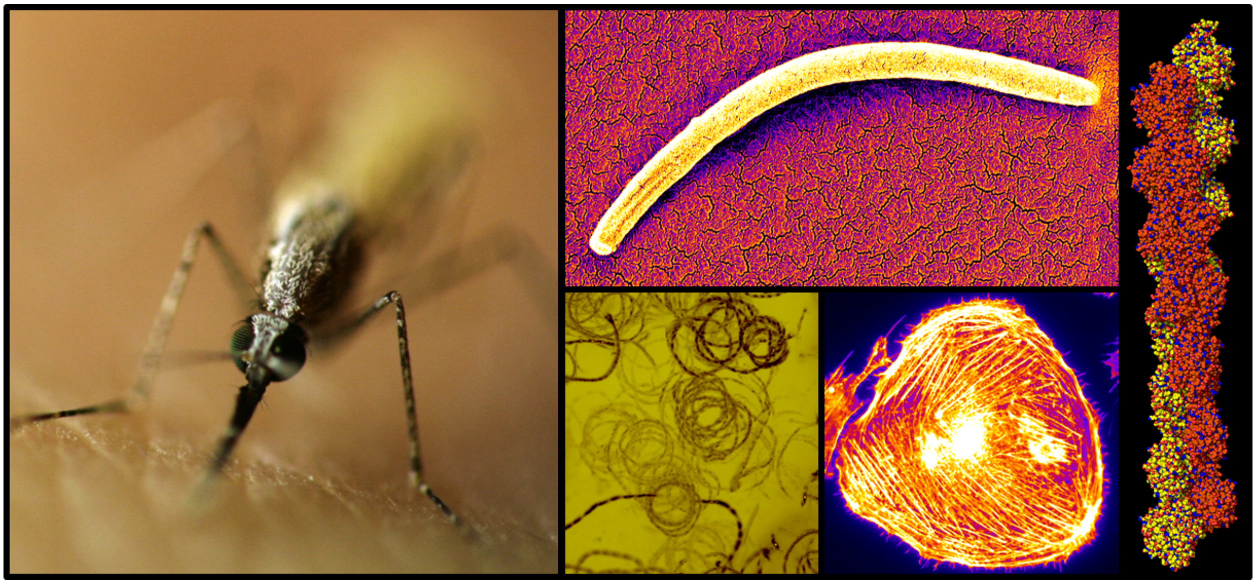


Detecting malaria using convolutional neural network

- ***Detecting Malaria with Deep Learning*:-**

Malaria is a deadly, infectious mosquito-borne disease caused by Plasmodium parasites. These parasites are transmitted by the bites of infected female Anopheles mosquitoes.
With regular manual diagnosis of blood smears, it is an intensive manual process requiring proper expertise in classifying and counting the parasitized and uninfected cells.
Deep Learning models, or to be more specific, Convolutional Neural Networks (CNNs) have proven to be really effective in a wide variety of computer vision tasks and image classification. Automated malaria detection using deep learning models like CNNs could be very effective, cheap and scalable especially with the advent of transfer learning and pre-trained models which work quite well even with constraints like less data.

Problem Statement :-

- Microscopic examination of blood is the best known method for diagnosis of malaria. A patient's blood is smeared on a glass slide and stained with a contrasting agent that facilitates identification of parasites within red blood cells.

- Based on the guidelines from the WHO protocol, this procedure involves intensive examination of the blood smear at a 100X magnification, where people manually count red blood cells that contain parasites out of 5000 cells.

- Thick blood smears assist in detecting the presence of parasites while thin blood smears assist in identifying the species of the parasite causing the infection (Centers for Disease Control and Prevention, 2012). The diagnostic accuracy heavily depends on human expertise and can be adversely impacted by the inter-observer variability and the liability imposed by large-scale diagnoses in disease-endemic/resource-constrained regions

- As you can imagine, manually counting 5,000 cells is a slow process. This can easily burden clinic staff, especially where outbreaks occur.

