



Malaria Infected Cell Detection

Anubhav Shrimal (MT18033), Vrutti Patel (MT18020)

Advisor: Dr. Richa Singh

Statistical Machine Learning (CSE 542)

ABSTRACT

In this project we aim to build and compare various machine learning models which classify a given cell image as uninfected or infected by malaria parasite. Our aim is to classify images using various methodologies such as Bayesian classifier, PCA and LDA to reduce dimensionality, CNN, and Pre-trained CNN and compare their performance on different evaluation metric to give detailed analysis on which method works better and why.

USE CASE OF THE PROBLEM

- Hundreds of thousands of people die every year due to malaria majorly in the underdeveloped or developing countries due to delayed diagnostics and unavailability of specialized doctors in this field. The goal of our project is to automate the detection of malaria parasite in a given blood sample image accurately.
- It is a challenge in Computer Vision & Machine Learning to handle sensitive cases like detecting cancerous cell and classifying whether a person is suffering from a disease or not. This project is a good example of solving such issues and can be extended to other diseases as well.

DATASET

- The dataset consists of 27,558 cell images; 13,780 images of infected and uninfected cells each and is taken from the official NIH Website.
- Link: <https://ceb.nlm.nih.gov/repositories/malaria-datasets/>

EVALUATION METRIC

- Confusion Matrix
- Receiver Operating Characteristic (ROC)
- Cumulative Matching Characteristic (CMC)
- Detection Error Trade-off (DET)
- Accuracy, Precision and Recall

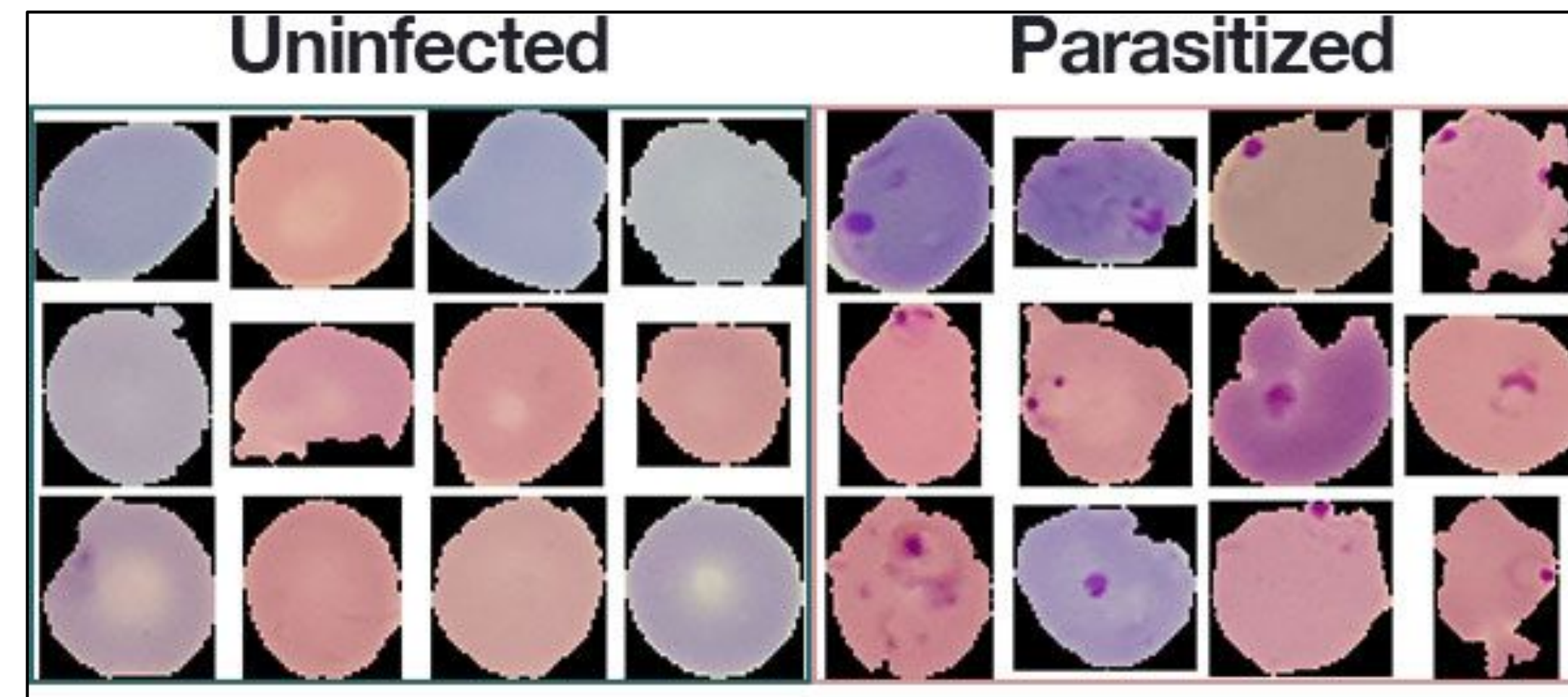


Figure 1. Dataset of uninfected and Malaria infected cells. (Source: Google Images)

