class of deep learning network. Now when we concentrate on a neural network ,we concentrate about matrix multiplication but this not happens with ConvNet. It uses a special technique called Convolution. ConvNet role is to reduce the images into a form that is easier to process, without losing the characteristics that are critical for getting a good prediction.

3.1.1 How does it work?

- An RGB image is a matrix of pixel values having three planes and on the other hand a grayscale image is the same but it has a single plane.
- Let's work with grayscale images as we try to understand how CNNs work.

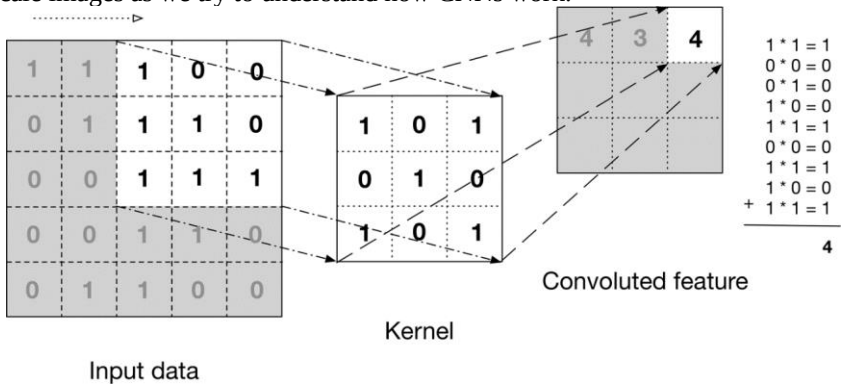


Fig 3: Convolution Process

- The image shows what a convolution is. We take a filter/kernel (3×3 matrix) and apply it to the input image to urge the convolved feature. This convolved feature is passed on to the next layer of our CNN model.
- Convolutional neural networks are combination of multiple layers of artificial neurons. Artificial neurons, a rough imitation of their biological counterparts, are mathematical functions that calculate the weighted sum of multiple inputs and gives an activation value. When you input an image during a ConvNet, each layer generates several activation functions that are passed on to the subsequent layer.
- The first layer usually extracts basic features like horizontal or diagonal edges. This output is passed on to the subsequent layer which detects more complex features like corners or combinational edges. As we move deeper into the network it can extract even more complex features like objects, faces, etc.
- Based on the activation value of the final convolution layer, the classification layer outputs a set of confidence scores (values between 0 and 1) that specify how likely the image is to belong to a "class."

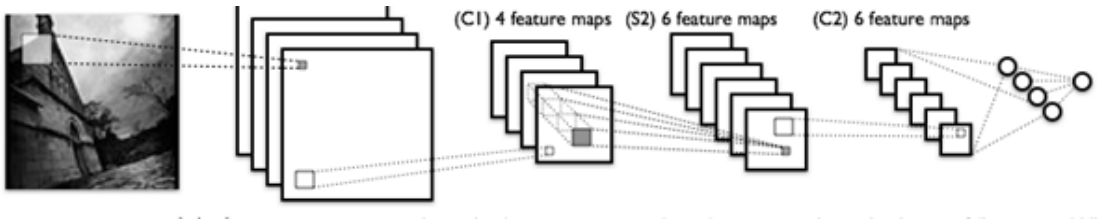


Fig 4: Convolutional Neural Network (CNN) architecture

Convolutional Neural Network Architecture is formed by a stack of distinct layers that transform the input volume into an output volume through a differentiable function. A convolutional neural network is made up of input layer, hidden layers and an output layer. In any neural network, any middle layers are called hidden because their inputs and outputs are masked by the activation function and final convolution.

3.2 System Modules

- **Dataset :** Malaria dataset contains cell images classified into two groups called parasitized and uninfected cells , where each cell contains an equal number of instances. In this data , parasitized images mean that there is the presence of Plasmodium whereas the uninfected images refer to the absence of Plasmodium .
- **Design and implement custom CNN model:** In Jupyter notebook, we design and implement CNN model by importing the libraries like numpy , mlxtend , keras , openCV , tensorflow etc. Using the libraries the dataset is trained and tested before model evaluation.

