

DETECTING MALARIA USING DEEP LEARNING MODELS

Priya Dalal¹, Tripti Sharma², Neetu Anand³, Umesh Kumar⁴

¹Assistant Professor, Maharaja Surajmal Institute of Technology, Janakpuri, Delhi, India
Email: priya@msit.in

²Associate Professor, Maharaja Surajmal Institute of Technology, Janakpuri, Delhi, India
Email: tripti_sharma@msit.in

³Assistant Professor, Maharaja Surajmal Institute, Janakpuri, Delhi, India
Email: neetuanand@msijanakpuri.com

⁴Assistant Professor, J. C. Bose University of Science & Technology, YMCA, Faridabad, Haryana, India
Email: umesh554@gmail.com

Abstract– Over the past few years, the corona has created havoc across the globe but still, malaria is holding the position of disease with the highest mortality rate in few parts of the world. Malaria is caused by the bite of female mosquitoes - Anopheles. According to WHO, in 2020 the total number of malaria cases reported was more than 240 million and about 627,000 deaths were reported [5]. It can be examined on time, so now the major concern is to identify if a person is affected by malaria or not. There are many traditional ways to test for malaria but either they require highly competent doctors or may give results in high time. Scaling of this old technique is very difficult and not having doctors with proper expertise in rural areas is also a problem. So, in this paper, we have used a Convolutional Neural Network (CNN) to classify the blood images as infected or not and get the results faster. Three different deep learning models were compared to find out the most accurate model which will automate the process and can be used by doctors in remote areas to get faster results.

Keywords –Convolutional Neural Network, Deep Learning, Malaria Detection, VGG- 16, RESNET-50, Inceptionv3, layers, transfer learning

I. INTRODUCTION

Malaria is a serious disease caused by Plasmodium bacteria. It is a mosquito-borne disease caused by the bite of a female Anopheles that transmits a type of infective single-celled organisms into the human body, which reproduces in the blood cells. Researches show that Malaria affects more than 500 million people every year, nearly 0.5% of which results in death; here children are the main victim of death due to malaria.

The symptoms of the person infected with malaria parasites can vary from mild to serious illnesses or even death. The most common symptoms are Cough, Cold, Fever, Nausea, Headache, Joint Pain, Muscle Pain, Fatigue, Vomiting, Diarrhoea, etc. Malaria is a common disease in tropical and subtropical areas like South and Southeast Asia, Pacific Islands, Central, and northern South America, and Sub-Saharan Africa.

It is often found in rural areas where the practices of medical diagnoses are generally poor and the access to health care methods is limited. It is a treatable disease that can be diagnosed and treated quickly and effectively. Diagnosis involves diagnosing malaria parasites or antigens/products in a patient's blood but Diagnosis of Malaria is a challenging step. The procedure does not work well

as the diagnosis is based on experience and the person with the knowledge required for the test. “It is eye- straining and over exhaustive work for pathologists when they examine voluminous samples [6].” We aim to build an accurate malaria detection model that can be used without relying on skilled professionals. To make the diagnostic process automatic, much research has been done on machine learning models. Machine learning models require more tuning, material analysis, and feature engineering. In this paper, we have compared 3 models to find out which Convolutional Neural Network between VGG-16, RESNET50, and Inception v3 performs better on the malaria dataset.

A Convolutional Neural Network (CNN) is a type of neural network that is used in image processing problems. It transforms an input image in to an output image with a class label. “Convolutional neural network (CNN) is a multi-stage deep architecture that alternates convolutional layers with pooling or subsampling layers, followed by one or several fully connected layers [4].” “CNN architecture offers end-to-end learning, it takes raw microscopic images as input and performs classification asks, with substantial computation power [7].”

“With the increase of the depth of the neural network, the performance of the neural network will indeed be improved, but this improvement is at the expense of time and computing resources. Therefore, to reduce the cost of training, deep learning based on transfer learning emerges [3].”

