MALARIA CELL CLASSIFICATION WITH VGG-19 AND INCEPTION-RESNET

ABSTRACT

Malaria is a life-threatening disease that affects millions of people worldwide. The diagnosis of malaria is typically done by examining blood smears under a microscope for the presence of malaria parasites. However, this process is time-consuming and requires skilled technicians to accurately identify the parasites. Manual examinations can be burdensome for large-scale diagnosis in endemic regions resulting in poor quality, and unnecessary medication, leading to severe economic impact on individual health programs. In recent years, there has been increasing interest in using machine learning algorithms to automate the diagnosis of malaria from digital images of blood smears. In this study, analysis was done for two classification models i.e., VGG-19 and Inception-Resnet, to see which model performs the best in detecting malaria cells. As a result, the VGG-19 and Inception-Resnet obtain a validation accuracy of 0.87 and 0.92 respectively.

INTRODUCTION

Malaria is an infectious disease caused by the single-celled Protozoan Parasite of the Plasmodium group. The disease is mainly spread through the bite of infected female Anopheles Mosquito. The disease puts nearly 40% of the world population to risk with nearly 240 million cases reported each year [1]. African countries got affected the most due to malaria. 92% of global malaria deaths occur in Africa where children below 5 years accounted for two-thirds of the malaria-infected deaths [2]. The typical malaria symptoms are fever, nausea, headaches and in severe cases, yellow skin, seizures, and coma, which leads to death. Several millions of blood films are examined every year by trained experts for the detection of malaria infection. The detection of malaria involves counting the parasites and infected red blood cells manually. However, it completely depends on the experience and skill of the microscopist [3]. While working in a limited resource set-up with no system helpful for the maintenance of the skill will affect the diagnostic quality [4]. Which leads to wrong diagnostic decisions. Deep learning can be implemented in the classification of cell images which can prevent the wrong diagnostic decisions. Deep learning is an area of machine learning, which performed exceptionally well in many non-medical fields. The applications of deep learning have been limited in the medical field due to a lack of expert knowledge in that field and privacy concerns. But, in recent years deep learning was used in many medical fields [5- 6].

In recent years, there has been increasing interest in using machine learning algorithms to automate the diagnosis of malaria from digital images of blood smears. The goal of malaria cell classification is to classify blood cell images into two classes: uninfected and parasitized. Uninfected cells are those that do not contain the malaria parasite in RBC, while parasitized cells are those in which RBC is infected with the parasite. By correctly identifying parasitized cells, a machine learning algorithm can provide a fast and accurate diagnosis of malaria, which can lead to better treatment outcomes and ultimately save lives.

In this study, the comparison was made between the performance of two popular deep learning models, VGG-19 and Inception-Resnet. Publicly available dataset of blood smear images containing both infected and parasitized cells. I did fine-tune pre-trained models on the malaria

cell dataset and evaluated their performance using various evaluation metrics such as accuracy, precision, recall and f1 score.

DATASET

The dataset used for the malaria cell classification [8], consists of 27560 images of parasitized and uninfected RBC cells. The average size of all the images was found to be (132,132). The difference between malaria-infected and uninfected cells from the dataset, chosen randomly, can be seen in Fig. 1
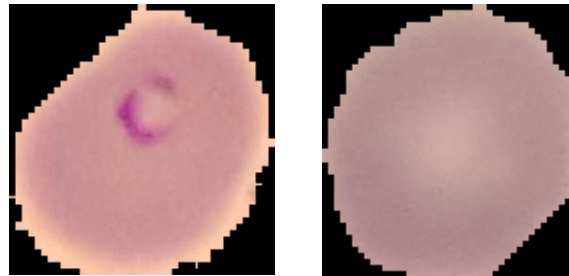

Fig 1: Image of parasitized and uninfected cell

TRANSFER LEARNING

Transfer learning generally refers to a process where a model is trained on one problem and is used in some way on the second related problem. It is one of the most popular methods in computer vision because it allows us to build accurate models in a time-saving way.