- Malaria detection is definitely an intensive manual process which can perhaps be automated using deep learning in which image classification forms the basis of my project.

Solution for clients

- My client can range from small pathology labs to public/private hospitals.

- With regular manual diagnosis of blood smears, it is an intensive manual process requiring proper expertise in classifying and counting the parasitized and uninfected cells. Typically this may not scale well and might cause problems if we do not have the right expertise in specific regions around the world. Some advancements have been made in leveraging state-of-the-art (SOTA) image processing and analysis techniques to extract hand-engineered features and build machine learning based classification models. However these models are not scalable with more data being available for training and given the fact that hand-engineered features take a lot of time.

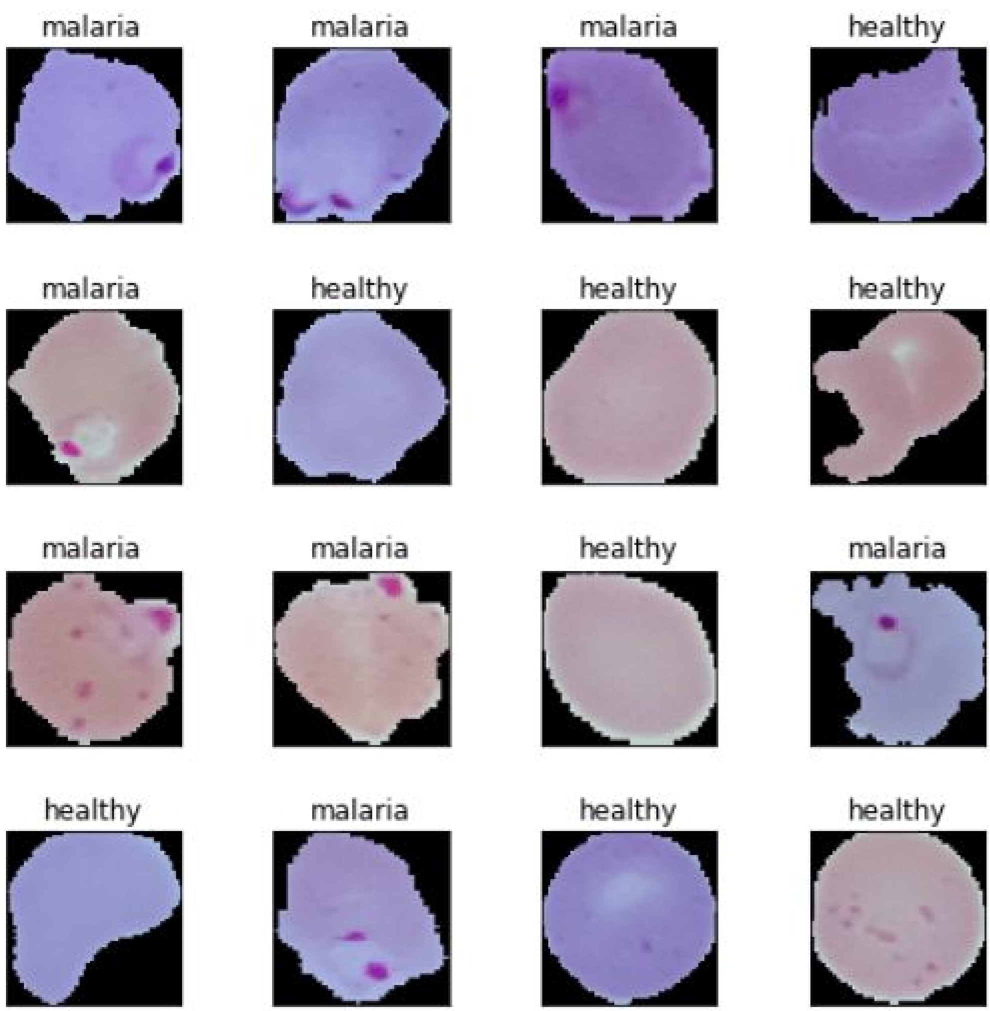
- Main aim to do this research is to reduce the time taken by prevailing methods of malaria detection which will prove to be a time and life saver in both normal and outbreak times.

