

# Malaria Parasite Detection Using Deep Learning

Naveen<sup>1</sup>, Chaganti Sai Sampath<sup>2</sup>, Erusu Tarun<sup>3</sup>, A.V.S.N.Murthy<sup>4</sup>,  
D.Uday Kiran<sup>5</sup>

Assistant Professor<sup>1</sup>, Raghu Institute of Technology, Visakhapatnam, A.P  
Student<sup>2,3,4,5</sup>, Raghu Institute of Technology, Visakhapatnam, A.P  
Department of Computer Science and Engineering.

Submitted: 01-06-2022

Revised: 05-06-2022

Accepted: 08-06-2022

**ABSTRACT:** Jungle fever is quite possibly the most irresistible illnesses on the planet, especially in creating mainland like Africa and Asia. Because of the great number of cases and absence of adequate demonstrative offices and experienced clinical staff, there is a requirement for cutting edge symptomatic methodology to supplement existing techniques. Consequently, this study proposes the utilization of AI models to recognize the jungle fever parasite in blood-smear pictures. Six distinct highlights — VGG16, VGG19, ResNet50, ResNet101, DenseNet121, and DenseNet201 models — were extricated. A Decision Tree, Support Vector Machine, Naïve Bayes, and K-Nearest Neighbor classifiers were prepared utilizing these six highlights. Broad execution examination is introduced as far as accuracy, review, f-1score, exactness, and computational time. The outcomes showed that computerizing the cycle can successfully identify the jungle fever parasite in blood tests with a precision of more than 94% with less intricacy than the past methodologies tracked down in the writing.

**Keywords:** Malaria, convolutional neural networks, deep learning, Plasmodium malaria, classical.

## I. INTRODUCTION

Malaria is a disease caused by mosquitoes that are bitten by female Anopheles mosquitoes. Various types of mosquito parasites infect humans through these bites. These include Plasmodium malaria, Plasmodium malaria, Plasmodium vivax, and Plasmodium falciparum. The 2020 Global Malaria Report [1] shows that these parasites are responsible for an estimated 229 million cases of malaria, killing a total of 409,000 people worldwide. Children under the age of 5 are the

most vulnerable group, accounting for 67% of all malaria deaths worldwide. Africa accounts for up to 94% of the world's cases, and six countries account for about half of the world's malaria deaths: Nigeria (23%), Burkina Faso (4%), Niger (4%), Democratic Republic of the Congo. (11%), United Republic of Tanzania (5%) and Mozambique (4%). Malaria is a disease caused by mosquitoes that are bitten by female Anopheles mosquitoes. Various types of mosquito parasites infect humans through these bites.

These include Plasmodium malaria, Plasmodium malaria, Plasmodium vivax, and Plasmodium falciparum. The 2020 Global Malaria Report [1] shows that these parasites are responsible for an estimated 229 million cases of malaria, killing a total of 409,000 people worldwide. Children under the age of 5 are the most vulnerable group, accounting for 67% of all malaria deaths worldwide. Africa accounts for up to 94% of the world's cases, and six countries account for about half of the world's malaria deaths: Nigeria (23%), Burkina Faso (4%), Niger (4%), Democratic Republic of the Congo. (11%), United Republic of Tanzania (5%) and Mozambique (4%). Other areas more affected by malaria are Southeast Asia and the eastern Mediterranean. Despite all efforts and financial resources to date, B. In 2019, more than \$3 billion was raised for malaria control and eradication, of which \$900 million was raised in malaria endemic countries. However, the disease is still fatal and life-threatening. A blood test assessment helps to conclude an infection and provides reliable results. Dilute blood stains help identify the type of parasite that causes contamination, and dark blood stains help identify the presence of parasites [2].

The accuracy of the survey depends on human experience and can actually be affected by changes in the observer. Blood samples are taken

in a pathology laboratory and the diagnosis of jungle fever is completed by a specialist who uses a magnifying glass to identify parasites on the blood section. A synthetic cycle called Giemsa stain is used to detect jungle heat parasites. In this cycle, blood tests detect and distinguish parasites. Giemsa stain stains red blood cells (RBC) and Plasmodium parasites. Colored objects are required to detect Plasmodium parasites [3].

To avoid misleading consequences, these contaminated items are further decomposed to determine if they are infested or healthy. According to the WHO Convention, there are various methods for detecting jungle fever, including a serious assessment. Manual testing requires the clinician to detect 5000 cells at 100x magnification, a long and exhausting cycle. Polymerase chain reaction (PCR) and rapid analytical test (RDT) were considered to speed up the analysis. These tests are quick, but less accurate. Identifying parasites using the Giemsa blood test requires a ready and experienced professional. Digitizing the cycle reduces the time required for screening. This is useful when working on decision consistency.

