Tuberculosis (TB) is an infectious disease caused by bacteria, whose scientific name is Mycobacterium Tuberculosis. It was first isolated in 1882 by a German physician named Dr. Robert Koch who received the Nobel Prize for this discovery. Tuberculosis is a chronic infectious disease which is transmitted by cough-propelled droplets that carry the etiologic bacterium, Mycobacterium tuberculosis [1]. TB may affect many body organs but primarily targets the lungs. TB is continuously spread through air. It is spread through the air from person to person through respiratory secretions such as sputum or aerosols released by coughing, sneezing, laughing, or breathing [2]. Each untreated person with active TB infects 10 –15 other people every year. The World Health Organization (WHO) estimates that one-third of the world’s population is currently infected with M. Tuberculosis and that a new person is infected every second [2]. Worldwide, TB is still the leading cause of death due to infection, killing about 2 million people a year.

**1.2 About Tuberculosis** (TB) is an infectious disease caused by Mycobacterium tuberculosis bacteria. Most of those who become infected with Mycobacterium tuberculosis manage to confine the mycobacteria to a few cells in their body, where they stay alive but in an inactive form.
Mycobacterium tuberculosis scientific classification

Kingdom - Bacteria
Phylum - Actinobacteria
Order  - Actinomycetales
Suborder- - Corynebacterineae
Family - Mycobacteriaceae
Genus  - Mycobacterium
Species - M. Tuberculosis

1.2.1 Microbiological characteristics
Mycobacterial cell wall contains 1-outer lipids, 2-mycolic acid, 3-polysaccharides (arabinogalactan), 4-peptidoglycan, 5-plasma membrane, 6-lipoarabinomannan (LAM), 7-phosphatidylinositol mannoside, 8-cell wall skeleton. Mycobacteria are aerobic and nonmotile bacteria (except for the species Mycobacterium marinum, which has been shown to be motile within macrophages) that are characteristically acid-alcohol-fast. Mycobacteria do not contain endospores or capsules and are usually considered Gram-positive. A recent paper in PNAS showed sporulation in Mycobacterium marinum and perhaps in M. bovis [3]. However, this has been strongly contested by other scientists.[4] While mycobacteria do not seem to fit the Gram-positive category from an empirical standpoint (i.e., in general, they do not retain the crystal violet stain well), they are classified as an acid-fast Gram-positive bacterium due to their lack of an outer cell membrane. All Mycobacterium species share a characteristic cell wall, thicker than many other bacteria, which is hydrophobic, waxy, and rich in mycolic acids/mycolates. The cell wall consists of the hydrophobic mycolate layer and a peptidoglycan layer held together by a polysaccharide, arabinogalactan. The cell wall makes a substantial contribution to the hardiness of this genus. The biosynthetic pathways of cell wall components are potential targets for new drugs for tuberculosis [5].

1.2.2 General characteristics of M.TB organism
Mycobacterium tuberculosis is a fairly large non motile, rod-shaped bacterium distantly related to the actinomycetes. Cells are rod shaped, measuring 0.3 to 0.6 by 0.5 to 4µ, straight or slightly curved, occurring singly and in occasional threads
[6]. Cells are slender rods, straight or slightly curved, occasionally slender filaments but branching forms rarely occur. They are non motile, aerobic, and non-spore-forming. Two obligate parasites have not been cultivated on culture media. Mycobacterium TB is acid-fast and gram-positive [7]. The bacterium is a facultative intracellular parasite, usually of macrophages, and has a slow generation time, 15-20 hours [8], physiological characteristic that may contribute to its virulence.

1.2.3 Cell wall structure of M.TB organism

The cell wall structure of Mycobacterium tuberculosis deserves special attention because it is unique among procaryotes and it is a major determinant of virulence for the bacterium. The cell wall complex contains peptidoglycan, else it is composed of complex lipids. Over 60% of the mycobacterial cell wall is lipid. The lipid fraction of M.TB’s cell wall consists of three major components.