- **Model evaluation:** Models are evaluated using the datasets i.e., parasitized and uninfected images, that are stored in the files . We start building a sequential model called "my_model". The first convolutional block consists of convolutional layer "Conv2D" and a "MaxPooling2D" layer. The convolutional layer uses 32 filters that are applied to each part of the image, returning 32 arrays of activation values called feature maps that indicate where certain features are located in the image. The max pooling layer decreases the dimensionality of the feature maps. The architecture of the model is shown before compiling the model . Next, to compile the model. We use the ADAM optimizer because it allows for the learning rate to get smaller over time, an advantage when estimating a large number of weights.
- **Prediction of data:** Predictions are made based on the model and predict command using the tested models. The predict command is model.predict and the confusion matrix is plotted.
- **Generation of URL:** In anaconda prompt using python language we will generate a URL to predict the cell image. To generate the URL first activate project environment and use the command used to generate the URL is python filename.py runserver –d.
- **Acquisition of cell image :** Copy paste the URL so that we can open the html page. Using the select image button, the image is selected from the folder. The selected image is stored temporarily in uploads file to evaluate and predict whether it is parasitized or not.
- **Display prediction of data :** Last the prediction about the cell image is displayed whether it is parasitized or not using the predict button below. And the output is displayed.

4. EXPERIMENTAL RESULTS

This section specifies the results obtained by conducting the experiment. We have adopted the subsequent approach so as to assess the performance of the proposed CNN model for the classification of uninfected and parasitized cell images. During the research we had tested more than 40 images. We have received a base accuracy of 96.46% with high precision and recall towards classifying the infected and normal cells which is reasonable. By investigating the confusion matrix as shown in figure above, we can see that the count for False Negatives (FN) is 31 which is pretty low for a disease identification problem. FN indicates that the model declares a malaria patient to be healthy whereas the patient is parasitized. This will severely hamper the patient treatment and should end in death. Our goal is to scale back this number with the proposed improved model. A reduced number of FN will ensure that our model is effective in identifying parasitized cell images.

In anaconda prompt using python program we will generate a URL to predict the cell image. To generate the URL first activate project environment and use the command used to generate the URL is:

```
python filename.py runserver –d.
```

Copy paste the URL so that we can open the html page. The admin opens a cell image, and therefore the deployed model provides the prediction label. Figure below is the html page when the URL is pasted. And below are the snapshots of selecting input images and their results whether they are parasitized or not.