The 3 Convolutional Neural Network models used in this research are:

- I. VGG16**
- II. InceptionV3**
- III. RESNET-50**

Models will be tested based on accuracy. We aim to meet the need for another reliable method that will help provide access to the standard high-quality diagnostics that are not currently available.

II. RELATED WORK

Malaria has been a life-threatening disease and significant research has been done to detect it quickly without the involvement of skilled professionals. In the early days, malaria was mostly diagnosed in the laboratory which requires the high expertise of humans. But now using machine learning and deep learning models can be

Used to overcome this problem. For Instance, Dongetal. For example, evaluated the performances of three popular Convolutional Neural networks [9], namely LeNet5[10], AlexNet [11], and GoogLeNet [12].

Raghuveer etal [13] said that variability and artifacts are important for taking microscopic images of malarial cells. The model shows that they have taken Leishman's blood smears for this project. So, understanding the concept of the Leishman blood smears and undergoing our project withthe same concept.

Weihong et al [14] proposed and introduced the advanced concept of a convolutional neural network called the VGG which is Visual Geometry Graphics. The VGG- 16 is used in their model. The concept of the VGG-16 model is used in our project.

Vadavallietal [15] focused on the concept of deep learning and the use of deep learning techniques to define the truth of a given corpus.

Hungetal pre-trained a model on Imagenet [16] but fine- tuned it on their own data for detecting

malaria parasites [17]. In this paper, we compare 3 deep convolutional neural networks i.e., VGG- 16, RESNET50, and Inception v3, and find which model gives the best accuracy on the dataset.

III. METHODOLOGY

Dataset and Preprocessing

The dataset is taken from the National Institute of Health's (NIH) official website. It is collected from the National Library of Medicine. It contains blood cell sample images of two categories- parasitized and uninfected. There are 27,557 samples in total out of which 13,778 are infected and 13,779 non- infected, added in two different folders labeled 'Parasitized' and 'Uninfected'. Fig1: shows parasitized and uninfected blood smears respectively.

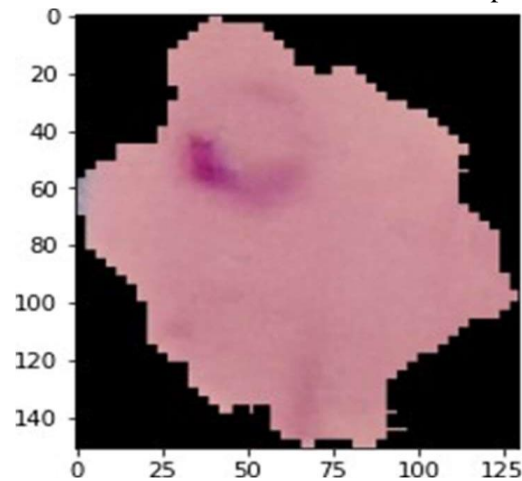


Fig1: Parasitized and Uninfected Cell

In this project we are dealing with image data, hence CNN is used. Here, we have used transfer learning using the following architecture:

- VGG-16
- Inception
- ResNet50

These 3 models will be compared on our dataset to check out the most accurate model. After acquiring all the images, we started the process of creating train and test data frames. To make sure that training and testing data frames contain an equal number of parasitized and uninfected images, firstly training and test data frames for parasitized images will be created and then for uninfected images. And now for creating final data frames, the train data frame of parasitized and uninfected images is concatenated and similarly, the test data frame is also concatenated. Provided 23426 images to the training dataset and 4134 for testing.

CNN works when all images have the same dimension. Fig shows the Plot for the distribution of dimensions of images in test_df. So, after observing and calculating the mean size of an image, we have fixed the length and breadth of all the images to 132.