In computer vision, transfer learning is usually expressed through the use of a pre-trained model. A pre-trained model is a model that was trained on a large benchmark dataset such as an imagenet, to solve problems similar to the one we want to solve. The popular ones among various models based on published literature are VGG, Inception, Resent and so on. These are. VGG-19 and Inception Resenet are the two models used in this study and their comparison was done.

VGG19 is a deep convolutional neural network architecture that was developed by the Visual Geometry Group at the University of Oxford. It was introduced in 2014 as an improvement over the earlier VGG16 model. The VGG19 model consists of 19 layers, including 16 convolutional layers and 3 fully connected layers. The convolutional layers are organized into five groups, with each group containing multiple convolutional layers followed by a max pooling layer. The fully connected layers are followed by a softmax activation function that produces the final output probabilities.

Inception-ResNet is a deep neural network architecture that combines the Inception and ResNet architectures. It was introduced by Google researchers in 2016 as an improvement over previous state-of-the-art models for image classification tasks. The Inception-ResNet model uses a combination of Inception modules and residual connections to extract features from input images. The Inception module is a convolutional neural network module that performs multiple parallel convolutions at different scales and concatenates their output. The residual connection is a skip connection that allows the model to learn residual functions that can be added to the input to produce the output.

FINETUNING VGG AND INCEPTION RESNET

For both VGG-19 and Inception Resnet models, the image will enter through these models. The output is then finetuned with the Average Pooling layer followed by a dense layer. The output is a single neuron with a sigmoid activation function for binary classification i.e., uninfected and parasitized cell images.

The reason to use sigmoid activation other than softmax is that the task which is at hand is the binary classification of images and sigmoid performs well with binary classification. Softmax is good for multi-class classification. The mathematical representation of sigmoid activation is shown in (1), where x is the dot product of each neuron value with the weights

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{1}$$

The experiment was carried out on Jupyter Notebook, in an environment of Keras with Tensorflow backend. The system in which the experiment was carried out runs on Google Colaboratory.

Splitting Data

The total number of infected and uninfected cell images in the dataset is 27,560. The split was done with 85, 15 per cent for train and testing. It is a well-known practice of splitting the data such that training data gets to be more than two-thirds of the whole data. Finally, after the split, training data has 23,426 cell images (both infected and uninfected). Whereas, training and validation data has 4,134 cell images each.

Training Model

The model was trained for a maximum of 20 epochs. The training was early stopped if the validation loss was not improved for 2 epochs. That means if the validation loss does not decrease for two consecutive epochs, training was stopped early. Optimizer, in simple terms, sets the learning rate of a neural network. The optimizer used for this model is ADAM(Adaptive Moment Estimation).ADAM optimizer has proven to be performing better than many other optimizers. [10]. Choosing a loss function can also be a difficult task. The loss function used to find the loss for this model is Binary-Cross Entropy.

	VGG-19	INCEPTION-RESNET
Training Accuracy	0.9169	0.9131
Validation Accuracy	0.8737	0.9156
Training Loss	0.2132	0.2229
Validation Loss	0.3287	0.2121

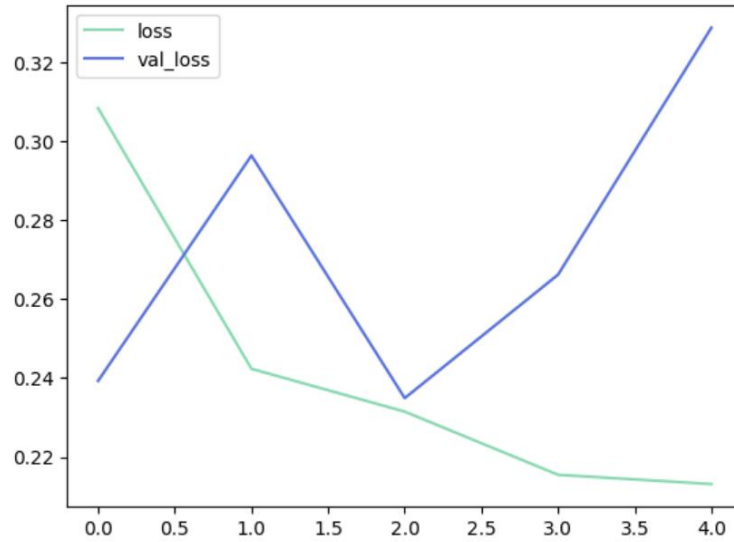Table 1: Comparison of VGG-19 and Inception-Resnet

