KANTIPUR ENGINEERING COLLEGE

(Affiliated to Tribhuvan University)

Dhapakhel, Lalitpur

[Subject Code: CT755]

A MAJOR PROJECT MID-TERM REPORT ON

MALARIA DETECTION

Submitted by:

AABISKAR SHRESTHA	**[089/BCT/2073]**
HEMANT SHRESTHA	**[105/BCT/2073]**
KISHAN THAPA	**[107/BCT/2073]**
MUKESH RAJBANSHI	**[112/BCT/2073]**

A MAJOR PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR IN COMPUTER ENGINEERING

Submitted to:

Department of Computer and Electronics Engineering

December, 2019

MALARIA DETECTION

Submitted by:

AABISKAR SHRESTHA [089/BCT/2073]

HEMANT SHRESTHA [105/BCT/2073]

KISHAN THAPA [107/BCT/2073]

MUKESH RAJBANSHI [112/BCT/2073]

A MAJOR PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR IN COMPUTER ENGINEERING

Submitted to:

Department of Computer and Electronics Engineering

Kantipur Engineering College

Dhapakhel, Lalitpur

December, 2019

CHAPTER 1
ABSTRACT

Malaria management is a challenging problem all over the globe particularly in Asian and African continents. Presently, even 110 years after the Nobel Prize of Ronald Ross for his work on malaria, people in the European region are also at risk from diseases carried by vectors both within the region and when traveling abroad. While treatment of malaria itself is a challenging problem its quick detection is also a problem with no less significance. There are mainly four species of malaria parasites infecting human beings namely, Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale and Plasmodium malaria. Plasmodium vivax, is found mainly in tropical and subtropical areas and has a severe clinical manifestation. Rapid detection of presence of the parasite in human blood and early institution of antimalarial drugs are the mainstay of management of the disease. WHO recommends that all cases of suspected malaria be confirmed using parasite-based diagnostic testing (either microscopy or rapid diagnostic test) before administering treatment . In the malaria detection test, microscopy based diagnosis has the central importance for species differentiation, parasite quantification, management of severe disease. Additionally, the method may be amenable to larger section of society because of its scalability and low running cost.

Two types of blood smears, thick and thin, are prepared from the blood of patients, who are clinically suspected to be suffering from malaria. The thick smear is more useful for parasite detection whereas the thin smear is particularly used for identification of malaria species. When the parasite load is low, malaria may be detected about 20 times more rapidly in thick smear than in thin smear. The methods based on which image analysis of blood smears have been made, broadly fall into two classes namely, analysis based on morphology and that based on colour. Some of the reported methods use supervised training sets and lack of availability of appropriate training sets may delimit the scalability of such methods.

Keywords: **Color based pixel discrimination**, **Morphological operation**, **Dilation**, **Erosion**, **Gray Scale Image**, **thin blood film smears**

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 2
INTRODUCTION

2.1 Background

Malarial disease is one of the most and dynamically occurring challenging problem all over the world mainly in Asian and African continents. Presently, even 110 years after the Nobel Prize of Ronal Ross for his work on malaria, people in European region are also at risk from these diseases. And mostly 1 million of the people are being dead in Africa due to malaria [1]. As it is all known to us that treatment of malarial is itself a challenging task its quick detection is also a problem with more obstacles. There are mainly four species of malaria parasites infecting human beings namely, plasmodium falciparum, plasmodium vivax, plasmodium ovale and plasmodium malaria [2]. Plasmodium vivax, is found mainly in tropical and subtropical areas and has several clinical demonstrations. Rapid detection of presence of Malaria parasites in human blood and early institute of Antimalarial drugs are the backbone of management of the disease. WHO recommends that all cases of suspected malaria be confirmed using parasite-based diagnostic testing before managing treatment. In the malaria detection test, microscopy based diagnosis has the major importance for species differentiation, parasite quantification, management of serious or critical disease. Additionally, the method may be tractable to large section of the society because of its scalability and low running cost. Mainly two types of blood smears are there as, thick and thin, are prepared from the blood of patients, those who are clinically reckon to be suffering from malaria. The thick smear blood is more useful for parasite detection whereas the thin smear is practically used for identification of malaria species. When the parasite load is low, malaria may be detected about 20 times more rapidly in thick smear then in thin smear. The methods on which image analysis of blood smears have been construct, especially fall into two types namely, analysis based on morphology and that based on color. Some of the reported methods use direct training sets and lack of availability of appropriate training sets may determine the scalability of such methods.

2.2 Problem Statement

Thick blood smear examination is a necessary part for rapid screening of malaria parasite. Primary diagnosis of malaria by thick smear examination is cheap and highly sensitive, advocated by the World Health Organization (WHO).For the examination of thick blood smear, manpower and time can be reduced by using automated computational techniques.

To provide a solution to the above problem we design a Image processing system using Matlab which processes the image to obtain a result whether a person having malaria or not.

2.3 Objectives

The objective of the project is to develop a fully automated image classification system to positively detect the parasites of malaria using thick blood smears identify malaria parasites present in thin blood smears, and differentiate the species.

2.4 Applications

1. The image based approach enabled by computer technology that helps in diagnosis of digital slide.
2. To obtain faster and efficient results when there is lack of experts.

2.5 Project Features

1. Detection of malaria
2. Classification of species

2.6 Feasibility Analysis

A highly brief summary of the report and results are given below.

2.6.1 Technical Feasibility

To develop such an application, the following technical skills and tools are required:

- MATLAB
-
-

All of the above technical requirements are fulfilled. The project is technically feasible.

2.6.2 Economic Feasibility

Considering that this is an under-grad project, ROI and profit in terms of money is not expected. However, we must and did take into consideration whether developing the product itself was within budget. The cost of libraries, software tools and other components was found to be within budget. The project is indeed economically feasible.