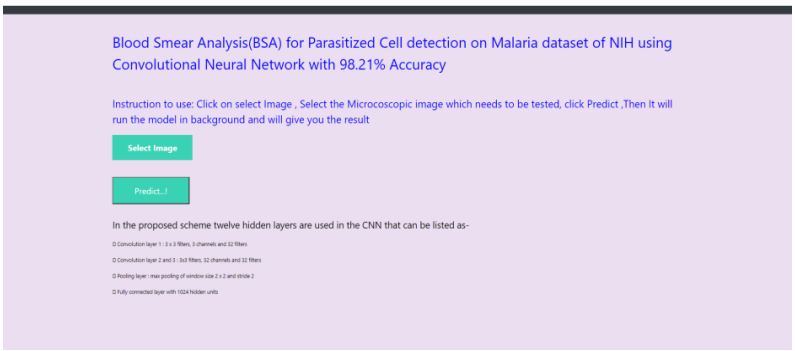


Fig 5: Initial home page

4.1 Test case 1 : Malaria Parasitized given as input

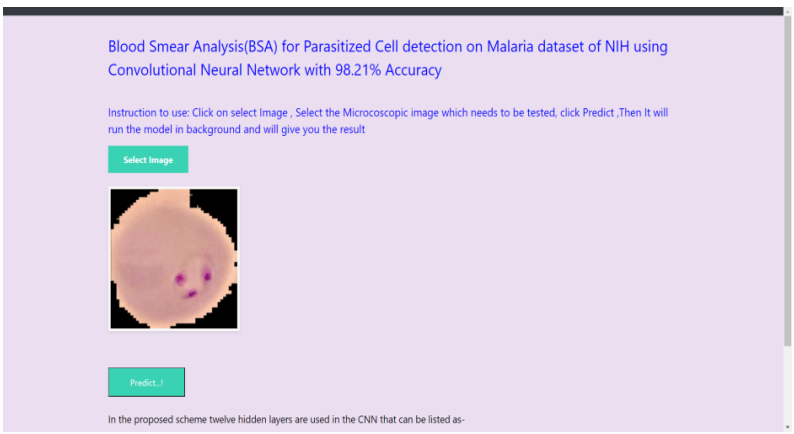


Fig 6 : Parasitized cell given as input

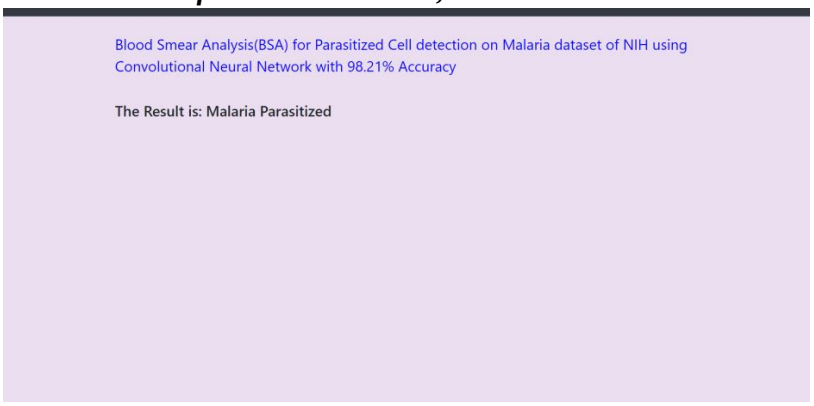


Fig 7: Output is malaria parasitized

4.2 Test case 2 : Uninfected cell given as input

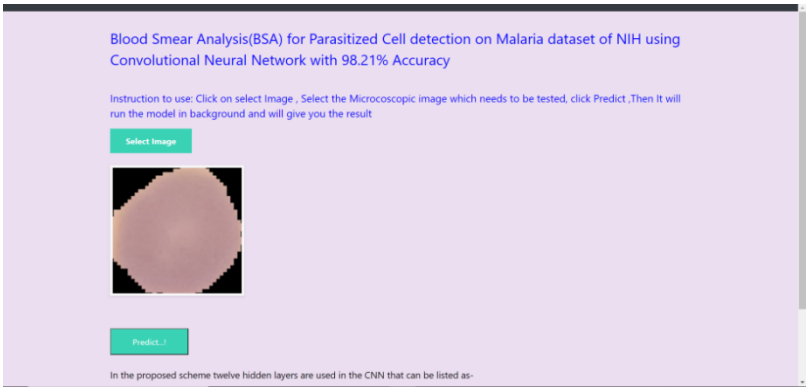


Fig 8: Uninfected cell given as input

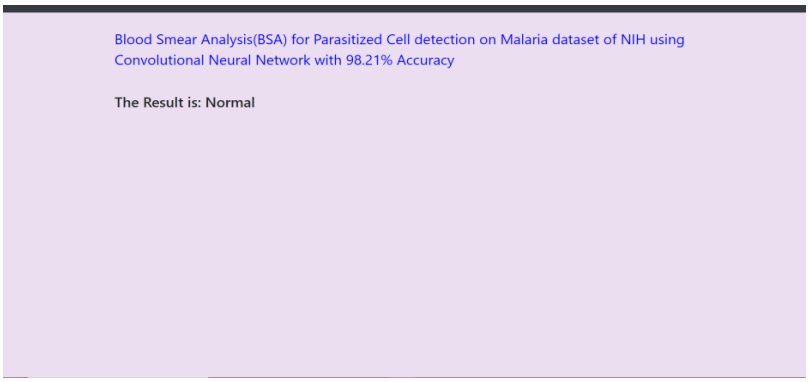


Fig 9: Ouput is normal

5. CONCLUSION

The main objective of this paper is to develop a "Malaria Detection system" that can be applied in real world environment. This can be used mainly in regions where there are more malaria cases and less resources. This helps to recognize the malaria parasite quickly and take actions accordingly. This may lead to less number of deaths all throughout the world. In this project, we looked at an interesting real-world medical imaging case study of malaria detection. Easy-to-build, open source techniques leveraging Deep Learning can give us state-of-the-art accuracy in detecting malaria, thus enabling AI for social good. The CNN model we have used in our project gives us an accurate result at a faster rate. This is really helpful in regions where malaria cases are more and the resources to detect this disease are less.

