



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 7, July 2022

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.118

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)



# Malaria Detection Using Hybrid Approach

Miss. Kolekar Priyanka Zumbar, Dr. Chaudhari S.B.

Department of Computer, JSPM's Jaywantrao Sawant College of Engineering, Hadapsar, Pune, India

**ABSTRACT:** Digestive ailment revelation is a disturbing situation for most trained professionals and it requires experiences and expertise. The machine learning (ML) procedure can be used to relieve this issue. This paper endeavor to find sensible model to help with recognizing wilderness fever with accuracy. The used dataset was conveyed by National Institute of Health in USA moreover contained a flat out number of 27,560 red platelet (RBC) pictures with indistinguishable instances of parasitized and uninfected RBCs pictures. A singular mystery layer feed-forward brain associations theory to be explicit Extreme Learning Machine (ELM) model was applied to bunch and predict regardless of whether a patient has been influenced by wilderness fever. In this paper we are proposing the convolutional brain organization and backing vector machine based jungle fever discovery. In this paper we get 89.75% accuracy on 50epochs.

**KEYWORDS:** Malaria Detection, Convolutional Neural Network, Neural Network, Deep Learning, SVM.

## I. INTRODUCTION

Mechanized analysis of jungle fever has gotten consideration in ongoing patterns. In the clinical field particularly through tiny pictures the conclusion of jungle fever is financially savvy and hearty which measure and group the tainted parasite all the more precisely. Jungle fever is a dangerous illness which prompts expansion in death rate. It is an irresistible sickness brought about by Plasmodium parasites which are communicated to human through tainted female mosquitoes known as the vectors (transporters). The plasmodium species which cause intestinal sickness incorporate falciparum, vivax, malariae, ovale. The determination of jungle fever requires costly procedures which are inaccessible in the country regions where the illness is predominant. Minute assessment of Giemsa stained blood spreads is ordinarily utilized for deciding the formative phases of jungle fever parasites and its development in human. The human eyebrain co-appointment framework is the broadly followed procedure for identification of tainted cells and parasites in tiny pictures and the last determination is finished. But since of more number of tests to be broke down and time intricacy, there can be blunder in identification and the precision will be brought down. Along these lines, computerized framework for recognition of parasites will be of extraordinary need.

## II. LITERATURE SURVEY

G. Karthik S. Muttan M. P. Saravanan R. Seetharaman et al. [1] stated that, this paper quantifies and classifies the red blood cells infected by Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale and Plasmodium malariae in thin blood smears. The experimentation was conducted in MATLAB in which image acquisition, pre-processing including segmentation, classification and a combination of other image processing techniques were performed.

T. A. Aris, S. A. Nasir [2] proposed that proposed fast kmeans clustering is used because of its advantage which is no need to retrain cluster center that causes time taken to train the image cluster centers is reduce. Meanwhile, different color models have been utilized in order to identify the most relevant color model that obviously highlight the parasites. Hussin K. Ragb, Ian T. Dover [3] proposed that The NIH Malaria Dataset from the National Institute of Health, a dataset of 27,558 images formed from microscopic patches of red blood smears, is used in these experiments. The dataset is segmented into 80% training set, 10% validation set, and 10% test set. The validation set is used as the decision metric for choosing ensemble network architectures and the test set is used as the evaluation metric for each model.

Mohammed Osman, Hassan Salih, Osama Salih and Nazar [4] stated that in this paper, The readings are divided into two parts: negative readings were thirty four and positive readings are 6 Readings. The scattering angle depends on the absorption of the malaria parasite. The detection accuracy of our system is determined to be 95%, these are Male results but unfortunately all reading for females is zero.

### III. PROPOSED METHOD AND ALGORITHM

#### A. Proposed Methodology

In a proposed system, we are proposing experiment on malaria disease with limited set of supervised data as shown in figure1. We are proposing a Convolutional neural network for feature extraction and classification by using support vector machine based model for disease risk prediction model for disease with higher accuracy. We are going to solve accuracy issue in diagnosis of disease with accurate stage predictions.

##### A. Dataset

We have collected the dataset from kaggle. In that total 460 images are present. Out of them we are split the dataset into two categories training and testing such as 400 and 60 respectively.