2.6.3 Operational Feasibility

The product has high applicability among a variety of users. Being lightweight and easy to use, a wide user base is expected.The operational environment specification is low. Simply stated, this means that the software will run on a lot of devices. Choosing a higher specifications level would have meant a richer, more efficient feature set, but that would reduce the possible user base by 40

2.7 System Requirement

2.7.1 Software Requirement

1. Operating System : Windows XP or Higher
2. Software : Matlab 2013a

2.7.2 Hardware Requirement

1. Processor: Pentium 4 +
2. RAM: 2GB
3. Hard Disk: 10GB
4. Speed: 1.2 GHz+

CHAPTER 3
LITERATURE REVIEW

Kazarine, Alexei et.al. suggested a approach called third harmonic generation image scanning cytometry combining with third-harmonic generation imaging, high-speed motorized scanning, and automated software processing. It is a non-linear process that uses the selectivity of this contrast mechanism to perform label-free image scanning cytometry of patient blood smears for automated malaria detection [3]. Kishor Roy et.al. using colour and segmentation based algorithms i.e. HSV segmentation and watershed segmentation. On using 100 real samples and 200 internet samples, watershed segmentation detected 85 and 164 respectively whereas HSV segmentation detected 70 and 160 respectively but the detection increased when both segments were mapped [4].Diaz in his research determines a color segmentation method for division of pixels into three distinct classes: red blood cell, parasite and background, on the basis of standard supervised classification algorithms[5]. The paper by Silvia Halim, designed a system for estimating parasitemia. Template matching approach is used for RBCs detection. Detection of parasites are done by means of the variance based system from grayscale images and next approach is based on color co-occurrence matrix which is on the basis of the individual color index of pixel and color indices of its eight neighboring pixel[6]. Deepa. A. Kurer et.al. designs a new approach for low-level image processing – SUSAN (Smallest Unvalued segment assimilating nucleus) Principle, that performs Edge and Corner detection. Image features depend on the geometry of the cells, texture and color and generation of parasites takes place, also features which uses a priori knowledge of the classification problem and mimic features used by human technicians[7]. Somasekar proposed a linear programming based Image segmentation and morphological operations for detection of malarial parasites. Two applications are presented: formulation of a linear programming based on the given data and solving given problem using graphical method approach for detecting parasites. But there was no complete system for detecting different parasites [8].

CHAPTER 4
METHODOLOGY

4.1 Main Algorithm

Here the proposed system includes two algorithms

1. Morphological Operation.
2. Colour based discrimination.

4.1.1 Algorithm1: Morphological Operation

MODULE

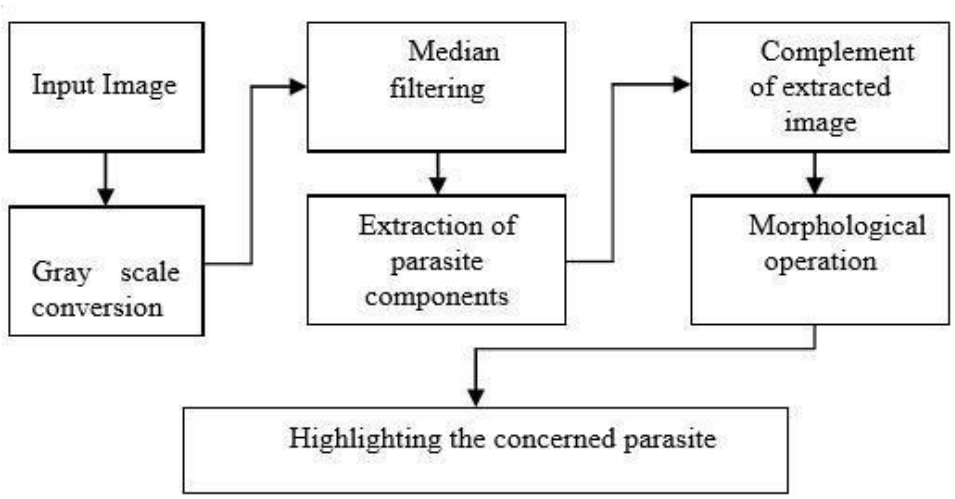

Figure 4.1: Module

Morphology is a broad set of image processing operations that process images based on shapes. Morphological operations apply a structuring element to an input image, creating an output image of the same size. In a morphological operation, the value of each pixel in the output image is based on a comparison of the corresponding pixel in the input image with its neighbors. By choosing the size and shape of the neighborhood, you can construct a morphological operation that is sensitive to specific shapes in the

input image.

The most basic morphological operations are dilation and erosion. Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries. The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. In the morphological dilation and erosion operations, the state of any given pixel in the output image is determined by applying a rule to the corresponding pixel and its neighbors in the input image. The rule used to process the pixels defines the operation as a dilation or an erosion.

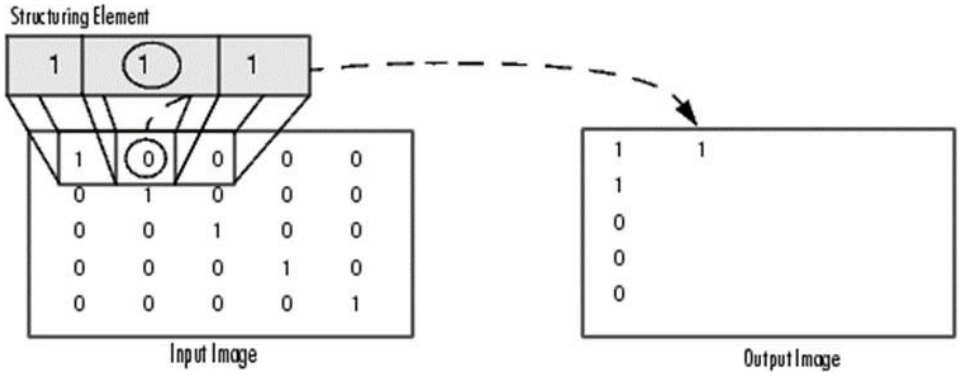

Figure 4.2: Morphological Operation

The above figure illustrates this processing for a grayscale image. The figure shows the processing of a particular pixel in the input image. Note how the function applies the rule to the input pixel's neighbourhood and uses the highest value of all the pixels in the neighbourhood as the value of the corresponding pixel in the output image.

Pre-Processing of Image

Pre-processing of image consists of converting the original image into grayscale image that can be done by using an inbuilt function rgb2gray(). RGB to gray conversion is done by averaging all the three components i.e. R, G and B which results in gray scale.

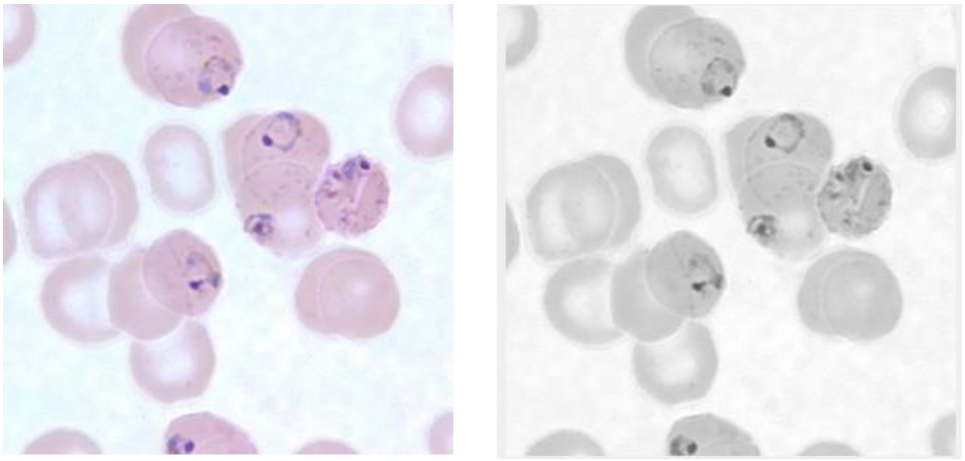

Figure 4.3: RGB to grayscale conversion

After the conversion of grayscale image we apply inbuilt function edge() that detect edges using the Canny method. The Canny method finds edges by looking for local maxima of the gradient of I. The edge function calculates the gradient using the derivative of a Gaussian filter. This method uses two thresholds to detect strong and weak edges, including weak edges in the output if they are connected to strong edges. By using two thresholds, the Canny method is less likely than the other methods to be fooled by noise, and more likely to detect true weak edges.

