بسم الله الرحمن الرحيم

Sudan University of Science and Technology

College of Graduates Studies

Biomedical Engineering Department

Automatic Method for Tuberculosis Bacilli Identification in Sputum Smear Microscopic Images Using Image Processing Techniques

طريقة تلقائية لتحديد عصيات السل في صور لطخة البصاق المجهرية باستخدام تقنيات معالجة الصور

A Thesis Submitted in partial fulfillment of the requirement for the M.Sc. Degree in Biomedical Engineering

Presented by:

Roa Osama Awad Altayeb

Supervised by:

Dr. Mawia Ahmed Hassan

June - 2016

الآية

بسم الله الرحمن الرحيم

﴿وَيَرَى الَّذِينَ أُوتُوا العِلْمَ الَّذِي أُنزِلَ إِلَيْكَ مِن رَّبِّكَ هُوَ الحَقّ وَيَهْدِي إِلَى صِرَاطِ العَزِيزِ الحَمِيدِ﴾

{سورة سبأ : الآية 6}

صدق الله العظيم

DEDICATION

I dedicated this thesis to my beloved greet mother. For her support, encouragement and prays that make me able to get such success and honor.

ACKNOWLEDGMENTS

First of all, I am grateful to almighty of god for completion of this master's thesis. I would like to thank my beloved family and friends for their encouragement, without their support it would be impossible to finish my master studies.

I would like to express my deepest gratitude to my supervisor, Dr **Mawia Ahmed hassan**, for his guides, advices help and correction.

Special appreciate and heartfelt thanks for my friend, Eng **Nura Abdallah** for her help in collecting the sputum smear images.

I would like also to thank all tuberculosis reference laboratory staff.

Table of contents

List of figures

List of tables

ABBREVIATIONS

ZN	Ziehl-Neelsen
TB	tuberculosis
HIV	human immunidefiency virus
AIDs	acquired immune deficiency syndrome
MTB	Mycobacterium tuberculosis
HBCs	high-TB burden countries
EPTB	extra-pulmonary tuberculosis
DST	Drug-susceptibility testing
MDR-TB	Multidrug-resistant tuberculosis
LTBI	latent TB infection
IGRAs	interferon-gamma release assays
FDA	Food and Drug Administration
BCG	bacilleCalmette–Guérin
EM	Electromagnetic
GUI	Graphic User Interface
CAD	Computer Aided Detection
ANNs	Artificial Neural Networks
SGLD	Spatial Gray Level Dependency

NBTF	the National Brain Tumor Foundation
PEs	Processing Elements
ADC	Analog-to-digital converter
CM YK	Cyan-magenta-yellow-black
HIS	hue, saturation, and lightness
HSV	hue, saturation, and value
PNNs	probabilistic neural networks
RBF	radial basis function
PDF	probability density functions
GRNN	General Regression Neural Networks
NN	Nearest neighbor
SVM	support vector machine

ABSTRACT

Tuberculosis is one of deadly disease in the world especially in developing countries. Sputum smear microscopy is the main diagnostic tool in developing countries and high TB burden countries for the detection of TB .previous studies show that the manual screening for sputum smear microscopic images lead to misdiagnosis and false result, Image processing techniques are applied in this research to enhance, segment and classify the sputum smear images for computerized process of TB bacilli identification. Image processing algorithms which used for sputum smear image include a series of enhancement techniques, segmentation methods and morphological operation. As the non-bacillus objects in sputum smear image can bias the detection, it should be suppressed from the smear image. This research employs color image segmentation technique for the segmenting the TB bacillus objects from the background.

TB bacillus objects are segmented from the background in two stages based on color space conversion and k-means clustering ,to identify the TB bacilli, the proposed method uses eleven shape feature descriptors, which are compactness, eccentricity ,area, perimeter and Hu moments ($M_1 to M_7$),and makes the judgment using a support vector machine .experimental result confirmed the superior performance of the proposed method .

المستخلص

السل واحد من الامراض المسببة للوفاة في العالم خاصة في الدول النامية. الوسيلة الاساسية للكشف عن السل في الدول النامية و في الدول ذات الانتشار العالي بالسل هي التصوير المجهري للطخة البصاق. الدراسات السابقة توضح ان الفحص اليدوي لصور لطخة البصاق المجهرية يؤدي الي خطأ في التشخيص و خطأ في النتائج ، تقنيات معالجة الصورة طبقت في هذا البحث لتحسين ، تقسيم و تصنيف صورة لطخة البصاق من اجل حوسبة عملية تحديد عصيات السل. خوارزميات معالجة الصورة المستخدمة لصورة لطخة البصاق تحتوي علي سلسة من تقنيات التحسين، طرائق التقسيم والعمليات المورفولوجيا.

الاجسام غير العصوية في صورة لطخة البصاق يمكنها تحيز الكشف، لذلك يجب ازالتها من صورة اللطخة. هذا البحث يستخدم تقنية تقسيم الصورة الملونة لاستخلاص الاجسام العصوية من الخلفية. الاجسام العصوية تستخلص من الخلفية في مرحلتين اعتمادا علي تحويل حيز اللون و ال K-means clustering لتحديد عصيات السل.

الطريقة المقترحة تستخدم احدى عشر واصف لملامح الشكل و هي عبارة عن (eccentricity ، مساحة العصيه، محيط العصيه، تراص العصيه، Hu moments ($M_1 to M_7$) ويتم الحكم باستخدام ال support vector machine. النتائج التجريبية توافق الاداء المسبق شرحه في الطريقة المقترحة.

CHAPTER ONE

INTRODUCTION

1.1 General overview

Tuberculosis is one of leading deadliest disease in the worldwide, caused by bacillus mycobacterium tuberculosis. And the diagnosis of this disease had different detecting methods depending on the tuberculosis type and the patient's status.

Early diagnosis of tuberculosis and initiating optimal treatment it not only enable a cure of an individual patient but also it will also curb the transmission of the infection and disease to another in the community.

For developing countries like our country the sputum smear microscopy still the main method for diagnosing tuberculosis. The main advantages of smear microscopy is simple and inexpensive also it is relatively easy to perform and read and detect transmitters of tubercle bacilli also its results can be reported within hours of receipt of sample and provide reliable epidemiological indicators for evaluation of patient status.

Smear microscopy is quick but requires a very high organism load for positivity as well as the expertise to read the satin sample. For avoiding the two above main limitation of using smear microscopy computer based method for identifying the TB bacilli in stain sample is required.

Automatic Image processing of the stained sputum smear digital images would reduce the burden on the pathologist or technician, reduce human error, and improve sensitivity of the test. It would reduce the time required

and need of lab technicians and allow preliminary detection of TB In remote or rural areas where pathologists or technicians are unavailable [1].

The rapid advancement of computer hardware, software, image processing algorithm and artificial intelligence has led to the development of various computer-aided systems for TB detection. The systems aim to assist medical technologist in the diagnostic process. Any automatic method for diagnosis TB disease was combination techniques between image processing and pattern recognition technique.

1.2 The Problem Statement

The Training and experience of the person examining the sputum smear are the one of the constraints that limits the identification of bacilli in smear microscopy. Which it can lead to misdiagnosis and false results, also the another limitations according to world health organization The number of ZN smears examined per microscopist per day should not exceed 20 as visual fatigue leads to a deterioration of reading quality; on the other hand, proficiency in reading ZN smears can only be maintained by examining at least 10-15 ZN smears per week.

1.3 Thesis Objectives

1.3.1 Overall Objective

- To implement an automatic method for tuberculosis bacilli identification in sputum smear images.

1.3.2 Specific Aims

- To assist the medical technologist in the diagnostic process.
- To convert to the process of bacilli identification from manual to automatic

- To reduce the time consumed in identifying the TB bacilli in smear image.

1.4 Methods

Design a computer based method for identifying the tuberculosis bacilli in sputum smear image to facilitating the process of diagnosis and reduces the time consumed when using the manual process. The automatic method for processing the sputum smear images was consisting of four stages. In the first stage a series of enhancement technique were performed for enhancing the ZN-stained images. Color space conversion and k-means clustering methods were used for color image segmentation step. In third stage a number of shape features were extracted to be an input for classification stage. The support vector machine was used for classification the extracted objects to two classes (TB bacillus objects and non-bacillus objects).

1.5 Thesis Organization

This Thesis is divided into six chapters, first chapter is an introduction, second chapter which reviews the previous studies and the existing literature in the field , third chapter will provide the theoretical background, forth chapter will describe the research methodology , the result obtained and discussion of these results given in the fifth chapter and in the last chapter will contain conclusion and recommendations future work should be considered.

CHAPTER TWO

2. THEORETICAL BACKGROUND

2.1 Introduction to the tuberculosis

Tuberculosis (TB) is second only to HIV/AIDS as the greatest killer worldwide due to a single infectious agent. In 2013 according to world health organization statistics 9 million people fell ill with TB and 1.5 million died from the disease.

Tuberculosis is a bacterial disease usually affecting the lungs (pulmonary TB) other parts of the body can also be affected, for example lymph nodes, kidneys, bones, joints, etc. (extra-pulmonary TB).

It can affect anyone of any age. People with weakened immune systems are at increased risk. Also it's spread through the air when a person with untreated pulmonary TB coughs or sneezes. Prolonged exposure to a person with untreated TB usually is necessary for infection to occur.

2.1.1 Mycobacterium tuberculosis Characteristics

Mycobacterium tuberculosis (MTB) is bar shaped straight or slightly tortuous. This bacterium has a measurement of width 0.3 – 0.6 mm and length 1-4 mm.

MTB is a small, slow-growing bacterium that can live only in people. It's is an aerobic bacterium, meaning it needs oxygen to survive. For this reason, during active TB disease, MTB complexes are always found in the upper air sacs of the lungs

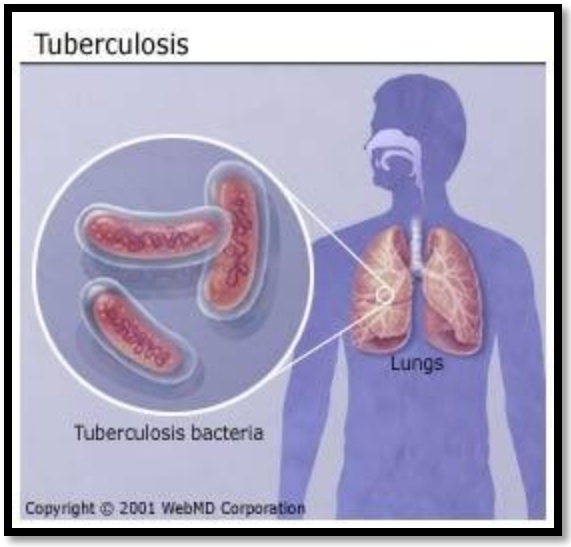

Figure 2.1: Mycobacterium tuberculosis shape

2.1.2Tuberculosis Infection

When a person breathes in MTB-contaminated air, the inhaled TB bacteria reach the lungs. This causes an MTB infection. However, not everyone infected with TB bacteria becomes sick. The bacteria can remain dormant (asleep) for years and not cause any TB disease. This is called latent TB infection. People who have latent TB infection do not get sick and do not spread the bacteria to others. But, some people with latent TB infection eventually do get TB disease. So it is important to get the appropriate treatment and get rid of the bacteria even in latent TB infection.

2.1.3 Tuberculosis symptoms

Pulmonary TB disease usually causes one or more of the symptoms indicated in Table 2.1. Extra-pulmonary TB disease may cause symptoms related to the part of the body that is affected. For example, TB of the spine may cause back pain; TB of the kidney may cause blood in the urine; TB meningitis may cause headache or confusion. Extra-pulmonary TB disease should be considered in the differential diagnosis of ill persons who have

systemic symptoms and who are at high risk for TB disease. Both pulmonary and extra-pulmonary TB disease symptoms can be caused by other diseases; however, they should prompt the clinician to consider TB disease.

Table 2.1

Symptoms of Pulmonary and Extra-pulmonary TB Disease

Symptoms of pulmonary TB disease	Symptoms of possible extra-pulmonary TB disease
• Cough (especially if lasting for 3 weeks or longer) with or without sputum production • Coughing up blood (hemoptysis) • Chest pain • Loss of appetite • Unexplained weight loss • Night sweats • Fever • Fatigue	• TB of the kidney may cause blood in the urine • TB meningitis may cause headache or confusion • TB of the spine may cause back pain • TB of the larynx can cause hoarseness • Loss of appetite • Unexplained weight loss • Night sweats • Fever • Fatigue

Most people infected with the germ that causes TB never develop active TB. If active TB does develop, it can occur two to three months after infection or years later. The risk of active disease lessens as time passes.

2.1.4 Tuberculosis diagnosis

There are several TB tests available to diagnosis TB. There are also TB tests to find out whether someone has TB bacteria that are susceptible to TB drug treatment or are drug resistant. TB tests to find out if someone has drug resistant TB are known as drug susceptibility tests. There are three main issues that impede effective TB diagnosis. First, many patients in high-TB burden countries (HBCs) do not have access to effective diagnosis. The WHO estimates that in 2013 only 5.6 million of 9 million incidents Global cases (64 %) were detected and notified [2]. Second, despite the availability of modern diagnostics, many of these tools have limited use in HBCs, which are mostly low and middle income countries. The costs and infrastructural requirements of many technologies can be prohibitive. Third, there are limitations associated with existing diagnosticscapacity to detect extra-pulmonary tuberculosis (EPTB), pediatric TB, TB in people living with HIV , and drug resistance. Drug-susceptibility testing (DST) is necessary for the detection of drug resistance, and initiation of appropriate treatment for MDR-TB. However, in 2013, only 8.3 % of new cases and 17 % of re-treatment cases underwent DST [3].

There are two kinds of tests that are used to determine if a person has been infected with TB bacteria: the tuberculin skin test and TB blood tests. A positive TB skin test or TB blood test only tells that a person has been infected with TB bacteria. It does not tell whether the person has latent TB infection (LTBI) or has progressed to TB disease. Other tests, such as a chest x-ray and a sample of sputum, are needed to see whether the person has TB disease.

Tuberculin skin test: The TB skin test (also called the Mantoux tuberculin skin test) is performed by injecting a small amount of fluid (called

tuberculin) into the skin in the lower part of the arm. A person given the tuberculin skin test must return within 48 to 72 hours to have a trained health care worker look for a reaction on the arm. The health care worker will look for a raised, hard area or swelling, and if present, measure its size using a ruler. Redness by itself is not considered part of the reaction.

The skin test result depends on the size of the raised, hard area or swelling. It also depends on the person's risk of being infected with TB bacteria and the progression to TB disease if infected.

- **Positive skin test:** This means the person's body was infected with TB bacteria. Additional tests are needed to determine if the person has latent TB infection or TB disease. A health care worker will then provide treatment as needed.
- **Negative skin test:** This means the person's body did not react to the test, and that latent TB infection or TB disease is not likely.

TB blood tests: TB blood tests (also called interferon-gamma release assays or IGRAs) measure how the immune system reacts to the bacteria that cause TB. An IGRA measures how strong a person's immune system reacts to TB bacteria by testing the person's blood in a laboratory.