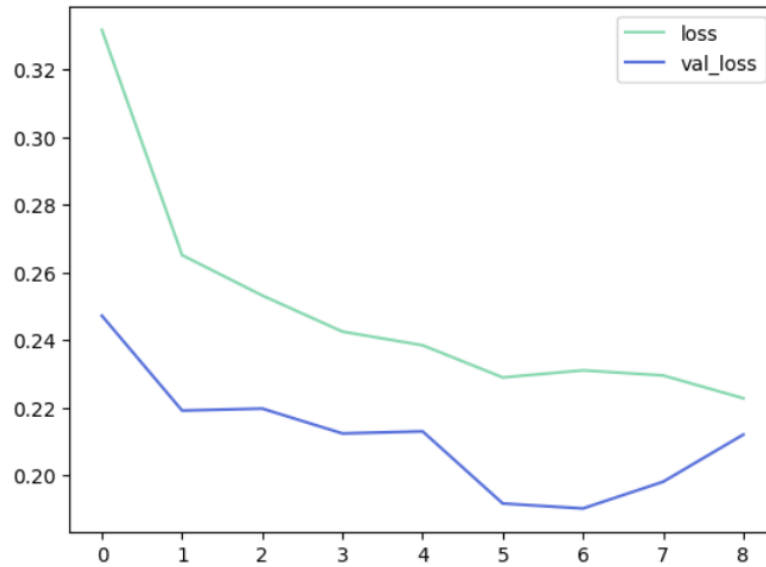Fig 2: Training Loss vs Validation Loss for VGG-19 model

Fig 3: Training Loss vs Validation Loss for Inception-Resnet model

EVALUATION METRICS

Malaria cell classification is a binary classification problem where the goal is to classify whether a given cell image is infected (parasitized) or not (uninfected). Accuracy, precision, recall, and F1 score are commonly used evaluation metrics for binary classification problems.

- Accuracy: measures the proportion of correctly classified samples out of the total samples. In malaria cell classification, accuracy represents the percentage of correctly classified parasitized and uninfected cell images out of the total number of images.

- Precision: measures the proportion of true positives (TP) out of all positive predictions (TP + FP). In malaria cell classification, precision represents the percentage of correctly classified parasitized cell images out of all the images predicted as parasitized.

- Recall: measures the proportion of true positives (TP) out of all actual positives (TP + FN). In malaria cell classification, recall represents the percentage of correctly classified parasitized cell images out of all the actual parasitized cell images.
- F1 score: is the harmonic mean of precision and recall. It combines both precision and recalls into a single metric. The F1 score is useful when the class distribution is imbalanced. In malaria cell classification, the F1 score represents the balance between precision and recall.