Programmed display cycle reduces symptom time. It can also be used very well as a second assessment by a pathologist. AI strategies are used to look up clinical information. It is commonly used as a tool for screening and distinguishing intestinal disorders in individuals. The profound learning model uses non-linear unit load layer grouping to find key attributes from coarse information. Very high quality elements are separated from the low-end highlights, helping to get the hang of dynamic and indirect dynamic cycles, providing end-to-end characterization and highlight extraction [4]. Convolutional neural networks (CNNs) have different layers, and each layer of the CNN grants activation capacity to a particular image. The underlying layer of CNN separates gross elements such as mass, edges, and tones. These elements are separated by a profound interior. A layer of CNN [5] to create high-level features that provide richer images.

## II. LITERATURE REVIEW

Various exploration projects have been carried out around the world to incorporate sound learning models for clinical use. Part of the work based on this work was AlexNet, a CNN-based deep learning model planned by Alex Krizhevsky in 2012, which consistently extended CNN's presentation to include regular image classification [6]. Numerous CNNs such as GoogLeNet [7],

VGGNet [8], and ResNet [9] have introduced significant improvements to achieve the annual difficulties of ILSVRC. The model devised by [10], Xception, used a very different convolution that governs ImageNet [11]'s Inception V3 model [7] in grouping information.

In 2016, Huang [12] proposed a model in which each layer of the model is connected to the entire next layer using network technology, called Densely Connected Convolutional Networks (DenseNet), a variation of CNN. I did. The model has undergone many upgrades and uses much less computation and scope to create the conditions for the craftsman model. CNN has promising results, complemented by accessibility to a lot of information. A mobile learning (TL) approach is performed and the DL model is fitted or used as a component extraction feature for the visual verification task [13]. Razavian et al. (2014) pointed out that CNNs built on large datasets can act as element extraction functions for a wide range of PC vision errands that lead to more advanced execution, as opposed to progressive methods [14].

Experts around the world are beginning to understand the errands that have applied the DL model to many clinical imaging studies and have shown promising results [15]. 2017, Dong et al. , 2017 [16] discontinued the publication of prepared DL models such as LeNet, AlexNet, and GoogLeNet that use SVMs to place between healthy and contaminated cells. Liang et al. [17] In 2017, we proposed a 16-layer CNN model ranked between spoiled and healthy cells.

## III. WORKING

This task aims to establish a framework for detecting jungle heat parasites in blood tests. The framework expects to detect solid and parasitic blood tests. This task consists of image recognition and image characterization, and many opportunities need to be accurately grouped to achieve the goal. This model was specific to Python 3.7.x and the JupyterLab IDE. Techniques for dealing with such issues are typically: The resulting image has undergone several preprocessing methods. Preprocessing has been completed to make the image more meaningful to subsequent systems. Preprocessing strategies include resizing, noise reduction, and image contrast. Image size standardization is essential to catch up with spatial targets for images from multiple sources.

The proposed model is shown in Figure 1

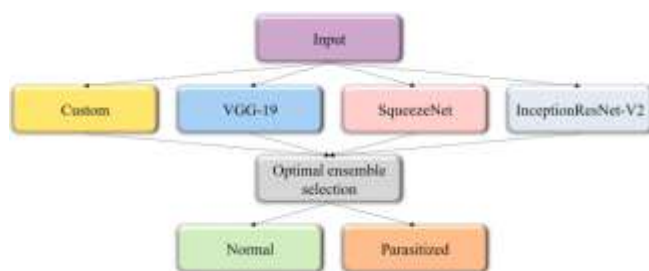


Fig. 1. Block Diagram

The highlights are sorted using the appropriate classifier and then split into different classes. Reasonable boundaries are selected that properly represent the image. Highlights are inferred from images or quantities such as histograms and markers.

In Figure 2, the image is shown as an example of parasitic (0) and uninfected (1). These names are used in the model to predict the outcome

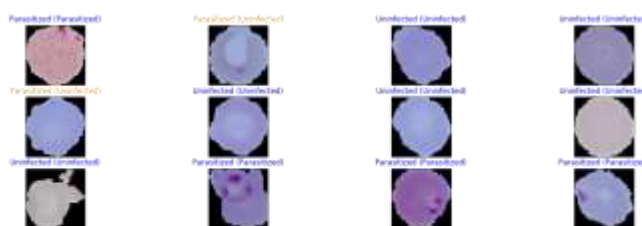


Fig. 2. Parasitized and unparasitized pictures

In addition, the image is split into train and test datasets. The preparation dataset contains 8000 images and the test dataset contains 2000 images. After that, a CNN model is created and the result of the model is expected. These results are from model support for determining contaminated and solid blood tests. In addition, the image is split

into train and test datasets. The preparation dataset contains 8000 images and the test dataset contains 2000 images. After that, a CNN model is created and the result of the model is expected. These results are from model support for determining contaminated and solid blood tests.

```

train_image_gen = ImageDataGenerator(rescale=1./255,
                                     shear_range=0.2,
                                     zoom_range=0.2,
                                     horizontal_flip=True,
                                     fill_mode='nearest')
    
```

Fig.3. Code piece for picture pre-handling

ImageDataGenerator rescales the image so that the pixel values of all images are 0 to 1, regardless of the RGB factor. Shear ranges are used to focus on the desired function of an image while ignoring unwanted environments. Here, the full bandwidth is assumed to be 20%. Similarly, zoom range and horizontal inversion are selected to increase image variation. This will improve the quality of your dataset and give you better results.