Two IGRAs are approved by the U.S. Food and Drug Administration (FDA) and are available in the United States:

- QuantiFERON®–TB Gold In-Tube test (QFT-GIT)
- T-SPOT®.TB test (T-Spot)

Positive IGRA means that the person has been infected with TB bacteria. Additional tests are needed to determine if the person has latent TB infection or TB disease. A health care worker will then provide treatment as needed.

Negative IGRA means that the person's blood did not react to the test and that latent TB infection or TB disease is not likely.

IGRAs are the preferred method of TB infection testing for the following:

- People who have received bacilleCalmette–Guérin (BCG). BCG is a vaccine for TB disease.
- People who have a difficult time returning for a second appointment to look for a reaction to the TST.

Newer LTBI tests are being developed to overcome the limitations of existing tests. In 2015, Qiagen announced QuantiFERON®-TB Gold Plus (QFT®-Plus), the fourth generation IGRA which incorporates CD8+ T cell response data. Autoimmune Diagnostika (Straßberg, Germany) is currently developing an IGRA and IL-2 assay, which can distinguish between active and latent TB. Protein Logic is developing ImmiPrint® (ProteinLogic, Cambridge, UK) , a multiplexed antibody assay which gives a finger print to active and latent infection antigens, facilitating the differentiation between the two forms of tuberculosis. None of these newer LTBI assays have been validated yet, nor approved for clinical use [3].

Chest x-ray TB test:

Acute pulmonary TB can be easily seen on X-ray .however ,what it shows is not specific and normal chest X-ray cannot exclude extra pulmonary TB.

Also in countries where resources are more limited, there is often a lack of X-ray facilities.

2.2 sputum smear microscopy

It was a first technique that was early developed by German researcher, RobertKoch, which he used to identify the rod-shaped bacterium. smear microscopy is the process that dealing with collecting a biological sample (usually or some other material), fixing it thinly on the glass slide and then staining it with a dye that binds specifically to mycobacteria and the stained slide is examined under a microscope for signs of the TB bacteria, also this method is sometimes called AFB testing, because mycobacteria are 'acid-fast bacilli' (AFB) . this mean that the certain dyes adhere to the waxy coat of mycobacteria and remain visible even after rinsing with water and being briefly treated with a solution of acid- alcohol that strips the dye from the rest of the smear [4].

Sputum smear microscopy is inexpensive and simple, and people can be trained to do it relatively quickly and easily. In addition the results are available within hours. The sensitivity though is only about 50-60%.In countries with a high prevalence of both pulmonary TB and HIV infection, the detection rate can be even lower, as many people with HIV and TB co-infection have very low levels of TB bacteria in their sputum, and are therefore recorded as sputum negative [5].

Two techniques are used for TB diagnostic with sputum smear microscopy: Fluorescence microscopy and conventional microscopy. Fluorescence microscopy uses an acid-fast fluorochrome dye (eg, auramine O or auramine-rhodamine), while conventional microscopy uses the

carbolfuchsinZiehl-Neelsen - ZN or Kinyoun acid-fast stains. While the first one uses an intense light source, such as a halogen or high-pressure mercury vapor lamp, the second one uses a conventional artificial light source.

A Sputum stain for mycobacteria is a laboratory test performed on a sample of sputum , or phlegm . it's also known as an AFB stain or a TB smear . a doctor typically ordered the test to determine if a person has TB or another type of mycobacterial infection .

There are two procedures commonly used for acid fast staining :

1. Carbolfuchsin methods which include the Ziehl-Neelsen and kinyoun methods (light/bright field microscope)
2. Fluorochrome procedure using auramine-O or auramine-rhodamine dyes (fluorescent microscope).

2.2.1 Principe of Ziehl-Neelsen method of acid-fast staining

Mycobacteria, which do not stain well by gram stain, are stained with carbolfuchsin combined with phenol. In the 'hot' ZN technique, the phenol-carbolfunchsin stain is heated to enable the dye to penetrate the waxy mycobacterial cell wall. In the 'cold' technique known as kinyounmethod , stain are not heated but the penetration is achieved by increasing concentration of basic fuchsin and phenol and incorporating a 'wetting agent' chemical .

The stain binds to the mycolic acid in the mycobacterial cell wall. After staining, an acid decolorizing solution is applied. This removes the red dye from the background cells, tissue fibres, and any organisms in the smear expect mycobacteria which retain (hold fast to) the dye and are therefore referred to as acid fast bacilli (AFB).

Following decolorization, sputum smear is counterstained with malachite green, or methylene blue which stains the background material, providing a contrast colour which the red AFB can be seen.

2.2.2 Sputum sample collection &preparation

Due to overnight accumulation of secretions, first morning specimens are more likely to yield better recovery of AFB. they are two types of smear

- **Direct smear:** smear prepared directly from a patient specimen prior to processing.
- **Indirect smear :** smear prepared from processed specimen after centrifugation (is used to concentrate the material)

Reagents required for of the ZN-staining:

1. Carbolfuchsin stain (filtered)
2. Acid alcohol 3% v/v (or 20%sulfuric acid)
3. Malachite green 5g/l (0.5% w/v) or methylene blue , 5g/l

A. Ziehl-Neelsen staining procedure :

The sputum is spread evenly over the central area of the slide using a continuous rotational movement. The recommended size off the smear is about 20 mm by 10 mm. the slides are Place on dryer with smeared surface upwards, and air dry for about 30 minutes, then the dried smear is Fix with heat and it is Cover will with carbolfuchsin stain, the smear is heat until vapor just begins to rise (i.e., about 60 degree Celsius). Do not overheat (boil or dry), Add additional stain if necessary. Allow the heated stain to remain on the slide for 5 minutes. The stain is Wash off with clean water and Cover with 3% v/v acid alcohol for 2-5 minutes (or 20% sulfuric acid) or until the smear is sufficiently decolorized , i.e. pale pink ,and after it Wash well with

clean water. After that the stain is cover with malachite green stain for 1-2 minutes

1. Wash off stain with clean water, the back of the slide is wipe , and it is place in a draining rack for smear to air dry ,The smear is Examine microscopically, using the 100x oil immersion objective (10x eye piece for a total 1000X magnification) and it is scan the smear systematically.

Results:

- **AFB:** Red, straight or slighty curved rods, occurring singply or in small groups, may appear beaded.
- **Cells :**green
- **Background material**: green

Reagent	**Acid fast**	**Non-acid fast**
Carbolfuchsin with heat	Red (hot pink)	Red (hot pink)
Acid alcohol	Red	Colorless
Methylene blue/malachite	Red	Blue/Green

2.3 Light Microscopy:

Studies of the fine structure of biological cells and tissues require significant magnification for visualization of the details of interest .Useful magnification of up to ×1,000 may be obtained via light microscopy by the use of combinations of lenses. However, the resolution of light microscopy is reduced by the following factors :

- **Diffraction:** The bending of light at edges causes blurring; the image of a pinhole appears as a blurred disc known as the Airy disc .
- **Astigmatism:** Due to non-uniformities in lenses, a point may appear as an ellipse .
- **Chromatic aberration:** Electromagnetic (EM) waves of different wave length or energy that compose the ordinarily used white light converge at different focal planes, thereby causing enlargement of the focal point. This effect may be corrected for by using monochromatic light.
- **Spherical aberration:** The rays of light arriving at the periphery of a lens are refracted more than the rays along the axis of the lens. This causes the rays from the periphery and the axis not to arrive at a common focal point, thereby resulting in blurring. The effect may be reduced by using a small aperture.
- Geometric distortion: Poorly crafted lenses may cause geometric distortion such as the pincushion effect and barrel distortion.

  Whereas the best resolution achievable by the human eye is of the order Of 0.1- 0.2 mm, light microscopes can provide resolving power up to about 0.2µm [6].

2.4 Automated TB Diagnosis

Automated TB Diagnosis Is the technique, method, or system of operating or controlling a process of tuberculosis diagnosing and the bacilli identification process by highly automatic means, as by electronic devices, reducing human intervention to a minimum.

Automated TB diagnostic methods bring about several benefits in the fight against the disease. Some of the advantages of computer-based diagnosis are:

- Computer-assisted diagnosis can be faster than manual screening.
- A large number of slides can be analysed and hence the number of patients tested in any given amount of time is increased.
- Human involvement in the screening process is reduced. This will result in three desirable effects; less experienced staff can conduct the screening, labour costs are reduced and technicians suffer less fatigue, since the need to view the slides on the microscope is eliminated or reduced.
- Accuracy may be improved; in areas of high TB incidence like sub-Saharan Africa where a large number of patients have to be tested, technicians may test slides less thoroughly than required and smear positive cases may go unnoticed. Automated detection would ensure that the entire slide I examined in a short space of time and would thus reduce cases of misdiagnosis.

In essence, an automated TB detection process may allow a greater number of patients to be tested at a faster rate by fewer technicians and with higher accuracy [7].

2.5 Image processing techniques

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.

Many researches in tuberculosis were used different image processing technique in the process of enhancing the sputum smear microscopic images to have better bacilli segmentation also it were used in corporation with pattern recognition to facilities the process of classifying in input images in categorical form .

2.5.1 Image enhancement techniques:

It is the process of improving the interpretability or perception of information in images for human viewers, or to provide 'better' input for other automated image processing techniques.

Image enhancement techniques can be divided into two broad categories:

1. Spatial domain methods, which operate directly on pixels.
2. Frequency domain methods, which operate on the Fourier transform of an image.

Unfortunately, there no general theory for determining what is 'good' image enhancement when it comes to human perception. if it looks good , it is good ! , however , when image enhancement techniques' are used as pre-processing tools for other image processing techniques , then the quantities measures can determine which techniques are most appropriate .

2.5.2 Image representation

A color space is a mathematical representation of a set of colors. The three most popular color models are RGB (used in computer graphics); YIQ, YUV, or YCbCr (used in video systems); and CMYK (used in color printing). However, none of these color spaces are directly related to the intuitive notions of hue, saturation, and brightness. This resulted in the

temporary pur- suit of other models, such as HSI and HSV, to simplify programming, processing, and end- user manipulation.

2.5.2.1 RGB color space

The red, green, and blue (RGB) color space is widely used throughout computer graphics. Red, green, and blue are three primary additive colors (individual components are added together to form a desired color) and are rep- resented by a three-dimensional, Cartesian coordinate system in (Figure2.2). The indicated diagonal of the cube, with equal amounts of each primary component, represents various gray levels.

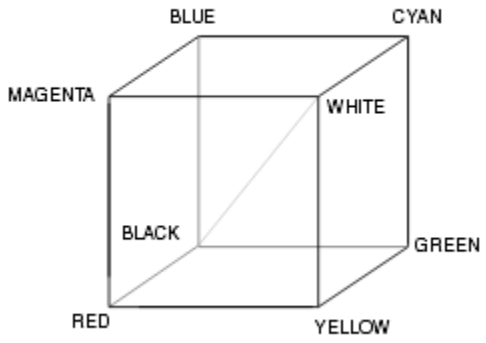

Figure2.2: The RGB Color Cube.

The RGB color space is the most prevalent choice for computer graphics because color displays use red, green, and blue to create the desired color. Therefore, the choice of the RGB color space simplifies the architecture and design of the system. Also, a system that is designed using the RGB color space can take advantage of a large number of existing software routines, since this color space has been around for a number of years. However, RGB is not very efficient when dealing with “real-world” images. All three RGB components need to be of equal bandwidth to generate any color within the

RGB color cube. The result of this is a frame buffer that has the same pixel depth and display resolution for each RGB component. Also, processing an image in the RGB color space is usually not the most efficient method. For example, to modify the intensity or color of a given pixel, the three RGB values must be read from the frame buffer, the intensity or color calculated, the desired modifications performed, and the new RGB values calculated and written back to the frame buffer. If the system had access to an image stored directly in the intensity and color format, some processing steps would be faster. For these and other reasons, many video standards use luma and two color difference signals. The most common are the YUV, YIQ,and YCbCr color spaces. Although all are related, there are some differences.

2.5.2.2 HSV Color model

The HSV (hue, saturation, value) color spaces was developed to be more "intuitive" in manipulating color and was designed to approximate the way humans perceive and interpret color. It was developed when colors had to be specified manually, and is rarely used now that users can select colors visually or specify Pantone colors.

The HSV color space is preferred for manipulation of hue and saturation (to shift colors or adjust the amount of color) since it yields a greater dynamic range of saturation. Figure 2.2 illustrates the single hexcone HSV color model the top of the hexcone corresponds to V = 1, or the maximum intensity colors. The point at the base of the hexcone is black and here V = 0. Complementary colors are 180°opposite one another as measured by H, the angle around the vertical axis (V), with red at 0 . The value of S is a ratio, ranging from 0 on the center line vertical axis (V) to 1 on the sides of

the hexcone. Any value of S between 0 and 1 may be associated with the point V = 0.The point S = 0, V = 1 is white. Intermediate values of V for S = 0 are the grays. Note that when S = 0, the value of H is irrelevant. From an artist's viewpoint, any color with V = 1, S = 1 is a pure pigment (whose color is defined byH). Adding white corresponds to decreasing S(without changing V); adding black corre-sponds to decreasing V (without changing S).Tones are created by decreasing both S and V.

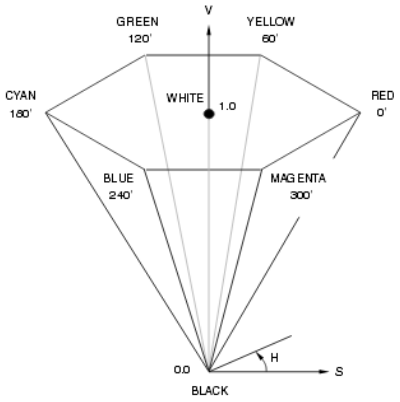

Figure2.3: SingleHexcone HSV Color Model 1

2.5.2.3 L*A*B color space

A Lab color space is a color-opponent space with dimension L for lightness and a and b for the color-opponent dimensions, based on nonlinearly compressed (e.g. CIE XYZ color space) coordinates.

The L*a*b* color space includes all perceivable colors, which means that its gamut exceeds those of the RGB and CMYK color models (for example, ProPhoto RGB includes about 90% all perceivable colors). One of the most important attributes of the L*a*b*-model is device independence. This means that the colors are defined independent of their nature of creation or

the device they are displayed on. The L*a*b* color space is used when graphics for print have to be converted from RGB to CMYK, as the L*a*b* gamut includes both the RGB and CMYK gamut. Also it is used as an interchange format between different devices as for its device independency. The space itself is a three-dimensional Real number space, that contains an infinite possible representations of colors. However, in practice, the space is usually mapped onto a three-dimensional integer space for device-independent digital representation, and for these reasons, the L*, a*, and b* values are usually absolute, with a pre-defined range. The lightness, L*, represents the darkest black at L* = 0, and the brightest white at L* = 100. The color channels, a* and b*, will represent true neutral gray values at a* = 0 and b* = 0. The red/green opponent colors are represented along the a* axis, with green at negative a* values and red at positive a* values. The yellow/blue opponent colors are represented along the b* axis, with blue at negative b* values and yellow at positive b* values. The scaling and limits of the a* and b* axes will depend on the specific implementation of Lab color, as described below, but they often run in the range of ±100 or −128 to +127. Unlike the RGB and CMYK color models, *Lab* color is designed to approximate human vision. It aspires to perceptual uniformity, and its L component closely matches human perception of lightness, although it does not take the Helmholtz–Kohlrausch effect into account. Thus, it can be used to make accurate color balance corrections by modifying output curves in the a and b components, or to adjust the lightness contrast using the L component. In RGB or CMYK spaces, which model the output of physical devices rather than human visual perception, these transformations can be done only with the help of appropriate blend modes in the editing application[8].

