MALARIA CELL-IMAGE CLASSIFICATION USING CNN

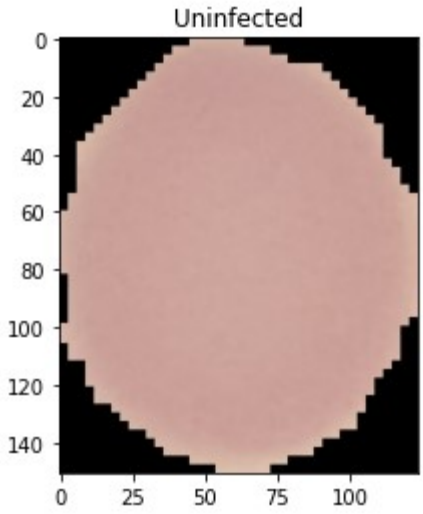
Adwaid R
adwaid.dhwani@gmail.com
+91-9072424645

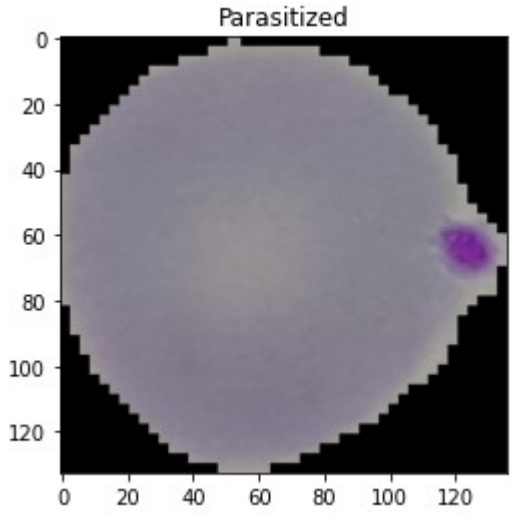
Introduction :

Malaria is a disease spread by female anopheles mosquito bites that is fatal. Plasmodium parasites are carried by the bite. According to the WHO (World Health Organization), 229 million cases of malaria were reported globally in 2019. In the same year, an estimated 409 000 people died from malaria. If diagnosed early and treated promptly, the infection will not spread. The disease is diagnosed via a microscopic examination of the patient's blood sample. The sample is thinly spread as a smear, with the cell images serving as the visual criteria. Diagnosis is a time-consuming process that requires the help of an expert. Various machine learning and deep learning strategies have been developed to avoid inaccurate conclusions caused by human error. I built a simple CNN to detect the malaria cell .The results reveal that the model had a 96.36 percent accuracy score.

Datasets :

There are 27,560 cell images in the dataset. Half of the photos are parasitized, while the other half are unaffected. Positive samples contained Plasmodium and negative samples contained no Plasmodium but other types of objects including staining artifacts /impurities . This dataset can be downloaded from the National Library of Medicine's official website (NLM). It's also available in Kaggle and Tensorflow datasets. The images that are stored in the dataset are taken from patients who are chosen at random. An instance image of each class label namely parasitized and uninfected is shown below.





Deep Learning :

" Deep learning attempts to mimic the human brain—albeit far from matching its ability—enabling systems to cluster data and make predictions with incredible accuracy. "

Deep learning is a subset of machine learning, which is essentially a neural network with three or more layers. These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—allowing it to "learn" from large amounts of data. While a neural network with a single layer can still make approximate predictions, additional hidden layers can help to optimize and refine for accuracy. Deep neural networks consist of multiple layers of interconnected nodes, each building upon the previous layer to refine and optimize the prediction or categorization. This progression of computations through the network is called forward propagation. The input and output layers of a deep neural network are called *visible* layers. The input layer is where the deep learning model ingests the data for processing, and the output layer is where the final prediction or classification is made.

Another process called backpropagation uses algorithms, like gradient descent, to calculate errors in predictions and then adjusts the weights and biases of the function by moving backwards through the layers in an effort to train the model. Together, forward propagation and backpropagation allow a neural network to make predictions and correct for any errors accordingly. Over time, the algorithm becomes gradually more accurate.

The above describes the simplest type of deep neural network in the simplest terms. However, deep learning algorithms are incredibly complex, and there are different types of neural networks to address specific problems or datasets. For example,

