Tuberculosis Diagnosis through Image Processing

NAHID KAMAL(12201058)
SUBRINA ISLAM(13301139)
AL FAYSAL BIN ASAD (13101095)
SUMAIYA KHATUN(12201099)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Supervisor:

Dr. Amitabha Chakrabarty

Assistant Professor

Department of Computer Science and Engineering

DECLARATION

We, hereby declare that this thesis is based on results we have found ourselves. Materials of work from researchers conducted by others are mentioned in references.

Signature of Supervisor   Signature of Authors

_____

Dr. Amitabha Chakrabarty   _____
Assistant Professor   Nahid Kamal (12201058)
Department of Computer Science &Engineering
School of Engineering & Computer Science
BRAC University

                                                  _____
                                                  Subrina Islam (13301139)

                                                  _____
                                                  Al Faysal Bin Asad (13101095)

                                                  _____
                                                  Sumaiya Khatun (12201099)

ABSTRACT

Tuberculosis is most common contagious disease. Nowadays, millions of human beings of the world are suffering from it. We will use the most used worldwide method Ziehl-Neelsen stain (ZN-stain) to detect Tuberculosis which is based on sputum examination microscopically. This method needs expert human resources and implicit examination. The main constraints are expertise human,time and cost to implement our process. We will use Thresholding, multi-stage, color-based Bayesian segmentation identified possible 'Tuberculosis objects', removed artifacts by shape comparison and color-labeled objects as 'definite', 'possible' or 'non-Tuberculosis', bypassing photomicrographic calibration.In our work, we will use an algorithm based on image processing is developed for identification of Tuberculosis.

Acknowledgment

We would like to express our gratitude to the Almighty who gave us the opportunity, determination, strength and intelligence to complete our work.

A very big and humble thank you goes to our supervisor Dr. Amitabha Chakrabarty who has constantly believed in us and has been there for us through thick and thins of the thesis and continuously pushed us to complete our work in time. We are fortunate and grateful to be able to work under his supervision.

Lastly, our gratitude goes to the faculty members of the Department of Computer Science and Engineering, BRAC University from whom we gained the knowledge, appreciation and help for the completion of our thesis work.

List of Figures

LIST of TABLES

CHAPTER 1

1.1   Introduction

Tuberculosis (TB) is a longstanding infectious disease and a major public health issue. TB has caused by Mycobacterium tuberculosis (M. tuberculosis) that is airborne. It spreads through the air from an infected person to susceptible ones. The risk of becoming infected depends principally on how long and how intense the exposure to the bacterium is. Our approach is to develop the algorithm that will automatically detect the TB viruses.

One third of the world's population has infected with TB [3]. New infections occur in about 1% of the population each year [3]. In 2014, 9.6 million cases of active TB that resulted in 1.5 million deaths. More than 95% of deaths occurred in developing countries. The number of new cases each year has decreased since 2000 [3]. About 80% of people in many Asian and African countries test positive while 5–10% of people in the United States population tests positive by the tuberculin test [3]. Tuberculosis has been present in humans since ancient times [3].

Tuberculosis control efforts hampered by a mismatch in diagnostic technology: modern optimal diagnostic tests are least available in poor areas. Lack of adequate early diagnostics and MDR detection is a critical problem in control efforts. Our intention is to make the process easier for rural area people [4].

1.2   Research Goal

The goal of our research is to detect Tuberculosis through images. Our approach is to develop the algorithm that will automatically detect the TB viruses. Our intention is to make the process easier for rural area people. Tuberculosis is one of the most important diseases in developing countries like Bangladesh. We will use the most used worldwide method Ziehl-Neelsen stain (ZN-stain) to detect Tuberculosis. The procedure based on sputum examination microscopically. This method needs expert human resources and implicit examination. The main constraints are expertise human, time and cost to implement our process. We will use automated multi-stage; color-based Bayesian segmentation identified possible "Tuberculosis objects", removed artifacts by shape comparison and color-labeled objects as "definite", "possible" or "non-Tuberculosis", bypassing photo micrographic calibration. In our work, we will use an algorithm based on image processing is developed for identification of Tuberculosis bacteria in sputum. We believe that the approach we are going to use Tuberculosis is one of the most important diseases in developing countries like Bangladesh.

1.3   Motivation

Every year many people diagnosed by Tuberculosis all over the world. The testing cost is overpriced and need a huge enough amount of time. There are many place where the processing of detecting Tuberculosis cannot done because of lacking resources. As Bangladesh is a developing country, there are many places where people live under the poverty line. In many part of our country, have no proper electricity supply and resources, as the machines are very costly. Many people who lives outside of Dhaka come to Dhaka for detecting Tuberculosis. Therefore, we thought that we could solve the problem by making a program that can detect Tuberculosis from an image of a spit. As mobile phone is a common in our country. Therefore, we can use mobile phone to take picture and by processing the image, it will detect Tuberculosis. Finally, by our program, one can easily detect Tuberculosis in a low cost and less time.

1.4   Methodology

The dataset has collected from slides using digital microscope. The slides are analyzed using Olympus binocular oil immersion light microscope with 100X magnification and their images are captured using Luminera infinity2 digital camera attached to the microscope. The images

are of 24-bit color images in size of 512x512 pixels. The steps involved in processing of input images for segmentation of mycobacterium tuberculosis in tissue sections using thresholding and clustering techniques showed in Figure 1.1
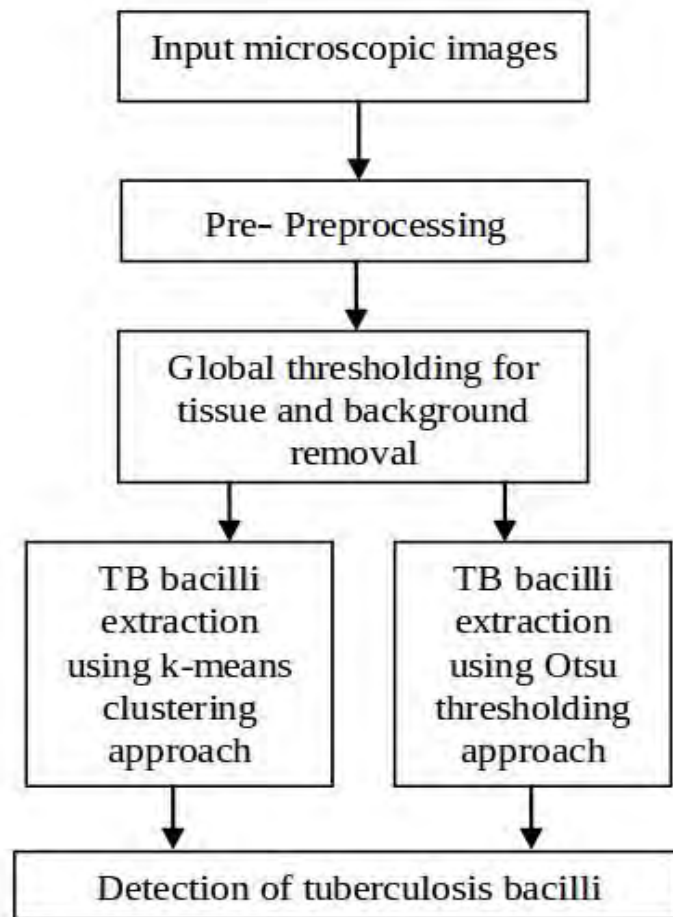
Fig 1.1 Steps of segmentation and detection of TB

The present study employs the global thresholding, to remove the tissues and background that stained by methylene blue dye, in order to initiate segmentation of the TB bacilli. For the second step, k-mean clustering used to extract the TB bacilli from the remaining background. The third step involves labeling and removing noises and large over stained regions using the region growing method. The study introduces a local thresholding and a median filter, as a fourth step, to refine the segmented regions. In the following sections, the methods are present and discussed in details.

1.4.1 Thresholding

In mathematical or statistical modeling a threshold model is any model where a threshold value or set of threshold values that has used to distinguish ranges of values where the behavior predicted by the model varies in some important way. A particularly important instance arises in toxicology, where the model for the effect of a drug may be that there is zero effect for a dose below a critical or threshold value, while an effect of some significance exists above that value. Certain types of regression model may include threshold effects.

1.4.2 Region Growing Method

Region growing is a simple region-based image segmentation method. It is classified as a pixel-based image segmentation method since it involves the selection of initial seed points.

1.4.3 K means Clustering

K-means clustering is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells.

1.4.4   Morphology

Morphology is a broad set of image processing operations that process images based on shapes. An essential part of the morphological dilation and erosion operations is the structuring element used to probe the input image.

1.5   Thesis Outline

Chapter 2

Describe the background research and basic review about the topic

Chapter 3

In this chapter, we demonstrate the implementation terminology and methodology of our approach.

Chapter 4

In this chapter, we analyze the present scenario of the patients In Bangladesh

Chapter 5

In this chapter, we demonstrate result

Chapter 6

In this chapter, we discussed limitations and future scope of the research along with conclusion

CHAPTER 2

2.1   Literature review:

TB is a curable disease and it can be controlled. But to control and cure the disease, correct diagnosis is needed. This disease needs continuous monitoring and proper guide lining. The guidelines of WHO suggests to diagnose TB by screening of Ziehl Neelson (ZN) stained specimen under light microscope or Rhodamine/ Auramine stained specimen under a fluorescence microscope [1]. But this technique is sensitive as it needs highly trained specialists and it has a high false negative rate.