Because Lab space is much larger than the gamut of computer displays, printers, or even human vision, a bitmap image represented as Lab requires more data per pixel to obtain the same precision as an RGB or CMYK bitmap. This color space is better suited to many digital image manipulations than the RGB space, which is typically used in image editing programs. For example, the Lab space is useful for sharpening images and the removing artifacts in JPEG images or in images from digital cameras and scanners.

2.5.3 Image segmentation

The process of dividing an image into regions that could correspond to structural units, objects of interest, or ROIs it's called image segmentation, it is an important prerequisite for most techniques for image analysis [6].

Every pixel in an image is allocated to one of a number of these categories. A good segmentation is typically one in which:

- pixels in the same category have similar grayscale of multivariate values and form a connected region,
- Neighboring pixels which are in different categories have dissimilar values.

Segmentation is often the critical step in image analysis , If was done well then all other stages in image analysis are made simpler. There are three general approaches to segmentation, termed thresholding, edge-based methods and region methods.

2.5.3.1 Thresholding method

The main principle behind thresholding is that image pixels, falling into a predefined range of intensity values, are assigned a single intensity value, and the remaining pixels are assigned a different intensity value. A thresholding function can be formally defined as,

$$g(x,y) = \begin{cases} iiff(x,y) \in [L_i, L_u] \\ jiff(x,y) \in [L_i, L_u]' \end{cases} \quad \text{..... Eq 2.1}$$

Where $g(x,y)$ is the thresholded version of image $f(x,y)$ in equation (2.1), i and j are the two intensities used to differentiate between two groups of pixels, and L_i and L_u are the lower and upper limits of the intensity range used to define the two groups [Gonzalez]. Definition of the intensity range varies depending on the images and task in hand. One can simply define a single intensity value as a threshold or define several intensity ranges. The resultant image only contains two pixel intensities, representing a binary image. For simplification purposes, in a binary image it is common to use intensity values 0 (black) and 1(white) for i and j respectively.

The basic principle in thresholding makes it highly suitable for segmenting ROIs that will have distinct intensities from the rest of the image.

The major drawback to threshold-based approaches is that they often lack the sensitivity and specificity needed for accurate classification. The problem gets severe in case of multi-modal histograms with no sharp or well-defined boundaries. It is often difficult to define functional and statistical measures only on the basis of gray level value (histogram).this problem was processed in Region Growing based segmentation techniques, such as: Region splitting, Region merging, Split and Merge and Region growing techniques.

Homogeneity of regions is used as the main segmentation criterion in region growing. The criteria for homogeneity:

- gray level
- color
- texture

- shape
- model

2.5.3.2 Edge-based methods

Edge-based techniques rely on discontinuities in image values between distinct regions, and the goal of the segmentation algorithm is to accurately demarcate the boundary separating these regions.

2.5.3.3 region-based methods Segmentation

Region-based techniques rely on common patterns in intensity values within a cluster of neighboring pixels. The cluster is referred to as the region, and the goal of the segmentation algorithm is to group regions according to their anatomical or functional roles.

Region-based segmentation methods attempt to partition or group regions according to common image properties. These image properties consist of:

• Intensity values from original images, or computed values based on an image operator
• Textures or patterns that is unique to each type of region
• Spectral profiles that provide multidimensional image data

2.6 K-means clustering segmentation method

The k-means algorithm partitions a collection of N vectors $x_i, i = 1, \dots, N$ into K groups $G_k, k = 1, \dots, K$, and finds a cluster center in each group such that a cost function of dissimilarity (or distance) measure is minimized. When the Euclidean is chosen as the dissimilarity measure between a vector x in the k^{th} group and the corresponding cluster center c_k, the cost function can be defined by

$J = \sum_{K=1}^{K} J_K = \sum_{k=1}^{k} \sum_{x \in G_k} d(x, c_k)$ ..... Eq 2.2

where J_K is the cost function within group k and d is a generic distance function

$$d(x, c_k) = (x - c_k)^T M (x - c_k)$$

Where M is the distance matrix. the partitioned groups are typically defined by a membership $[K \times M]$ matrix U where the element u_{ki} is 1 if the i^{th} data point x_i belongs to group k , and 0 otherwise . the matrix U satisfies the following condition :

$\sum_{k=1}^{K} u_{ki} = 1 \qquad \forall_i = 1, \dots, N \sum_{K=1}^{K} \sum_{i=1}^{N} u_{ki} = N$ .....Eq 2.3

Once the cluster centers c_k are fixed, the terms u_{ki} which minimize Eq 2.2 are :

$$u_{ki} = \begin{cases} 1 & if d(x_i, c_k) \leq d(x_i, c_j), foreach j \neq k \\ 0 & otherwise \end{cases}$$ .....Eq 2.4

This means that x_i belongs to the group k if c_k is the closet center among all centers.

2.7 Morphological Operations

Morphological operations used in digital image processing are a way of extracting image components that can be used to express details about a regions shape, its boundaries, its area and so on[9].

Morphological operations are affecting the form, structure or shape of an object .it was applied on binary images (black & white images with only 2 colors). They are used in pre or post processing (filtering, thinning, and pruning) or for getting a representation or description of the shape of objects/regions (boundaries, skeletons convex hulls).

The two principal morphological operations are dilation and erosion. Dilation allows objects to expand, thus potentially filling in small holes and connecting disjoint objects. Erosion shrinks objects by etching away (eroding) their boundaries. These operations can be customized for an application by the proper selection of the structuring element, which determines exactly how the objects will be dilated or eroded

2.7.1 Dilation

Dilation is one of the two basic operators in the area of mathematical morphology, the other being erosion. It is typically applied to binary images, but there are versions that work on grayscale images. The basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels (*i.e.* white pixels, typically). Thus areas of foreground pixels grow in size while holes within those regions become smaller.

The dilation operator takes two pieces of data as inputs. The first is the image which is to be dilated. The second is a (usually small) set of coordinate points known as a structuring element (also known as a *kernel*). It is this structuring element that determines the precise effect of the dilation on the input image.

2.7.2 Erosion

This operation is the sister of dilation. What this does is to compute a local minimum over the area of the kernel.

Erosion operation is a morphological operation for reducing the foreground area. The effect of this operation is shrunk foreground. The foreground is reduced from its outer edge to inside its area. If there is a hole inside the foreground area, the hole enlarges. It uses a structuring element and it is

done with a convolution operation between the image and the structuring element. This operation is for binary images.

The erosion process will set the foreground pixel to be background if there is part of the structuring element that reaches the background while the center of the structuring element reach the foreground edge.

2.7.3 Region filling

Region filling is the process of "coloring in" a definite image area or region. Regions may be defined at the pixel or geometric level. At the pixel level, we describe a region either in terms of the bounding pixels that outline it or as the totality of pixels that comprise it. In the first case the region is called boundary-defined and the collection of algorithms used for filling such a region is collectively called boundary-fill algorithms. The other type of region is called an interior-defined region and the accompanying algorithms are called flood-fill algorithms. At the geometric level a region is defined or enclosed by such abstract contouring elements as connected lines and curves.

There are two ways in which pixels are considered connected to each other to form a "continuous" boundary. One method is called 4-connected, where a pixel may have up to four neighbors; the other is called 8-connected, where a pixel may have up to eight neighbors.

2.7.4 Extraction of connected component

Extraction of connected components from a binary image is important for many automated image analysis applications. Let A be a set containing one or more connected components. We form an array X_0(of the same size as the array containing$A$), whose elements are zeros (background values), except at each location known to correspond to a point in each connected

component in A, which we set to one (foreground value). The objective is to start with X_0 and finds all the connected components by the following iterative procedure:

$X_k = (X_{k-1} \oplus \mathrm{B})$……..……Eq 2.5

2.8 feature extraction

In machine learning, pattern recognition and in image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is related to dimensionality reduction.

When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed into a reduced set of features (also named a features vector). This process is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.

One very important area of application is image processing, in which algorithms are used to detect and isolate various desired portions or shapes (features) of a digitized image or video stream. It is particularly important in the area of optical character recognition. can be divided in the following types:

1. Low-level

    - Edge detection
    - Corner detection
    - Blob detection
    - Ridge detection
    - Scale-invariant feature transform

2. Curvature

    - Edge direction, changing intensity, autocorrelation.

3. Image motion

    - Motion detection. Area based, differential approach. Optical flow.

4. Shape based

    - Thresholding
    - Blob extraction
    - Template matching
    - Hough transform
        - Lines
        - Circles/ellipses
        - Arbitrary shapes (generalized Hough transform)
        - Works with any parameterizable feature (class variables, cluster detection, etc..)

5. Flexible methods

    - Deformable, parameterized shapes
    - Active contours (snakes)

2.8.1 Shape feature extraction

Shape is known as an important cue for human beings to identify and recognize the real-world objects, whose purpose is to encode simple geometrical forms such as straight lines in different directions. Shape feature extraction techniques can be broadly classified into two groups contour based and region based methods. The former calculates shape features only from the boundary of the shape, while the latter method extracts features from the entire region.

In addition, spatial relationship is also considered in image processing, which can tell object location within an image or the relationships between objects. It mainly includes two cases: absolute spatial location of and relative locations of regions [10].

The following Figure shows the hierarchy of the classification of shape feature extraction approaches excerpted from the corresponding literature [11].

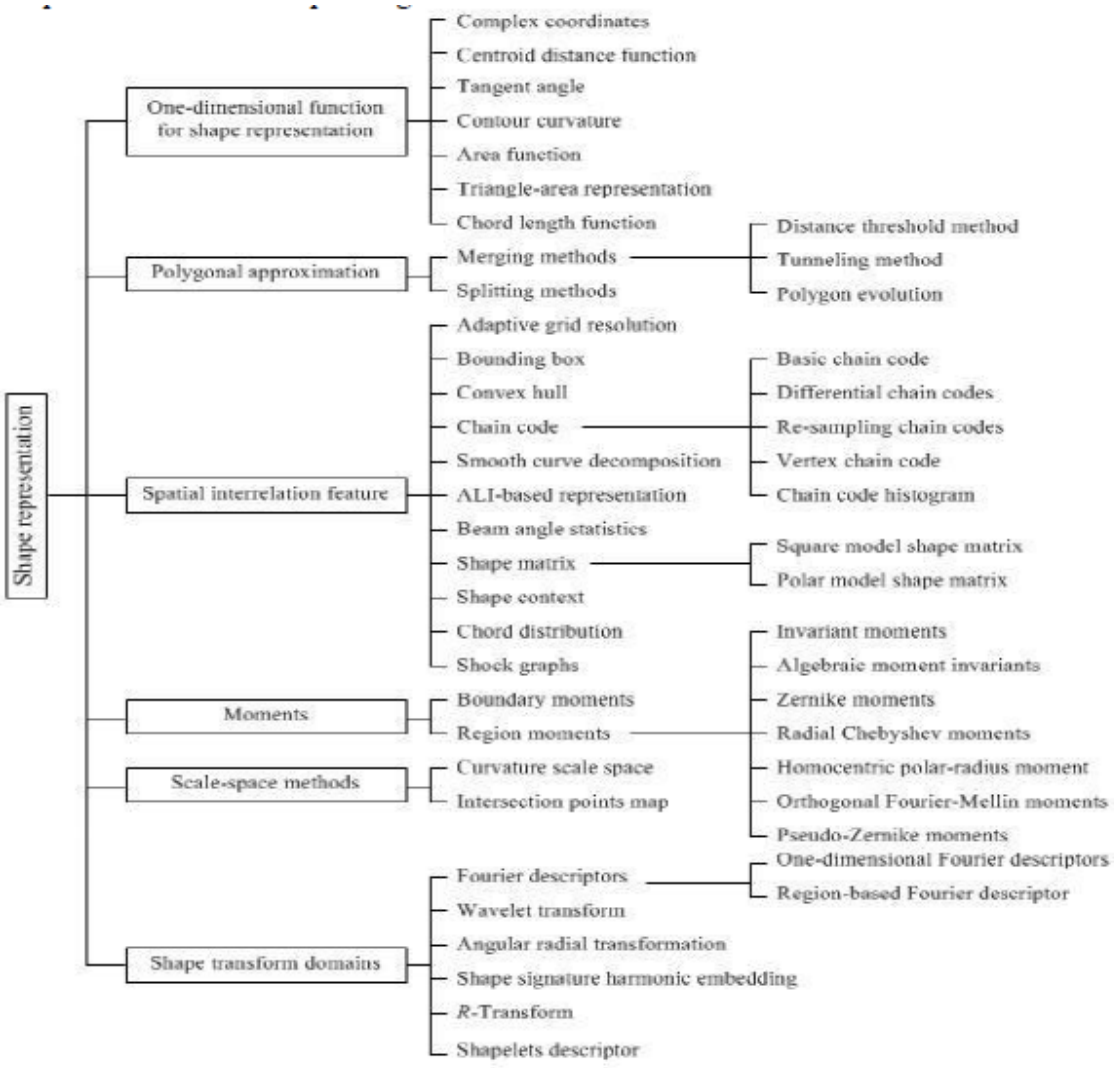

Figure 2.4: an overview of shape description techniques [11]

2.9 Image classification

Classification includes a broad range of decision-theoretic approaches to the identification of images (or parts thereof). All classification algorithms are based on the assumption that the image in question depicts one or more features (*e.g.*, geometric parts in the case of a manufacturing classification system) and that each of these features belongs to one of several distinct and exclusive classes. The classes may be specified a priori by an analyst (as in supervised classification) or automatically clustered (i.e. as in unsupervised

classification) into sets of prototype classes, where the analyst merely specifies the number of desired categories. (Classification and segmentationhave closely related objectives, as the former is another form of component labeling that can result in segmentation of various features in a scene.

Image classification analyzes the numerical properties of various image features and organizes data into categories. Classification algorithms typically employ two phases of processing: training and testing. In the initial training phase, characteristic properties of typical image features are isolated and, based on these, a unique description of each classification category, i.e. training class, is created. In the subsequent testing phase, these feature-space partitions are used to classify image features.

The description of training classes is an extremely important component of the classification process. In supervised classification, statistical processes (i.e. based on an a priori knowledge of probability distribution functions) or distribution-free processes can be used to extract class descriptors. Unsupervised classification relies on clustering algorithms to automatically segment the training data into prototype classes. In either case, the motivating criteria for constructing training classes are that they are:

- independent, i.e. a change in the description of one training class should not change the value of another,
- discriminatory, i.e. different image features should have significantly different descriptions, and
- reliable, all image features within a training group should share the common definitive descriptions of that group.

In machine learning and statistics, classification is the problem of identifying to which of a set of categories (sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known.

A large number of algorithms for classification can be phrased in terms of a linear function that assigns a score to each possible category k by combining the feature vector of an instance with a vector of weights, using a dot product. The predicted category is the one with the highest score.