Segmentation

Segmentation process starts with the **strel()** function which takes the pre-processed image as an input. A strel object represents a flat morphological structuring element, which is an essential part of morphological dilation and erosion operations.
SE = strel('disk',R,N) creates a disk-shaped structuring element, where R specifies the radius. N specifies the number of line structuring elements used to approximate the disk shape. Morphological operations using disk approximations run much faster when the structuring element uses approximations. After applying strel function, the output of this function is taken as input for **imdialte()** function. **IM2 = imdilate(IM,SE)** dilates the grayscale, binary, or packed binary image IM, returning the dilated image, IM2. The argument SE is a structuring element object, or array of structuring element objects, returned by the strel or offset strel function.

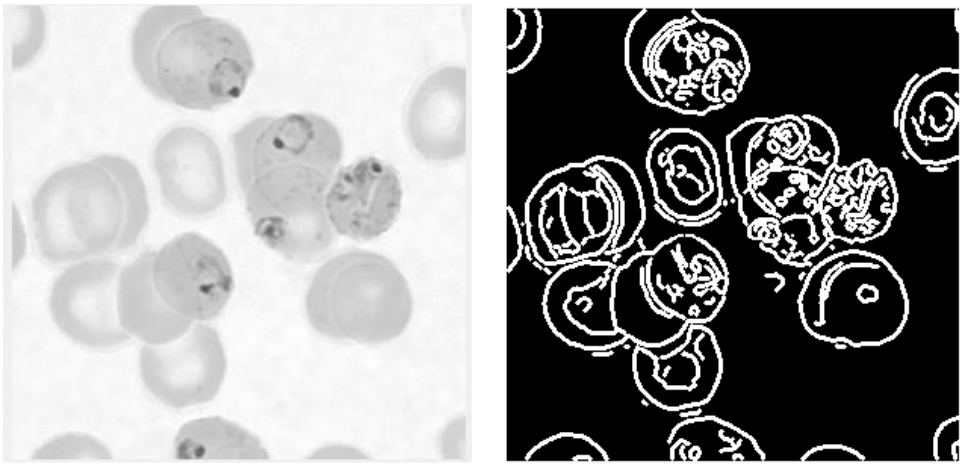

Figure 4.4: Conversion of grayscale to dilated image

Next, **BW2= imfill(BW,'holes')** fills holes in the input binary image BW using imfill(). In this syntax, a hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image.

Figure 4.5: Filling the holes

Next, **BW2 = bwareaopen(BW,P)** removes all connected components (objects) that have fewer than P pixels from the binary image BW, producing another binary image, BW2. This operation is known as an area opening.

Classification

After applying bwareopen we apply regionprops, graycomatrix and graycoprops to measure properties of image regions and apply red rectangular box for each detected region using bounding box.

glcms = graycomatrix(I,Name,Value,...) returns one or more gray-level co-occurrence matrices, depending on the values of the optional name/value pairs. Parameter names can be abbreviated, and case does not matter. **stats = graycoprops(glcm,properties)** calculates the statistics specified in properties from the gray-level co-occurrence matrix glcm. glcm is an m-by-n-by-p array of valid gray-level co-occurrence matrices. If glcm is an array of GLCMs, stats is an array of statistics for each glcm.

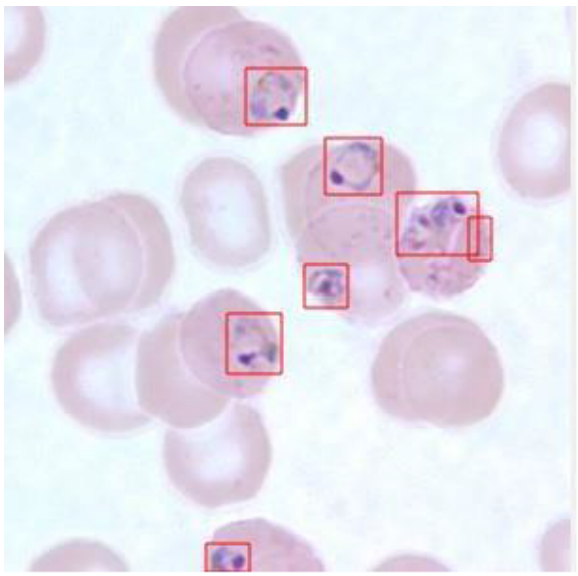

Figure 4.6: Detection of malaria infected cells

4.1.2 Algorithm2: Colour Based Discrimination

Colour based discrimination is finding out malaria cells using difference between each pixels value and finding out the accuracy of the using mathematic classifier like SVM. This process can be explained by the series of process.

- Segmentation
- Water segmentation
- Finding number of RCBs cells in a image
- Intensity adjustment and identification of infected malarial cells

Segmentation

When we ”segment” an image, we distinguish the regions of interest (ROIs) from the non-ROI portion, generally creating a binary mask of what we want to qualify, quantify, track, etc. This segmentation done by an app called image segmenter which is inbuilt in matlab and able to export code.

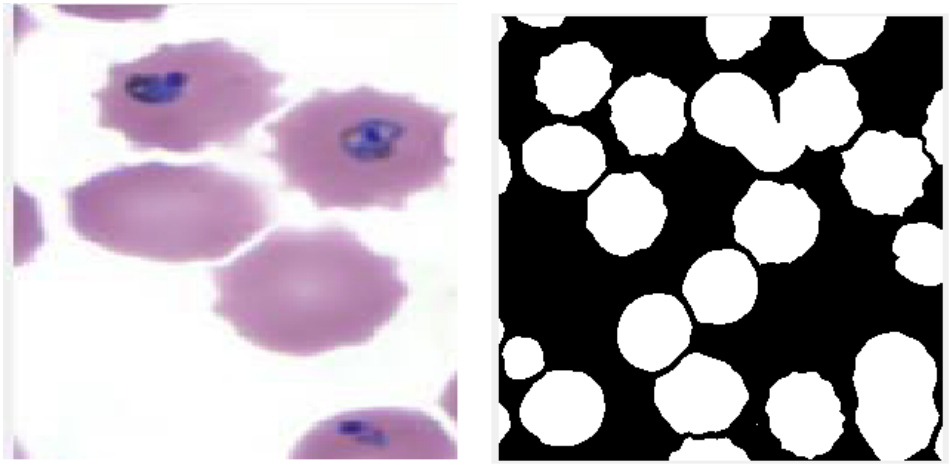

Figure 4.7: segmentation of cells from original image

You can improve the results by using in built function bwareaopen() on the segmentation image.