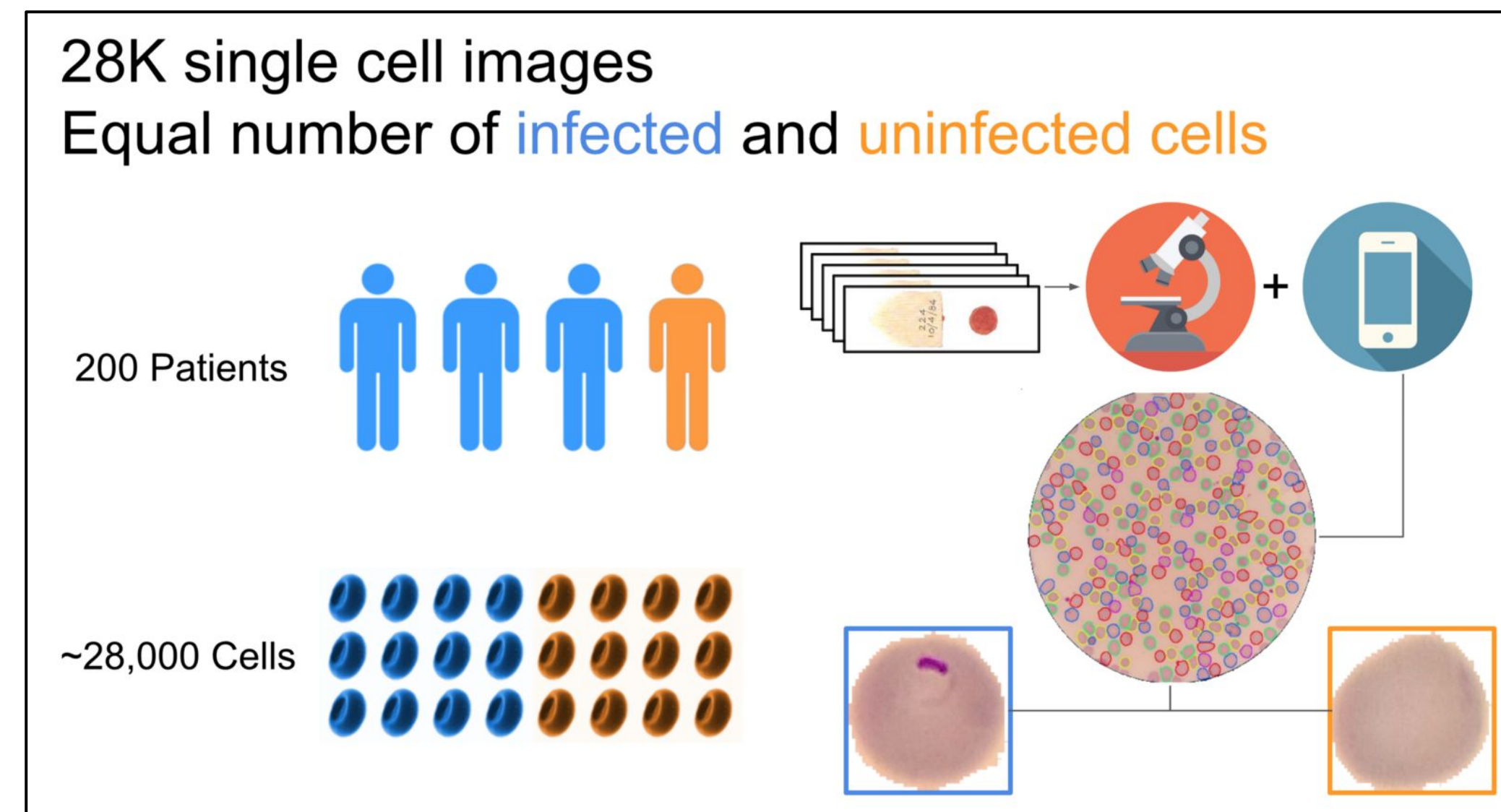
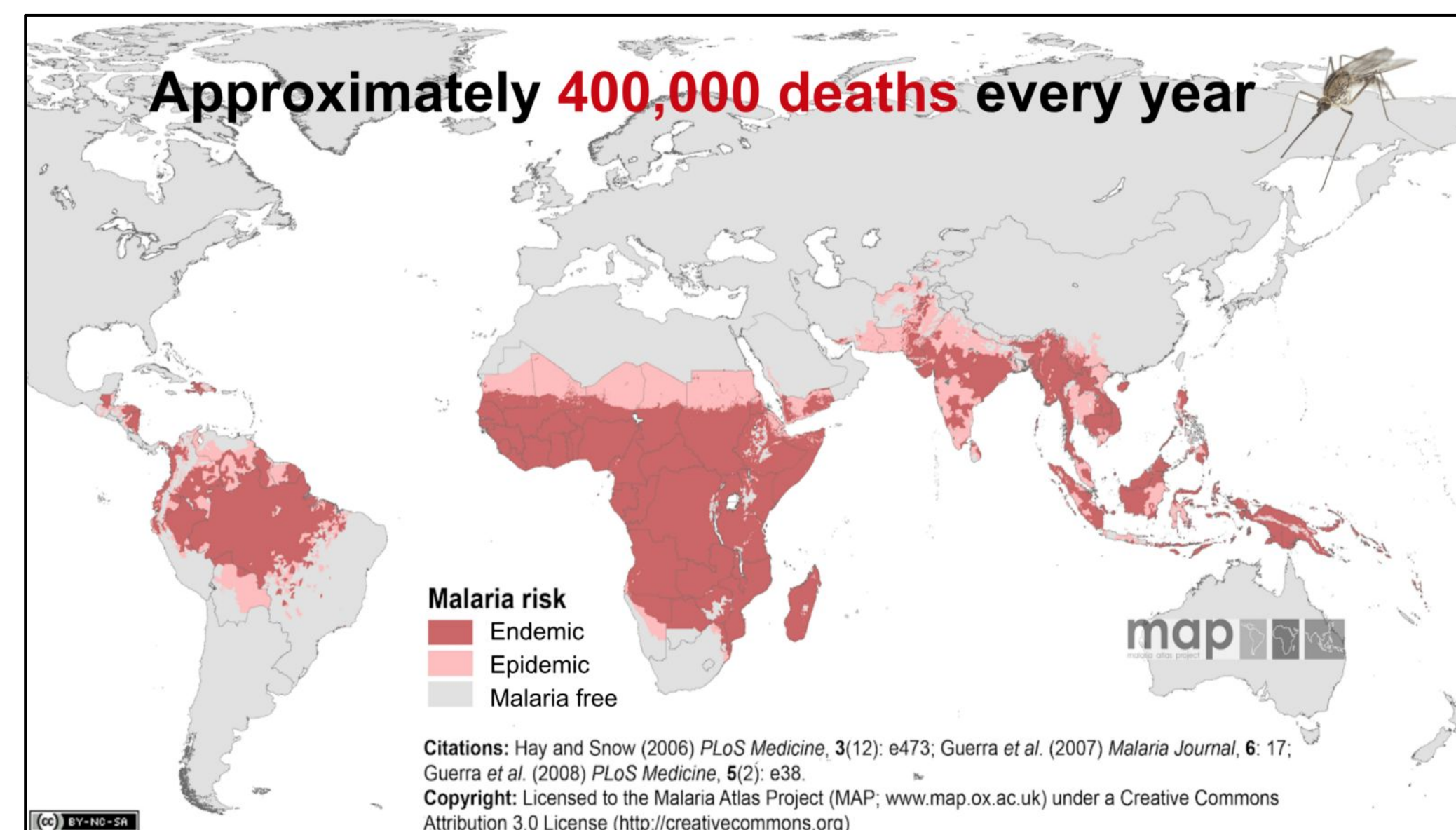