Examples of such algorithms are

- Logistic regression and Multinomial logistic regression
- Probit regression
- The perceptron algorithm
- Support vector machines
- Linear discriminant analysis.

Brief descriptions of number of classifiers that are used in tuberculosis bacilli classification are presented as follow:

2.9.1 PNN classifier

PNNs (probabilistic neural networks) are radial basis function (RBF) networks that replace each data point by a kernel. The probability density functions (PDF) of the training data are estimated from these kernels. The class whose PDF dominates at the position of a query object is estimated and its label is assumed by the query object (Khutlang, 2009). Support vector machines map the given dataset into a high-dimensional feature space.

Probabilistic (PNN) and General Regression Neural Networks (GRNN) have similar architectures, except there is a fundamental difference. General Regression neural networks perform regression where the target variable is continuous, whereas Probabilistic networks perform classification where the target variable is categorical. All PNN networks have four layers:

1. Input layer: There is one neuron in the input layer for each predictor variable. For

the case of categorical variables N-1 neurons are used where N is the number of categories. Input neurons (or processing before the input layer) standardizes the range of the values by subtracting the median and dividing by the inter quartile range. Then input neurons feed the values to each of the neurons in the hidden layer.

2. Hidden layer: This layer has one neuron for each case in the training data set. Neuron stores the values of the predictor variables along with the target value for the case. When presented with the x vector of input values from the input layer, the Euclidean distance of the test case from the neuron's center point is computed by hidden neuron and then using the sigma value(s) apply the RBF kernel function. Resulting value is passed to the neurons in the pattern layer.

3. Pattern layer / Summation layer: The next layer in the network is different for PNN and for GRNN. For PNN networks for each category of the target variable there is one pattern neuron. The actual target category of each training case is stored with each hidden neuron; the weighted value coming out of a hidden neuron is fed only to the pattern neuron that corresponds to the hidden neuron's type. Pattern neurons add the values for the class they represent (hence, it is a weighted vote for that category). For GRNN networks, in the pattern layer there are only two neurons. One neuron is the

denominator summation unit the other is the numerator summation unit. The weight values coming from each of the hidden neurons adds up by denominator summation unit. Numerator summation unit adds up the weight values multiplied by the actual target value for each hidden neuron.

4. Decision layer: The decision layer is different for PNN and GRNN. For PNN networks the decision layer compares the weighted votes for each target category accumulated in the patter n layer and uses the largest vote to predict the target category [12].

2.9.2 Linear classifier

The linear regression classifier establishes a linear mapping between stored points and their labels. The mapping is used to predict labels of query points. The error between classes in the least square sense is minimized by the mapping using the Euclidean distance. The logistic linear classifier is attained by iteratively reweighting the least squares solution to the plot of a line separating the two classes.

2.9.3 Quadratic discriminant classifier

A quadratic mapping is established between objects and their labels using the information like mean and covariance estimated from the training data points. The quadratic classifier presumes the classes of the dataset to have normal density functions. It is a density based classifier which uses covariance matrix and mean vector to estimate the density distribution for each class.

2.9.4 nearest neighbor classifier

The nearest neighbor classifier compares the Euclidean distance between each query point and the stored objects to predict the labels of the query

points. The query point gets assigned the label of the stored point that has the lowest Euclidean distance. When k points are compared, the NN becomes kNN classifier. Then the query point is assigned the label of the majority of its classifiers.

2.10 Classifier validation

Cross validation is used to evaluate learning algorithms by dividing data into a training set and test set. A classifier would use the training set to learn a model and the test set to validate the classifier performance. For example, a dataset containing 100 objects can be divided into 10 subsets where one subset is used for training and tested on another. The test set is then added on to the training set and the new training set can be validated on another testing set.

When this is done 10 times, it is called 10-fold cross-validation. In general, k-fold cross-Validation may be used to evaluate the classifier.

2.11 SVM classifiers

The statistical learning theory is the backbone of support vector machine , it' Wipe s provides a new framework for modeling learning algorithms ,merges the fields of machine learning and statistics , and inspires that overcome all of the above difficulties .

The SVM technology is one of the advanced statistical methodologies such as neural network. Neural network suffer from a limited ability to handle data and can only analyze the form two or three dimensions. Support vector machines, however, are able to process infinite amounts of data and to analyze the data to find separations and delineations high dimensionality [13].

Support vector machines map the given dataset into a high-dimensional feature space. They then try to locate in that space a plane that separates the two classes

CHAPTER THREE

3. LITERATURE REVIEW

Tuberculosis is a bacterial disease that can attack any part of the body, most frequently the lungs. The bacterium, Mycobacterium tuberculosis, is an airborne bacterium, meaning that it is passed from person to person through the air, i.e. via coughing, sneezing, laughing or even just talking. So that the number of people that have been infected with TB is rapidly increase , according to world health organization (WHO)annual report on global control of tuberculosis published in 2014, there were an estimated 9.0 million new cases of tuberculosis (TB) incidence in 2013 and about 1.5 million TB deaths.[2]

Gaps in TB case finding and the emergence of drug-resistant TB have created a pressing need for robust and accurate diagnostics. Newer tests such as Xpert MTB/RIF are having an impact in increasing case detection and reducing time to treatment [14] . The first methods for automatic bacilli screening in conventional microscopy were published only in 2008 by:

Costa, et al [15] they have developed an automatic identification method of mycobacterium tuberculosis with conventional microscopy images based on Red and Green color channels using global adaptive threshold segmentation. Differing from fluorescence microscopy, in the conventional microscopy the bacilli are not easily distinguished from the background. The key to the bacilli segmentation method employed in their work is the use of Red Minus Green (R-G) images from RGB color format, the bacilli appear as white regions on a dark background. Some artifacts are present in the (R-G)

segmented image. To remove them we used morphological, color and size filters. The best sensitivity achieved was about 76.65%. The main contribution of this work was the proposal of the first automatic identification method of tuberculosis bacilli for conventional light microscopy.

P. **Sadaphal ,et al [16]** proposed an Automated, multi-stage, color-based Bayesian segmentation identified possible 'TB objects', has been used to remove artifacts by shape comparison and color-labeled objects as 'definite', 'possible' or 'non-TB', by passing photomicrographic calibration. Superimposed AFB clusters, extreme stain variation and low depth of field were challenges. Their novel method facilitated electronic diagnosis of TB, permitting wider application in developing countries where fluorescent microscopy is currently inaccessible and unaffordable.

Khutlang et. al [17] The research illustrate that the Screening for tuberculosis (TB) in low- and middle-income countries is centered on the microscope. We present methods for the automated identification of Mycobacterium tuberculosis in images of ZN stained sputum smears obtained using a bright-field microscope. They segment candidate bacillus objects using a combination of two-class pixel classifiers. The algorithm produced results that agree well with manual segmentations, as judged by the Hausdorff distance and the modified Williams index. The extraction of geometric-transformation-invariant features and optimization of the feature set by feature subset selection and Fisher transformation follow. Finally, different two-class object classifiers are compared.

M. K. Osman , et al [18] presented a procedure for automated detection of tuberculosis bacilli in Ziehl-Neelsen-stained tissue slides using image

processing and neural network. Image segmentation using CY-based color filter and k-mean clustering procedure is used to separate objects of interest from the background. A number of geometrical features are then extracted from the segmented images. A modified training algorithm called Extreme Learning Machine (ELM) was used to train a hybrid multilayered perceptron network (HMLP) for the classification task.

Shan-e-Ahmed Raza , et al [19] A key component of their proposed algorithm is the enhancement of raw input image using anisotropic tubular filtering (ATF) which suppresses the background noise while simultaneously enhancing strong anisotropic features of AFBs presented in the image. For segmentation they were used color features and candidate AFBs were identified. Finally, a support vector machine classifier using morphological features from candidate AFBs decided whether a given image was AFB positive or not.

KusworoAdi, et al [20] In their research, represent an algorithm to identifying and counting the number of tuberculosis bacteria has been developed by microscope imaging. Color segmentation was done by way of extracting the Saturation channel of NTSC (Luminance, Hue, saturation) color model. Thresholding process was using Otsu Method. Feature extraction for bacteria shape identification process using two parameters i.e. eccentricity and compactness. The training and object recognition using Support Vector Machine algorithm.

Rachna H.B., et al [21] in their work ,developed an algorithm based on image processing for identification of TB bacteria in sputum . the method based on Otsu thresholding and k-means clustering and show good accuracy and efficiency .

CHAPTER FOUR

4. METHODOLOGY

Tuberculosis mycobacteria is spread primarily as airborne aerosol from an individual who is in infection stage of TB .so that the process of early and accurate diagnosis remains a significant challenge, particularly in resource poor setting , although the new TB tests are becoming available , they are too expensive for developing countries , also the another limitation it is require significant laboratory facilities ,including the availability of highly trained staff, hence came the need for updating of computer-based methods for the diagnosis of tuberculosis .

Smear microscopy is not a very sensitive technique and requires that the patient comes at least twice to the clinic – but it is also relatively fast, inexpensive, and specific for TB in high incidence areas. Thus, in the absence of better alternatives, it is a useful tool in the basic laboratories common in developing countries. Many researchers have focused on new tools and methods to improve the sensitivity and/or efficiency of microscopy [2].

There are two main reasons why sputum smear microscopy is appropriate for TB diagnosis. Special dyes allow for differentiating the bacillus from the background, and there is a positive correlation between the number of bacilli in the smear and the probability of them being identified by microscopy.

the proposed method to automate the TB bacilli identification using sputum smear microscopy images . It consists of five main steps: image acquisition,

image enhancement, image segmentation, feature extraction and classification. Figure 4.1 illustrate the five steps respectively.

4.1 Image Acquisition

A total of 280 (negative and positive) sputum smear images was collected using the both of two microscopic diagnostic techniques (fluorescence and conventional microscopy) .

Images were taken using a ZEISS iLED microscope and NIKON D3100 digital camera. The pixel resolution was 3456×2304. The images were stored in JPEG file format, with 24 bits per pixel, in color.

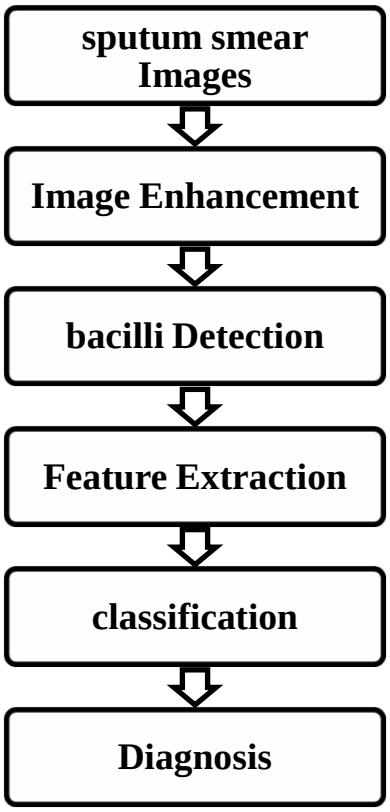

Figure 4.1: Block diagram of our proposed method

Sputum smear slides were prepared by tuberculosis reference Laboratory (TRL)at national laboratory of public health in Khartoum, a total of 144 auramine-stained sputum smear was excluded the reason of bad image resolution (motion artifacts) that had been occur when capturing the image .

Figure 4.3.: Nikon camera and ZEISS iLED microscope that were used for image acquisition

All the following algorithms were implemented on matlab 8.1.0.604 (R2013a) which runs on a pc computer.

4.2 Image Preprocessing

Image pre-processing step was been perform in order to reduce the image information to be analyzed as it would be an extremely slow and very computationally expensive for matlab to process the whole image as it is originally given by the camera .

All unnecessary information such as the object of no interest are removed .after implementing this task the images were also resized to 500×750 according to the large image dimensions so that to be easily fitted and processed .

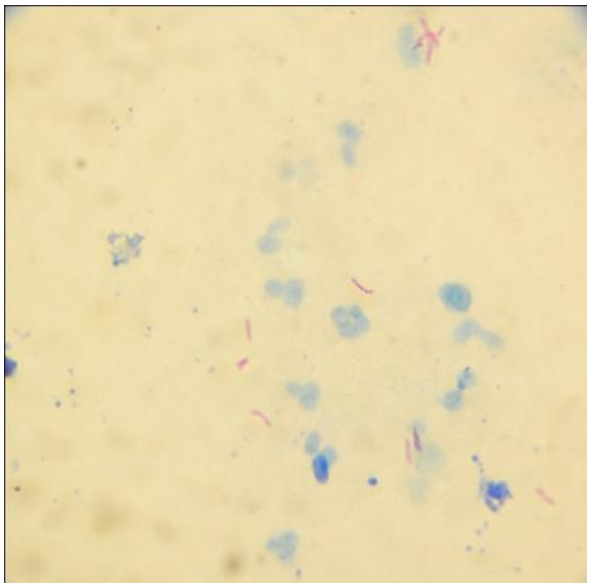

Figure 4.2: Example ZN-stained sputum smear image**.**

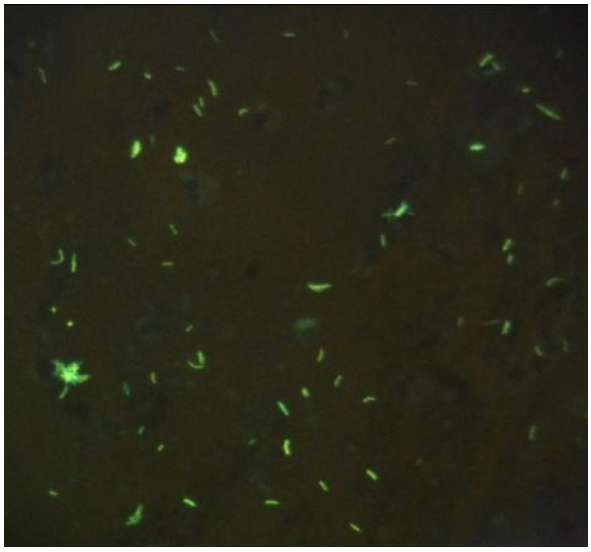

Figure 4 3: Example auramine-stained sputum smear image .

4.3 Image Enhancement

The image enhancement step is the backbone of our proposed method. it concern with adjusting the sputum smear images so that the result are more suitable for further image analysis steps and identifying the bacilli features .

Three techniques were used for enhancing the Ziehl-Neelsen Stainimage De-correlation Stretching, Gaussian filter and contrast stretching respectively, figure 4.4 below represented all the enhancement steps .

