

Project Proposal

The Problem:

Naturally diagnosing diseases requires effort, time, and people specialized in the field, and the possibility of human error in diagnosis is very high. Therefore, the need for an automated system for diagnosis has increased. In recent years, a lot of research has been produced where it combines artificial intelligence and machine learning with the medical field. One of the most dangerous diseases world around is Malaria, which could infect humans with parasites that are transmitted through the bites of infected female Anopheles mosquitoes. where the estimated deaths in 2019 reached about 409000 cases [1]. Based on the purpose of early diagnosing of malaria this project will propose building a model using deep learning methods such as the convolution neural network (CNN) to classify the cells into, infected with malaria or unaffected. This project produces important information for those who are interested in diagnosing malaria using deep learning methods and for the medical community in general.

The Dataset:

The used dataset is microscopic cells images (27,558 cell images), labelled to parasitized(abnormal) and uninfected cells(normal), this dataset was collected by developing a mobile application that runs on a standard Android smartphone attached to a conventional light microscope. The dataset was proposed first in[2] and uploaded in Kaggle[3]. Figure 1 below shows a sample of the dataset. The dataset is a balanced dataset where the number of image samples of the parasitized class is the same as the number of uninfected class which equal 13779 images.

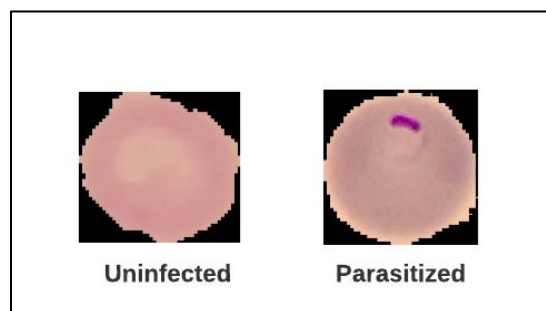


Figure 1:Dataset Sample

The Design:

This project will aim to diagnose malaria based on detecting the parasitized cells using the deep learning method (CNN). Figure2 below provide an overview of the main methodology used in this project where the first step is preprocessing the dataset by resizing the images to fit the model, after that dividing the dataset into a training set to train the model, validation set, and testing set to evaluate the model.

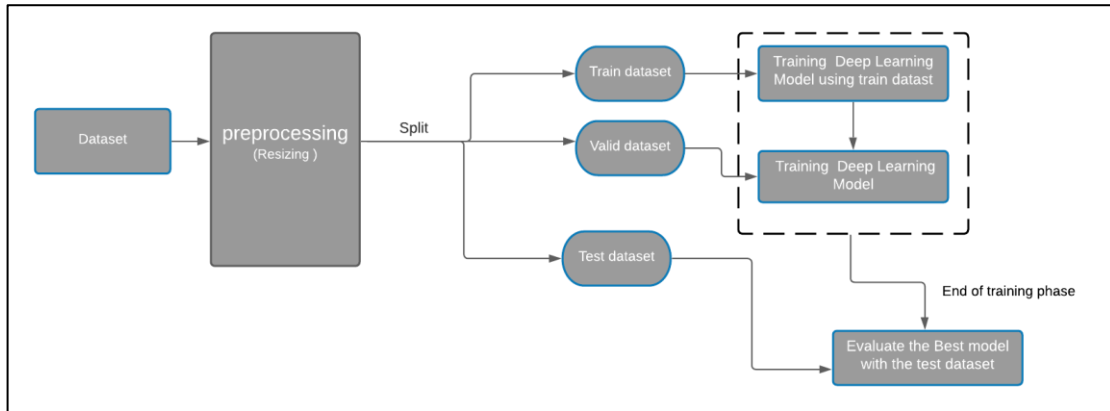


Figure2 :Project Design

The Model:

The used model is a convolutional neural network (CNN) using a transfer learning strategy, where the CNN model is pre-trained in bigger dataset such as (ImageNet) and used the learned weights to implement the classification in another dataset, in this case, the malaria dataset. Based on the fact the dataset is balanced so, the performance evaluation will be calculated by the accuracy.

Tools:

- Preprocessing:
OpenCV, TensorFlow, keras, sklearn, PIL, numpy
- The model:
keras.applications
- Visualization:
matplotlib

References:

- [1] "Malaria." <https://www.who.int/news-room/fact-sheets/detail/malaria> (accessed Nov. 27, 2021).
- [2] S. Rajaraman *et al.*, "Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images," *PeerJ*, vol. 2018, no. 4, 2018, doi: 10.7717/PEERJ.4568.
- [3] "Malaria Cell Images Dataset | Kaggle." <https://www.kaggle.com/iarunava/cell-images-for-detecting-malaria> (accessed Nov. 27, 2021).