BW2 = bwareaopen(BW,P) removes all connected components (objects) that have fewer than P pixels from the binary image BW, producing another binary image, BW2. This operation is known as an area opening. But this results of segmentation is not

enough so we go for water segmentation.

Water Segmentation

A watershed is a transformation defined on a grayscale image. The name refers metaphorically to a geological watershed, or drainage divide, which separates adjacent drainage basins. The watershed transformation treats the image it operates upon like a topographic map, with the brightness of each point representing its height, and finds the lines that run along the tops of ridges.

To apply water segmentation on a grayscale image, we use inbuilt function called **watershed()**.

By applying this function we get a results as fig 5.7, but this image has lots of watershed line which is over segmenting the cells. To get a watershed line in between the each cells for proper segmentation we need to further process it using IMHMIN.

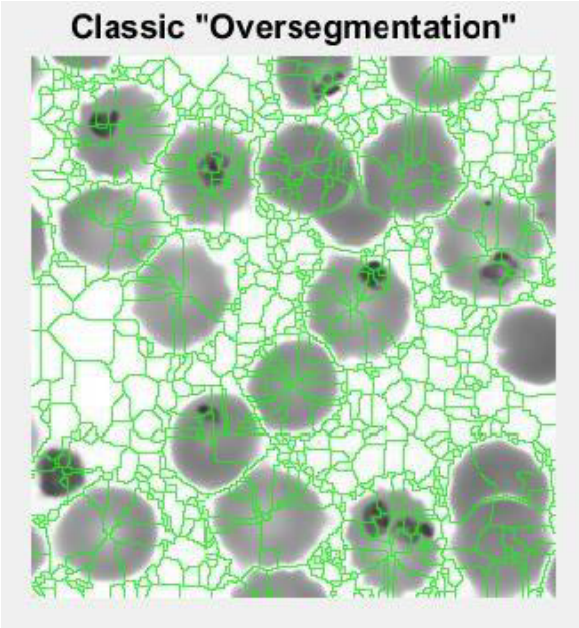

Figure 4.8: Over segmentation of watershed

Improving that result, **I2 = imhmin(I,h)** suppresses all minima in the intensity image I whose depth is less than h, where h is a scalar. Regional minima are connected components of pixels with a constant intensity value, t, whose external boundary pixels all have a value greater than t. After applying imhmin function we get a suppressed grayscale image.

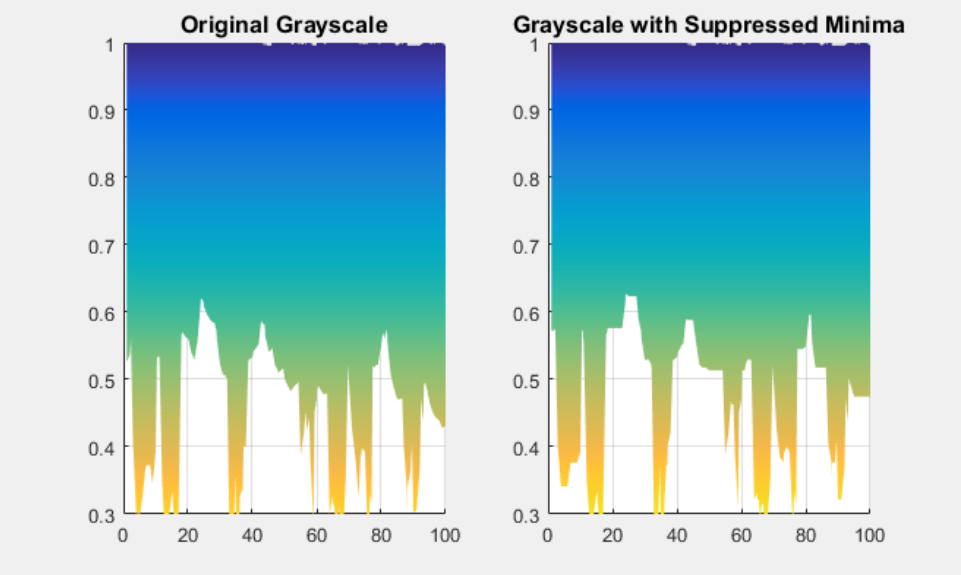

Figure 4.9: Suppressed minima in the intensity of grayscale image

After applying suppressed minima on a grayscale image if we apply watershed we get a better segmentation of cells using water segmentation.

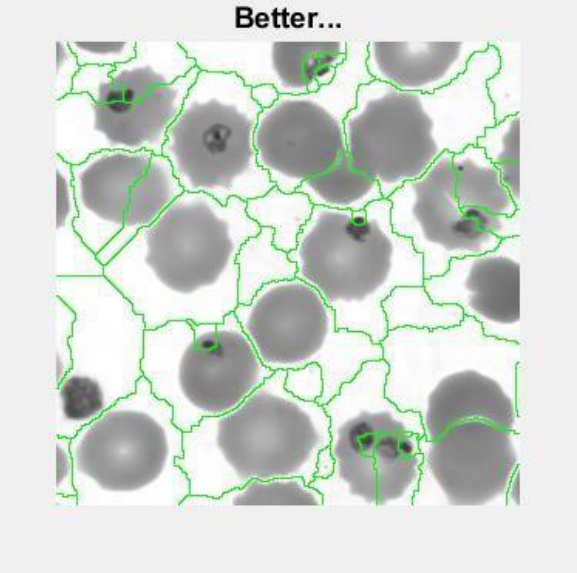

Figure 4.10: Better segmentation of wartershed

Finding number of RBCs cells in a image

To find the RBC cells in a image we need to start with how questioning how many circles are there in a cells. So by finding out how many circles are there in a image we can determine the number of cells on a image. There are a lot of technique to find

a circle like imfindcircles-Find circles using circular Hough transform inbuilt function. But in our project we are using an app called circleFinder.