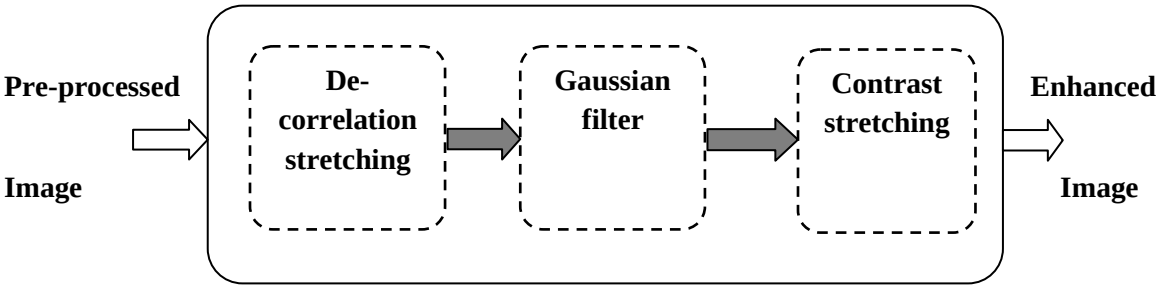

Fig.4.4 : Block diagram of image enhancement step

ZN- Stained *image* has more complex ground detail, so that the object (TB Bacteria) must be separated by the ground before identifiable. For that reason the De-**correlation Stretching** was used toincrease color separation of an image significantlyin the multi-channel image. This way increase the visualoriginal color is charted to a new set of color value with a wider range.[9] after that the **Gaussian filter** was implemented to diminish the effects of camera noise and spurious pixel value also it was used to reduce the sputum smear image details , further more it proposed as preprocessing stage in our computers vision algorithm in order to enhance the image structure at different scales . Finally the **contrast stretching** was used to improve the contrast in an image without distorting relative level of intensities [22].

4.4 Color image segmentation and bacilli detection:

The ZN-stained bacilli are reddish in colour on a blue background. Thereforecolour features may be useful to distinguish TB bacilli in ZN-sputum images, So that the color image segmentation had been used.

The segmentation of ZN-stained image was implemented is two stage .the figure below illustrate the two stages .

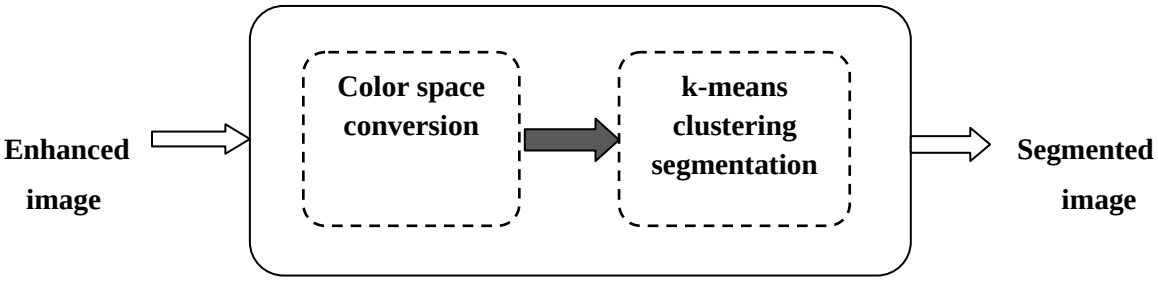

Figure 4.5: illustrate Segmentation Stage

4.4.1 Color image processing

The color image segmentation stage began with conversion the RGB image to another color model. Depending on the staining procedure the TB bacilli assume different colors may vary from light fuchsia to purple so that, There were different types of color spaces used in color image segmentation for our sputum smear images .

L*A*B and HSV color space were the two frequently chosen color space for color image segmentation. Our proposed method was converted the RGB enhanced image to the LAB when the background of enhanced image was off-white to identify different colors in the image so to be easily segmented in next stage, otherwise the image was converted to HSV color space.

4.4.2 Segmentation

Segmentation is an important step in image processing which aids extraction of information and attributes from images for image understanding and interpretation.

The distinct characteristics of ZN-stained sputum smears, which contain red bacilli against a blue and off-white background, present a useful property that can be exploited for segmentation of bacilli.

previous researches showed that the a more convincing segmentation performance has been achieved by using the clustering methods mainly k-means clustering for segmenting the bacilli [21] , so that our proposed method use k-means clustering for segmenting the ZN-stained sputum smears image .

4.5 Post processing stage

After completing the segmentation process ,RGB image with black background and pink bacilli was obtained then it was converted to black and white image to facilitate the feature extraction process ,also the Edge detection method was used to calculate the difference between the corresponding pixel intensities of an black and white image for detection the image edges .region filling morphological operation was used to fill the area of edges that suppressed in previous step .

4.6 bacilli feature extraction:

After the segmentation step is finished not only the bacilli were segmented, different structure were appeared which have the same colour properties as bacilli in conventional microscopy.

The following shape descriptors were investigated to choose the best characteristic of Tb bacilli and each of them were saved in format.mat,

further more used as input features in the support vector machine training process. These Shape descriptors were eccentricity, perimeter, area, compactness, Hu moments.

A. Eccentricity

It defined as a ratio between the distance of foci ellipse and length of major axis**.**

$$e = \sqrt{1 - (\frac{b}{a})^2} \quad \text{.....Eq 4.1}$$

Where a is the length ofmajor axis, and b is the length of minor axis.

B. Perimeter (p)

C. Area (A)

D. Compactness (cf)

It is a simple and popular measure of the efficiency of a contour to contain a given area, and defined as:

$$Co = \frac{p^2}{A} \quad \text{.....Eq 4.2}$$

Where the p and A are the contour perimeter and area enclosed respectively, Compactness provides a measure of how closely the shape of the bacilli approaches a circle. The definition has been modified as follow in order to restrict and normalize the range of the parameter to [0, 1] [6].

$$cf = 1 - \frac{4\pi A}{p^2} \quad \text{.....Eq 4.3}$$

E. Hu Moments

An image moment is certain particular weighted average (mean) of the image pixels' intensities. The image moments are useful feature to describe TB bacilli after the segmentation process. The equations below represent the

mathematical description of the moment that used in extracting the TB bacilli shape feature.

$M_1 = v_{20} + v_{02}$……Eq 4.4

$M_2 = (v_{20} - v_{02})^2 + 4{v_{11}}^2$……..Eq 4.5

$M_3 = (v_{30} - 3v_{21})^2 + (3v_{21} - v_{03})^2$……..Eq 4.6

$M_4 = (v_{30} - v_{21})^2 + (v_{21} - v_{03})^2$……..Eq 4.7

$M_5 = (v_{30} - v_{21})(v30 \quad + v_{21})[(v_{30} - 3v_{12})^2 - 3(v_{21} - v_{03})^2] +$
$(3v_{21} - v_{03})(v_{21} + v_{03})[3(v_{03} + v_{12})^2 - (v_{21} + v_{03})^2]$……....Eq 4.8

$M_6 = (v_{20} - v_{02})[(v_{30} + v_{12})^2 - (v_{21} + 3v_{03})] + 4v_{11}(v_{30} +$
$v_{12})(v_{21} + v_{03})$……..…Eq4.9

$M_7 = (3v_{21} - v_{03})(v_{30} + v_{21})[(v_{30} + v_{12})^2 - 3(v_{21} - v_{03})^2] +$
$(3v_{21} - v_{03})(v_{21} + v_{03})[3(v_{30} + v_{12})^2 - (v_{21} + v_{03})^2]$……..Eq 4.10

4.7 bacilli Classification method:

Classification is the final stage in any pattern recognition system where each unknown pattern is assigned to a category. In the previous studies there were different classification methods used in tuberculosis bacilli identification Bayes , linear, quadratic, and *k*NN classifiers, as well as probabilistic neural networks (PNNs) and SVMs.

The our proposed system was used a support vector machine (SVM) to identify Tb bacilli in sputum smear image , which classified the eleven computed shape features (eccentricity, compactness, $, M_1, M_2 M_3, M_4, M_5, M_6, M_7$) from sputum slide in two class TB and non-TB bacilli

CHAPTER FIVE

5. RESULTS AND DISCUSSION

In this section, we will show the results of the all steps of automatic method of identify TB bacilli in sputum smear image. The 40 ZN-stained images are processed using different enhancement and segmentation techniques.

5.1 Image Enhancement and segmentation results and discussion

The results of enhancement step using de-correlation stretching, Gaussian filter and contrast stretching for sputum smear microscopic image respectively are presented for 2different input ZN-stained images in the figures below.

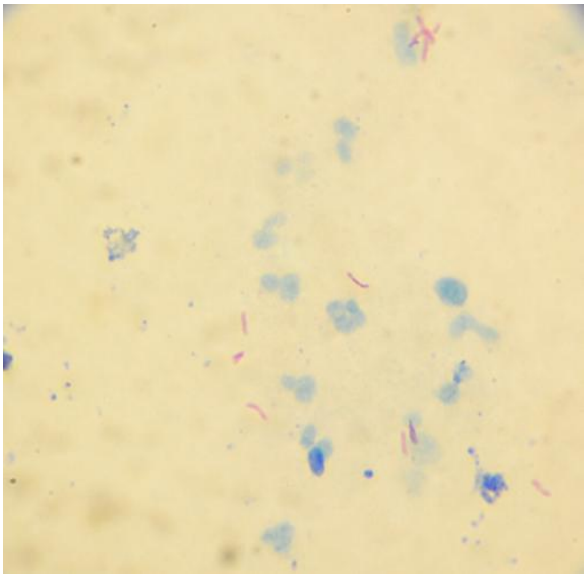

Figure 5.1: sample of input microscopic image

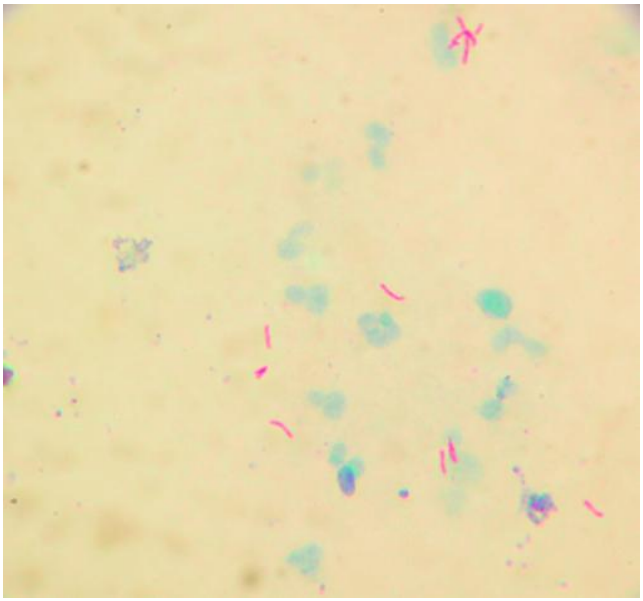

Figure 5.2: represent the original image after implementing the de-correlation stretching