Figure 2. Dataset description. (Source: Google Images)

LITERATURE REVIEW

The authors of the paper [1] show many different methods for classification which could be applied for malaria detection based on the thick or thin smear. In this paper [2] the authors give an in depth review of the how malaria detection can be done using machine learning and clever image pre-processing. In the paper [3] the authors describe techniques to morph the cell images to have better boundary detection and then compared different techniques such as Naive bayes and neural networks. The authors [4] show the effects of image transformations such as segmentation to get better results.

PROJECT TASK & MILESTONES

Mid-term Evaluation:

- Visualizing & understanding the dataset.
- Bayesian classifier.
- Convolutional Neural Network based classifier.

End-Sem Evaluation:

- Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) for dimensionality reduction.
- Extending CNN with Pre-trained deeper models such as ResNet34 & ResNet50.
- Comparing performance of models by evaluating it across different measures such as ROC, CMC, DET, Confusion Matrix, Accuracy, Precision and Recall.

REFERENCES

1. Poostchi M, Silamut K, Maude R, Jaeger S, Thoma G (2018) "Image analysis and machine learning for detecting malaria" Transl Res 194
2. Jan Z, Khan A, Sajjad M, Muhammad K, Rho S, Mehmood I (2017) "A review on automated diagnosis of malaria parasite in microscopic blood smears images" Multimedia Tools Appl 77:1–26
3. Das DK, Maiti AK, Chakraborty C (2015) "Automated system for characterization and classification of malaria-infected stages using light microscopic images of thin blood smears" J Microsc 257(3):238–252
4. Suryawanshi MS, Dixit V (2013) "Improved technique for detection of malaria parasites within the blood cell images" Int J Sci Eng Res 4:373–375