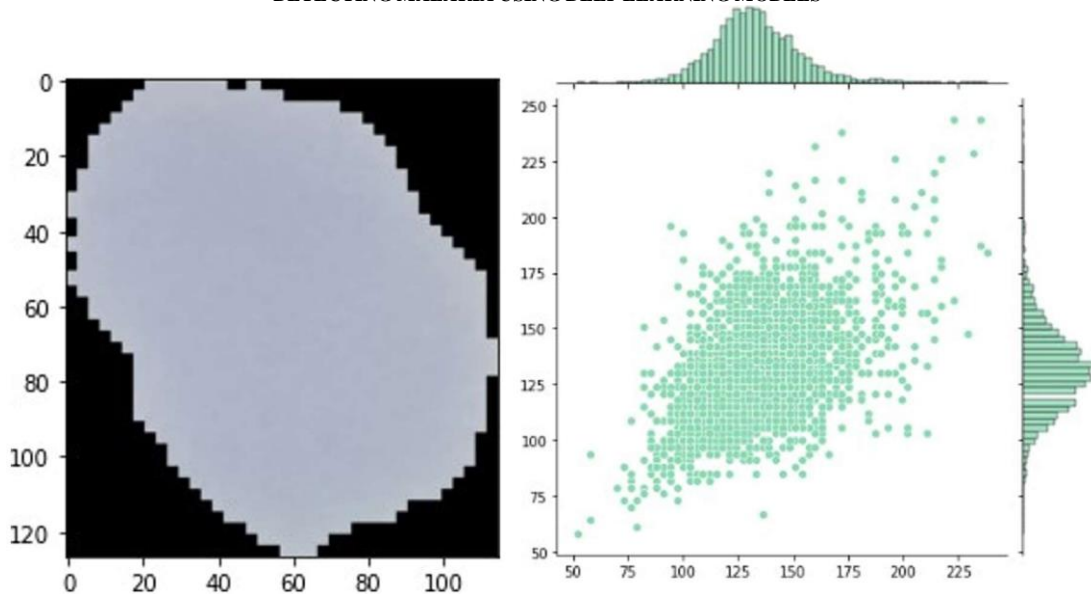


Fig2: Plot for distribution of dimension of images in test_df

After this, we can start building our models.

IV. MODELS

As discussed earlier, in this project three different deep learning models were compared named VGG16, InceptionV3, and Resnet50. In this project transfer learning (i.e already trained model) is used.

VGG 16

The CNN model AlexNet has many convolutions, max pools, and fully connected dense layers. In AlexNet the total number of Convolution layers is 5 and it contains 3 max pools and fully connected layers. In AlexNet there are a lot of variations as the kernel sizes used are not specific 16 is preferred as it has a very uniform architecture and can be more easily understandable than AlexNet VGG stands for visual geometric group and vgg16 contains 16 different layers which is why it is named VGG 16. To avoid these variations, the visual geometric group of Oxford University proposed this architecture in 2014 and secured the 2nd position in the visual recognition challenge. In vgg16 architecture, the number of convolution layers is 13, max pool layers are 5 and it has 3 fully connected layers and the last layer is softmax i.e., output layer, having 1000 output.

Features of VGG 16

It accepts color images of size 224 by 224 pixels, so there are three different channels at the input, red, green, and blue. And the layer is sent through many convolutional layers.

In VGG 16 every convolution layer has a filter size 3 by 3 and stride equal to 1 and padding is the same, which indicates that the receptive field of every convolution kernel is an image of size 3 by 3. Row and Column padding is used by every convolution layer to keep the size of the input and output feature map the same.

The Maxpool operation is carried out with a filter size of 2 by 2 and with a stride equal to 2 and padding is also the same, no max pooling is performed after every convolution layer. There are 3 fully connected dense layers, the first and second fully-connected layers have 4096 units, and the last layer i.e., Softmax layer has 1000 channels, as ImageNet has 1000 classes. And every hidden layer or every convolution layer has ReLu (Rectified Linear Unit) activation function to stop forwarding the negative values through the network.