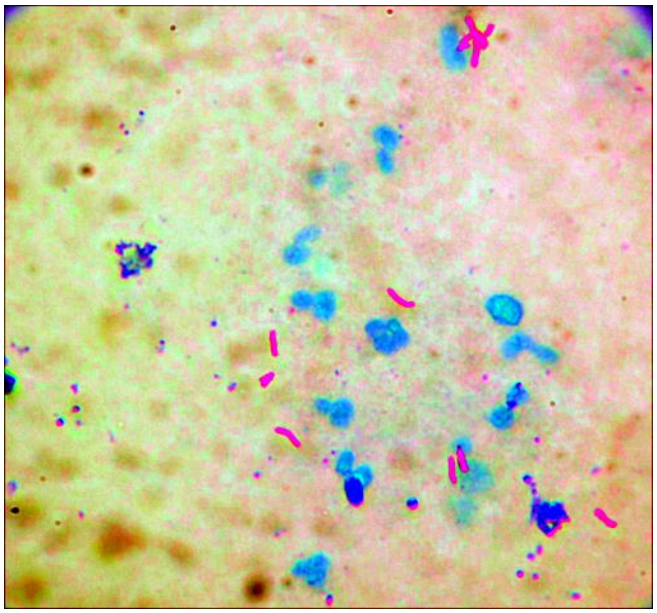

Figure 5.3 : illustrate the smoothed image using Gaussian filter

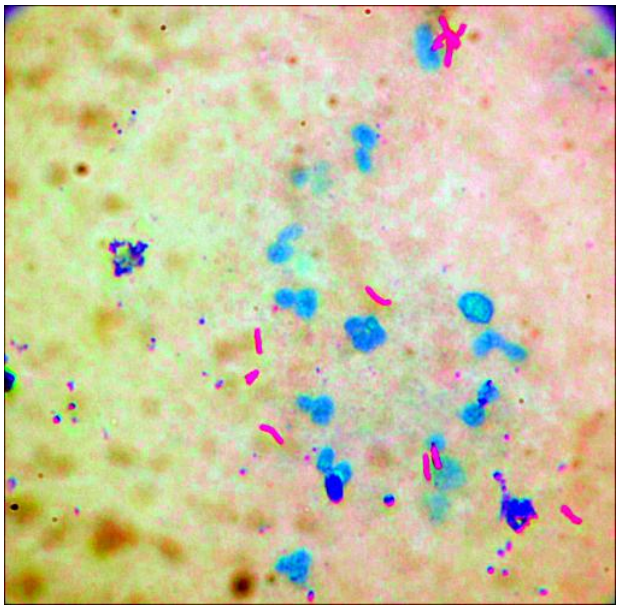

Figure 5.4: image enhancing using contrast stretching

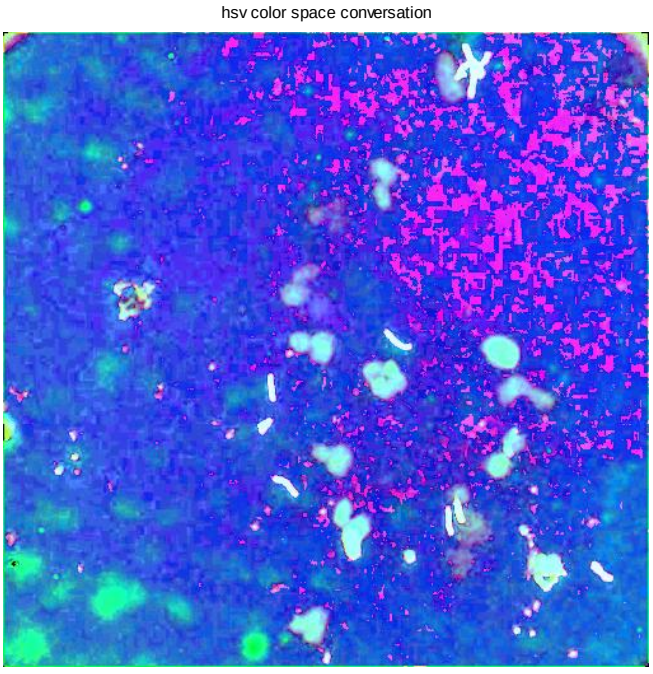

Figure 5.5: represent the HSV color space conversion

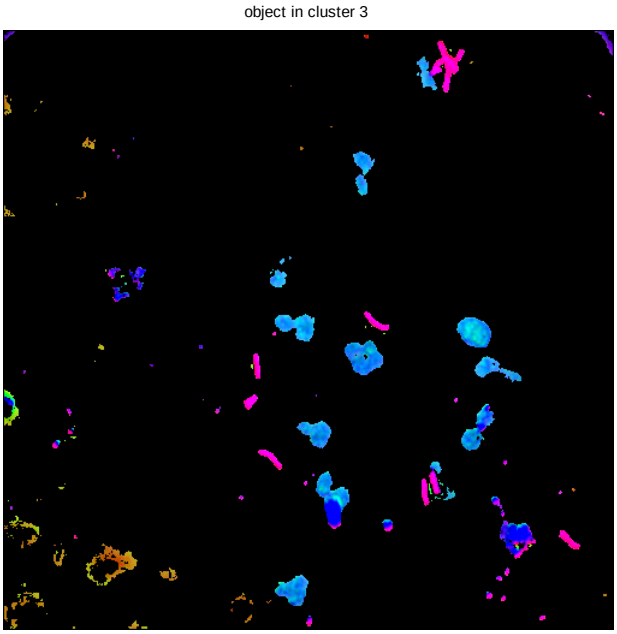

Figure 5.6 :illustrate the image after k-means segmentation

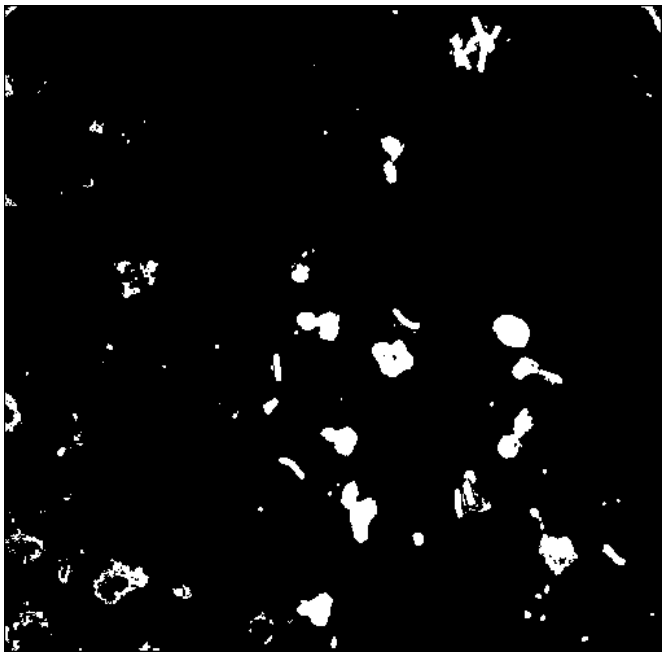

Figure 5.7: represent the image after implementing the morphological operation

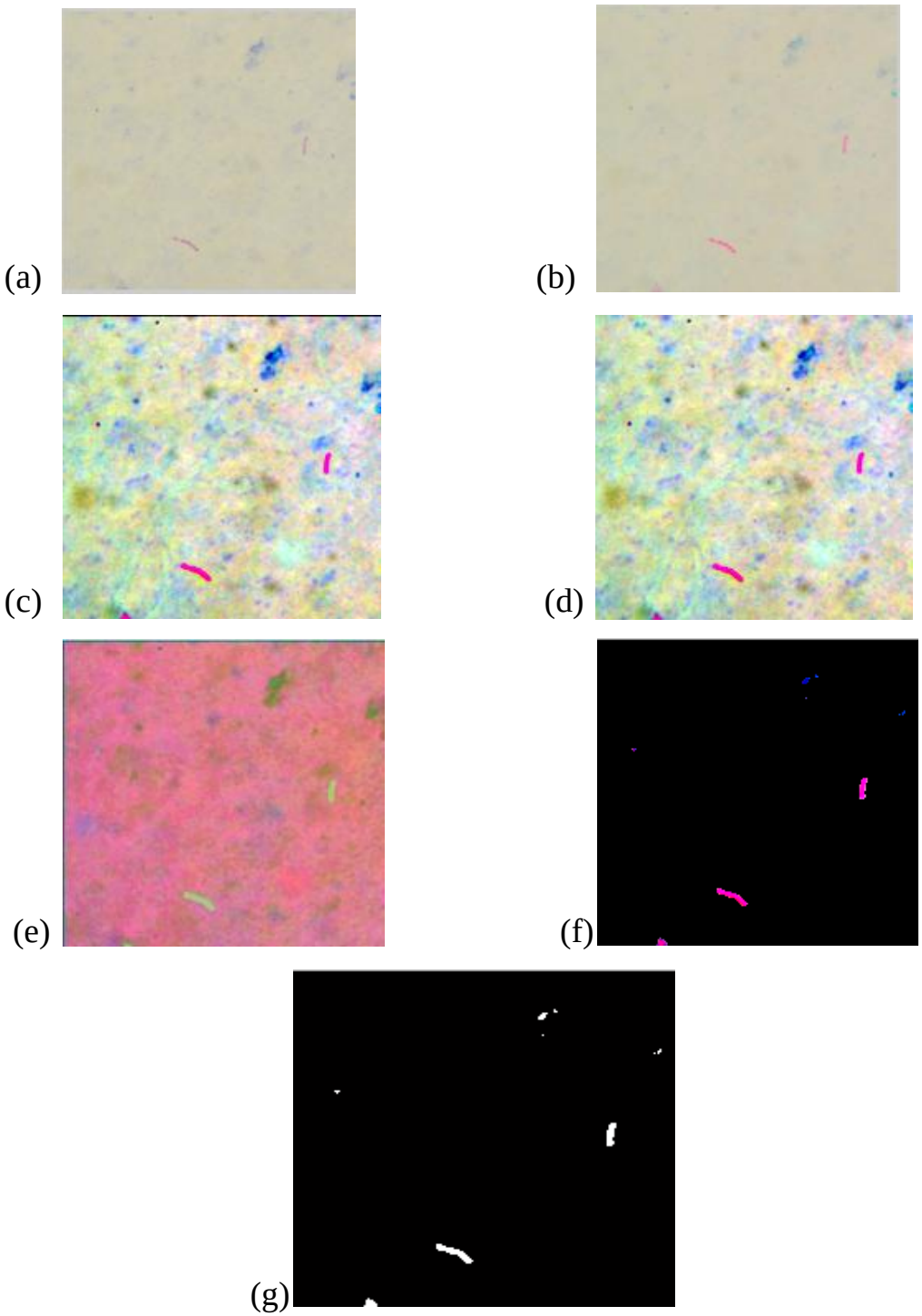

Figure 5.8: summarized the all process of proposed method (a)the original image (b)after applied de-correlation stretching (c)after applied Gaussian filter (d)after applied contrast stretching (e)l*a*b color conversion (f)segmented image using k-means clustering (d)after morphological operation

In first example of input image in fig 6.1 a series of enhancement techniques were implemented. HSV color space conversion process was applied; segmentation was done using k-means clustering. The example of input image in fig 6.8 was processed using the same techniques in first example expect the color conversion technique. L*a*b color space was used in the second example. an appropriate color space is essential to segment sputum smear microscopic images , the process of choosing the suitable color space for better automatic detection of TB bacilli affect with the staining of smear slides and the procedure that was used for acquit ion the image .

In summary the results of enhancement and segmentation step indicate that the process of choosing the most compatible series of enhancement technique depend on the uniformity of the sputum smear images, so that the manual method that we are used for capturing the image using external camera it's not good enough to be an followed method for acquisition the sputum smear image .

Different status of identifying TB bacillus object and non-TB bacillus objects had been illustrated in the figure 6.9 and figure 6.10 below. In first stage sputum smear color was converted to black and white image, next the examined object had been selected, after that region filling morphological operation was implemented. all shape feature that proposed in research method were extracted to investigate if the object is TB bacillus object or not .the method was successfully classify the TB bacillus and non-TB bacillus object

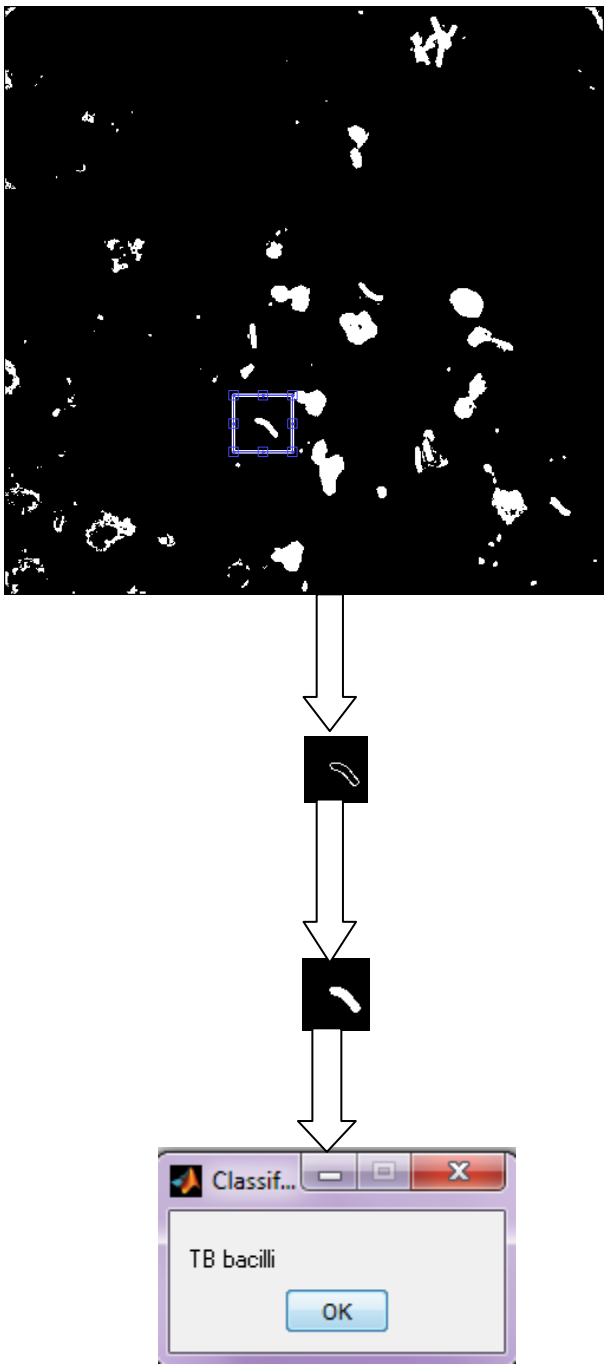

Figure 5.9: illustrated the process of identifying the Tb bacillus object in first example

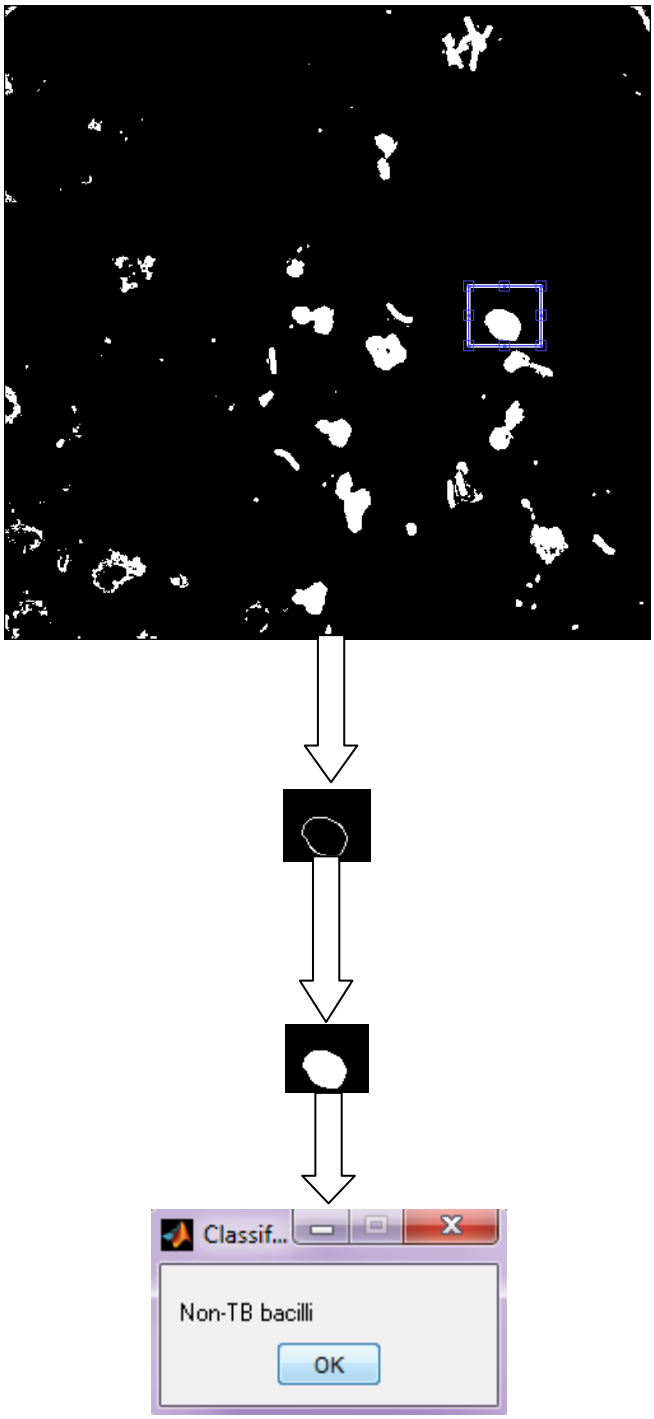

Figure 5.10: illustrated the process of identifying the non-Tb bacillus object in first example

6.2 SVM Classification Results

Classification is the final stage of any image-processing system where each unknown pattern is assigned to a category. The framework developed in this research for the classification of TB bacillus objects and non-TB bacillus objects is tested using a HP Compaq pc, with a 3.00 GHz Intel Pentium Processor and 8.00 GB of RAM.

About 70 sputum smear images are used as training imags and the remaining ones as test images. TB bacillus objects are crroped from location that indicate with color there was tb bacilli . non-TB bacillus are manually selected from sputum smear images . in training phase number of shape features are computed for both TB bacillus objects and non-TB bacillus objects.results of TB bacillus objects shape features are shown in figures below from 6.11 to 6.15 .