Sadaphal et al. [2] gave an approach which is based on the idea of using clear Mycobacterium Tuberculosis broth images to characterize Acid-Fast Bacilli color. Bayesian segmentation was used to predict the probability of a pixel presenting a TB object. An algorithm was used to recognize Acid-Fast Bacilli under wide latitudes of staining, magnification and resolution.
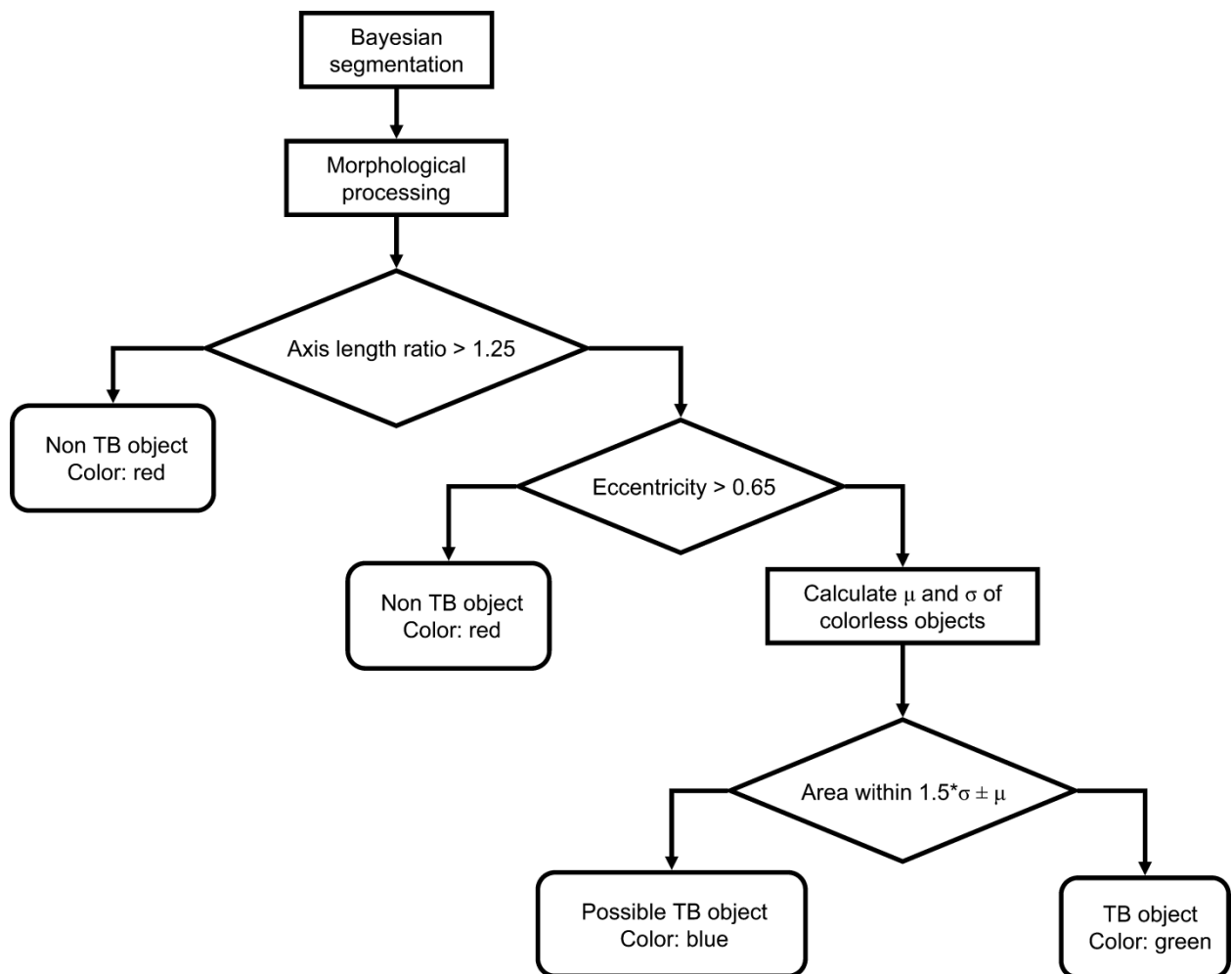
Fig 2.1 Classification steps for automatic identification and labeling of bacilli.

M.G. Forero et al. [3] gave an approach of obtaining a probabilistic estimation of the bacillus set by using an optimum classifier which was based on the Bayesian decision theory. Respiratory and non respiratory clinical specimens were collected and these specimens were stained with Fluorochrome Auramine O and a fluorescence microscope was used to capture images. Presence of positive fluorochrome smears was confirmed by the growth of Mycobacterium Tuberculosis bacilli from culture of specimens. Leica Leitz was used to analyze the sample slides. This method faces problems as specialists decision was needed to be taken to consider an object as a bacillus. They had faced many challenges to automatically evaluate the sample of sputum as the images were occupied with debris and smudges or fingerprints.
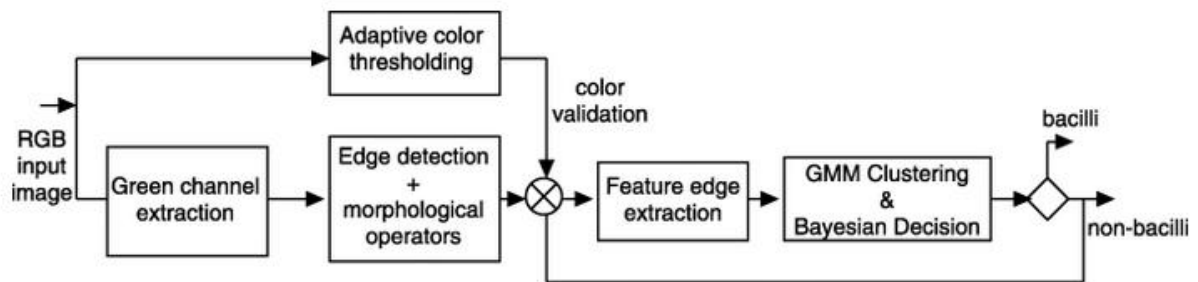
Fig 2.2 Schematic procedure of the stages required for bacillus/ non bacillus discrimination. RGB; GMM, Gaussian mixture model.

Another approach given by Osman et al. [4] which gives the idea of combining k-means clustering and thresholding algorithm to segment the TB bacilli in tissue sections. Three color models RGB, HIS and C-Y was compared to the work and better segmentation performance was found.

Rachna et al. [5] gave an approach of comparing clustering and thresholding algorithms to find the best method for bacilli segmentation and automatic identification of Tb bacilli from sputum specimens.
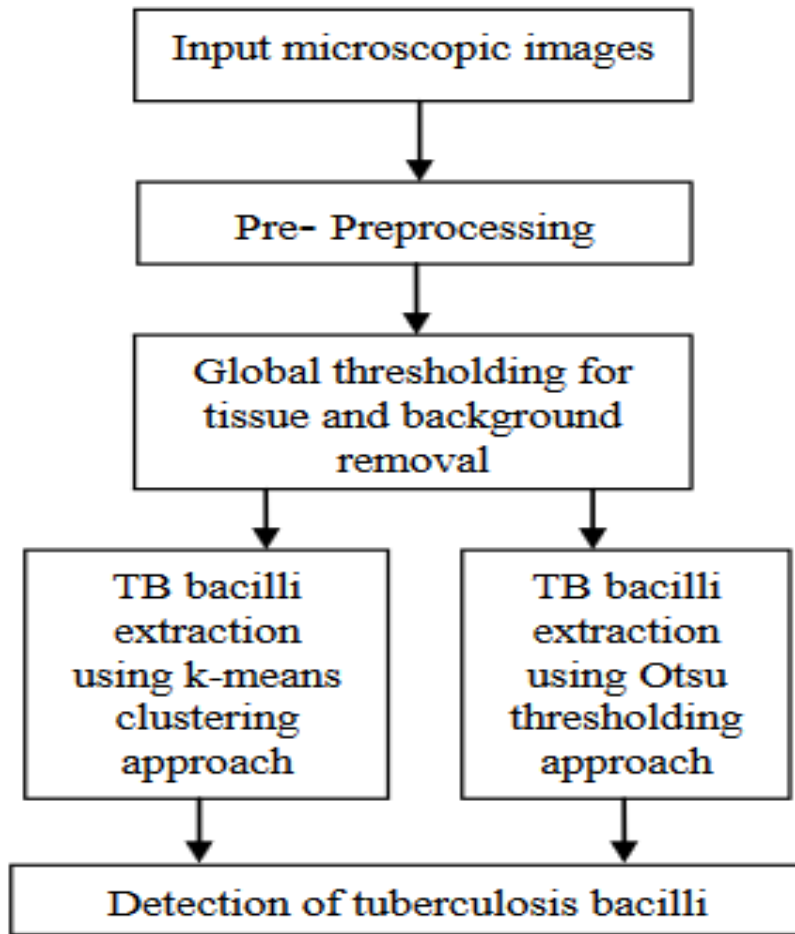
Fig 2.3 Steps of segmentation and detection of TB.

Costa Filho et al. [6] gave an approach of identifying bacillus which is composition of image acquisition, image segmentation and post processing. SVM and neural network classifier techniques were used in the segmentation step. Combinations of pixel color characteristics are the input variables of these classifiers. Characteristics were selected from four color spaces and the best ones were selected by a scalar features selection technique. The outputs are objects that can be bacilli or artifacts. Artifacts were eliminated by the post-processing step.

CHAPTER 3

3.1   Implementation

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. Nowadays, image processing is among rapidly growing technologies. It forms core research area within engineering and computer science disciplines too.

Image processing includes the following three steps:

- Importing the image via image acquisition tools;
- Analysing and manipulating the image;
- Output in which result can be altered image or report that is based on image analysis.

There are two types of methods used for image processing namely, analogue and digital image processing. Analogue image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. Digital image processing techniques help in manipulation of the digital images by using computers. The three general phases that all types of data have to undergo while using digital technique are pre-processing, enhancement, and display, information extraction.

MATLAB is one of the greatest image processing tools in the world nowadays. For that reason, we use this tool as our solving equipment. We also use some of the noteworthy methods of MATLAB. The methods are given below:

3.2   Methodology

3.2.1   Morphological Method

Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. Morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. Morphological operations can also be applied to greyscale images such that their light transfer functions are unknown and therefore their absolute pixel values are of no or minor interest.

Morphological techniques probe an image with a small shape or template called a structuring element.

Many morphological operations are represented as combinations of erosion, dilation, and simple set-theoretic operations such as the **complement** of a binary image:

$$f^c(x,y) = 1 \text{ if } f(x,y) = 0, \text{ and } f^c(x,y) = 0 \text{ if } f(x,y) = 1,$$

the **intersection** $h = f \cap g$ of two binary images f and g:

$$h(x,y) = 1 \text{ if } f(x,y) = 1 \text{ and } g(x,y) = 1, \text{ and } h(x,y) = 0 \text{ otherwise,}$$

and the **union** $h = f \cup g$ of two binary images f and g:

$$h(x,y) = 1 \text{ if } f(x,y) = 1 \text{ or } g(x,y) = 1, \text{ and } h(x,y) = 0 \text{ otherwise:}$$
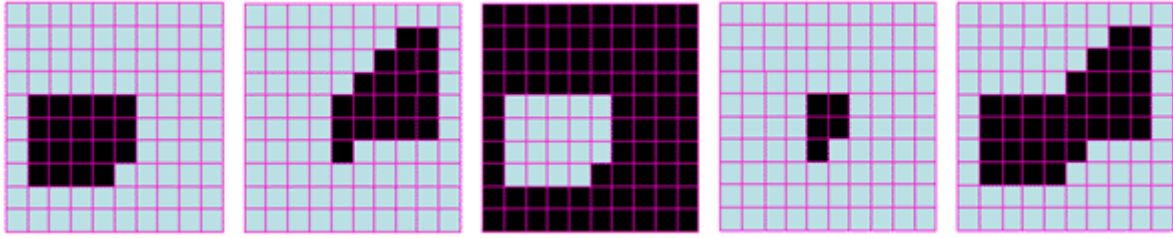
Fig 3.1 Set operations on binary images: from left to right: a binary image f, a binary image g, the complement f^c of f, the intersection $f \cap g$, and the union $f \cup g$.