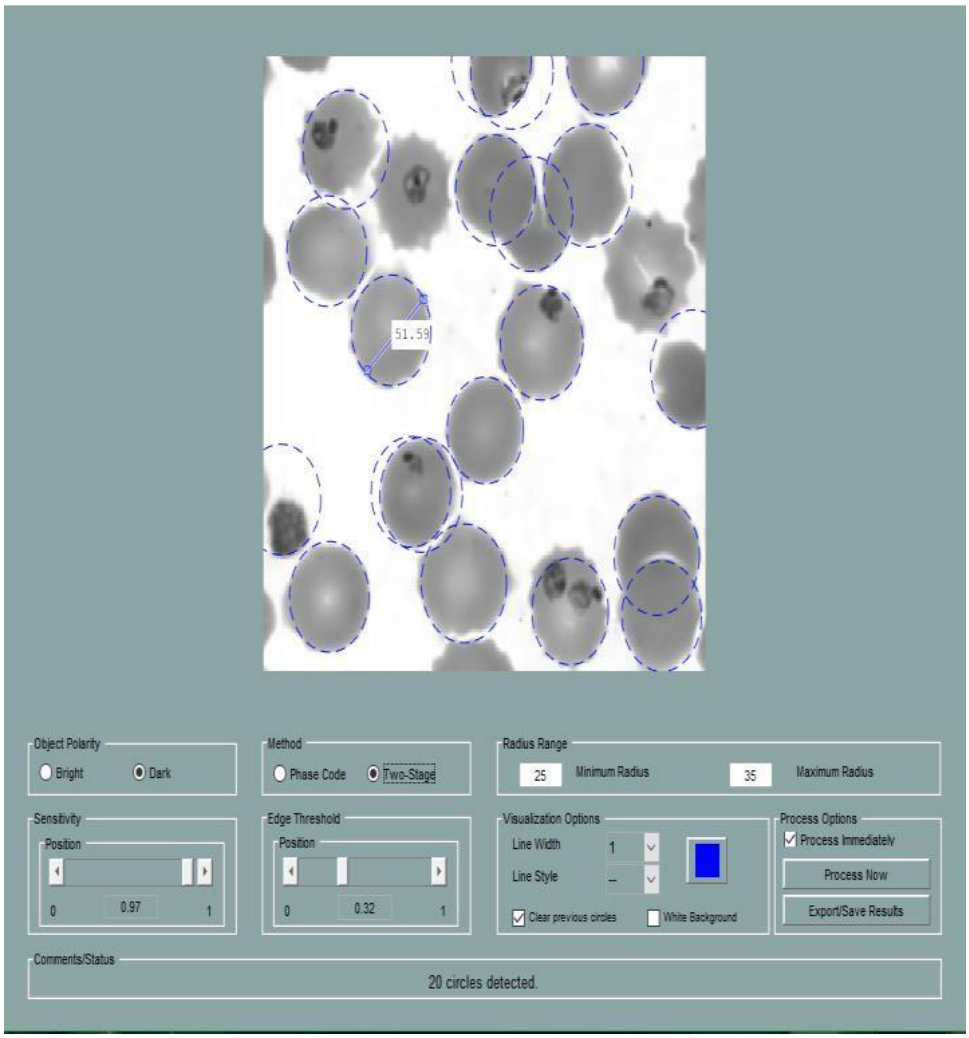

Figure 4.11: CircleFinder App in matlab

Which can be setup as

- Object polarity which in this case is a dark i.e. dark cells on bright background.
- Radius range by measuring
- Sensitivity and method to detect a cells on a image as shown above.

After exporting a results from a circleFinder app from single image you use the same results to find the cells on all the images.

The results generated by cirlceFinder app :

detectCircles = @(x) imfindcircles(x,[20 35], ... 'Sensitivity',0.89, ...

'EdgeThreshold',0.04, ...

'Method','TwoStage', ... 'ObjectPolarity','Dark');

[centers, radii, metric] = detectCircles(grayscale);

When we use the exported value from cirlceFinder app in our algorithm to find the cells in a image, We get Cells.

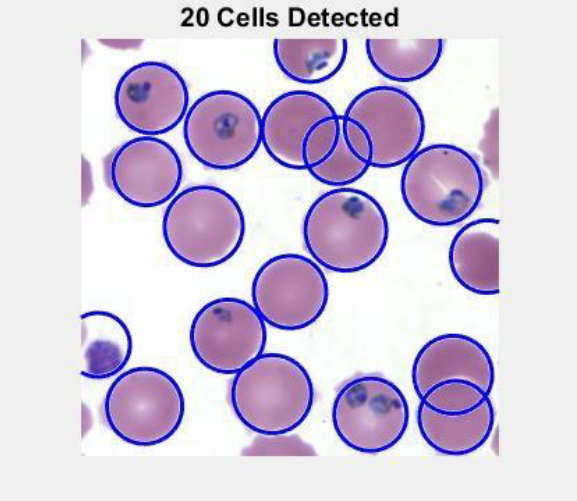

Figure 4.12: Cells Detection

Intensity adjustment and Identification of infected malarial cells

Before identification of infected cells on a image, we need to adjust the intensity of a image because there will be heterogeneous data that can have varying in intensity. Adjusting a image intensity helps in identifying a cells very fast and effective because lots of infected images will have darker regions which has high threshold value.

Adjusting a intensity is done by inbuilt function called imhistmatch-Adjust image to match its histogram to that of another image.

b = imhistmatch(A, REF) transforms the input grayscale or TrueColor image A so that the histogram of the output image B approximately matches the histogram of the reference image REF, when the same number of bins are used for both histograms. For TrueColor images, each color channel of A is matched independently to the corresponding color channel of REF.

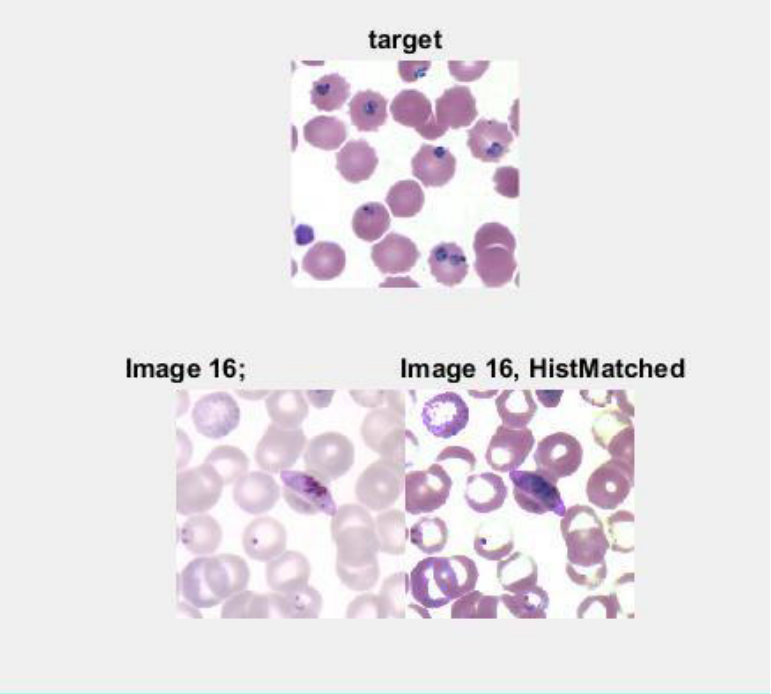

Figure 4.13: Intensity Adjustment

By adjusting intensity we can able to find the threshold value of infected cells by finding out the pixel value of the infected pixel cells. This is done by matlab app called imtool which helps in identifying each pixel value in the image.