For both VGG-19 and Inception Resnet, the confusion matrix along with the values of precision, recall, and f1-score are given in the figure below. Here, 0 indicates parasitized and 1 indicates uninfected cells.
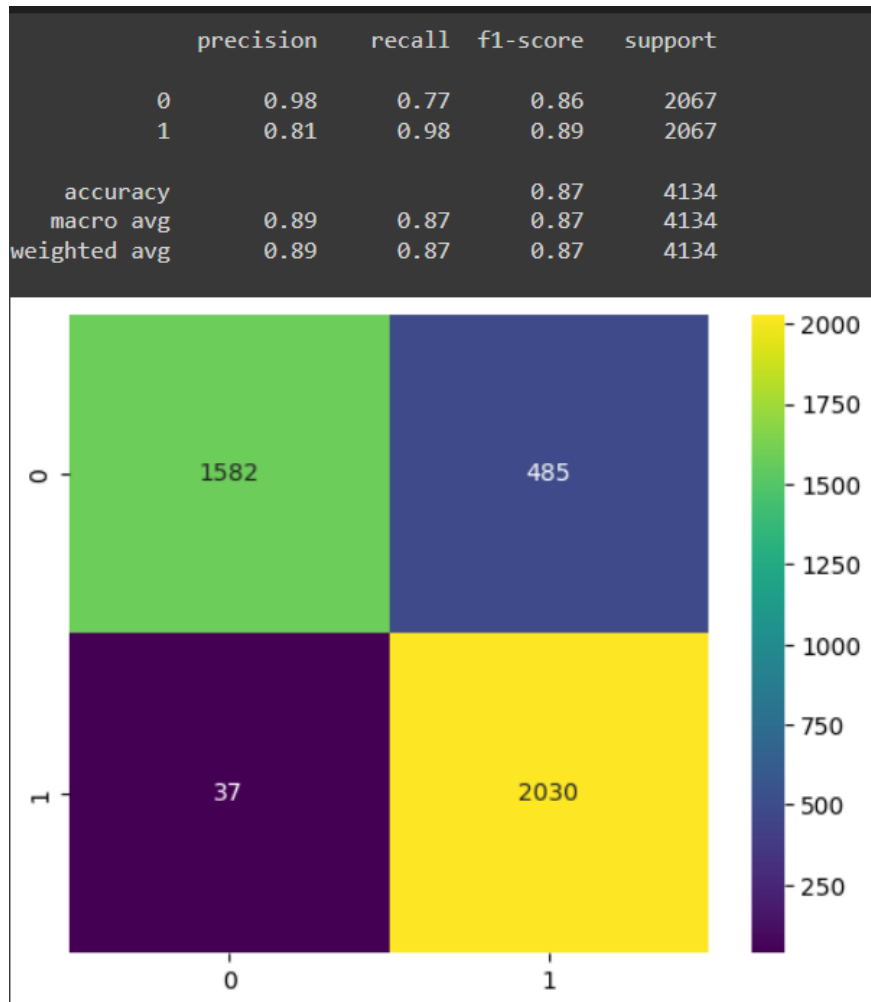

Fig 4: Different values of evaluation metrics for the VGG-19 model along with the confusion matrix
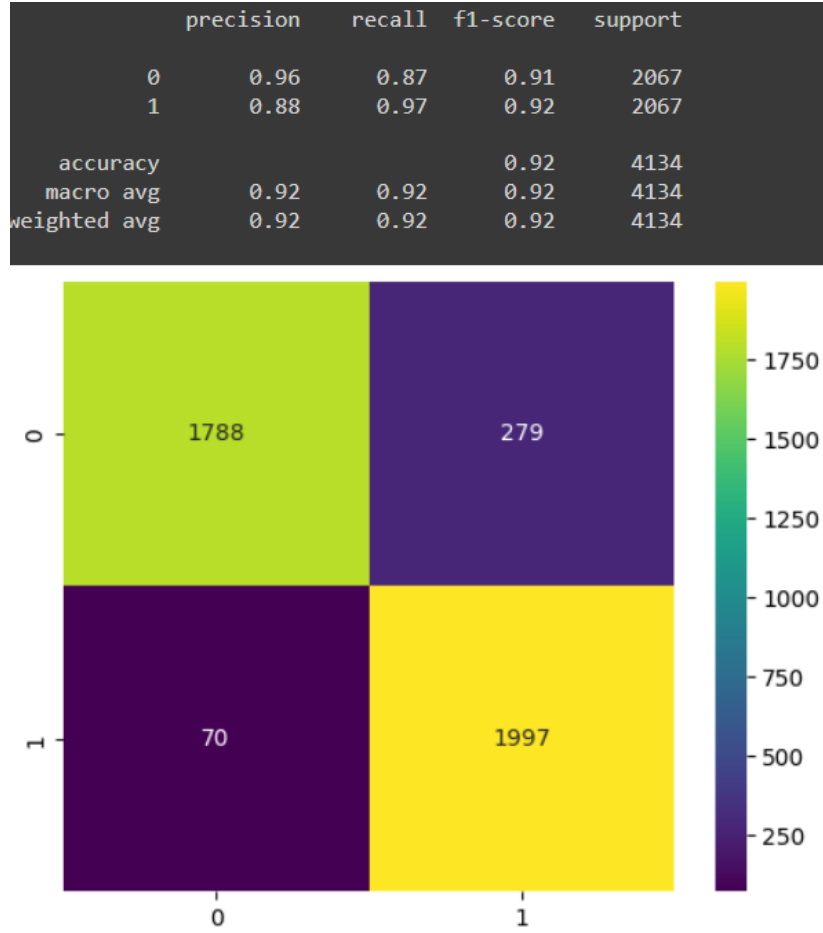
```
              precision    recall  f1-score   support

           0       0.96      0.87      0.91      2067
           1       0.88      0.97      0.92      2067

    accuracy                           0.92      4134
   macro avg       0.92      0.92      0.92      4134
weighted avg       0.92      0.92      0.92      4134
```