1. Mycolic acids are unique alpha-branched lipids found in cell walls of Mycobacterium and Corynebacterium. They make up 50% of the dry weight of the mycobacterial cell envelope. Mycolic acids are strong hydrophobic molecules that form a lipid shell around the organism and affect permeability properties at the cell surface. Mycolic Acids are thought to be a significant determinant of virulence in M.TB. Probably, they prevent attack of the mycobacteria by cationic proteins, lysozyme and oxygen radicals in the phagocytic granule. They also protect extracellular mycobacteria from complement deposition in serum.

2. Cord Factor is responsible for the serpentine cording mentioned above. Cord factor is toxic to mammalian cells. Cord factor is most abundantly produced in virulent strains of M.TB.

3. Wax-D in the cell envelope is the major component of Freund's complete adjuvant (CFA).

In summary, the high concentration of lipids in the cell wall of Mycobacterium tuberculosis has been associated with these properties of the bacterium:

- Impermeability to stains and dyes.
- Resistance to many antibiotics.
• Resistance to killing by acidic and alkaline compounds.
• Resistance to osmotic lysis via complement deposition.
• Resistance to lethal oxidations and survival inside of macrophages.

1.2.4 Staining characteristics

Mycobacteria are classical acid-fast organisms [9]. Acid-fastness is a physical property of certain bacteria, specifically their resistance to depolarization by acids during staining procedures [10]. Stains used in evaluation of tissue specimens or microbiological specimens include Fite's stain, Ziehl-Neelsen stain, and Kinyoun stain.

The high mycolic acid content of certain bacterial cell walls, like those of Mycobacteria, is responsible for the staining pattern of poor absorption followed by high retention. The most common staining technique used to identify acid-fast bacteria is the Ziehl-Neelsen stain, in which the acid fast bacilli are stained bright red and stand out clearly against a blue background. Another method is the Kinyoun method, in which the bacteria are stained bright red and stand out clearly against a green background. Acid-fast bacteria can also be visualized by fluorescence microscopy using specific fluorescent dyes (auramine-rhodamine stain, for example) [11]. Some bacteria may also be partially acid-fast.

The acid-fast staining method for Mycobacterium tuberculosis is the Ziehl-Neelsen stain. When this method is used, the M.TB. smear is fixed, stained with carbol-fuchsin (a pink dye), decolorized with acid-alcohol. The smear is counterstained with methylene-blue or certain other dyes. Acid-fast bacilli appear pink in a contrasting background. In order to detect Mycobacterium tuberculosis in a sputum sample, in excess of 10,000 organisms per ml of sputum are needed to visualize the bacilli with a 100X microscope objective. One acid-fast bacillus/slide is regarded as "suspicious" of an M.TB. infection.
Pulmonary Tuberculosis is the origin of all extra-pulmonary TB. It is detected by sputum smear microscopic examination and it is rapidly detected from sputum sample. Sputum images are easily available and can be easily analyzed by image processing tools. So in the proposed method only pulmonary tuberculosis is detected using sputum sample image. Identification of tubercle bacilli are routinely done in sputum smears using microscope that is done by laboratory technician and the doctor analyzes the disease. The sputum is taken from the patient in the morning which is the first sputum out. Before analyzing under microscope, the sputum specimen is colored using Ziehl – Nelsen stained. After this staining, tuberculosis bacteria gives color in red because of acid resistance and the background give color in blue. The clinicians identify and count the amount of mycobacterium manually. The clinicians have a lot of tuberculosis patient every day. Each sputum specimen needs 15 minutes to be examined and scanned with 100 fields under microscope. This condition gives heavy workload for the clinicians that can reduce the accurateness of tuberculosis diagnose and cause high error rate detection.