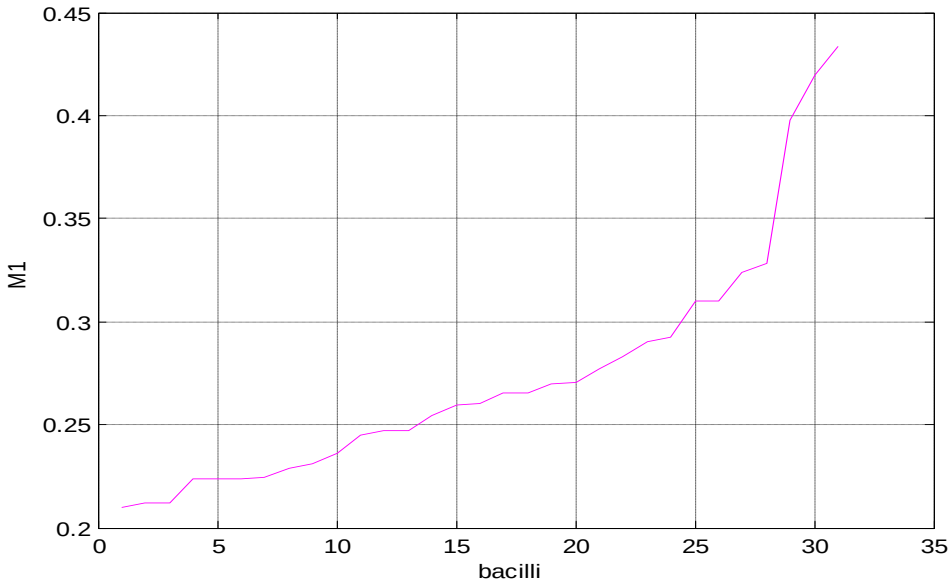

Figure 5.11: plot of M1 features extracted from cropped TB bacilli of training images

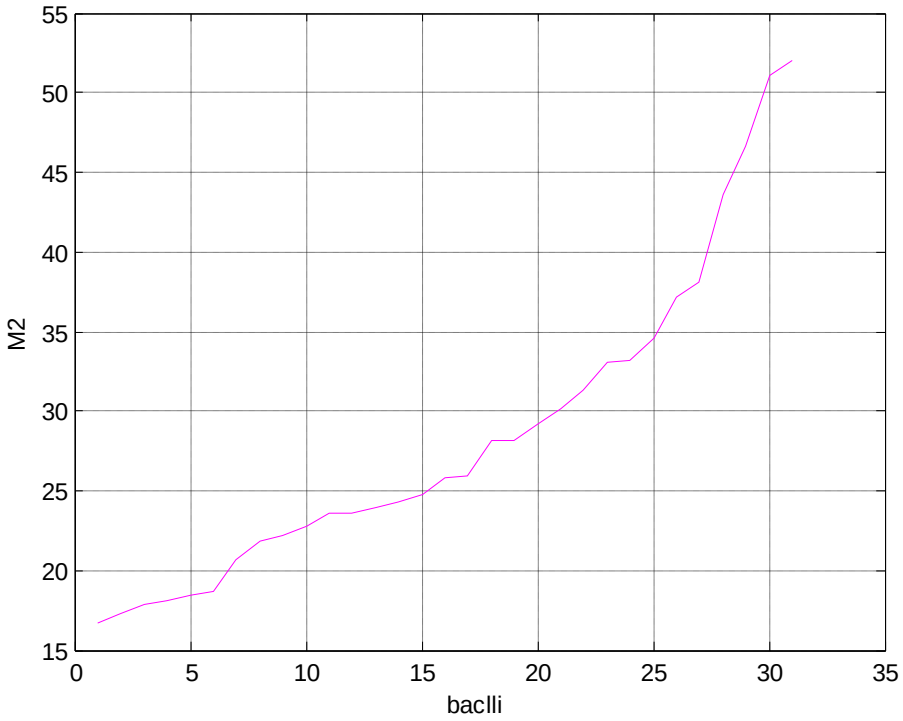

Figure 5.12: plot of M2 features extracted from cropped TB bacilli of training images

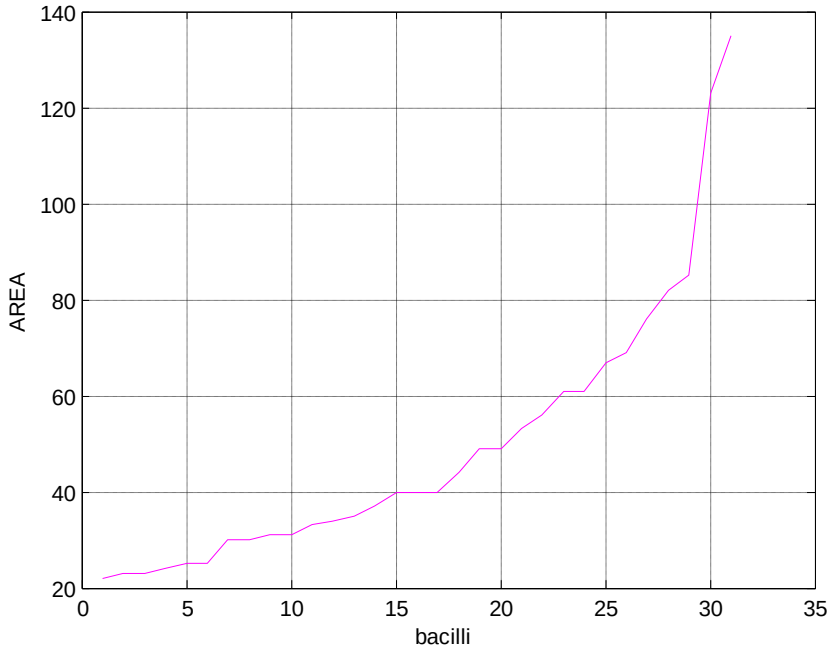

Figure 5.13: plot of Area features extracted from cropped TB bacilli of training images

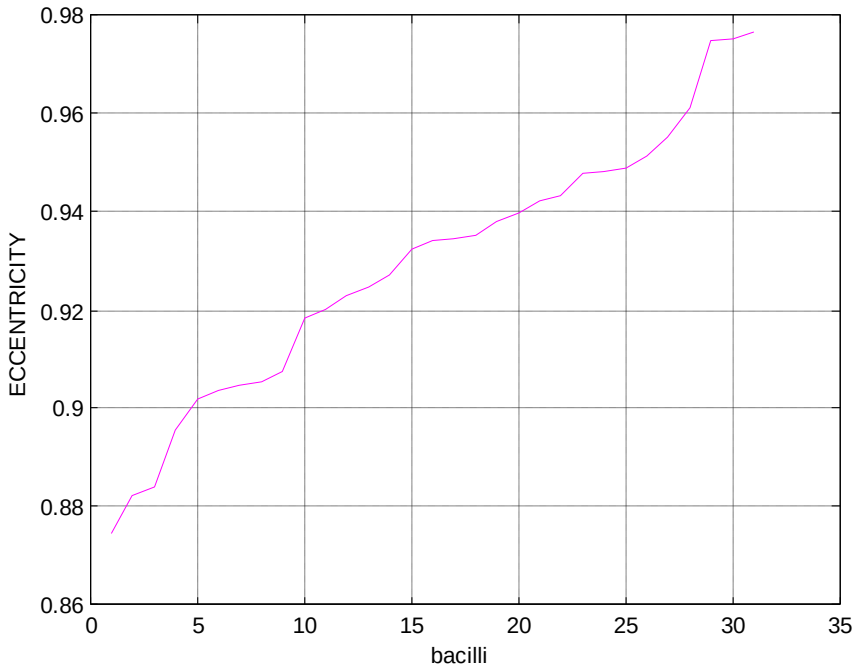

Figure 5.13: plot of Eccentricity features extracted from cropped TB bacilli of training images

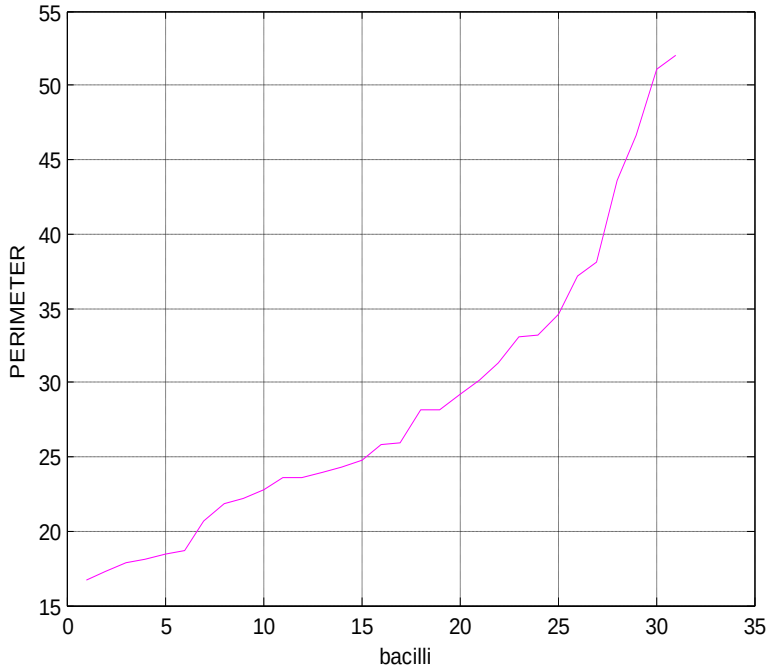

Figure 5.14: plot of perimeter features extracted from cropped TB bacilli of training images

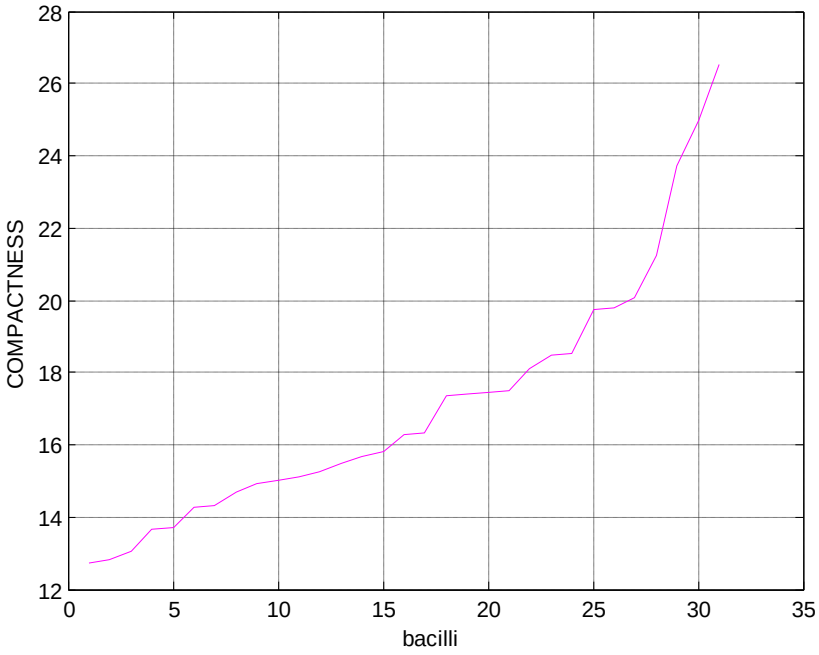

Figure 5.15: plot of compactness features extracted from cropped TB bacilli of training images

In this research, the SVM classification engine is developed using 70 training samples (31 different TB bacilli shapes and 8 non TB bacilli) for a binary classification problem, and for testing the whole method 69 different bacilli shapes were used. Confusion matrix can be derived, as indicated in table 5.1

Table 5.1: testing whole method results

Bacilli status	System result		Total
	Bacilli	Non-TB bacilli	
TB bacilli	**54**	**11**	**65**
Non-TB bacilli	**2**	**4**	**6**
Total	**56**	**15**	**71**

The SVM testing results obtained in the binary confusion matrix in Table 5.1 show that, 56 out of the 69 tested samples are classified correctly by the SVM, which give a binary classification accuracy of 81.15 percent as indicated. Using the confusion matrix results in Figure table 5.1, the four binary classification performance metrics namely the TP, FP, TN and FN are computed as shown in Table 5.2

Table 5.2: represent the classification performance metrics

52 TP	11 FN
2 FP	4 TN

The accuracy of proposed method for automatic TB bacilli identification was computed using the following equation

$$Accuracy = \frac{T_c}{T_s} \times 100\% \quad \text{.....Eq 5.1}$$

Where,

T_c= total number of samples correctly classified

T_s =total number of samples used for training

$$\text{Accuracy} = \frac{58}{71} = 0.81$$

$$\text{Sensitivity} = \frac{\text{Number of true positives}}{\text{Number of true positives+Number of false negatives}} \quad \text{.....Eq (5.2)}$$

$$\text{Specificity} = \frac{\text{Number of true negatives}}{\text{Number of true negatives+Number of false positives}} \quad \text{..... Eq (5.3)}$$

$$\text{Specificity} = \frac{4}{4+2} * 100 = 66.66\%$$

The *sensitivity* and *specificity* metrics (in equations (5.2) and (5.3)) from the confusion matrix performance metrics are computed to be 0.83 and 0.66 respectively, where the minimum and the optimum values of both are 0 and 1 respectively. The classification accuracy is computed using equation (5.1), which is found to be *81.0* percent, where 58 out of the total 71 tested samples are classified correctly by the SVM. Since the sensitivity, specificity and accuracy values are less than 0.95 (95 percent), thus, the performance of the developed framework is acceptable.

In comparing the values of eccentricity and compactness for our method with it's values in previous study. The table5.3 below represents how the value of each pair of eccentricity and compactness in our method are very close to the values of previous other method.

In comparing with some previous studies table 5.3 below illustrate that the proposed method and their algorithm gave acceptable results according for it is sensitivity and accuracy percentage and anther methods.

Table 5.3: comparison between the (eccentricity and compactness) values of our result and the values of previous study

Our method		**Previous method [24]**	
Eccentricity	**compactness**	**Eccentricity**	**Compactness**
0.97656	26.51361	0.96904	24.89174
0.97479	24.9626	0.94706	18.58705
0.975166	23.72402	0.95395	19.51604
0.96107	19.7994	0.96391	21.96085
0.9482	20.0617	0.91198	15.89329
0.95138	18.49671	0.97303	25.41394
0.95517	19.72894	0.977	29.53841
0.9489	21.2399	0.91443	17.13787
0.92475	18.08388	0.9447	18.88333
0.92008	15.65328	0.91171	15.76337

Table 5.4: comparison of the proposed method with the previous studies

Author	***microscopy***	***Enhancement technique***	***Bacilli segmentation***	***Bacilli classification***	***Results***
Our method	ZEISS iLED microscope and NIKON D3100	De-correlation and Stretching Guassian filter Contrast stretching	Color conversion from RGB to HSV &L*A*B color space Segmentation k-means clustering	Features : Eccentricity, compactness, area , perimeter and hu moments from 1 to 7 . Classifier : Support vector machine	Accuracy :81% Sensitivity:83.07% Specificity:66.66%
Shan-e-AhmedRaza et.al 2015	Light microscopy	Novel anisotropic tubular filtering	• Color based segmentation using k-means clustering	• Fourier descriptors and Hu's moment • Classifier : support vector machine	Sensitivity : 90,44% Specificity : 84,08% Accuracy : 86,80%
M. K. Osman et. al.,2012	Light microscope	• CY-based color filter • a 5×5 median filter	k-mean clustering, with the cluster number, k =2	• geometrical features extraction Size and parameter ,Minimum and maximum radius ,Shape factor , Eccentricity , Zernike moments • Classifier : HMLP-ELM network	• Average accuracy for training and testing(94.51), (96.47) % respectively • Average Training time (0.1272s) • Hiddennodes 9

Author	***microscopy***	***Enhancement technique***	***Bacilli segmentation***	***Bacilli classification***	***Results***
Khutlang et al. 2010 July	Bright-field microscope		• combination of two-class pixel classifiersBayes', quadratic, and logistic linear classifiers	• Shape Descriptors:Fourier features, color moments, eccentricity, compactness • kNN, PNN,and SVM classifiers	Accuracy: 98.55% Sensitivity: 97.77% Specificity: 99.13%
Sadaphal et. al., 2008	Light microscopy		• Bayesian segmentation; • Color space**:** RGB	• Shape Descriptors:Axis ratio, eccentricity; • Classifier:Classificatio n tree	No information
IbnuSienaet.,al.2012	Light microscope	De-correlation and Stretching	• Morphological process (dilation ,erosion) • Color segmentation K-means clustering method	• Regional Descriptor and feature extraction (Eccentricity, Compactness) • Classifier : neural network	the network work properly using 15 hidden layer with accuracy about 88%.