The **opening** of an image f by a structuring element s (denoted by $f \circ s$) is an erosion followed by a dilation:

$$f \circ s = (f \ominus s) \oplus s$$

Fig 3.2 Binary image

Fig 3.3 Opening: a 2×2 square structuring element

Opening is so called because it can open up a gap between objects connected by a thin bridge of pixels. Any regions that have survived the erosion are restored to their original size by the dilation:

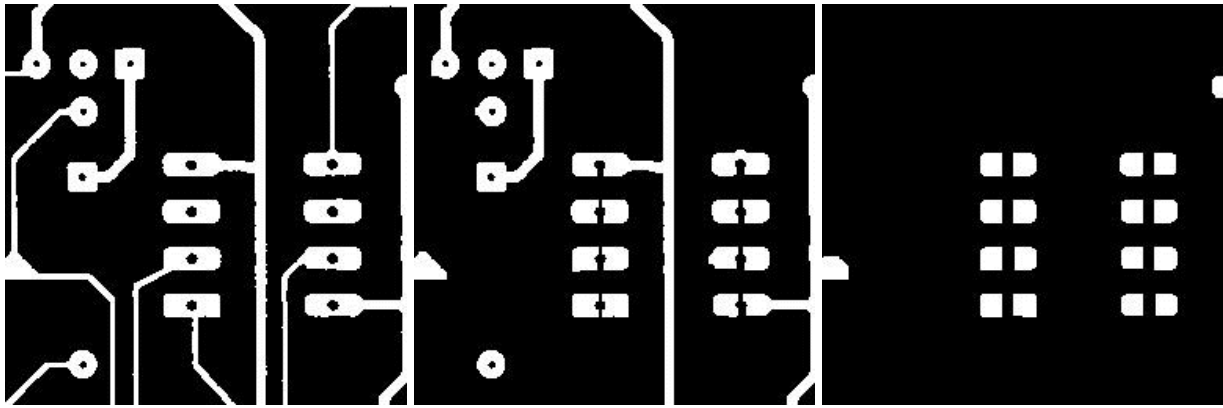| Binary image f | $f \circ s$ (5×5 square) | $f \circ s$ (9×9 square) |

Fig 3.4 Results of opening with a square structuring element

Opening is an **idempotent** operation: once an image has been opened, subsequent openings with the same structuring element have no further effect on that image:

$$(f \circ s) \circ s) = f \circ s.$$

The **closing** of an image f by a structuring element s (denoted by $f \bullet s$) is a dilation followed by an erosion:

$$f \bullet s = (f \oplus s_{rot}) \ominus s_{rot}$$

Fig 3.5 Binary Image and Closing Set

In this case, the dilation and erosion should be performed with a rotated by 180₀ structuring element. Typically, the latter is symmetrical, so that the rotated and initial versions of it do not differ.
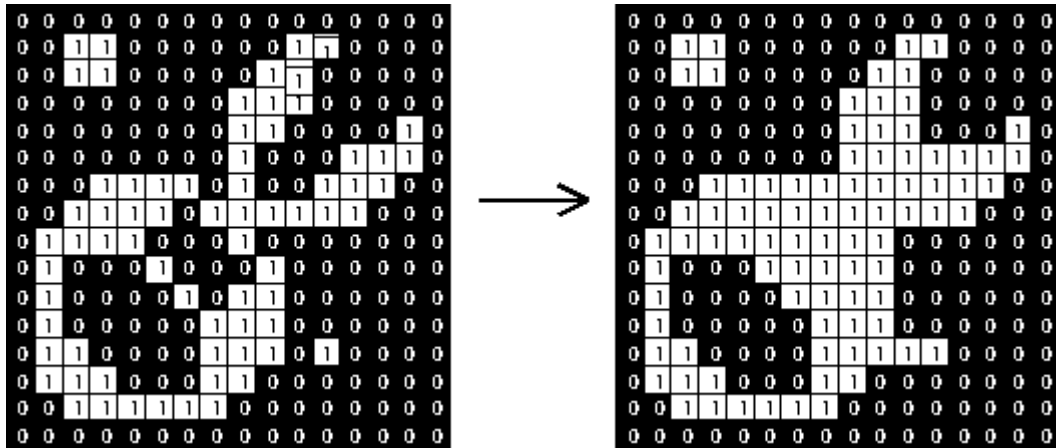

Fig 3.6 Closing with a 3×3 square structuring element

Closing is so called because it can fill holes in the regions while keeping the initial region sizes. Like opening, closing is idempotent: $(f \bullet s) \bullet s = f \bullet s$, and it is dual operation of opening (just as opening is the dual operation of closing):

$$f \bullet s = (f^c \circ s)^c; \quad f \circ s = (f^c \bullet s)^c.$$

In other words, closing (opening) of a binary image can be performed by taking the complement of that image, opening (closing) with the structuring element, and taking the complement of the result.

The **hit and miss transform** allows to derive information on how objects in a binary image are related to their surroundings. The operation requires a matched pair of structuring elements, $\{s_1, s_2\}$, that probe the inside and outside, respectively, of objects in the image:

$$f \circledast \{s_1, s_2\} = (f \ominus s_1) \cap (f^c \ominus s_2).$$

Fig 3.7 Binary Image and Hit and miss transform of

an elongated 2×5 structuring element

A pixel belonging to an object is preserved by the hit and miss transform if and only if s_1 translated to that pixel fits inside the object AND s_2 translated to that pixel fits outside the object. It is assumed that s_1 and s_2 do not intersect, otherwise it would be impossible for both fits to occur simultaneously.

It is easier to describe it by considering s_1 and s_2 as a single structuring element with 1s for pixels of s_1 and 0s for pixels of s_2; in this case the hit-and-miss transform assigns 1 to an output pixel only if the object (with the value of 1) and background (with the value of 0) pixels in the structuring element exactly match object (1) and background (0) pixels in the input image. Otherwise that pixel is set to the background value (0).

The hit and miss transform can be used for detecting specific shapes (spatial arrangements of object and background pixel values) if the two structuring elements present the desired shape, as well as for thinning or thickening of object linear elements.

Morphological filtering of a binary image is conducted by considering compound operations like opening and closing as filters. They may act as filters of shape. For example, opening with a disc structuring element smooth's corners from the inside, and closing with a disc smooth's corners from the outside. In addition, these operations can filter out from an image any details that are smaller than the structuring element, e.g. opening is filtering the binary image at a scale defined by the size of the structuring element. Only the filter passes those portions of the image that fit the structuring element; smaller structures are blocked and excluded from the output image. The size of the structuring element is most important to eliminate noisy details but not to damage objects of interest.

3.2.2 K-Means Clustering

k-*means* clustering is a partitioning method. The function k-means partitions data into *k* mutually exclusive clusters, and returns the index of the cluster to which it has assigned each observation. Unlike hierarchical clustering, k-means clustering operates on actual observations (rather than the larger set of dissimilarity measures), and creates a single level of clusters. The distinctions mean that k-means clustering is often more suitable than hierarchical clustering for large amounts of data.

K-means treats each observation in your data as an object having a location in space. It finds a partition in which objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. You can choose from five different distance measures, depending on the kind of data you are clustering.

Each cluster in the partition is defined by its member objects and by its centroid, or center. The centroid for each cluster is the point to which the sum of distances from all objects in that cluster is minimized. K-means computes cluster centroids differently for each distance measure, to minimize the sum with respect to the measure that you specify.

You can control the details of the minimization using several optional input parameters to k-means, including ones for the initial values of the cluster centroids, and for the maximum number of iterations. By default, k-means uses the K-means++ algorithm for cluster center initialization and the squared Euclidean metric to determine distances.

3.2.3 Approaches towards Implementation

First, we collected raw data of Patients from TB hospital of Mohakhali. We took pictures by using Microscope at 20x Zoom and 40x Zoom. After getting the huge amount of data, we

started processing the raw data. We build a software using MATLAB which can tell us whether a patient is TB affected or not. By using our software, we can also find out in which stage the patient condition is.