Looking at the fig u can tell that the infected cells have higher pixel value and other region in which we have no interest has lower pixel value from that we decide the threshold for identification of infected cells in a image.

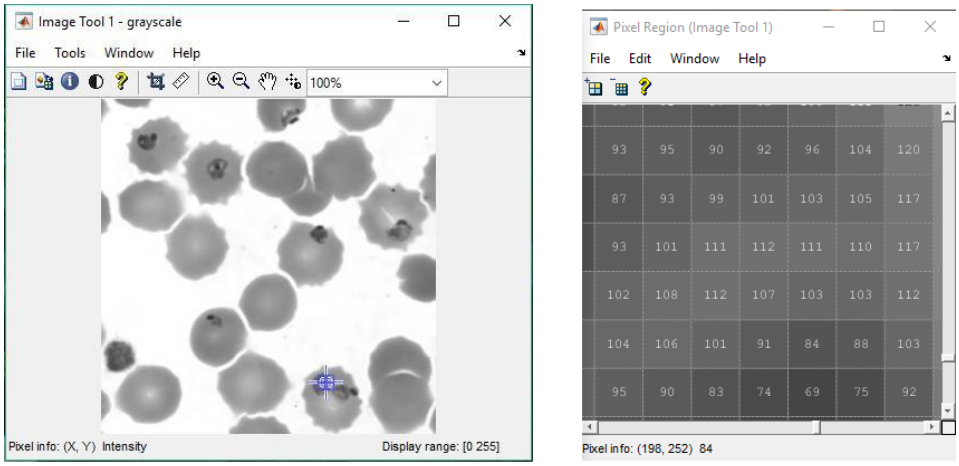

Figure 4.14: Threshold for infection using imtool

After applying every step explained as above we can able to identify a infected cell in a

image using this algorithm.

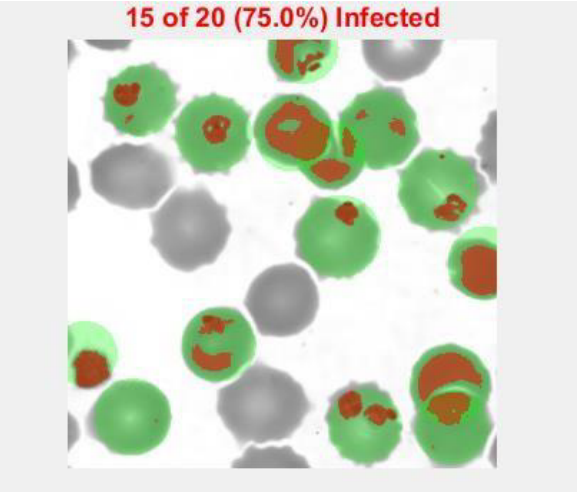

Figure 4.15: Identification of infected cells

4.2 System Design

Systems design is the process of defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements. System design consists of use case diagrams, data flow diagrams and sequence diagrams.

4.2.1 Proposed Architecture

Architecture Diagram is a graphical representation of the concepts, their principles, elements and components that are part of an architecture. For system developers, they need system architecture diagrams to understand, clarify, and communicate ideas about the system structure and the user requirements that the system must support. It's a basic framework can be used at the system planning phase helping partners understand the architecture, discuss changes, and communicate intentions clearly.

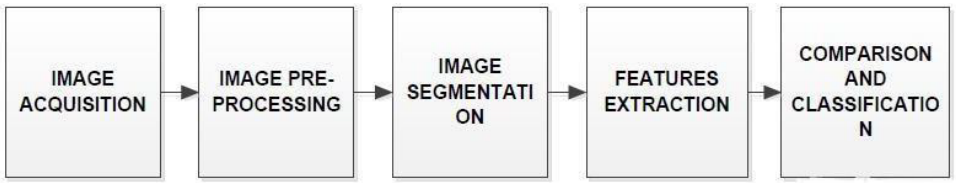

Figure 4.16: System Architecture

As it can be seen, the image received from user is processed followed by segmentation and feature extraction and finally the results are compared and classified.

4.2.2 Usecase Diagram

A use case diagram is a graphical representation of a user's interaction with the system and depicting the specifications of a use case and the relationship between the user and the different use cases in which the user is involved. A use case diagram can portray the different types of users of a system and the various ways that they interact with the system. This type of diagram is typically used in conjunction with the textual use case and will often be accompanied by other types of diagrams as well.

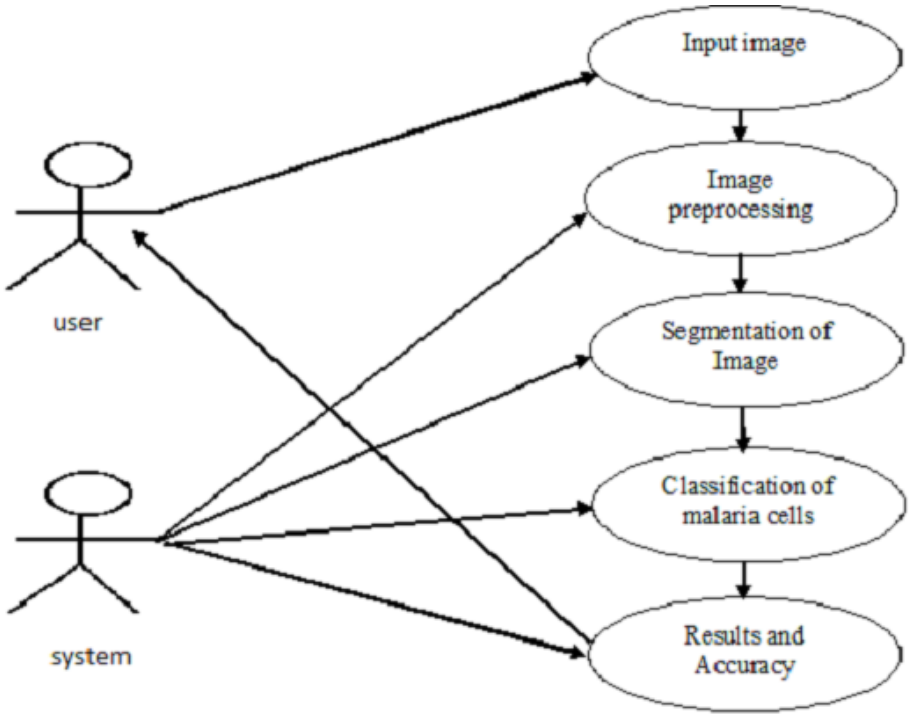

Figure 4.17: Use Case Diagram for Proposed System

In this system, sequence diagram there are two actors i.e. the user and system where the image is accepted by the system and it is further preprocessed, segmented and classified to give the results.

