MALARIA CELL IMAGE DETECTION

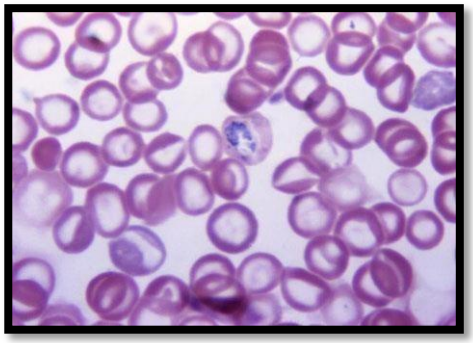
Domain Background

Malaria is a life-threatening disease caused by parasites that are transmitted to people through the bites of infected female Anopheles mosquitoes. It is preventable and curable.

- In 2017, there were an estimated 219 million cases of malaria in 90 countries.
- Malaria deaths reached 435 000 in 2017.
- The WHO African Region carries a disproportionately high share of the global malaria burden. In 2017, the region was home to 92% of malaria cases and 93% of malaria deaths.

Malaria is caused by Plasmodium parasites. The parasites are spread to people through the bites of infected female Anopheles mosquitoes, called "malaria vectors." There are 5 parasite species that cause malaria in humans, and 2 of these species – P. falciparum and P. vivax – pose the greatest threat.

Diagnosis of malaria can be difficult



- Where malaria is not endemic any more (such as in the United States), health-care providers may not be familiar with the disease. Clinicians seeing a malaria patient may forget to consider malaria among the potential diagnoses and not order the needed diagnostic tests. Laboratories may lack experience with malaria and fail to detect parasites when examining blood smears under the microscope.
- Malaria is an acute febrile illness. In a non-immune individual, symptoms usually appear 10–15 days after the infective mosquito bite. The first symptoms – fever, headache, and chills – may be mild and difficult to recognize as malaria. If not treated within 24 hours, P. falciparum malaria can progress to severe illness, often leading to death. **Microscopic Diagnosis**

Malaria parasites can be identified by examining under the microscope a drop of the patient's blood, spread out as a "blood smear" on a microscope slide. Prior to examination, the specimen is stained to give the parasites a distinctive appearance. This technique remains the gold

standard for laboratory confirmation of malaria. However, it depends on the quality of the reagents, of the microscope, and on the experience of the laboratories.

Identifying the Malaria detection by using computer Vision architecture is not much accuracy while finding the malaria cells in the human body. So to overcome this problem then Deep learning came into the picture by using it we can identify the uninfected image cell.

Convolutional neural networks have the ability to automatically extract features and learn filters. In previous machine learning solutions, features had to be *manually* programmed in — for example, size, color, the morphology of the cells. Utilizing Convolutional neural networks (CNN) will greatly speed up prediction time while mirroring (or even exceeding) the accuracy of clinicians. Reference Link:

https://towardsdatascience.com/detecting-malaria-using-deep-learning-fd4fdcee1f5a

Inspiration

Save humans by detecting and deploying Image Cells that contain Malaria or not!

Reference Links:

1. https://www.cdc.gov/malaria/diagnosis_treatment/diagnosis.html
2. https://www.who.int/news-room/fact-sheets/detail/malaria

Personal motivation

- From the very early age ,I always wanted to contribute something to the medical field.
- In this project I can learn how to handle the image data and pre-processes datasets for the training the data.
- In this project I am going to learn how to use the kaggle kernels.
- In this project I am going to learn how to detect cell image data by using keras.

Problem Statement

The Aim of this project is to detect the Malaria by using the Cell Image Dataset. In this project I am going to use keras and for improving the Accuracy for detecting the cell Image.

Datasets and Inputs

Dataset was Downloaded form https://www.kaggle.com/iarunava/cell-images-for-detectingmalaria

Content

The dataset contains 2 folders:

1. Infected
2. Uninfected

And a total of **27,558** image

- **Dimension 148 x 148**
- The dataset contains a total of 27,558 cell images with equal instances of parasitized and uninfected cells. An instance of how the patient-ID is encoded into the cell name is shown herewith: "P1" denotes the patient-ID for the cell labeled "C33P1thinF_IMG_20150619_114756a_cell_179.png". The data is a balanced data.
- I made data and labels list where data will be image to array implementation which contains RGB values of each image and label will be class of cells.
- I want to set up the data transformations for each set of data. In general, we want to have the same types of transformations on the validation and test sets of data. However, with the training data, we can create a more robust model by training it on rotated, flipped, and cropped images.

Reference Link:

- This data is from the Official NIH Website: https://ceb.nlm.nih.gov/repositories/malariadatasets/. The data is open – sourced and can be downloading for education purpose with no citation.

Solution

We will use keras and tensorflow frameworks for building our convolution neural networks. Adding convolution layers and adam as optimizer , I am going to identify the malaria cell.

Benchmark Model

- In this malaria cell detection, I want to set worst benchmark model by adding only one convolution Layer and Max pooling layer and dense layers with 2 units.
- I will try to improve the benchmark model accuracy by adding sufficient layers to the model.

Evaluation Metrics

As the data I am using is balanced, I want to use Accuracy as a evaluation metric which will be given by the (correct images/total number of images), and loss metric as 'binary_crossentropy' and adam optimizer since our model has two classes for both validation and testing set.

PROJECT DESIGN:

STEP 1: DATA PREPARATION

- I made data and labels list where data will be image to array implementation which contains RGB values of each image and label will be class of cells.
- With the training data, we can create a more robust model by training it on rotated, flipped, and cropped images. So, in the data preparation data is resized and flipped in different angles for the better performance of the model to get the best accuracy. The data is divided into training_ set and testing_set.

STEP 2:- ONE HOT ENCODING

I am using one hot encoding for each class in order to classify them into binary format.

STEP 3:- CREATING SEQUENTIAL MODEL

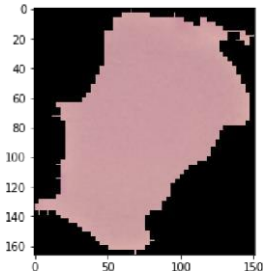
- For the Benchmark model, there will be only one convolution layer, max pooling layer and dense layer.
- To improve these benchmark model accuracy , I will add more layers with activation function 'softmax' since there is more than one class to classify and relu function to calculate the performance on various activation functions and 'dropout' to reduce the overfitting.

STEP 4:- VISUALIZATION OF RESULTS

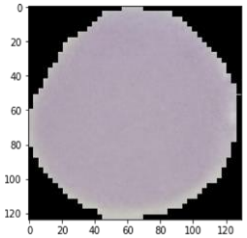
 The model is fitted, After the model is completed ,I will visualize the testing results.

Finally identifying the Image cell corresponds with class either Parasetized or Uninfected.

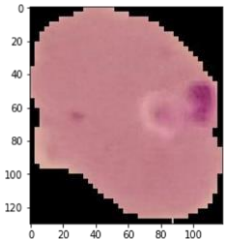
Sample Images



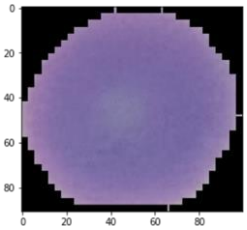
Uninfected



Uninfected



Parasitized



Uninfected

G. RAJESWARI,

MACHINE LEARNING NANO DEGREE,

CAPSTONE PROJECT PROPOSAL,

MALARIA CELL IMAGE DETECTION USING KERAS,

Saturday, February 16, 2019.