After using morphological process in MATLAB, we are getting results like this:
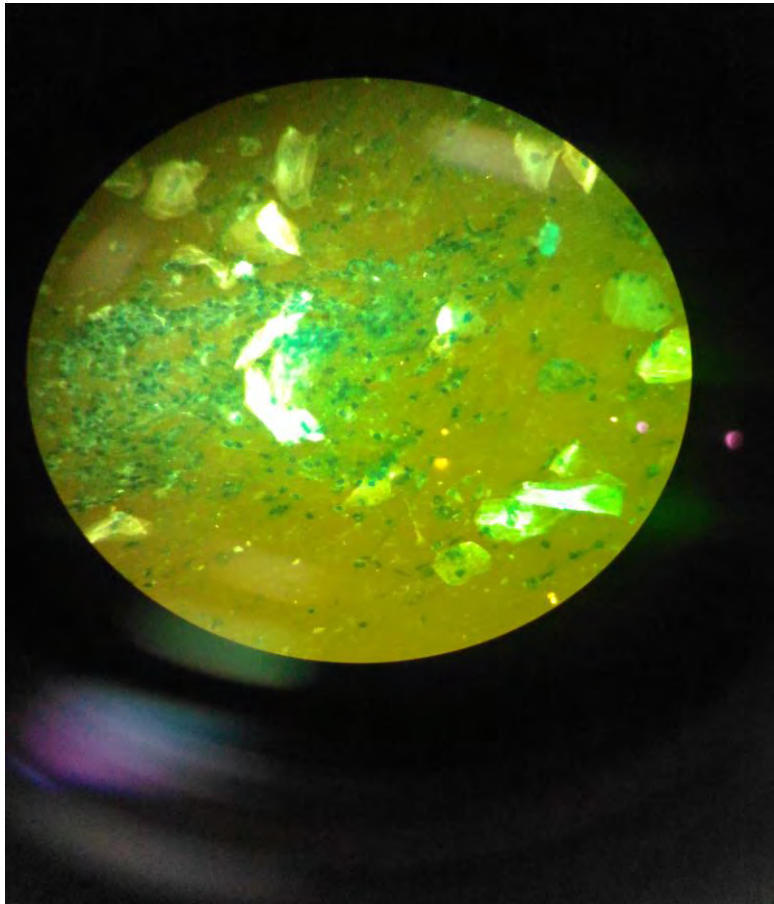
Fig 3.8 Raw Data (Sample No.16440)
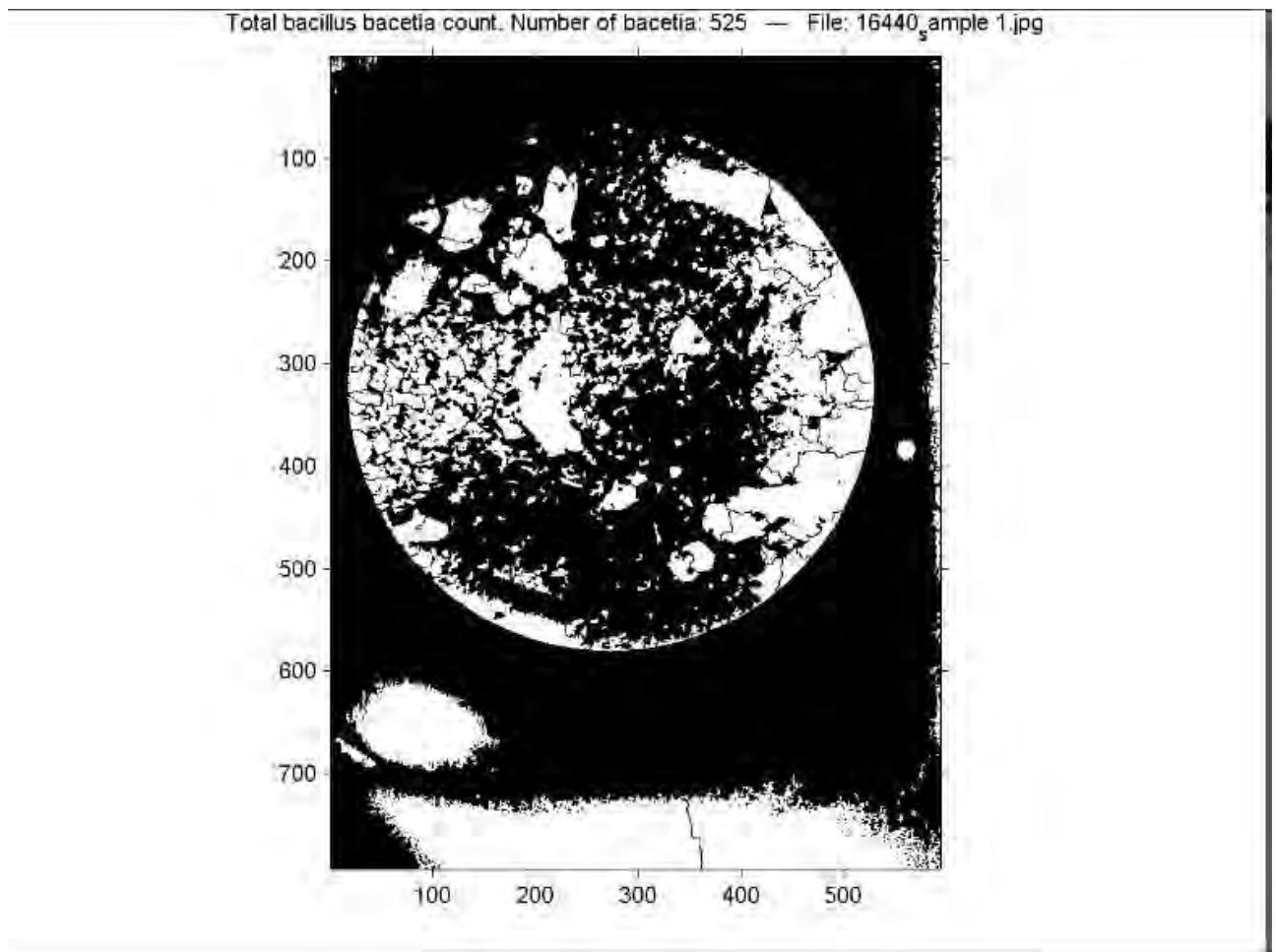
After Processing the sample data:

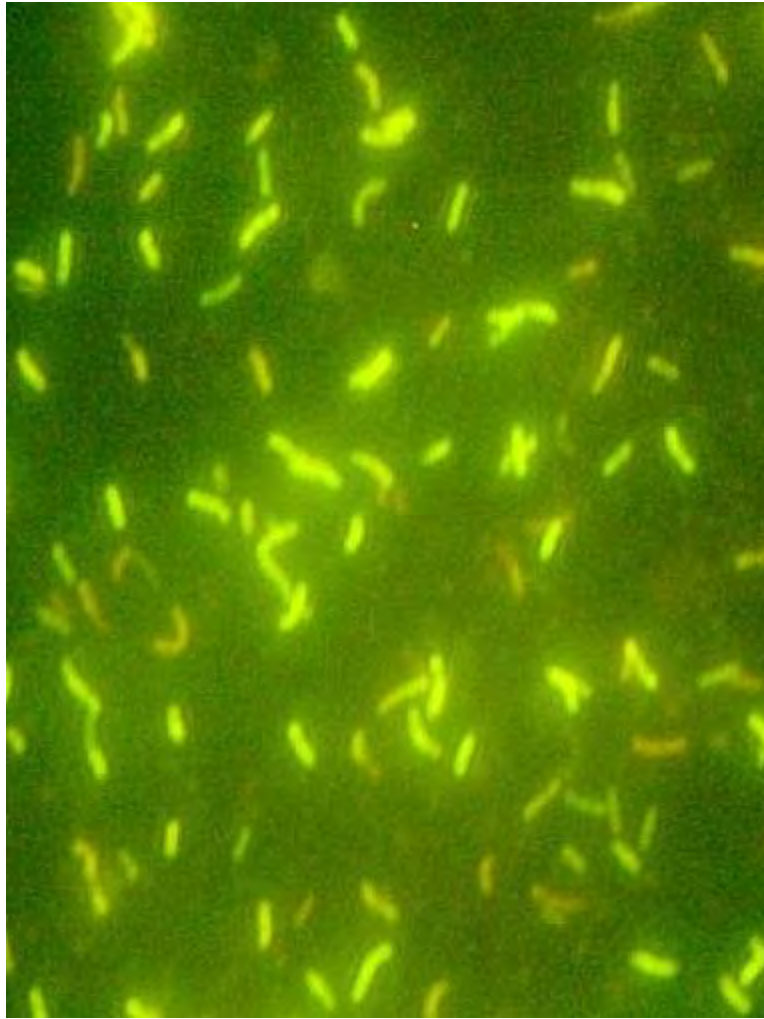Fig 3.9 After processing the sample data no. 16440
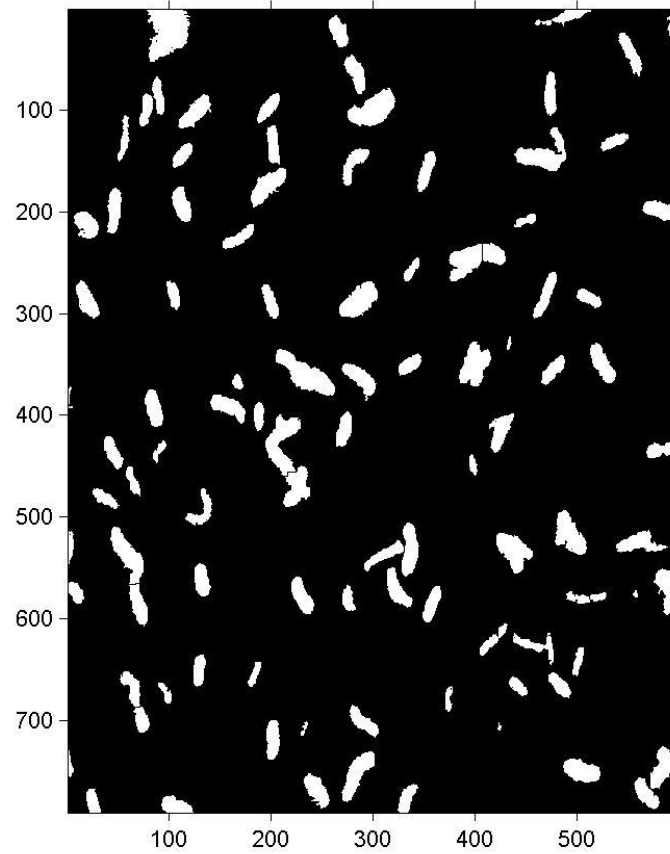
Fig 3.10 Raw Data Sample Data 16441(Cropped)

Fig 3.11 After processing the sample data no. 16441(Cropped)

CHAPTER 4

4.1 Analysis:

The Ziehl-Neelsen is the most usually utilized strategy for the showing of tubercle bacilli in histological areas. As far as we can tell, be that as it may, this stain yields a disquietingly low rate of positive outcomes in segments from injuries that are histological regular of tuberculosis. In such cases, a determination of tuberculosis is frequently made disregarding the inability to show the living being, in spite of the fact that a slight uncertainty about the genuine idea of the sore may stay in the pathologist's brain. Any method, which creates preferable outcomes over those acquired with the Ziehl-Neelsen stain would in this way, be welcome to histologists. This examination includes a correlation of the Ziehl-Neelsen recolor with two option fuchsin stains and with a fluorescent procedure trying to figure out which is of the most incentive in the standard examination of histological areas.

Seventy cases which had been analyzed as "tuberculosis" were chosen aimlessly from the records of the Department of Pathology, University of Manchester. Segments from every one of these cases demonstrated epithelioid cell granulomata and regions of caseation. The tissues were from an assortment of anatomical locales including lymph hubs (21), kidney (9), gastrointestinal tract (6), epididymus (5), liver (4), omentum (4), synovium (3), cerebrum (1), vagina (1), heart (1), spleen (1), and skin (1).