4.2.3 Sequence Diagram

A sequence diagram is an interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario.

Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams, event scenarios. A sequence diagram shows, as parallel vertical lines (lifelines), different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them, in the order in which they occur. This allows the specification of simple runtime scenarios in a graphical manner.

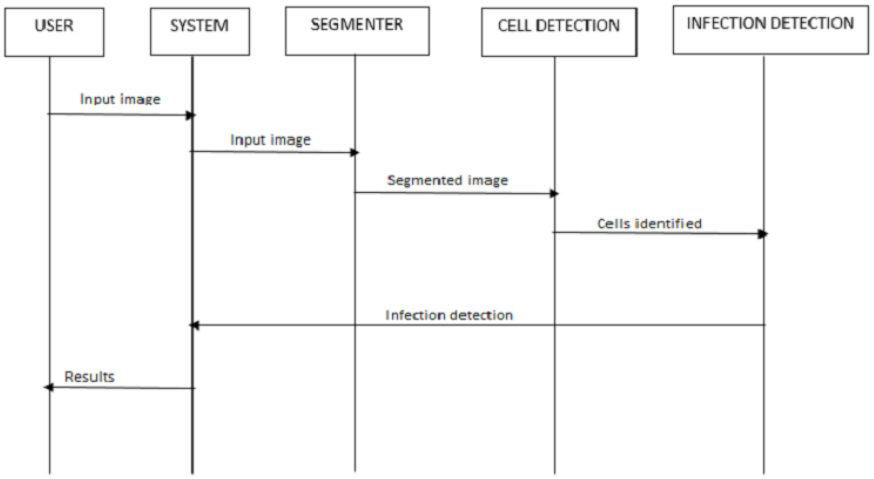

Figure 4.18: Sequence Diagram of Proposed System

4.3 Software Development Model

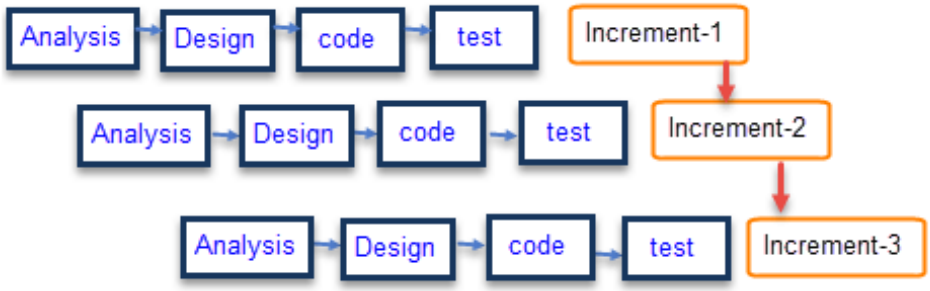

Figure 4.19: Incremental Software Development Model

CHAPTER 5
RESULT AND DISCUSSION

5.1 Output

The detection of malaria parasites is done by pathologists manually using microscopes. so, the chances of false detection due to human error are high, which in turn can result into fatal condition. this seminar curbs the human error while detecting the presence of malaria parasites in the blood sample by using image processing and automation.
Our project focuses on detection and classification of malaria.