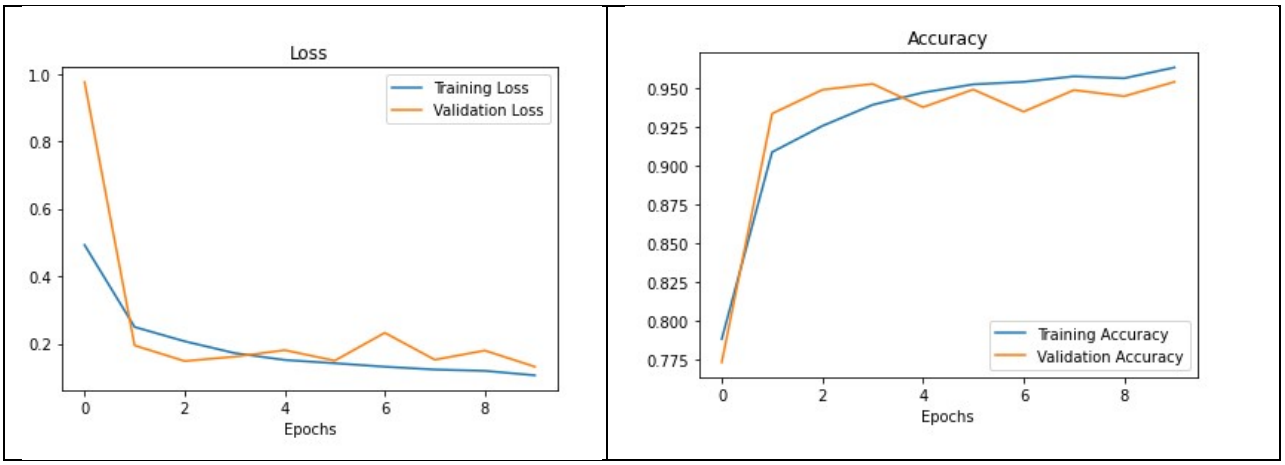
- ***Convolutional neural networks (CNNs),*** used primarily in computer vision and image classification applications, can detect features and patterns within an image, enabling tasks, like object detection or recognition.
- ***Recurrent neural network (RNNs)*** are typically used in natural language and speech recognition applications as it leverages sequential or times series data.

Model :

The proposed model receives images, produced by microscopic inspection of a patient's blood sample as input. Each image is of RGB (Red, Green, and Blue) color model. The images are augmented using tf.keras.utils.image_dataset_from_directory and sent to the model. Based on the architecture in the preceding code, our CNN model has two convolution and pooling layers followed by two dense layers and four dropout for regularization. We get a validation accuracy of 95.43 and training accuracy of 96.36 . From plotting loss and accuracy curves we get know that the model is stable after 8 epochs.

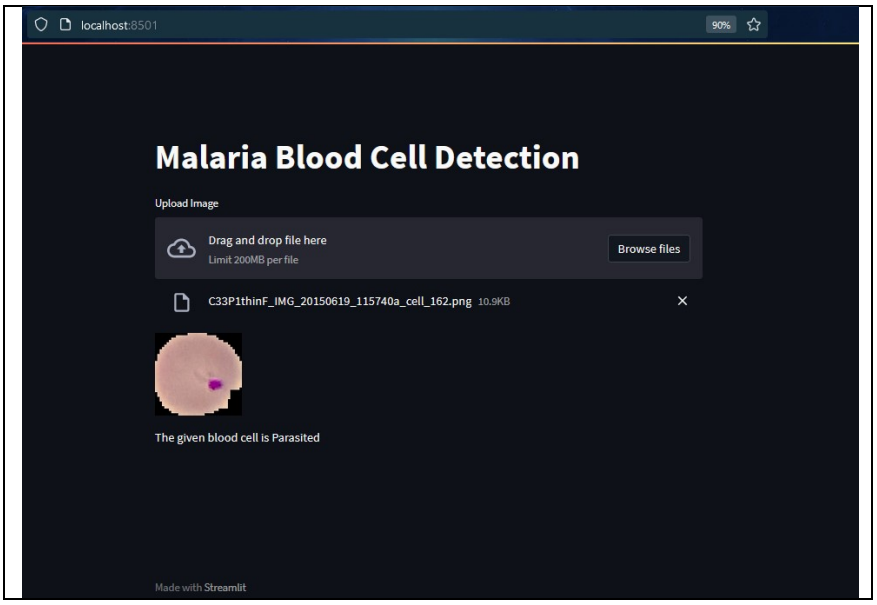
```
accuracy: 0.9636 - val_loss: 0.1314 - val_accuracy: 0.9543
```

Plotting loss training & validation curve :



Streamlit :

Streamlit is an open-source python library for creating and sharing web apps for data science and machine learning projects. The library can help you create and deploy your data science solution in a few minutes with a few lines of code. Streamlit can seamlessly integrate with other popular python libraries used in Data science such as NumPy, Pandas, Matplotlib, Scikit-learn and many more. I created a simple web app using streamlit which allows to upload an image of a blood cell and predicts whether the cell is infected with malaria.


localhost:8501
90%
Malaria Blood Cell Detection
Upload Image
Drag and drop file here
Limit 200MB per file
Browse files
C33P1thinF_IMG_20150619_115740a_cell_162.png 10.9KB
The given blood cell is Parasited
Made with Streamlit