All areas were analyzed at first utilizing a x 40 goal and x 10 eyepieces; they were then reconsidered utilizing x 100 oil-submersion target and x 10 eyepieces. Areas recolored by the fluorescent strategy were likewise inspected utilizing a x 25 target and x 10 eyepieces. Areas recolored by the fluorescent technique were analyzed under a fluorescent magnifying lens with proper warmth and obstruction channels and BG12 excitation channel.

According to our research, we found that the mostly affected patients are from industrial area and in residential area the amount of affected patients are lesser in number. We found almost 90% affected patients are living in industrial area and 10% are living in residential area. We

also found that the middle aged people (age range: 21-40 years) are mostly affected by Tuberculosis and in that case the number of female patients are greater than male patients. Additionally, we analyzed the total number of patients. Almost 85% patients are male and 15% patients are female. So, we can say that in total the male patients are mostly affected. Again, we compared the result between positive result and negative result that we found that only 15% patients are Tuberculosis affected among the total patients who came to test their cough.

Table 4.1 A comparison between Ziehl-Neelsen Method and Fite's Method

Ziehl-Neelsen Method	Fite's Method

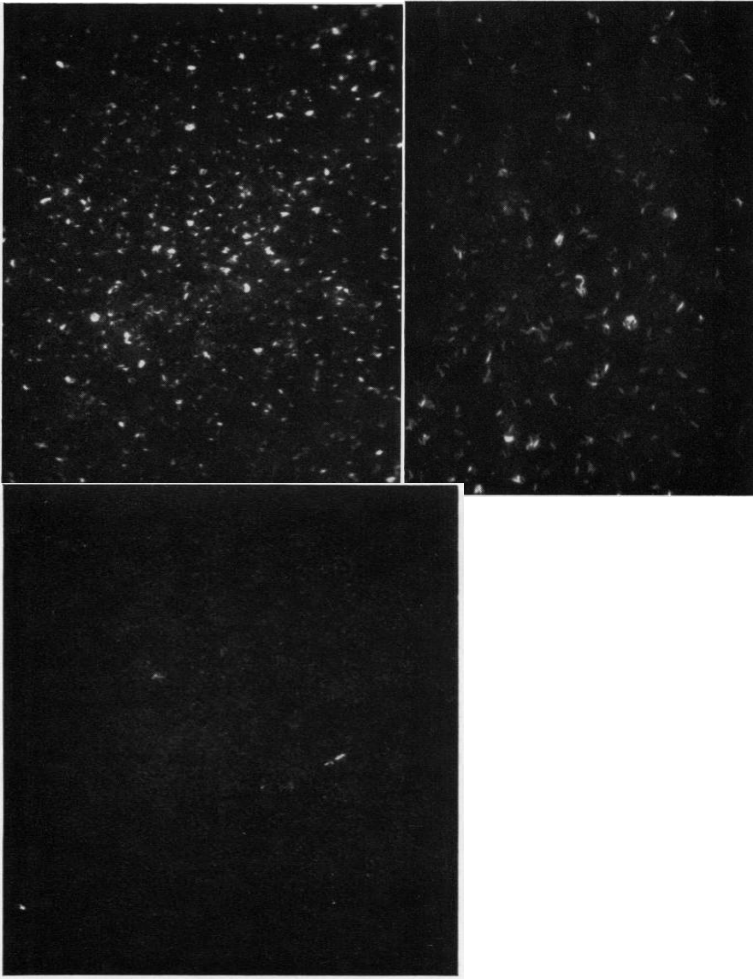|	
1) Convey segments to water.	1) Convey segments to water.
2) Stain in hot carbol-fuchsin in a coplin bump at 50°C for 30 minutes.	2) Stain in new fuchsin 0 5 g, phenol precious stones 5 0 g, ethanol 10. ml, and refined water to make 100 mlat 60°C (paraffin broiler) for 12 to 24 hours.
3) Wash in water to expel overabundance recolor.	
4) Separate in 3 % HCI in 70% liquor until pale pink for five to 10 minutes.	3)Reagent review, 40 %formaldehyde for five minutes.
5) Wash in water.	4) Hydrochloric corrosive 2% out of 95% liquor for 10 minutes.
6) Counterstain in 0-1 % methylene blue for 10 to 15 seconds.	5) Potassium permanganate 1 % watery for two to five minutes.
7) Wash in water.	6) Oxalic acid 2 % for one minute.
8) Dehydrate, clear, and mount.	7))Harris' haematoxylin for two minutes.
	8)Van Gieson's stain for three minutes.
	9) Get dried out, clear, and mount. Each of the above strides is gone before by washing in water.
Result: Tubercle bacilli stain red.	Result: Tubercle bacilli stain blue.

Fig 4.1 A comparison of methods for staining tubercle bacilli in histological sections:

The outcomes are appeared in table I. In table II the outcomes got with the option stains are contrasted and those of the Ziehl-Neelsen recolor.

Technique	Number of Positive Cases (x 250 Magnification)	Number of Positive Cases (x 400 Magnification)	Number of Positive Cases (x 1000 Magnification)
Zeihl-Neelsen	---	28 (40%)	33 (47.1%)

Fite's	---	22 (31.4%)	28 (40%)
Armstrong Price	---	19 (27.1%)	24 (34.3%)
Auramine phenol	42 (60%)	42 (60%)	42 (60%)

Table 4.2 Incidence of positive results for tubercle bacilli in 70 case of historically typical tuberculosis

Technique	Magnification	Zeihl-Neelsen Positivity (Positive Cases Read as Negative)	Zeihl-Neelsen Positivity (Negative Cases Read as Positive)
Fite's	400	11	1
	1000	5	1
Armstrong Price	400	14	0
	1000	8	0
Auramine phenol	400	0	9
	1000	0	9

Table 4.3 Relationship between Ziehl-Neelsen positivity and results obtained by other techniques

Evaluted test	DZM/DFM* (Positive Negative)	Gold standard DZM/LJ (Positive Negative)	Gold standard DFM/LJ (Positive Negative)
Positive	25 7	24 8	53 8
Negative	36 276	42 270	13 270
------------		Estimate (95%CI)	Estimate (95%CI)
Sensitivity (%)		36.4 (24.9-49.1)	80.3 (68.7-89.1)
Specificity (%)		97.1 (94.4-98.7)	97.1 (94.4-98.7)
Positive predictive value (%)		75.1 (56.6-88.5)	86.9 (75.8-94.2)
Negative predictive value (%)		86.5 (82.2-90.1)	95.4 (92.3-97.5)
Kappa agreement measure (%)		41.6 (36.6-46.6)	79.7 (74.3-85.1)

Table 4.4 A Comparison of Tools Used for Tuberculosis Diagnosis in Resource-Limited Settings

TB demonstrative yield was computed by isolating the quantity of TB cases recognized (in view of the full grouping of tests performed) by the aggregate number of hypothetical TB cases screened by calculation and time-point. Adherence to the spread/culture-based calculation was evaluated by ascertaining the extent of cases with two spread tests and the extent of spread

negative cases with culture tests. Adherence to the Xpert-based calculation was surveyed by ascertaining the extent of cases with a Xpert test and the extent of Xpert-negative cases with culture tests.

The extent of hypothetical TB cases tried in the five sub-areas was ascertained by isolating the quantity of people tried in the quarter by the mid-year populace gauges from Statistics South Africa [12].

Elucidating insights by calculation are exhibited utilizing frequencies, means and standard deviation. We utilized a binomial relapse demonstrate, balanced for site-level grouping, to appraise the distinctions in TB yield amongst calculations and survey fleeting patterns. This model was utilized to survey the impact of MDR-TB hazard on TB yield and to assess slants in testing after some time. All examinations were embraced utilizing STATA 12 (StataCorp).

4.2   Classifications of TB Patients Data in Bangladesh

For our thesis purpose, we have collected around 800 cough samples from National Institute of Diseases of Chest and Hospital. Most of the people go there every day to test their cough. We have analyzed these samples and found various types of informations from those patients' data. From these informations, we can clearly understand the current situation of TB affected patients of Bangladesh.

Here are some classifications of patients' data:

1. Percentage of patient affected by TB
2.   Percentage difference between male & female patients
3. Age distribution of male &female patients
4. Information about patients living area.

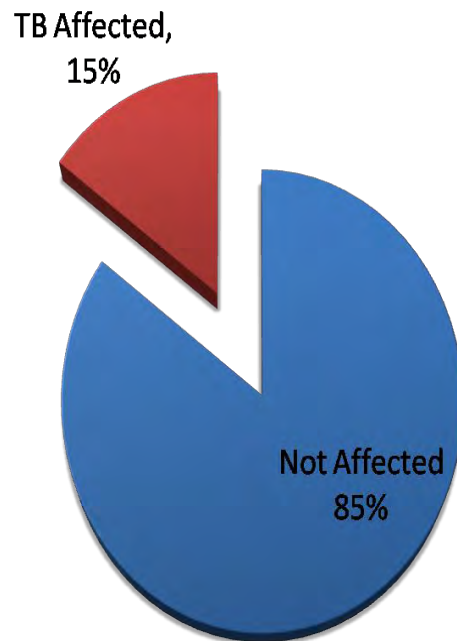### 4.2.1 Percentage of TB affected patients

Fig 4.2 Percentage of TB affected patients

From the collected cough samples, we found that 85% of the patients were not affected by TB. Only 15% of the patients were affected. Tuberculosis bacteria was found in their cough samples on different scales.

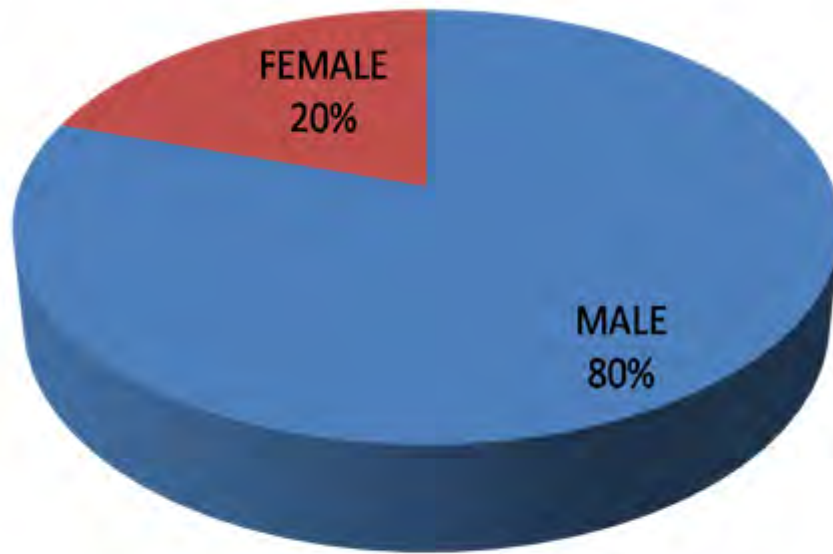### 4.2.2 Percentage difference between male and female patients

Fig 4.3 Difference between male and female patients

From the above figure we can see that male patients are huge in number. They are the 80% of the whole TB affected patients. Female patients are less in number. There are 20% female patients who are affected in TB.