##### B. Pre-processing

In pre-processing we convert the every image into 224\*224.

##### C. Data Augmentation

In data augmentation we are simply increase the dataset of training directory. We generate every image different format such as rotation, zoom and change the brightness of image.

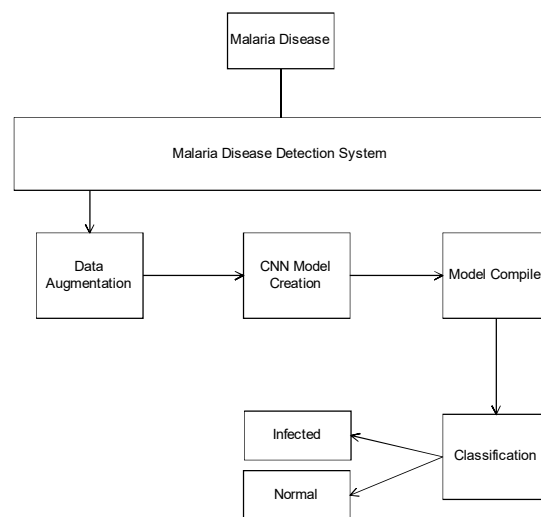


Fig1. Proposed Architecture

#### D. Algorithms

##### 1. CNN

The principle utilization of the Convolution activity if there should be an occurrence of a CNN is to recognize fitting highlights from the picture which goes about as a contribution to the primary layer. Convolution keeps up the spatial interrelation of the pixels This is finished by fulfillment of picture highlights utilizing miniscule squares of the picture. Convolution equation. E very picture is seen as a network of pixels, each having its own worth. Pixel is the littlest unit in this picture grid. Allow us to take a 5 by 5(5\*5) framework whose qualities are just in twofold (for example 0 or 1), for better agreement. It is to be noticed that pictures are by and large RGB with upsides of the pixels going from 0 - 255 i.e 256 pixels.

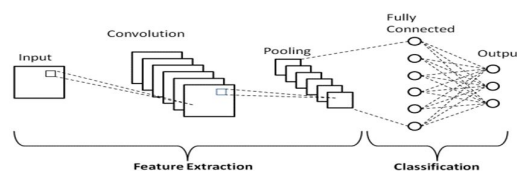


Fig2. CNN Architecture



(ii) Activation Functions:

1. Rectified Linear Unit (Relu)

ReLU follows up on a rudimentary level. All in all, it is an activity which is applied per pixel and overrides every one of the non-positive upsides of every pixel in the component map by nothing.

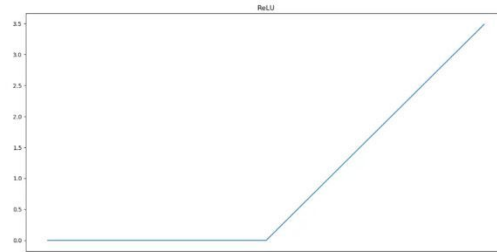


Figure3: Relu Activation

It is represented as:

$$f(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{otherwise} \end{cases} \dots\dots\dots (1)$$

(iii). Pooling or sub-sampling

Spatial Pooling which is likewise called sub-sampling or down sampling helps in lessening the elements of each element map yet even at the same time, holds the most important data of the guide. Subsequent to pooling is done, in the long run our 3D element map is changed over to one dimensional component vector.

2. Support Vector Machine

Generally the use of machine learning algorithm is used for the recommendation. In this paper, the support vector machine will suggest the diet. SVM is a supervised machine learning algorithm which works based on the concept of decision planes that defines decision boundaries. A decision boundary separates the objects of one class from the object of another class. Support vectors are the data points which are nearest to the hyper-plane. Kernel function is used to separate non-linear data by transforming input to a higher dimensional space. Gaussian radial basis function kernel is used in our proposed method. Given a set of labelled training examples, each belonging to one category. Support Vector Machine builds a model that assigns new examples to one of the category.