1.2.5 Mycobacterium tuberculosis infection and tuberculosis

The development of tuberculosis (TB) requires infection by bacteria of the Mycobacterium tuberculosis complex (M.TB). Infection with M.TB usually occurs through a potential host being exposed to the bacilli in airborne droplets. Such infectious airborne droplets are produced by someone with active (infectious) pulmonary tuberculosis disease when they cough. The bacilli when inhaled by the potential host can establish an infection in the lung (primary focus) [12]. Bacilli may spread from this focus, via the lymphatic or blood circulatory
system, to other parts of the body. Host defaces form a granuloma around the infecting bacilli in the primary focus, which become caseous and necrotic at the centre. In the majority of the cases the immune competent host is able to arrest the growth of the bacilli within this primary focus, with no obvious signs of illness [12]. However, in some cases the infection is not contained and disease ensues either at the site in the lung, at other sites to which the bacilli spread, or at both. Infectious potential depends upon the site and extent of disease with advanced lung disease (involving cavity formation) being the most common form with high infectious potential. In most cases, infection does not result in disease and the initial lesion resolves and eventually calcifies. Such old lesions may, however, still harbor viable bacilli and the host is considered to have a latent TB infection (LTBI). These latent infections may reactivate later, sometimes many years later, and result in disease [13]. Exposure to the bacilli does not necessarily lead to infection, and M.TB infection does not necessarily lead to disease (either at the time of infection or through reactivation). Disease does not always lead to infectiousness, nor death. Risk factors are important and a number of these are described later [14].

1.2.6 Predisposing factors for TB infection

The predisposing factors are close contact with large population of people, poor nutrition, alcoholism and HIV infection. This is 400-times the rate associated with the general public.

1.2.7 Main source of tuberculosis

A person can become infected with tuberculosis bacteria when he or she inhales minute particles of infected sputum from the air. The bacteria get into the air when someone who has a tuberculosis lung infection coughs, sneezes, shouts, or spits (which is common in some cultures).

1.3 TB infection Rate and Survey

Every day -

1. More than 40,000 individuals get infected with the TB bacillus.
2. More than 5000 people develop TB disease.
3. More than 1000 people die due to TB (i.e. 2 person die every 3 minute).
Every year -

1. About 3 lakh children drop out from the school because of their parents had TB
2. More than 1 lakh women are rejected by their families of stigma due to TB
3. It is estimated that in a year, about 200 TB cases may occur in a population of 1 lakh.
4. Annually, about 18 lakh new cases of TB occur of which about 8 lakh are sputum positive infectious cases.

Tuberculosis is becoming a world-wide problem. War, famine, homelessness, and a lack of medical care all contribute to the increasing incidence of tuberculosis among disadvantaged persons. Since TB is easily transmissible between people, then the increase in TB in any segment of the population represents a threat to all segments of the population. This means that it is important to institute and maintain appropriate public health measures, including screening, vaccination (where deemed of value), and treatment. A laxity of public health measures will contribute to an increase in cases. Failure of adequate treatment promotes the development of resistant strains of tuberculosis.

An estimated 9.3 million new cases and 1.8 million deaths due to TB occurred in 2007, of which 1.4 million cases and 0.5 million deaths were in HIV-positive people. An estimated 4.1 million (44%) of the new cases in 2007 would be sputum smear-positive [15]. The prevalence of TB in 2007 was estimated at 13.7 million cases globally [1] Half a million cases are estimated to be multi-drug resistant TB cases [15]. Twenty-two high-burden countries (HBCs), all low and middle-income countries (LMICs) collectively account for 80% of the global TB burden.

In 2010, there were 8.8 million (range, 8.5–9.2 million) incident cases of TB, 1.1 million (range, 0.9–1.2 million) deaths from TB among HIV-negative people and an additional 0.35 million (range, 0.32–0.39 million) deaths from HIV-associated TB [16].
Important new findings at the global level are:

1. The absolute number of TB cases has been falling since 2006 (rather than rising slowly as indicated in previous global reports).
2. TB incidence rates have been falling since 2002 (two years earlier than previously suggested).
3. Estimates of the number of deaths from TB each year have been revised downwards.
4. In 2009 there were almost 10 million children who were orphans as a result of parental deaths caused by TB.

There were an estimated 9.2 million new TB cases in 2006; 1.7 million people died from TB in 2006 more than 2 billion people equal to one third of world’s population are infected with TB bacilli [17]. Worldwide, TB is still the leading cause of death due to infection, killing about 2 million people a year.