The proposed CNN model consists of three layers of convolution. There are 32 filters in the first convolution layer, 64 filters in the second convolution layer, 128 filters in the third convolution layer, the kernel size of each layer is 3x3, and the activation function is ReLU. The sandwich design and the corresponding weight initialization performed by Adam Optimizer improve the model training process.

The model contains three pooling layers. It consists of a 2\*2 pixel pooling window. In addition, the model has three fully connected layers that take the output from the previous layer and use the labels to start classifying the image. The first two fully connected layers have a ReLU as an activation

function and a 50% dropout to avoid over fitting. The output layer has a sigmoid as an activation function. We evaluated the execution of the CNN model with a focus on accuracy. Figure 4 shows the accuracy of the model at each epoch.

epoch	accuracy	loss	val_accuracy	val_loss
0	0	0.712000	0.578206	0.8475 0.592168
1	1	0.873375	0.349373	0.9100 0.229852
2	2	0.913000	0.265330	0.9380 0.037783
3	3	0.929750	0.203314	0.9390 0.024114
4	4	0.939000	0.183909	0.9610 0.053621
5	5	0.948500	0.160007	0.9610 0.191077
6	6	0.949875	0.152722	0.9575 0.017607
7	7	0.951125	0.151227	0.9590 0.102460
8	8	0.950500	0.156431	0.9575 0.010689
9	9	0.951500	0.140251	0.9625 0.182639
10	10	0.954500	0.134922	0.9625 0.076045
11	11	0.955125	0.128416	0.9390 0.136788
12	12	0.955625	0.123483	0.9615 0.042004
13	13	0.957000	0.127172	0.9615 0.165423
14	14	0.953875	0.128704	0.9610 0.028344

Fig. 4. Accuracy at each epoch

A specially trained CNN prototype identifies the presence of Plasmodium parasites and helps detect malaria disease. A collection of 17,460 images was taken to test the functionality of the model and determine its accuracy. In this

collection, 8,760 images were infested and the other 8,700 images were normal. The model was able to make correct predictions for 8,345 parasitic samples and 8,202 healthy samples, as shown in Figures 5 and 6

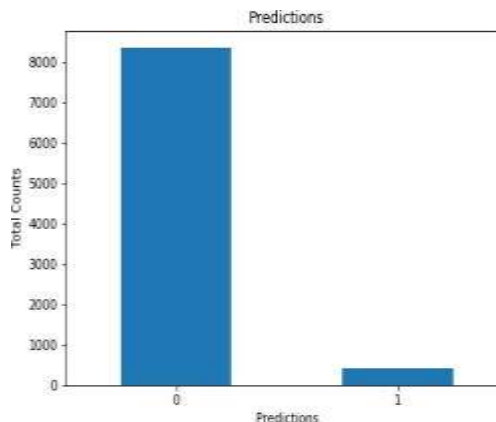


Fig. 5. Parasitized forecasts

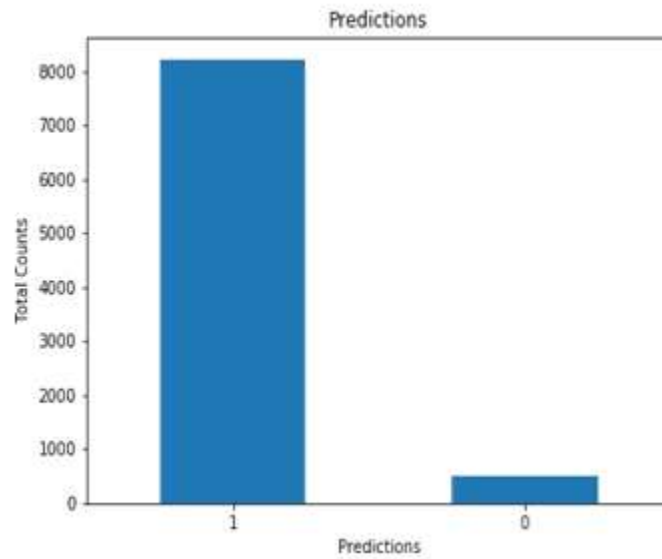


Fig. 6. Uninfected forecasts

Models were tested on randomly occluded images, which were predicted and classified as parasitic ("0") and non-infected ("1").