Author	***microscopy***	***Enhancement technique***	***Bacilli segmentation***	***Bacilli classification***	***Results***
KusworoAdi et.al 2013	Light microscope	color model which originally convert RGB to NTSC(Luminance, Hue, Saturation)	• Saturation component extraction • auto-thresholding process using Otsu methods	• feature extraction (eccentricity and compactness) • classifier : Support Vector Machine	result of development counting is equal to the manual counting result.
Sotaquiret. al. , 2009	Light microscopy		• First derivative of Histogram • Color space**:** YCbCr, Lab		Accuracy : 96.3% False detection:9.78%

CHAPTER SIX

6. CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

In this project an automatic method for tuberculosis bacilli identification had been built up. This automated method reduces fatigue and screening time by providing images on the screen and avoiding visual inspection of microscopic .the system has a acceptable degree of accuracy, specificity and sensitivity. the system has been use various enhancement technique, k-means clustering methods for segmenting the microscopic sputum smear images, segmentation algorithm is developed to automate the process of detection of TB using digital microscopic images of different subjects. Shape features extraction technique had been implemented to extract eleven shape features and finally for classification the support vector machine was used to be an pattern recognition tool for classify the object in sputum smear images to TB bacillus object and non-TB bacillus object.

6.2 Recommendations and Future Work

- Apply sputum smear image obtained by digital microscope as input to the system.
- Increase the sample size of the research that will give a good feedback for investigation our method.
- Extract another shape features example Fourier descriptors

- Staining procedure is recommended to be automated ,for evenly smear samples with good staining that gives a good contrast between bacilli and background
- Apply other methods of classification like neural network method and compared it with the research method.

REFRENCES

[1] VanDeun A, Salim AH, Cooreman E, Hossain MA, RemaA,etal,"Optimal Tuberculosis case detection by direct sputum smear microscopy: how much better is more?".Int J Tuber Lung Dis 6, (2002) , P 222-230

[2] World Health Organization, Global Tuberculosis Report 2014.

[3] UNITAID, Tuberculosis: diagnostic, technology and market landscape. 2014, WHO: Geneva. p. 1–42.

[4] Sputum Gram stain – Overview, University of Maryland Medical Center www.umm.edu/ency/article/

[5] www.tbfacts.org/tb-tests

[6] Biomedical image analysis book

[7]RonaldDendere, segmentation of candidate bacillus objects in images of ziehl-neelsen-stained sputum smears using deformable models, University of Cape Town February 2009

[8]https://en.wikipedia.org/wiki/Lab_color_space accessed on 3/5/2016

[9] R.Gonzalez and R. WoodsDigital Image Processing, Addison-Wesley Publishing Company, 1992

[10] Dong ping Tian, AReview on Image Feature Extraction and Representation Techniques,International Journal of Multimedia and Ubiquitous Engineering Vol. 8, No. 4, July,2013

[12] M. Yang, K. Kpalma and J. Ronsin. "A survey of shape feature extraction techniques", Pattern Recognition, **(2008)**, pp. 43-90.

[13]Sawantet al., International Journal of Advanced Research in Computer Science and Software Engineering

[14]www.healthdiscoverycorp.com/svm.php

[15] MarziehGhiasi&Tripti Pande1 &MadhukarPai , Advances in Tuberculosis Diagnostics

[16] Costa, M. G.F., Costa Filho, C. F. F., Sena, J. F., Salen, J. & Lima, M. O., Automatic identification of mycobacterium tuberculosis with conventional light microscopy, Proceedings of the 30th Annual International Conference of the IEEE EMBS,(2008), pp. 382- 385,Vancouver, British Columbia, Canada

[17]Sadaphal, P, Rao, J., Comstock, G.W. & Beg, M.F., Image processing techniques for identifying Mycobacterium tuberculosis in Ziehl-Neelsen stains. International Journal of Tuberculosis Lung Disease, (2008),Vol. 12, n. 5, pp. 579-582, ISSN 1027-3719.

[18] Khutlang, R., Krishnan, S., Dendere, R., Whitelaw, A., Veropoulos, K., Learmonth, G. Douglas, T. S. (2010) , Classification of Mycobacterium tuberculosis in Images of ZN-Stained Sputum Smears , IEEE Trans InfTechnol Biomed. 2010 July; 14(4): 949–957. doi:10.1109/TITB.2009.2028339.

[19] M. K. Osman , M. Y. Mashor, H. Jaafar , Detection of Tuberculosis Bacilli in Tissue Slide Images using HMLP Network Trained by Extreme Learning Machine ,electronics and electrical engineering , 2012. No. 4(120) , ISSN 1392 – 1215, T 125 .

[20]Shan-e-Ahmed Raza, M. QaisarMarjan, Muhammad Arif, Farhana Butt , Faisal Sultan, Nasir M. Rajpoot , Anisotropic Tubular Filtering for Automatic Detection of Acid-Fast Bacilli in Digitized Microscopic Images of Ziehl-Neelsen Stained Sputum Smear Samples

[21]Kusworo Adi1, RahmadGernowo, ArisSugiharto , K. Sofjan F , Adi P , Ari B ,tuberculosis (TB) identification in the zihel-nelseen sputum sample in NTSC channel and support vector machine (SVM) classification , International Journal of Innovative Research in Science, Engineering and Technology ,Vol. 2, Issue 9, September 2013 , ISSN: 2319-8753

[22]Rachna H.B., M.S.MallikarjunaSwamy ,detection of tuberculosis bacilli using image processing techniques ,international journal of soft computing and engineering .issn:2231-2307,volume-3,issue-4 ,September 2013

[23]JadhavMukti , Kale K.V , analysis of ZN-stained sputum smear enhanced images for identification of mycobacterium tuberculosis bacilli cells ,international journal of computer applications,vol.23 , no.5 , June 2011

[24] Ibnu Siena, Kusworo Adi, Rahmat Gernowoand Nelly Mirnasari, Development of Algorithm Tuberculosis Bacteria Identification Using Color Segmentation and Neural

Networks, International Journal of Video & Image Processing and Network Security IJVIPNS-IJENS Vol:12 No:04

Appendix

Features of The Data: This section present the eleven features of sputum smear images which used in this project.

The bacilli shape features are:Eccentricity, area, perimeter, compactness and Hu moments ($M_1 to\ M_7$),

M1	M2	M3	M4	M5	M6	M7	Area	Eccentricity	Perimeter	Compactness
0.2602	0.04264	0.0000524201	0.0000063108466333	0.0000000001111186	-0.0000001416778288	-0.0000000000085226	67.0000	0.9382	33.071	16.324
0.26561694591177 3	0.040172484646515	0.0004832 28205661	0.0000619156157 4 9	-0.0000000112056 5 1	-0.0000117648 2 3634	0.0000000 01938723	61.0000	0.92472513 9258274	33.21320 34355964 34	18.0838833 18923192
0.26511362492027 8	0.04388 1936304 893	0.0005901 02146891	0.0001882 7104660 4	-0.0000000 3670468 3	0.0000332243 9 7061	0.0000000 57287896	31.0000	0.93428492 1527292	22.14213 56237309 55	15.8152958 05796625
0.39758393874377 1	0.13168 2254722 427	0.0001867 67007297	0.0000239 1694896 6	-0.0000000 0063993 3	-0.0000062007 3 2985	0.0000000 00098505	61.0000	0.97516605 7347312	38.04163 05603426 18	23.7240271 42452350
0.29028721026103 1	0.05739 1162864 001	0.0000472 61184529	0.0000037 2376814 1	-0.0000000 0002064 9	-0.0000008603 8 0925	-0.0000000 00026949	49.0000	0.94789353 2796097	29.21320 34355964 34	17.4166
0.22398355689596 1	0.02459 5292633 395	0.0001735 65625103	0.0000312 2267714 7	0.0000000 0197096 0	0.0000033564 0 0855	-0.0000000 01183185	53.0000	0.90382980 5706843	28.14213 56237309 55	14.9430
0.24679687500000 0	0.03583 6041259 766	0.0001103 75261307	0.0000254 2238044 7	0.0000000 0120284 1	0.0000029275 5 7270	-0.0000000 00558240	40.0000	0.92733493 8097134	24.24264 068711	14.6926

M1	M2	M3	M4	M5	M6	M7	Area	Eccentricity	Perimeter	Compactness
0.2926841707111183	0.055853610200657	0.001277371815839	0.000157110407594	-0.000000059088434	-0.00001990416519 6	0.000000075537190	49.0000	0.9422170304373 6	30.142135623 73096	18.5418
0.2832	0.0540	0.0004	0.0000	-0.0000	-0.0000	0.0000	135.0000	0.9482	52.0416	20.0617
0.3099	0.0650	0.0027	0.0003	-0.0000	-0.0000	0.0000	123.0000	0.9489	51.1127	21.2399
0.2365	0.0302	0.0000	0.0000	0.0000	0.0000	0.0000	85.0000	0.9184	37.2132	16.2920
0.2242	0.0252	0.0000	0.0000	0.0000	0.0000	- 0.0000	40.0000	0.9056	23.8995	14.2796
0.2289	0.0267	0.0001	0.0000	0.0000	0.0000	0.0000	24.0000	0.9049	18.1421	13.7140

M1	*M2*	*M3*	*M4*	*M5*	*M6*	*M7*	*Area*	*Eccentricity*	*Perimeter*	*compactness*
0.2706	0.0464	0.0000	0.0000	0.0000	0.0000	0.0000	23.0000	0.9346	18.7279	15.2494
0.2548	0.040	0.0000	0.0000	0.0000	0.0000	0.0000	31.0000	0.9324	21.8995	15.4706
0.2309	0.0276	0.0004	0.0001	0.000	0.0000	0.0000	25.0000	0.9074	18.4853	13.6682
0.2121	0.0198	0.0001	0.0000	0.0000	0.0000	0.0000	22.0000	0.8840	16.7279	12.7192
0.21237774307 5532	0.0185 053504 153983	0.000103 4752777 97173	1.73060 133474 093e-05	6.76731 549454 568e-10	1.8087 80459 41241 e-06	- 6.864476 4539889 9e-10	23.0000	0.8746267 07417940	17.31370 8498984 8	13.033239 2168616
0.31025 925925 9259	0.0681 653347 050754	0.000655 1754492 70951	8.23847 161509 948e-05	4.26399 702694 079e-09	- 8.1051 10743 17579 e-06	- 6.105911 5997604 6e-10	30.0000	0.9513837 33258334	23.55634 9186104 1	18.496719 5659222
0.32396 193771 6263	0.0761 038897 755197	0.000858 0366613 40953	0.00013 127479 456463 2	1.19508 711202 412e-08	1.0044 45049 03332 e-05	- 4.341251 8993686 5e-08	34.0000	0.9551790 61168889	25.89949 4936611 7	19.728936 4109286

M1	M2	M3	M4	M5	M6	M7	Area	Eccentricity	Perimeter	Compactness
0.247210395970727	0.0342786511649055	0.000584172026371533	0.000126851165522131	-2.23453011762982e-08	1.0317953882182 3e-05	3.64031027605774e-08	33.0000	0.920086453678994	22.7279220613579	15.6532860977926
0.328093750000000	0.0812984775390625	0.000604847492797847	0.000271151027954098	4.14610965694696e-08	7.6754515311274 4e-05	8.17206103032220e-08	40.0000	0.961072284892925	28.1421356237310	19.7994949366117
0.223740740740741	0.0247437050754458	3.77161172585481e-05	2.96980419651480e-06	-1.13823724875605e-11	-4.4707335727577 9e-07	-1.80056496235719e-11	30.0000	0.902093585874687	20.7279220613579	14.3215584327242
0.277037900874636	0.0510415484704502	8.53760051354042e-05	9.63332740550291e-06	-1.86558269507399e-10	-5.3565744500811 3e-07	-4.79087323021345e-11	35.0000	0.943370506840632	24.7279220613579	17.4705751277882
0.209920000000000	0.0190644224000000	1.98192660480004e-05	9.63235020800040e-06	1.59279260517950e-10	1.1633096759706 3e-06	-1.72351625040197e-10	25.0000	0.882397793489530	17.8994949366117	12.8156767594315
0.223994630130496	0.0234144637515792	0.000270435185735574	4.28551476209998e-05	1.29144176377289e-09	6.7946374468838 1e-07	-7.69305290616419e-10	37.0000	0.895630986941519	23.5563491861041	14.9973401885855
0.433768372484439	0.15777796702725 0	0.002177235648505 79	0.000288870627846737	2.28055177394141e-08	-5.6848823253652 9e-05	-2.04289081963056e-07	82.0000	0.976563594881399	46.6274169979695	26.5136099500309

M1	M2	M3	M4	M5	M6	M7	Area	Eccentricity	Perimeter	compactness
0.269588192419825	0.0465047331601841	0.000564262439271867	0.000140322066522552	1.676941485602 53e-08	2.2075431416519 3e-05	-9.06881130757268e-10	56.0000	0.939905514061206	31.3137084989848	17.5097917849891
0.420022962530981	0.1460555306272 46	0.001434764126717 14	0.000336244696197133	1.4970310 7061247e-07	2.6161434468335 4e-05	-2.44445418857228e-07	76.0000	0.974791402211629	43.5563491861041	24.9625730844977
0.259518004072948	0.0415055421267507	0.000376808301954068	0.000153895471419398	-1.2483200 8787090e-08	3.1314776221886 6e-05	4.31926866091459e-08	69.0000	0.935418302442289	34.6274169979695	17.3776522891488
0.244905146506386	0.0340335758486795	0.000203219059756371	3.2813066 2396511e-05	1.6909364 2290997e-09	5.5918941312475 5e-06	-1.09617268374276e-09	44.0000	0.922911872272146	25.7989898732233	15.1269972381518

M1	M2	M3	M4	M5	M6	M7	area	Eccentricity	Perimeter	Compactness
0.1624	0.0004	0.0002	0.00000	0.0000	0.0000	-0.0000	214.0000	0.4629	53.4558	13.3529
0.1981252955	0.01299580469	0.0004128683915	5.681583922e-05	3.243528145e-09	4.525473105e-06	1.614816692e-09	238	0.8537196156	60.62741699	15.44404912
0.1601292697	0.00015954008	3.959701502e-05	4.522820548e-08	-6.747113012e-14	-5.290308165e-10	3.817454652e-14	243	0.381636792	55.69848480	12.766753
0.1680022808	0.00255848087	8.794483842e-05	1.655891753e-06	-1.793743414e-11	-7.606564944e-08	8.620188216e-14	299	0.6794360725	63.35533905	13.42441132
0.2174998767	0.01912551550	0.00121937549	8.213581345e-05	-1.281255037e-08	-9.297326817 5e-06	2.375695059e-08	230	0.8807985955	63.69848480	17.64129116
0.1801621472	0.006843709102	0.000110251633	7.0665837601e-06	1.494303741e-10	4.222109020 8e-07	1.287489538e-10	53	0.7886234866	25.55634918	12.32315063
0.1738979802	0.004659183303	0.00017936231	8.055675320e-06	2.863518813e-11	4.596365344 1e-07	1.729719788e-10	71	0.747222275	29.313708498	12.10272543