Updates to estimate of disease burden follow the completion of a series of consultations with 96 countries between 2009 and 2011, including China, India and 17 African countries in the past year, and much greater availability and use of direct measurements of TB mortality. Ongoing efforts to further improved measurement of TB cases and deaths under the umbrella of the WHO Global Task Force on TB Impact Measurement, including impressive progress on TB prevalence surveys and innovative work to strengthen surveillance, are summarized.

At country level, dramatic reductions in TB cases and deaths have been achieved in China. Between 1990 and 2010, prevalence rates were halved, mortality rates fell by almost 80% and TB incidence rates fell by 3.4% per year. Methods used to measure trends in disease burden in China – nationwide prevalence surveys, a sample vital registration system and a web-based case notification system – provide a model for many other countries.

Without treatment, mortality rates are high. In studies of the natural history of the disease among sputum smear-positive and HIV-negative cases of pulmonary TB, around 70% died within 10 years; among culture-positive (but smear-negative) cases, 20% died within 10 years [18]. Treatment using combinations of
anti-TB drugs developed in the 1940s and 1950s can dramatically reduce mortality rates. In clinical trials, cure rates of above 90% have been documented; the treatment success rate among smear-positive cases of pulmonary TB reported to WHO reached 87% at the global level in 2009.

In 1993, the World Health Organization (WHO) declared TB as a global public health emergency, at a time when an estimated 7–8 million cases and 1.3–1.6 million deaths occurred each year. In 2010, there were an estimated 8.5–9.2 million cases and 1.2–1.5 million deaths (including deaths from TB among HIV-positive people).2 TB is the second leading cause of death from an infectious disease worldwide (after HIV, which caused an estimated 1.8 million deaths in 2008).

1. There were an estimated 8.8 million incident cases of TB (range, 8.5 million – 9.2 million) globally in 2010, 1.1 million deaths (range, 0.9 million–1.2 million) among HIV-negative cases of TB and an additional 0.35 million deaths (range, 0.32 million–0.39 million) among people who were HIV-positive.

2. In 2009, there were an estimated 9.7 million (range, 8.5–11 million) children who were orphans as a result of parental deaths caused by TB.

3. Globally, the absolute number of incident TB cases per year has been falling since 2006 and the incidence rate (per 100 000 population) has been falling by 1.3% per year since 2002. If these trends are sustained, the MDG target that TB incidence should be falling by 2015 will be achieved.

1.4 Motivation

1.4.1 Time consuming Manual System

Many computer science researchers have been working on developing and designing an automated system for fast & accurate recognition of tuberculosis (TB), which would ensure speed up of treatment process. By observing literature survey some drawbacks of the existing manual systems for detection of pulmonary tuberculosis are observed.

1. Manual screening technique is time consuming and tedious as it needs experienced technologist.
2. The non-invasive characteristics of the sputum procedure are important because repeated examinations are needed for an early detection of the disease [19].

3. Manual screening for the bacillus identification is a labor-intensive task with a high false negative rate.

4. Exact counting of overlapped Mycobacterium tuberculosis bacilli cells or accurate counting of cells from bunches of MTB bacilli is very critical, time consuming there may be error prone.

5. If the grading is assigned to the TB patient he or she must be treated till completely recover from Tuberculosis.

6. Every month sputum examination is suggested to pulmonary TB patient, and the Lab technician must keep track of every patient by comparing previous slide to the new slide and check whether the grade of patient goes decreasing from 3+ to 2+….00. If the grade is not changed, means the given drug treatment is not suited to the patient and the patient’s TB is converted to MDR-TB or XDR–TB and need to be changed treatment with heavy drugs.

To narrate this problem of manual TB detection system, proposed system is developed for substantial reduction in clinician’s workload. The proposed system is able to improved test sensitivity and a better diagnostic accuracy by screening the number of acid fast stained sputum smear images.