4.2.3 Age distribution of male and female patients

Patients Age Distribution
■ MALE   ■ FEMALE

73%

58%

20%   20%

12%   10%

2%   5%
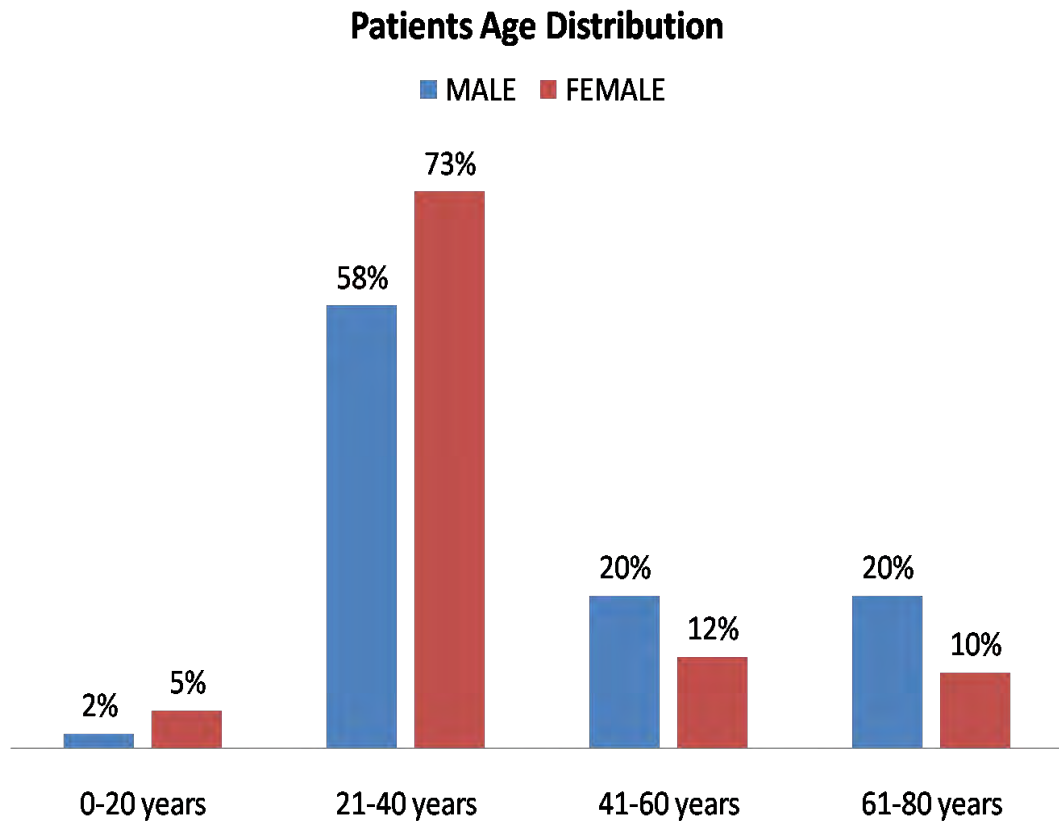
0-20 years   21-40 years   41-60 years   61-80 years

Fig 4.4 Patients age distribution

After analyzing the data we found that there are age differences between the patients. Both male and female patients TB affecting rate is high at the age of 21 to 40 years. Then in the second position is the range of 41 to 60 years.

4.2.4 Information about patients living area

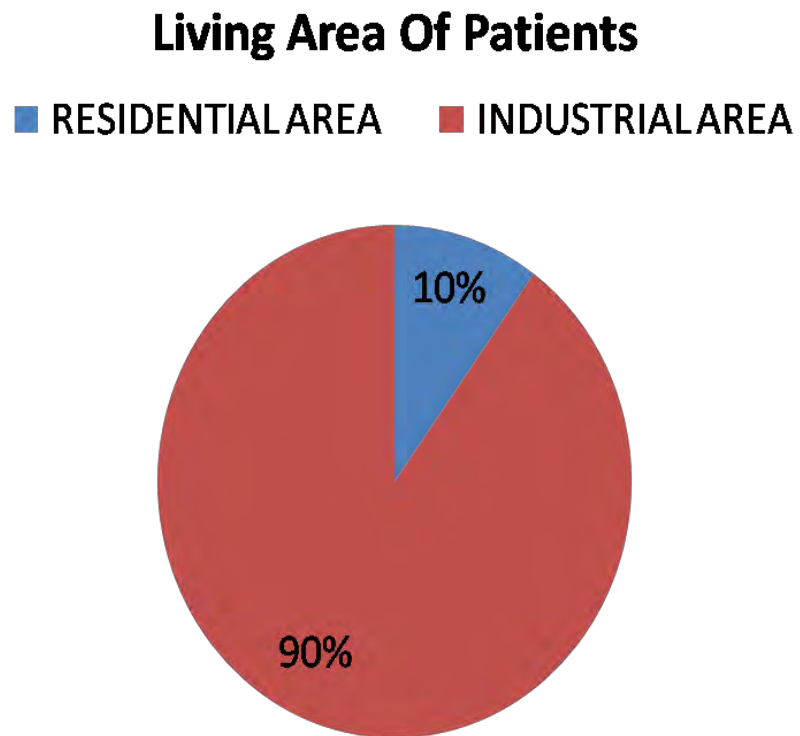## Living Area Of Patients

■ RESIDENTIAL AREA   ■ INDUSTRIAL AREA

Fig 4.5 patients living area.

After the analyzing of the living places of TB patients, we get the information that mostly TB affected patients are from industrial area. They have a high risk of affected by TB and other air-carried diseases because of the air pollution and for smoking cigarette.

CHAPTER 5

5.1 Result

After testing all data we gathered from laboratory of TB Hospital, our application accuracy rate is 90%. Even our application is capable of detecting the stage of Patient such as 1st stage, 2nd stage, 3rd stage etc. Here is some screenshots of how our application works and some sample of the results produced by our application.
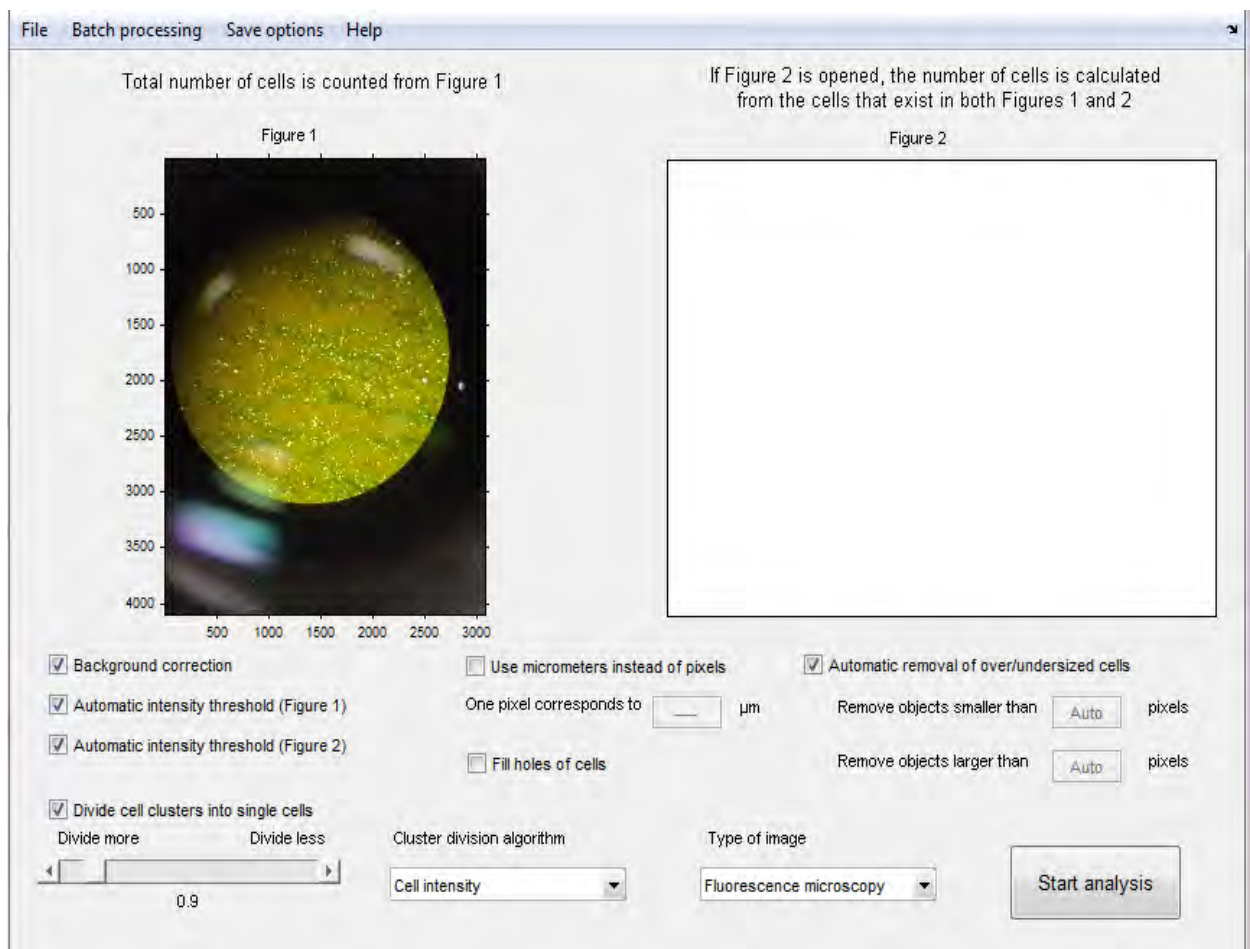