Description of the dataset, how you obtained, cleaned, and wrangled it

- The dataset contains 2 folders - Infected - Uninfected and a total of 27,558 images. So there was not much cleaning required as such because the objective is train and test model on images which either contain parasite or not.
- Data for this project is gathered from kaggle and is based on an article from the website towards Data Science.

  https://towardsdatascience.com/detecting-malaria-with-deep-learning-9e45c1e34b60

  A Glimpse of the Data Set:

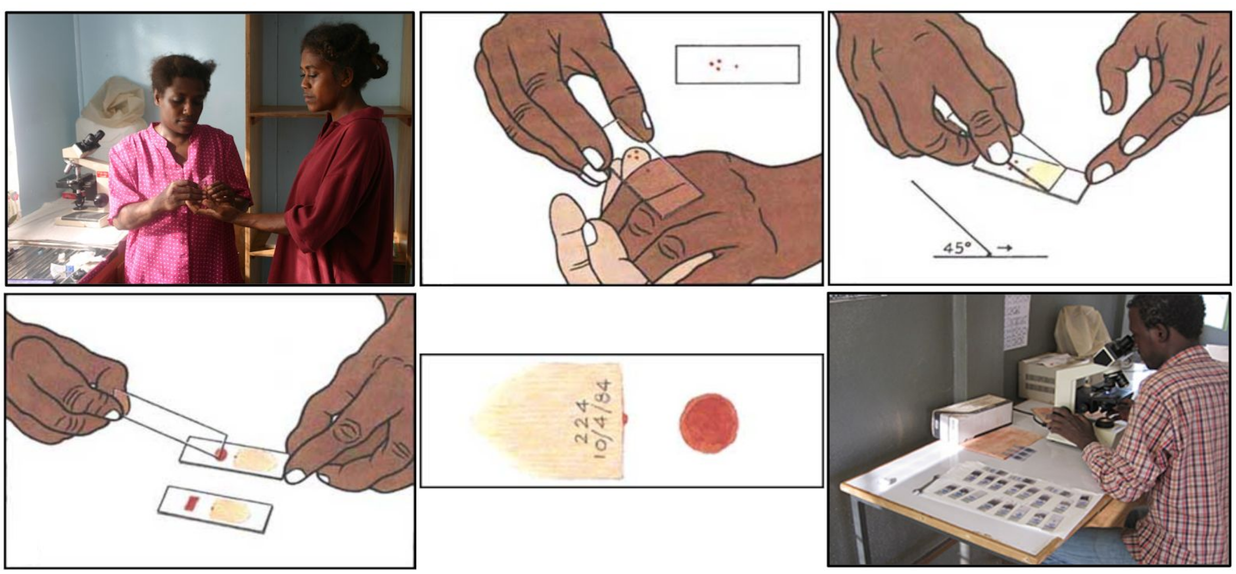


Summary of findings:-

- Methods for Malaria Detection

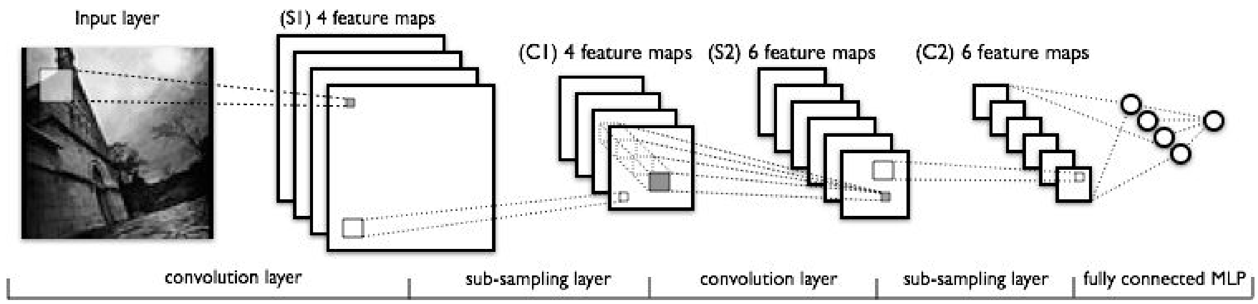
  There are several methods and tests which can be used for malaria detection and diagnosis. The original paper on which our data and analysis is based on, ‘ Pre-trained convolutional neural networks as feature extractors toward improved Malaria parasite detection in thin blood smear images’ by S Rajaraman et. al. introduces us briefly to some of these methods. These include but are not limited to, thick and thin blood smear examinations, polymerase chain reaction (PCR) and rapid diagnostic tests (RDT). While we won’t cover all the methods here in detail, an important point to remember is that the latter two tests are alternative methods typically used an an alternative particularly where good quality microscopy services cannot be readily provided.
  We will discuss briefly about a standard malaria diagnosis, based on a typical blood-smear workflow, thanks to this wonderful article by Carlos Ariza on Insight Data Science, which I got to know from Adrian Rosebrock’s excellent article on malaria detection on pyimagesearch, so my heartfelt thanks to both of them for such excellent resources, giving me more perspective in this domain.