5.2 Work Completed

- **Homepage**

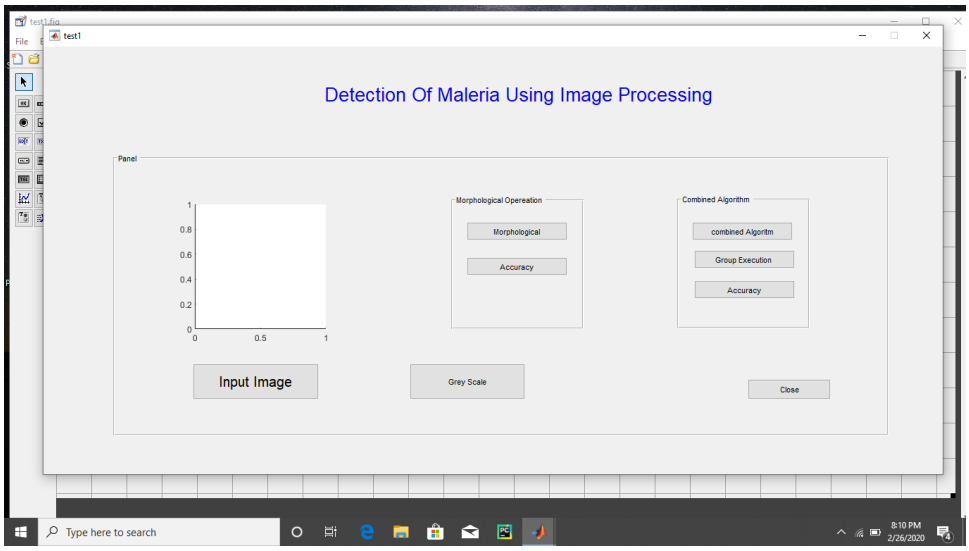

Figure 5.1: Homepage

- **Image Input Selection**

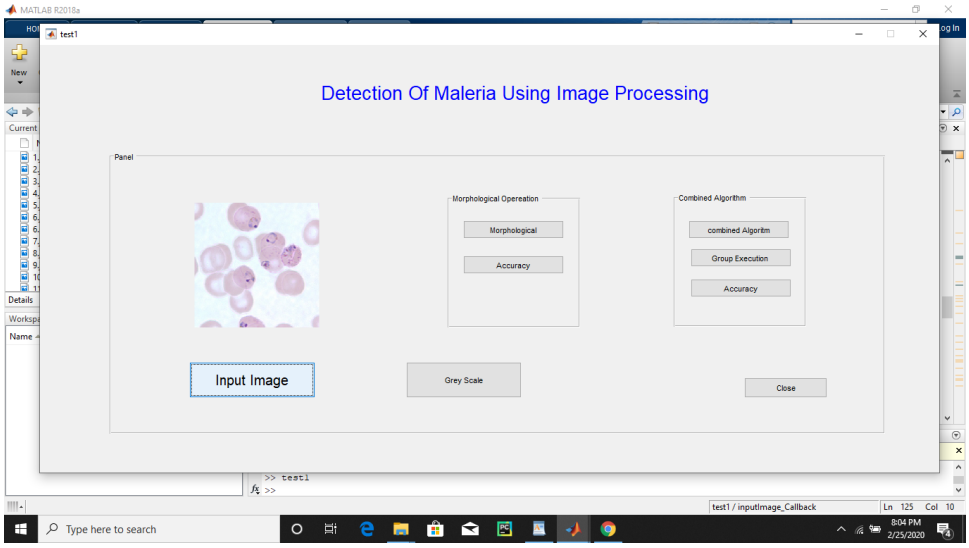

Figure 5.2: Image Input Selection

- **Morphological Operation**

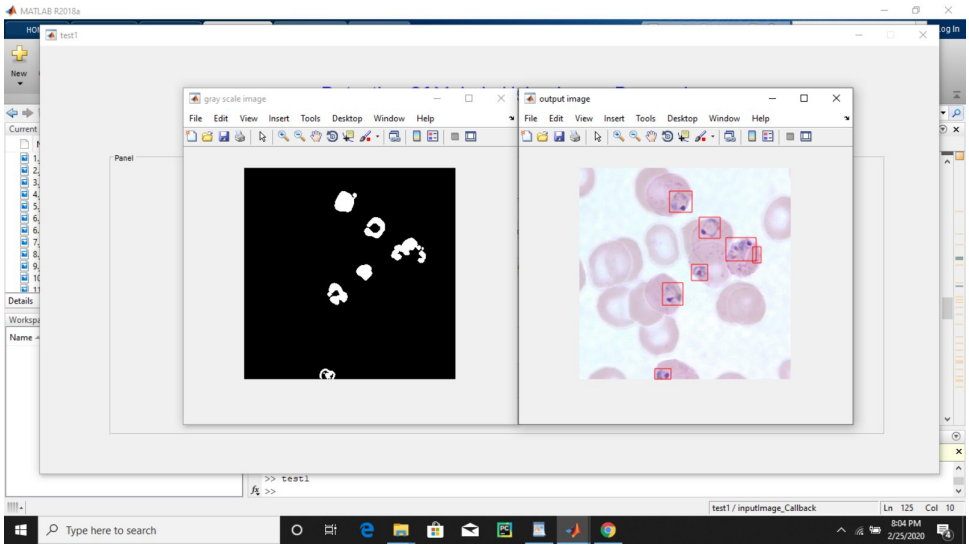

Figure 5.3: Morphological Operation

5.3 Work Remaining

- Accuracy calculation of morphological operation.
- Color based discrimination.
- Accuracy calculation of color based discrimination.
- Classification of malaria.
- Multiple image execution.

5.4 Problem Faced

-
-

5.5 Works Schedule

We have planned our schedule as per the requirements.Our proposed project is expected to be completed as accordance to the gantt chart shown.

5.6 Gantt Chart

Gantt Chart for our project is given below:

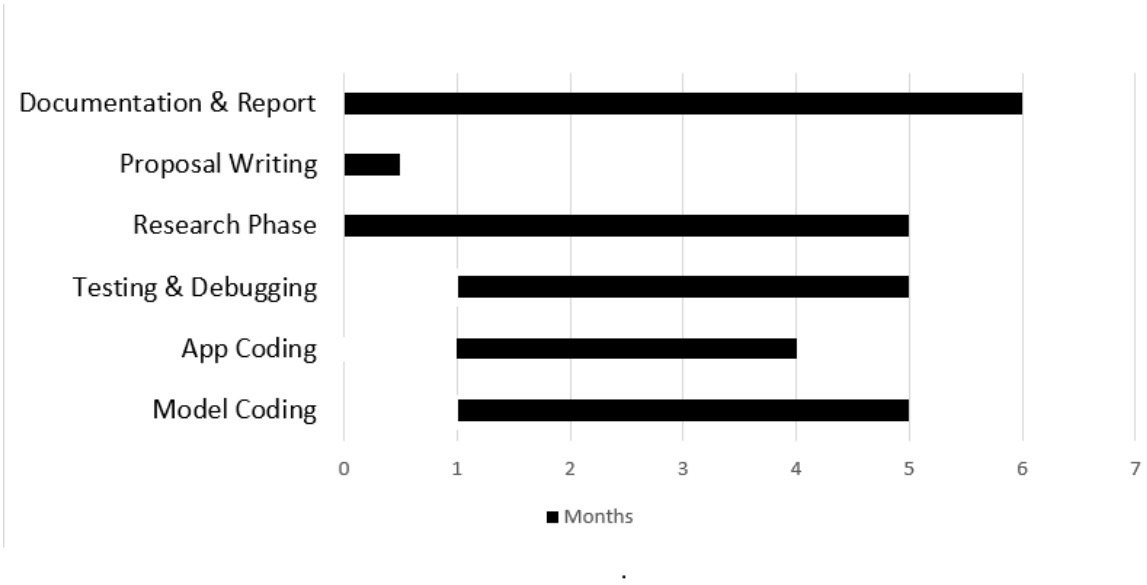

Figure 5.4: Gantt Chart

REFERENCE

[1] J. G. Breman, M. S. Alilio, and A. Mills, "Conquering the intolerable burden of malaria: what's new, what's needed: a summary," *The American journal of tropical medicine and hygiene*, vol. 71, no. 2_suppl, pp. 1–15, 2004.

[2] T. E. Kwenti, T. D. B. Kwenti, L. A. Njunda, A. Latz, K. A. Tufon, and T. Nkuo-Akenji, "Identification of the plasmodium species in clinical samples from children residing in five epidemiological strata of malaria in cameroon," *Tropical medicine and health*, vol. 45, no. 1, p. 14, 2017.

[3] A. Kazarine, F. Baakdah, A. A. Gopal, W. Oyibo, E. Georges, and P. W. Wiseman, "Malaria detection by third-harmonic generation image scanning cytometry," *Analytical chemistry*, vol. 91, no. 3, pp. 2216–2223, 2019.

[4] K. Roy, S. Sharmin, R. B. M. Mukta, and A. Sen, "Detection of malaria parasite in giemsa blood sample using image processing," *International Journal of Computer Science and Information Technology*, vol. 10, no. 1, pp. 55–65, 2018.

[5] G. Díaz, F. Gonzalez, and E. Romero, "Infected cell identification in thin blood images based on color pixel classification: comparison and analysis," in *Iberoamerican Congress on Pattern Recognition*. Springer, 2007, pp. 812–821.

[6] S. Halim, T. R. Bretschneider, Y. Li, P. R. Preiser, and C. Kuss, "Estimating malaria parasitaemia from blood smear images," in *2006 9th international conference on control, automation, robotics and vision*. IEEE, 2006, pp. 1–6.

[7] D. A. Kurer and V. P. Gejji, "Detection of malarial parasites in blood images," *International Journal of Engineering Science and Innovative Technology (IJESIT)*, vol. 3, no. 3, pp. 651–656, 2014.

[8] J. Somasekar, B. Reddy, E. Reddy, and C. Lai, "Computer vision for malaria parasite classification in erythrocytes," *International Journal on Computer Science and Engineering*, vol. 3, no. 6, pp. 2251–2256, 2011.