#### IV. RESULT

Detection of parasites in bowel disease was performed using a custom CNN model. When the

model was built, it showed the development of the correct answer rate, and six years later, it gave a constant correct answer rate of about 95%.

As shown in Table 1, tissue failure rates were 4.74% and 5.72% separately for parasitic and non-infected images.

Table 1. Results

True label	0	1
0	1297	66
1	30	1362

```
In [0]: accuracy = model.evaluate(x_test, y_test, verbose=1)
print('\n', 'Test_Accuracy:-', accuracy[1])

2755/2755 [*****] - 0s 87us/step

Test_Accuracy:- 0.9586206096551724
```

Figures 7 and 8 show indicators of accuracy and loss during the training and validation stages of the model.

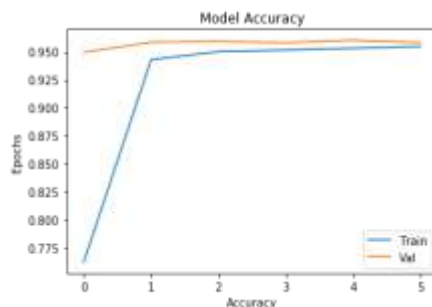


Fig.7 Model Accuracy

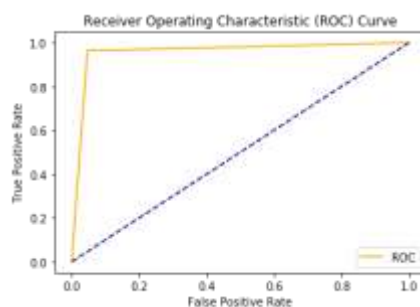


Fig 8.ROC Curve

After that, the overall identification accuracy of the model was about 95%. The results of the proposed model are very large in terms of accuracy. The current model contains a huge number of layers and provides the same accuracy as the reference model. In this proposed model, the convolution layer and the thick layer are well combined to achieve high accuracy in essence.

### V.CONCLUSION

This framework reduces the potential consequences of human error when detecting jungle heat parasites using image manipulation and sound learning strategies. Because pathologists physically detect jungle fever parasites with magnifying glass, there is the potential for human error and false parasite detection, which can lead to further patient management problems. With serious computational power, we acknowledge that the model can perform better than the current results. You can also extend the model to analyze various infections from blood samples.

We used convolution brain tissue and tagged datasets to drive an image clustering model. The proposed model provides about 95% accuracy. The framework is powerful and various variables do not affect the framework.

### REFERENCES

- [1]. World Health Organization. WHO Malaria Policy Advisory Group (MPAG) Meeting: Meeting Report; April 2021; World Health Association: Geneva, Switzerland, 2021.
- [2]. Abubakar, A.; Ugail, H.; Bukar, A.M. Evaluation of human skin consumes: A profound exchange learning approach. *J. Drug. Biol. Eng.* 2020, 40, 321-333. [Ref]
- [3]. Lowd, D.; Davis, J. Further developing Markov network structure getting the hang of utilizing choice trees. *J. Mach. Learn. Res.* 2014, 15, 501-532.
- [4]. Anuradha, J.; Ramachandran, V.; Arulalan, K.; Tripathy, B. Analysis of ADHD Using SVM Algorithm. In Proceedings of the Third Annual ACM Bangalore Conference, Bangalore, India, 22-23 January 2010; Association for Computing Machinery: New York, NY, USA, 2010; pp. 1-4.
- [5]. Zulfikar, W.B.; Irfan, M.; Alam, C.N.; Indra, M. The Comparison of Text Mining with Naive Bayes Classifier, Nearest Neighbor, and Decision Tree to Detect Indonesian Swear Words on Twitter. In 2017 fifth International Conference on Cyber and IT Service Management (CITSM), Denpasar, Indonesia, 8-10

- August 2017; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2017; pp. 1-5.
- [6]. Poostchi, M.; Silamut, K.; Maude, R.J.; Jaeger, S.; Thoma, G. Picture investigation and AI for identifying jungle fever. *Transl. Res.* 2018, 194, 36-55. [Ref]
- [7]. Rajaraman, S.; Antani, S.K.; Poostchi, M.; Silamut, K.; Hossain, A.; Maude, R.J.; Jaeger, S.; Thoma, G.R. Pre-prepared convolutional brain networks as component extractors toward further developed jungle fever parasite recognition in slim blood smear pictures. *PeerJ* 2018, 6, e4568. [Ref] [PubMed]
- [8]. Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, S.; Hasan, M.; Van Essen, B.C.; Awwal, A.A.S.; Asari, V.K. A best in class study on profound learning hypothesis and designs. *Hardware* 2019, 8, 292. [Ref].