DETECTING MALARIA USING DEEP LEARNING MODELS

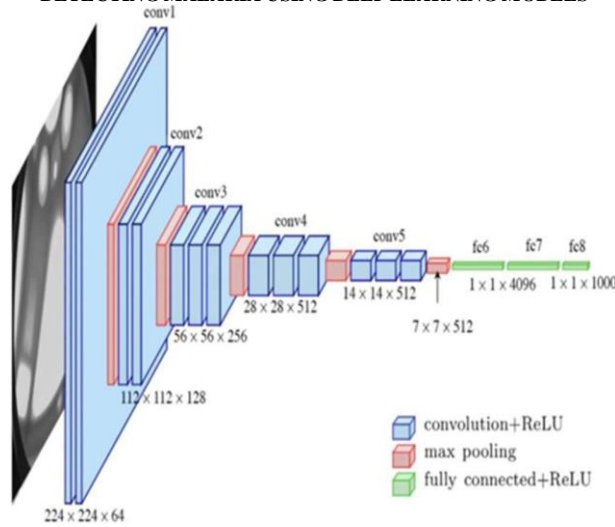


Fig3: VGG 16 Architecture

RESNET-50

Resnet-50 is a convolutional neural network with 50 layers, that is trained on more than a million images from the ImageNet dataset. It has 48 convolutional layers, 1 MaxPool layer, and 1 Average Pool layer. “It can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals [8].”

RESNET-50 Architecture

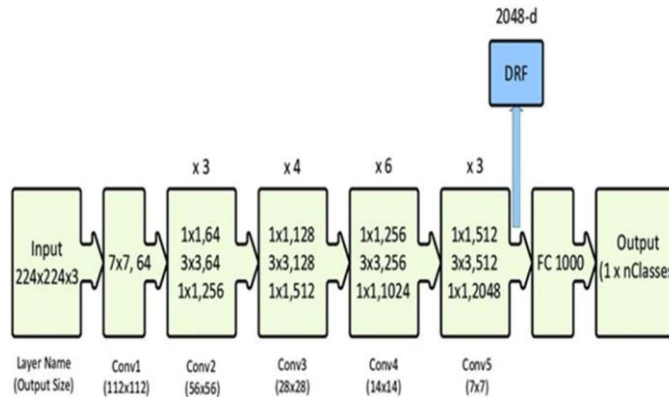


Fig4: RESNET-50Architecture

The above picture depicts RESNET-50 architecture and it has the following elements:

- a. Layer 1 is made up of 64 different kernels, each with a size 7*7 and with a stride of size 2.
- b. Next is max pooling is striding size same as layer 1.
- c. Next, there is convolution with 64 kernels which are 1*1, the next 64 are 3*3 and the last 256 kernels are also 1*1. This is repeated 3 times, so this gives us 9 layers.
- d. Next, there is convolution 3 with 128 kernels of 1*1, 128 kernels of 3*3, and 512 kernels of 1*1. This is repeated 4 times, so in total 12 layers.
- e. Next, there is convolution 4 with 256 kernels of 1*1, 256 kernels of 3*3, and 1024 kernels of 1*1. This is repeated 6 times, so it totals 18 layers.
- f. After that, there is convolution 5 with 512 kernels of 1*1, 512 kernels of 3*3, and 2048 kernels of 1*1. This is repeated 3 times, giving us 9 layers.

g. At last, “we do an average pool and end it with a fully connected layer containing 1000 nodes and at the end a softmax function so this gives us 1 layer [1].”

So, in total $1 + 9 + 12 + 18 + 9 + 1 = 50$ layers deep RESNET convolutional network.

This model can be used for both computer vision and non-computer vision tasks. The building block of RESNET-50 is of bottleneck design.