5.1 Future Work

As an immediate extension of this work, we will consider using image augmentation on the training data with the hope to further alleviate overfitting problem and different adaptive variants of the ADAM optimizer to observe their impact on the performance results. Another model we can see here is Pre-trained model as a feature extractor. For building this model, we'll leverage TensorFlow to load up the VGG-19 model and freeze the convolution blocks so we will use them as an image feature extractor. We will plug in our own dense layers at the end to perform the classification task.

6. REFERENCES

[1]. "Leveraging deep learning techniques for Malaria Parasite Detection using Mobile application". Available at: https://www.hindawi.com/journals/wcmc/2020/8895429/

[2]."Malaria Detection using Machine learning and image processing". Available at : https://www.researchgate.net/publication/322819026_Malaria_Detection_Using_Image_Processing_and_Machine_Learning

[3].“Automation Detection of Malaria Parasite from blood images”. Available at: https://www.kaggle.com/iarunava/cell-images-for-detecting-malaria.
[4].Suman Kumar, Manchana Shrestha, Rojesh Man Shikhrakar proposed Advanced Image Image Analysis Based System for Automatic Detection and Classification of Malaria Parasite in Blood Images.
[5]. Amit Kumar, Prof. Arun Choudhary, Enhanced Identification of Malarial Infected Objects using K-nearest neighbour algorithm and decision tree Algorithm from Thin Smear Digital Images.
[6].Machine Learning and models of machine learning reference links are available at:https://www.w3schools.com/python/python_ml_getting_started.asp
Image processing and techniques, image recognition and open cv reference links are available at: https://likegeeks.com/python-image-processing
[7].K. S. Makhija, S. Maloney, and R. Norton, “The utility of serial blood film testing for the diagnosis of malaria,” Pathology, vol. 47, no. 1, pp. 68–70, 2015.
[8]. WHO, Malaria Micropscopy Quality Assurance Manual, World Health Organization, 2020.
[9].S. Rajaraman, S. K. Antani, M. Poostchi et al., “Pre-trained convolutional neural networks as feature extractors toward improved plasmodium detection in thin blood smear images,” PeerJ, vol. 6, article e4568, 2018.
[10]. World Health Organization Report 2020.
[11]. M. Masud, M. S. Hossain, and A. Alamri, “Data interoperability and multimedia content management in e-health systems,” IEEE Transactions on Information Technology in Biomedicine, vol. 16, no. 6, pp. 1015–1023, 2012.
[12]. O. B. Leal Neto, C. M. Albuquerque, J. O. Albuquerque, and C. S. Barbosa, “The schisto track: a system for gathering and monitoring epidemiological surveys by connecting geographical information systems in real time,” JMIR Mhealth Uhealth, vol. 2, no. 1, article e10, 2014.
[13].S. Kaewkamnerd, C. Uthaipibull, A. Intarapanich, M. Pannarut, S. Chaotheing, and S. Tongsima, “An automatic device for detection and classification of malaria parasite species in thick blood film,” BMC Bioinformatics, vol. 13, Supplement 17, p. S18, 2012.
[14].D. Anggraini, A. S. Nugroho, C. Pratama, I. E. Rozi, A. A. Iskandar, and R. N. Hartono, “Automated status identification of microscopic images obtained from malaria thin blood smears,” in Proceedings of the 2011 International Conference on Electrical Engineering and Informatics, Bandung, Indonesia, July 2011.
[15].C. Mehanian, M. Jaiswal, C. Delahunt, C. Thompson, M. Horning, and L. Hu, “Computer-automated malaria diagnosis and quantization using convolutional neural networks,” in Proceedings of the IEEE International Conference on Computer Vision, pp. 116–125, Venice,Italy,2017.
[16].Afkhami, S., Rashidi Heram-Abadi, H., (2017).Detection of Malarial Parasite in Blood Images by two classification Methods:Support Vector Machine (SVM) and Artificial Neural Network (ANN), Int. J. of Comp. & Info. Tech. (IJOCIT), 5(2): 81-100.
[17]. Meimban, Raymond Joseph & Fernando, Alexa Ray & Monsura, Armil & Ranada, Jonrey & Apduhan, J.C. (2018). Blood Cells Counting using Python OpenCV. doi:50-53. 10.1109/ICSP.2018.8652384.