SVM cost function is given as:

$$\text{Min } C \sum [y^{(1)} \text{cost}_1(\theta^T x^{(i)}) + (1 - y^{(i)}) \text{cost}_0(\theta^T x^{(i)})]$$

$$h(x) = \begin{cases} 1, & \text{if } \theta^T x \geq 0 \\ 0, & \text{otherwise} \end{cases} \dots\dots\dots (2)$$

Where

x: Training example (feature vector)

m: Number of training examples

n: number of features

θ: parameter vector

C: constant

y: class label.

**IV. RESULTS AND DISCUSSION**

Learning model is partially implemented. It is trained for 2 classes. For training and testing purpose total 225 images are used per class. Out of which 550 images are used for training dataset and 80 images are used for testing dataset. So 90% and 10% distribution is used for training and testing dataset respectively. Images are resized to 224\*224 matrix and used as input to CNN.

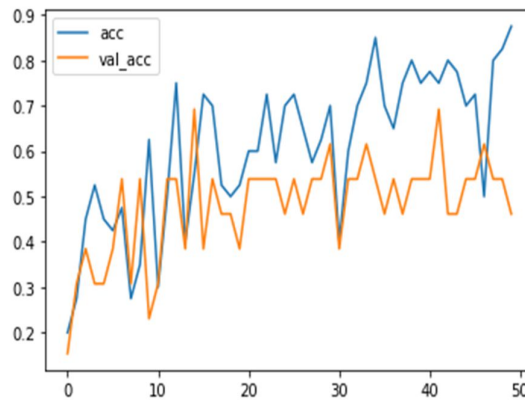


Fig.4:Accuracy graph

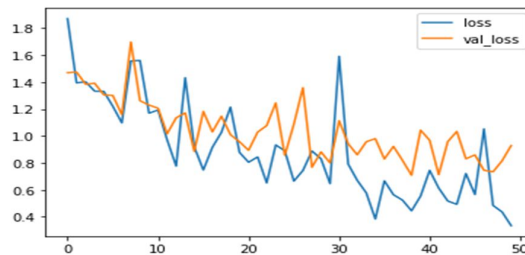


Fig.5:Loss graph

Table 1

Images	Train	Test	Accuracy
460	400	60	89.75%

Table1: Comparative Analysis

Once the CNN is trained with features of the training dataset, it performs the process of feature extraction in appropriate manner. The output of this trained CNN model is passed to multiclass linear SVM, which performs the classification process. Batch size for training and testing are kept as 400 and 86 respectively. Training and testing accuracy of model for 50 epochs is shown in figure 4 and 5. It is observed that accuracy increases with number of epochs. In this paper we have compare the results with one activation functions like Relu. But Relu gives the better accuracy as shown in Table1. The speed of building models based off on ReLU is very fast opposed to using Sigmoid activation function.

## V. CONCLUSION

We implemented malaria disease detection system over machine learning and CNN techniques which solves existing accuracy problem as well as reduce death rates by disease like malaria. In this paper we used hybrid approach CNN with Svm. We get 89.75% accuracy on 50 epochs. For future work, we can implement this technique on some more diseases with rich dataset. Increasing the number of diseases and dataset used for the process can improve the accuracy.

## REFERENCES

- 1] G. Karthik<sup>1</sup> S. Muttan<sup>1</sup> M. P. Saravanan<sup>2</sup> R. Seetharaman “AUTOMATED MALARIA DIAGNOSIS USING MICROSCOPIC IMAGES” International Conference on Inventive Systems and Control (ICISC 2019).
- 2] T. A. Aris, A., S. A. Nasi “Fast k-Means Clustering Algorithm for Malaria Detection in Thick Blood Smear” 2020 IEEE 10th International Conference on System Engineering and Technology (ICSET), 9 November 2020, Shah Alam, Malaysia.
- 3] Mohammed Osman, Hassan Salih, Osama Salih and Nazar “Design and Implementation of Non-Invasive Malaria



Detection System” 2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE).

4] Hussin K. Ragb, Ian T. Dover “Deep Convolutional Neural Network Ensemble for Improved Malaria Parasite Detection” 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR).

5] Octave Iradukunda , Haiying Che “Malaria Disease Prediction Based on Machine Learning” In 2016 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC). 2019 IEEE.



**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.118**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details