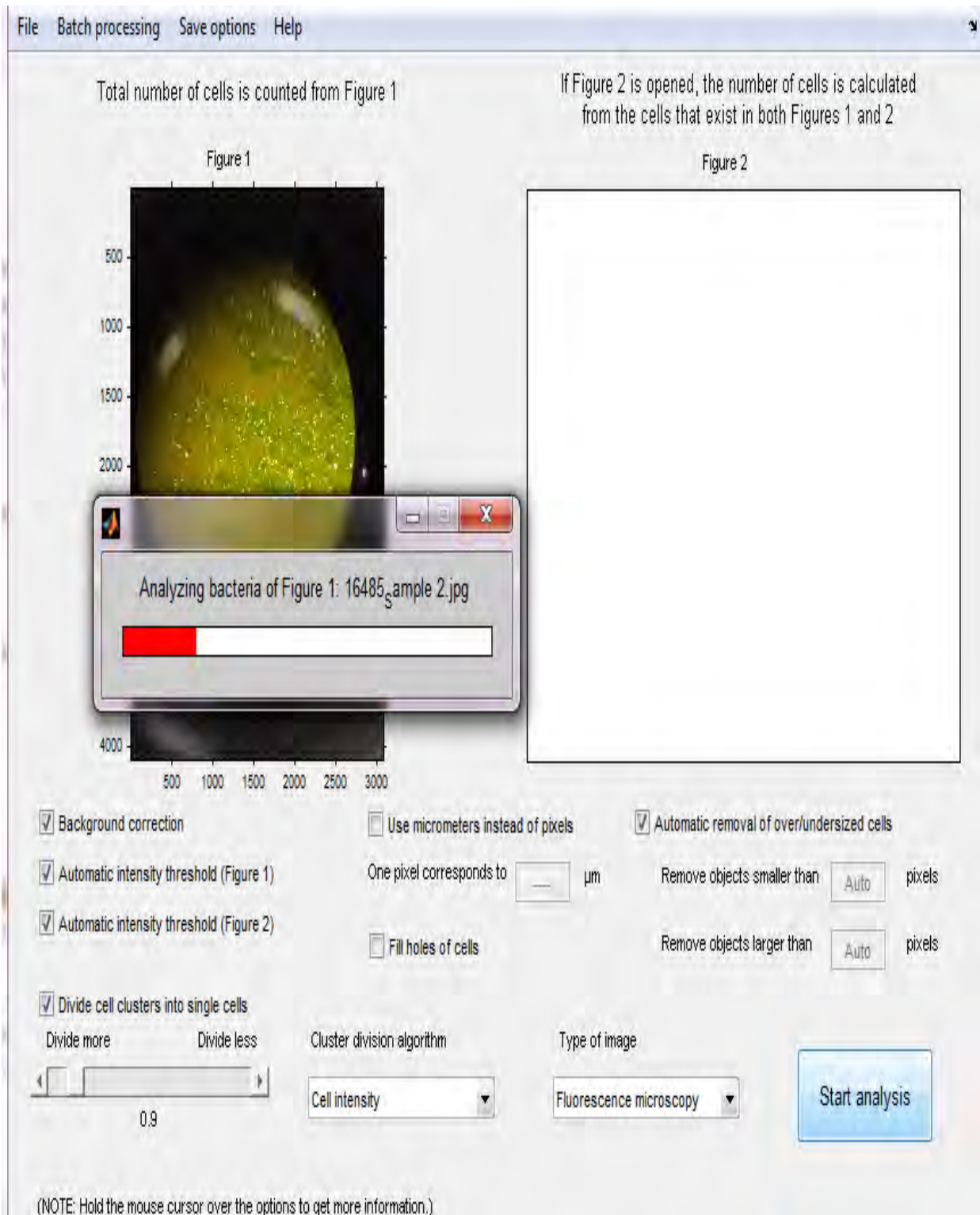Fig 5.1 Import the data set in the application
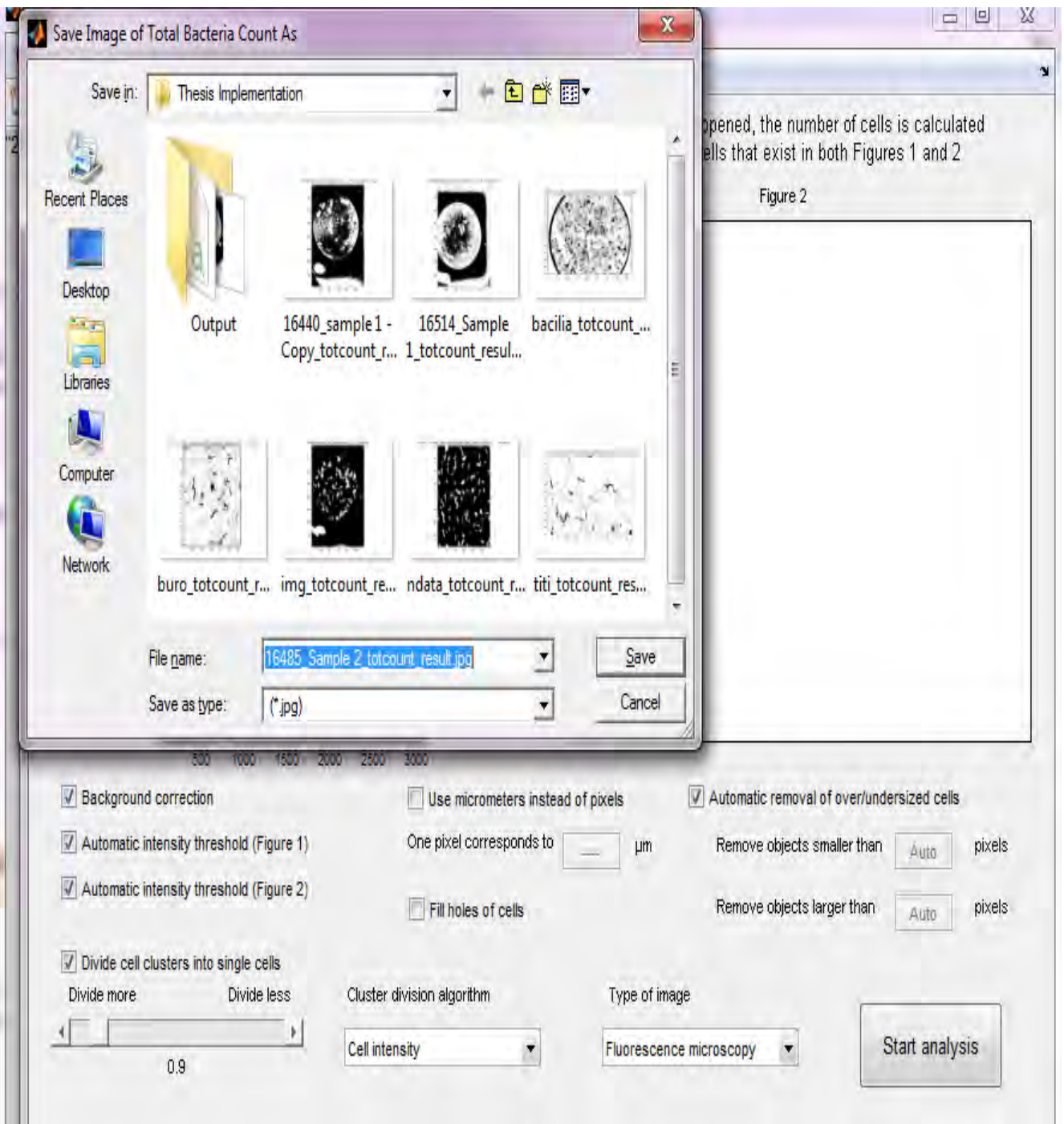
Fig 5.2 Analyzing the sample data

Fig 5.3 Prompting the user to save the output

Total bacillus bacteria count. Number of bacteria: 525 --- File: 16440$_s$ample 1 - Copy.jpg
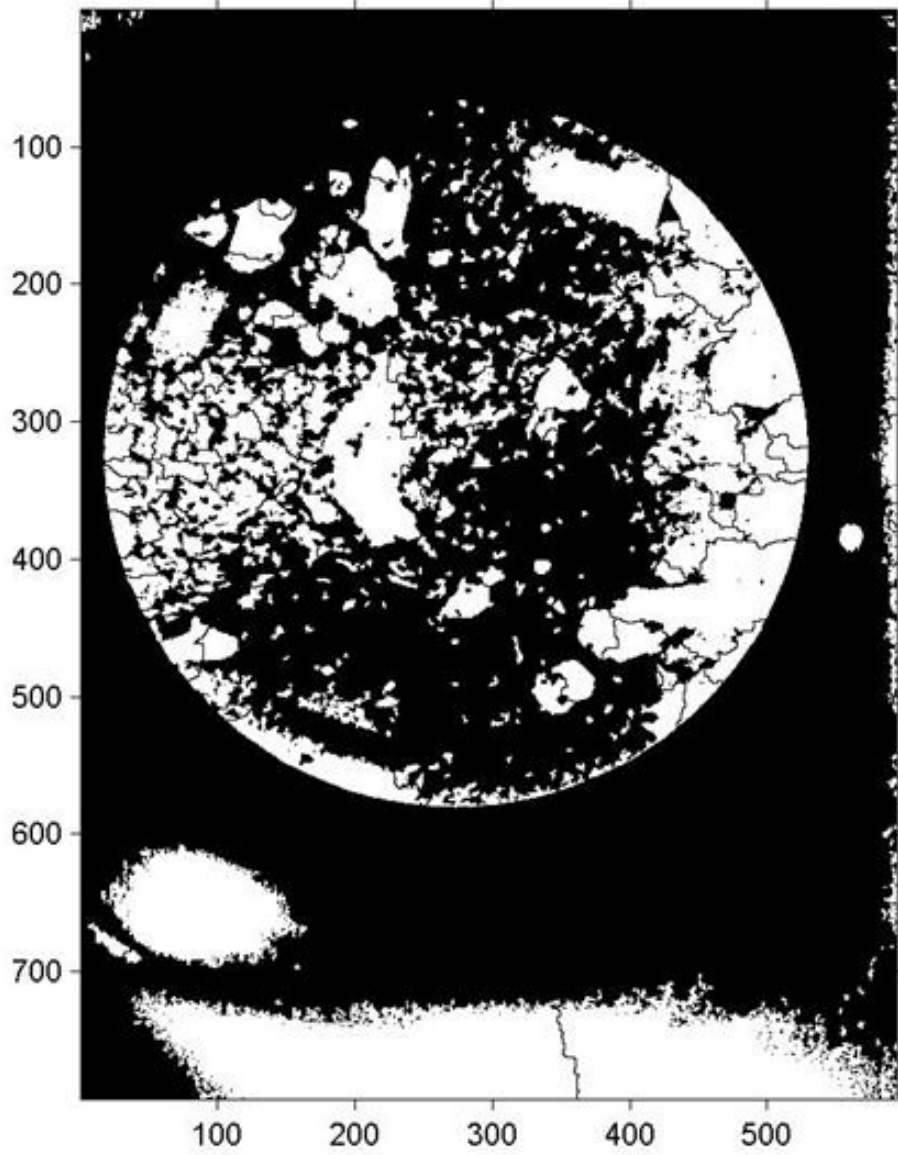