  A blood smear workflow for Malaria detection (Source)
  Based on the guidelines from the WHO protocol, this procedure involves intensive examination of the blood smear at a 100X magnification, where people manually count red blood cells that contain parasites out of 5000 cells. In fact the paper by Rajaraman et. al. which we mentioned previously, talks about the exact same thing and I quote the following exerpt from the paper to make things clearer.



- By using Deep Learning models, or to be more specific, Convolutional Neural Networks (CNNs) have proven to be really effective in a wide variety of computer vision tasks.

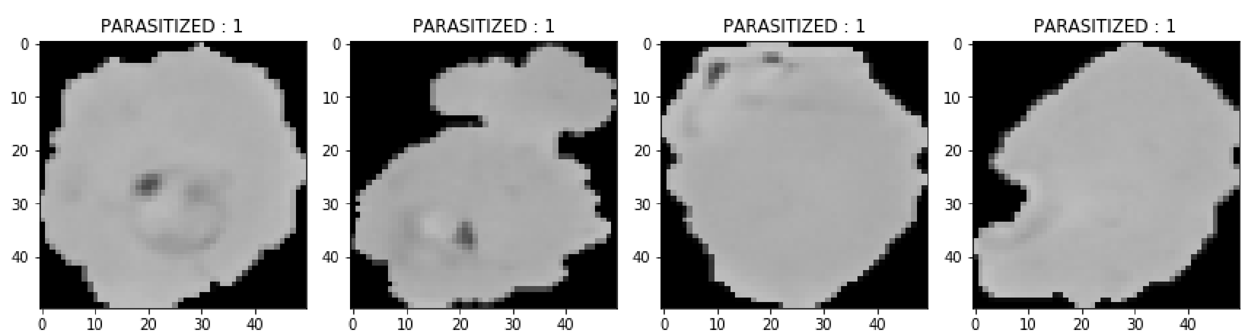
- The key layers in a CNN model include convolution and pooling layers as depicted in the following figure.



- Convolution layers learn spatial hierarchical patterns from the data, which are also translation invariant. Thus they are able to learn different aspects of images. For example, the first convolution layer will learn small and local patterns such as edges and corners, a second convolution layer will learn larger patterns based on the features from the first layers, and so on. This allows CNNs to automate feature engineering and learn effective features which generalize well on new data points. Pooling layers help with downsampling and dimension reduction.
- Thus, CNNs help us with automated and scalable feature engineering. Also, plugging in dense layers at the end of our model enables us to perform tasks like image classification.

VISUALS TO SUPPORT THE FINDINGS

Parasitized Cells -



Uninfected Cells:-