Inception v3

Inception v3 is a Deep Convolutional Neural Network and is trained on more than a million images from the ImageNet dataset. This dataset contains 1000 different classes such as “keyboard”, “computer”, “pen”, and many more., “Inception-v3 model includes three parts: the basic convolutional block, improved Inception module and the classifier[4].”

This model has been able to attain an accuracy greater than 78.1% on the ImageNet dataset in about 170 epochs. This model is made up of several symmetric and asymmetric building blocks, which include convolutions, concatenations, dropouts, average pooling, max pooling, and fully connected layers.

Transfer learning is applied to Inception v3 and then re-train it to classify a new set of images to get the desired results.

“Inception v3 focuses on burning less computational power. This was achieved by modifying the architecture of the previous Inception model i.e Inception v1 and v2. This idea was proposed in the paper *Rethinking the Inception Architecture for Computer Vision*, which was published in 2015. It was co-authored by Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, and Jonathon Shlens.[2]”

Inception Networks are more computationally efficient, in terms of the number of parameters generated by the network and the economic cost incurred i.e. memory and other resources.

In the Inception v3 model, techniques like factorized convolutions, regularization, dimension reduction, and parallelized computations are used for optimizing the network and making it easier to adapt.

“In summary, Inception-v3 has the state-of-the-art performance on object recognition, which benefits from its unique Inception architecture. Therefore, this model is widely used for transfer learning [4].”

Inception v3 architecture

This architecture contains the following modifications:

- [1] **Factorized Convolutions:** It keeps check on network efficiency and helps in reducing the computational efficiency by reducing the number of parameters involved in a network.
- [2] **Smaller Convolutions:** to get faster training speed, bigger convolutions are replaced by smaller convolutions. For eg: a 5×5 filter has 25 parameters, but two 3×3 filters replacing a 5×5 convolution have only 18 ($3 \times 3 + 3 \times 3$) parameters.

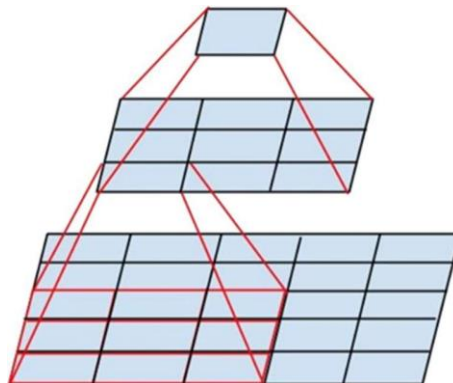


Fig5: Smaller Convolutions

[3] **Asymmetric convolutions:** Symmetric 3×3 convolutions could be replaced by 1×3 asymmetric convolutions followed by a 3×1 asymmetric convolution. If a 3×3 convolution is replaced by a 2×2 convolution, the number of parameters would be slightly higher than the asymmetric convolution proposed.

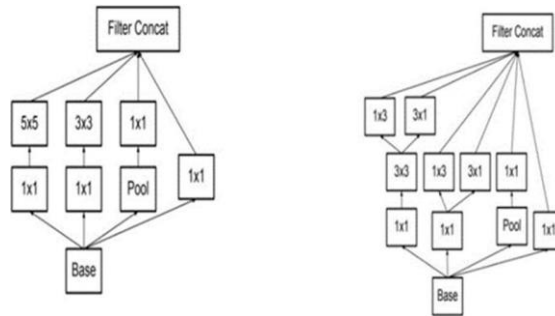


Fig6: Asymmetric Convolutions

[4] **Auxiliary classifier:** It acts as a regularizer. It is a small CNN inserted during training between layers, and the loss is added to the main network loss.

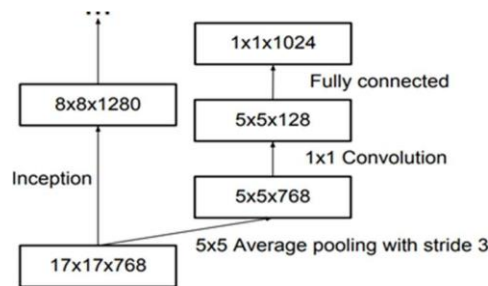


Fig7: Auxiliary classifier

[5] **Grid size reduction:** It is done by pooling operations.