Fig 5: Different values of evaluation metrics for the Inception Resnet model along with the confusion matrix

VISUAL CLASS ACTIVATION MAP

The class activation mapping (CAM) for CNNs with global average pooling enables classification-trained CNNs to learn to perform object localization without using anu bounding box annotations. Class activation maps allow us to visualize the predicted class scores on any given image, highlighting the discriminative object parts detected by the CNN [7].

The class activation map for the VGG-19 model and Inception Resnet was visualized in which the uninfected cell shows a similar colour map while the parasitized cell shows the high-intensity colour i.e., red in the infected area while the other shows the low-intensity blue colour.

CONCLUSION

The final result provides satisfactory results for VGG-19 as compared to Inception-Resnet. Both used transfer learning From the data obtained, Inception Resnet gave better accuracy for validation sets i.e., 92% as compared to 87% for VGG-19 along with resulting in low loss than the VGG-19 model. Both models were compared with other evaluation metrics such as precision, recall and f1 score as well.

REFERENCES

[1] World Health Organization (2012) malaria fact sheet no. 94. Geneva: WHO.

[2] WHO. World malaria report 2021. World Health Organization; 2021.

[3] WHO. Malaria microscopy quality assurance manual-version 2.World Health Organization; 2016.

[4] Moeskops P. et al. (2016) Deep Learning for Multi-task Medical Image Segmentation in Multiple Modalities. In: Ourselin S., Joskowicz L., Sabuncu M., Unal G., Wells W., (eds) Medical Image

[5] Computing and Computer-Assisted Intervention – MICCAI 2016. Lecture Notes in Computer Science, vol9901. Springer, Cham.

[6] Maninis KK., Pont-Tuset J., Arbelaez P., Van Gool L. (2016) Deep Retinal Image Understanding. In: Ourselin S., Joskowicz L., Sabuncu M., Unal G., Wells W. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. Lecture Notes in Computer Science, vol9901. Springer, Cham.

[7] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, Antonio Torralba Computer Science and Artificial Intelligence Laboratory, MIT (2015), Learning Deep Features for Discriminative Localization: https://arxiv.org/pdf/1512.04150.pdf

[8] Malaria Cell Images Dataset: https://www.kaggle.com/datasets/iarunava/cell-images-for-detecting-malaria

[9] https://link.springer.com/article/10.1007/s11042-022-13820-0

[10] Hassan, E., Shams, M.Y., Hikal, N.A. et al. The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study. Multimed Tools Appl **82**, 16591–16633 (2023). https://doi.org/10.1007/s11042-022-13820-0

[11] A. Sai Bharadwaj Reddy and D. Sujitha Juliet (2019) Transfer Learning with ResNet-50 for Malaria Cell-Image Classification https://www.researchgate.net/profile/Sai-Bharadwaj-Reddy-Arrabelly/publication/332678948_Transfer_Learning_with_ResNet-50_for_Malaria_Cell-Image_Classification/links/5f294409299bf134049ee5f8/Transfer-Learning-with-ResNet-50-for-Malaria-Cell-Image-Classification.pdf