Fig 5.4 Sample output of dataset 1

You are in "2+" condition. Please visit a doctor Within a Week. Number of bacteria: 1645 --- File: 16485$_S$ample 2.jpg
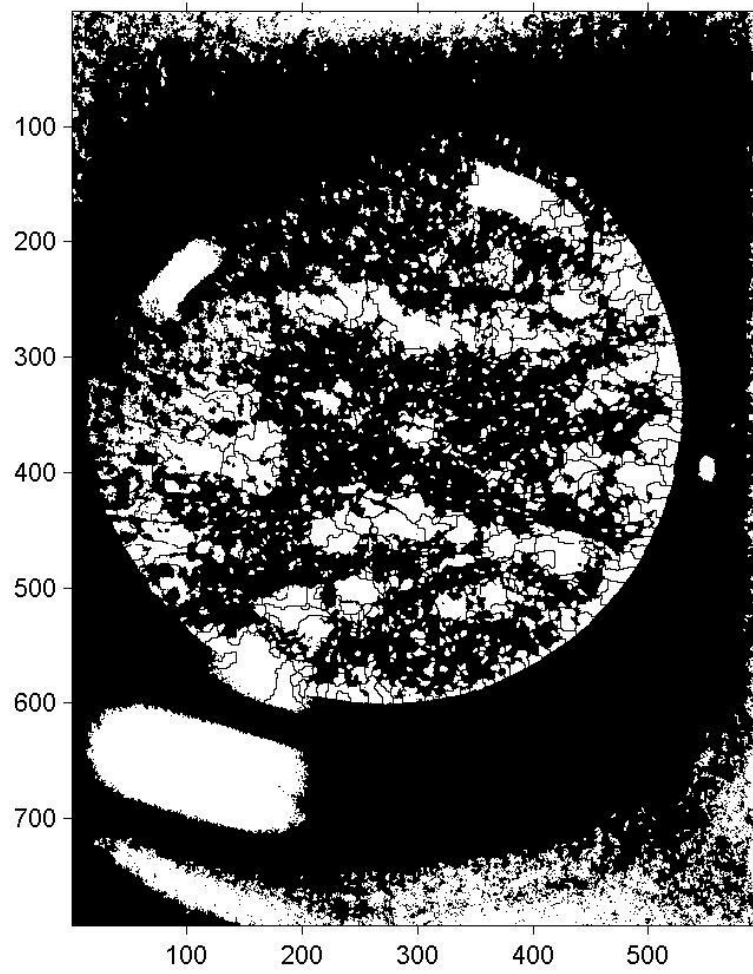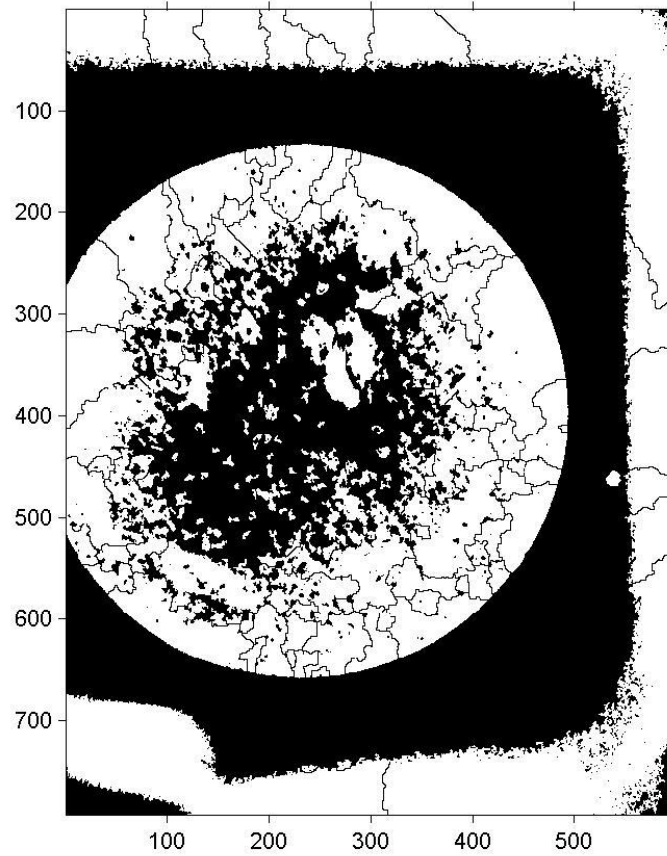
Fig 5.5 Sample output of dataset 2

Fig 5.6 Sample output of dataset 3

CHAPTER 6

6.1 Conclusion

We started our thesis about ten months ago and we are almost get our desire result in our project. However, there is always have some portion to improve. In our research program, there is many ups and down but we fight with them and accomplished our research. We have also discovered the scope for further improvement and extension of our work. In this final chapter, we have talked about the difficulties and future-plan before we draw a conclusion to this report.

In the conclusion, we can say that our program can detect Tuberculosis through morphology method and can give suggestion what should one do after their result. This process needs properly trained technicians and special equipment. However, in the developing countries, like Bangladesh, there are very few technicians and there is lack of proper equipment. Moreover, this process takes a quite long time to examine the samples. In rural areas, this kind of examination is time consuming and very much costly. In addition, there are quite chances of error in detection process. Therefore, keeping these things in mind we have decided to implement an algorithm that will help to detect TB through analyzing the image of spit of an ill people. The structure of our research has been program in this way that correct dataset and with minor alteration, it can easily find out the possibility of Tuberculosis of one person.

6.1.1 Difficulties

In our research time, we face some difficulties. The most important was about collecting data and images of coughs. Firstly, for collecting data we must get a permission from the heads of the National Institute of Diseases of the Chest and Hospital. Moreover, getting that permission

was not that easy. We have to wait for many hours in front of their offices. After one and half month later, we got our permission and after that, we started our data collection from their laboratory.

In time of collecting, we also faced problem. The image we were taking was not clear for processing. In addition, it was a big problem for us as it was the main data for our research. If we do not get that data properly than we will not able to start our research. Therefore, we started to find the solution and after searching, we found a solution. We found a special lens for the camera and after that, we could get a clear image of the coughs.

Another problem that, we face is that the images we get is not same size. We tried to resize it through MATLAB program but not all the pictures are in the center. Therefore, we have to resize it manually.

Though we get few problems, we successfully recover from it, finish our research in time, and make a program that can detect tuberculosis through image processing.

6.1.2 Future work

We develop a program that can easily detect Tuberculosis through image processing. It is unique in this area. There is very little research in this matter. It will be a great impact if we publish a mobile application. Therefore, our future-plan is that we will make a mobile application that can be both Android and IOS application. This application will be very helpful for people of all over the world. Through this app who are very poor, cannot go to the TB hospital that is in Dhaka, can have basic detection. The developing countries are going to be benefited by this application.

We are also planning for more research about Tuberculosis. We will collect more data from National Institute of Diseases of the Chest and Hospital so that we can give detection that is more correct. If we rich our database with image sample than we can make a better program. The more we train machine with data the more correct information we will get.

Therefore, our main future work is to develop a mobile application with more data collection.

REFERENCES

1. Kupper TH, Steffen K, Wekle G, Richartz G, Pfitzer P, Morphologicla study of bacteria of the respiratpry system using fluorescence microscopy of Papanicolaou stained smears with special regard to identification of Mycobacteria in sputum. Cytopath 1995; 6:388-402.

2. Sadaphal P, Rao J, Comstock GW, Beg M.F. Image processing techniques for identifying Mycobacterium tuberculosis in Ziehl-Neelsen stains. The International Journal of Tuberculosis and Lung Disease 2008: 12(5): 579-82.

3. Forero M.G, Cristobal G, Desco M .Automatic identification of Mycobacterium tuberculosis by Gaussian mixture models. Journal of Microscopy 2006: 120-132.

4. Osman M.K, Mashor M.Y, Jaafar H. Combining Thresholding and Clustering Techniques for Mycobacterium tuberculosis Segmentation in Tissue Sections. Australian Journal OF basic and Applied Sciences. 5, (12), pp. 1270. 1279. 2011

5. H.B. Rachna, Mallikaranjan Swamy M.S. Detection of Tuberculosis Bacilli using Image Processing Techniques. International Journal of Soft Computing and Engineering. 2013; ISSN: 2231-2307, Volume 3, Issue 4.

6. Costa Filho C.F.F, Levy P.C, Xavier C.M, Fujimito L.B.M, Costa Fernandes M.G. Automatic identification of tuberculosis mycobacterium. Research on Biomedical Engineering. 2015; Volume 31, Number 1, p 33-43.

7. Doi:https://sisu.ut.ee/imageprocessing/book/1

8. Doi:https://www.mathworks.com/help/stats/k-means-clustering.html

9. Er, O., Temurtas, F., and Tanrıkulu, A.C., Tuberculosis disease di- agnosis using artificial neural networks. J. Med. Syst. 34:299–302,2010. doi:10.1007/s10916-008-9241-x.

10. En.wikipedia.org. (2017). Tuberculosis. [online] Available at:

https://en.wikipedia.org/wiki/Tuberculosis#cite_note-9 [Accessed 14 Aug. 2017]

11. Alva, A., Aquino, F., Gilman, R., Olivares, C., Requena, D., Gutiérrez, A., Caviedes, L., Coronel, J., Larson, S., Sheen, P., Moore, D. and Zimic, M. (2017). Morphological Characterization of Mycobacterium tuberculosis in a MODS Culture for an Automatic Diagnostics through Pattern Recognition

.