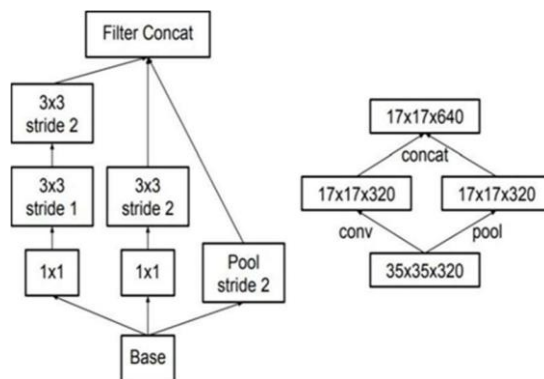


Fig8: Grid size-reduction

All these above concepts are combined to get this final architecture shown below:

DETECTING MALARIA USING DEEP LEARNING MODELS

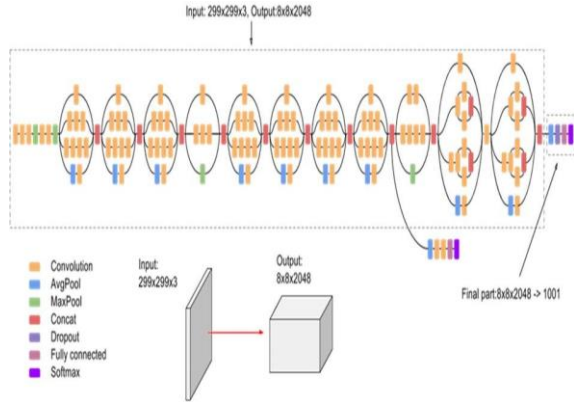


Fig9: Inception v3 Architecture

Inception v3 has the lowest error rates and is more efficient when compared to other models. It is made up of 42 layers in total.

V. RESULT

In this paper, research is done on human red blood cell smear images to classify them as parasitized or uninfected. It uses three different pre-trained Convolutional Neural Network models with transfer learning. These models were run on a malaria dataset from the NIH official website and their performances were recorded. We compared their performances to know which model gives the best accuracy on the dataset. The three models compared are named VGG16, InceptionV3, and Resnet50. VGG16 achieves an accuracy of 92%, Inception V3 achieves an accuracy of 91% and Resnet50 achieves an accuracy of 59%. It was evaluated that the highest accuracy is given by VGG16 followed by InceptionV3 and Resnet50.

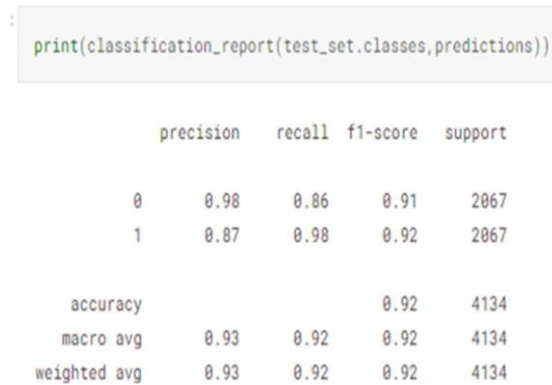
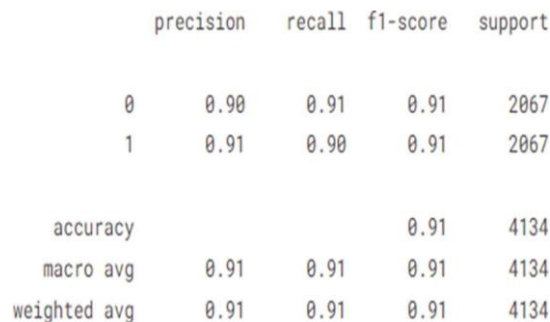


Fig10: VGG16 model Classification report



DETECTING MALARIA USING DEEP LEARNING MODELS

Fig11:Inceptionv3 model Classification report

	precision	recall	f1-score	support
0	0.73	0.30	0.42	2067
1	0.56	0.89	0.69	2067
accuracy			0.59	4134
macro avg	0.64	0.59	0.55	4134
weighted avg	0.64	0.59	0.55	4134

Fig12:Resnet50 model Classification report

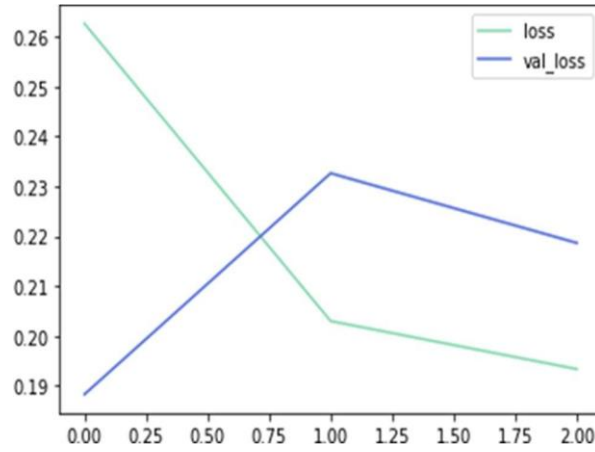


Fig13: Training and validation loss of VGG16

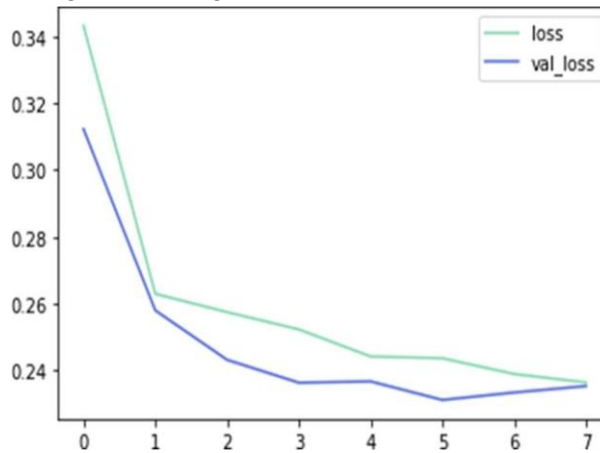


Fig14: Training and validation loss of Inception v3

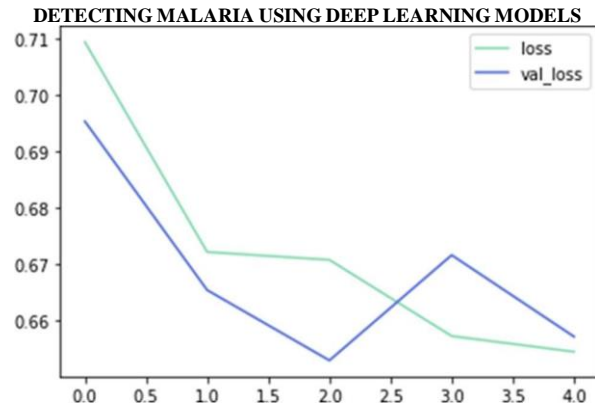


Fig 15: Training and validation loss of Resnet 50

```
vgg_model_accuracy = losses['val_accuracy'].iloc[-1].round(2) * 100
vgg_model_accuracy
```

92.0

Fig16: Final model accuracy of VGG 16

```
inception_model_accuracy = losses['val_accuracy'].iloc[-1].round(2) * 100
inception_model_accuracy
```

91.0

Fig 17: Final model accuracy of Inception v3

```
resnet_model_accuracy = losses['val_accuracy'].iloc[-1].round(2) * 100
resnet_model_accuracy
```

59.0

Fig 18: Final model accuracy of Resnet 50

The performance of VGG 16 is the best with 92% accuracy.

VI. CONCLUSION

In this work, we have first taken the dataset from the National Institute of Health's (NIH) official website. It consists of 27,557 images of blood cell samples in total. Out of which 13,778 are infected and 13,779 are non-infected, added in two different folders labeled 'Parasitized' and 'Uninfected'. We then created train and test data frames and changed the dimensions of each image to 132. Thereafter we ran 3 different models of the dataset namely, VGG16, Inception V3, and RESNET 50. We further compare and analyze the performance of these 3 models on our dataset and conclude that VGG 16 shows the best result with 92% accuracy. We have extensively evaluated and compared the results of the 3 models.

VII. FUTURE SCOPE

In the future, we would like to extend this model to smartphones using cloud technology, so that it will be easy for doctors in remote areas to detect malaria easily and decrease the damage caused by malaria.

REFERENCES

- [1] OpenGenus, Understanding ResNet50 architecture.
<https://iq.opengenus.org/resnet50-architecture/#:~:text=ResNet50%20is%20a%20variant%20of,explored%20ResNet50%20architecture%20in%20depth.>
- [2] A Review of Popular Deep Learning Architectures: ResNet, InceptionV3, and SqueezeNet.
<https://blog.paperspace.com/populardeeplearningarchitectures-resnet-inceptionv3-squeezenet/>
- [3] Wang, C., Chen, D., Hao, L., Liu, X., Zeng, Y., Chen, J. and Zhang, G., 2019. Pulmonary image classification based on inception-v3 transfer learning model. *IEEE Access*, 7, pp.146533-146541.
- [4] Lin, C., Li, L., Luo, W., Wang, K.C., and Guo, J., 2019. Transfer learning-based traffic sign recognition using the inception-v3 model. *Periodica Polytechnica Transportation Engineering*, 47(3), pp.242-250.
- [5] WorldHealthOrganizationWorldMalariaReport<https://www.who.int/news-room/fact-sheets/detail/malaria>
- [6] Poostchi M, Silamut K, Maude RJ, Jaeger S, Thoma G (2018) Image analysis and machine learning for detecting malaria. *Trans Res: J Lab Clin Med* 194:36–55
- [7] Rajaraman S et al. (2018) Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *Peer J*
- [8] MathWorks, ResNet-50 convolutional neural network MATLAB <https://www.mathworks.com/help/deeplearning/ref/resnet50.html;jsessionid=b1f6299bb45df84834cee735d5dc>
- [9] Dong, Yuhang, Zhuocheng Jiang, Hongda Shen, W. David Pan, Lance A. Williams, Vishnu VB Reddy, William H. Benjamin, and Allen W. Bryan. "Evaluations of deep convolutional neural networks for automatic identification of malaria-infected cells." In *2017 IEEE EMBS international conference on biomedical & health informatics (BHI)*, pp. 101-104. IEEE, 2017.
- [10] LeCun, Yann. "LeNet5, convolutional neural networks." *URL: http://yann. lecun. com/exdb/lenet* 20, no. 5 (2015): 14.
- [11] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems* 25 (2012).
- [12] Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1-9. 2015.
- [13] Makkapati, Vishnu V., and Raghuvver M. Rao. "Segmentation of malaria parasites in peripheral blood smear images." In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1361-1364. IEEE, 2009.
- [14] Liu, Shuying, and Weihong Deng. "Very deep convolutional neural network based image classification using small training sample size." In *2015 3rd IAPR Asian conference on pattern recognition (ACPR)*, pp. 730-734. IEEE, 2015.
- [15] Vadavalli, Adilakshmi, and R. Subhashini. "Deep Learning based truth discovery algorithm for research the genuineness of given text corpus." *International Journal of Recent Technology and Engineering* (2019).