1.4.2 Proposed System

In this proposed work pulmonary type of tuberculosis is considered, because pulmonary tuberculosis is the origin of extra pulmonary tuberculosis. Sputum smear microscopy is used for analysis of pulmonary tuberculosis. It is worldwide accepted technique in medical science for early diagnosis of tuberculosis diseases.

1. In the proposed system, computerized analysis of Acid Fast Stain sputum smear images is used for diagnosis of pulmonary tuberculosis. These images are live, and are captured from Govt. Tb hospital using Digital camera and light microscope under 100 x objectives.
2. Images of acid fast stain, sputum smear are input data for this experiment, bacilli cells of Mycobacterium Tuberculosis is detected, recognized and counted per oil immersion fields

3. Images are preprocessed, enhanced and segmented to separate M.TB bacilli cells from sputum smear on the basis of color property.

4. The output of segmentation, of M.TB bacilli cells is recognized using an appropriate pattern recognition approach.

5. Shape is the second discriminating property of M.TB bacilli. Recognition of M.TB bacilli cells is done using shape base feature & classifier like K-NN, SVM, Gaussian, PCA.

6. The recognized M.TB bacilli cells are exactly counted by calculating area of each cell because grading depends on the number of M.TB bacilli cells.

7. The results are analyzed i.e. positive or negative & grading is assigned to the patient as per chart provided by WHO (which is discussed in Table 2.1).

Manual screening for TB identification involves a labor-intensive task with poor sensitivity and specificity. To improve the diagnostic process we have developed an automated system for TB identification, which consists of a microscope with digital camera. The microscope is a light microscope with a 100x oil, 1.25 numerical apertures (N.A) objective lens and a moveable slide stage. A digital camera with 2 mega pixel and 2/3 inch format is attached to the eyepiece of the microscope and used to capture images of the sputum sample.

The system can capture a large number of images on sputum sample using digital camera and process all the images in real time to identify the bacilli, recognize and count their number. In order to speed up image acquisition while guaranteeing the image quality, an efficient method for capturing the images is proposed.

In the proposed work 60 sputum smear slides of patient are analyzed, 500 images are captured for this study. These 500 images of acid fast sputum smear are preprocessed enhanced and quality of images is measured by using quality
measures mean, standard deviation and histogram. The enhanced images are used as input data.

Due to Ziehl – Nelsen, stained sputum smear images appear blue background on which tuberculosis bacteria gives red color because of acid resistance. Color is the first parameter to identify tuberculosis bacteria. An image-based segmentation algorithm is applied on the images. The tuberculosis bacteria (Mycobacterium Tuberculosis (M.TB)) are separated from this image.

The enhanced images are used as input data.

The extracted objects after segmentation contain M.TB bacilli, stain residues & some Non TB bacilli. After segmentation from 500 images, 1000 objects are extracted. The edge detection algorithm is used to extract boundary of objects because shape is the discriminate property of M.TB bacilli. It is the second parameter to recognize M.TB bacilli.

Shape based features are calculated from extracted objects. For 1000 objects 11,000 features are extracted and inputted to various classifiers for recognition of M.TB bacilli. From the literature survey, it is observed that several researchers worked for identification and recognition of bacilli, but no one has assigned grading to patient sputum sample which depends on counting of bacilli.

In the proposed system a model has been designed for identification and recognition of M.TB bacilli cells. The exact counting of bacilli cells is performed which is necessary to assigns grade to the patient as per WHO chart (Table 2.1).

All images which are used in the proposed system are live and captured from two Govt. Tuberculosis Hospitals. The results of TB diagnosis by manual method and proposed method are verified by medical expertise.

1.4.3 Preparation of data set

Any type of research is not possible without data. It is critical for a computer researcher to prepare Ziehl-Neelsen stain sputum smear slide to generate data for this type of research work. It has been prepared by setting up image acquisition technique on simple light microscope. We have prepared acid fast stain sputum smear images data set using light microscope with digital camera at the eye piece.
of microscope, and analyzed it for diagnosis of pulmonary tuberculosis. While analyzing some image processing & pattern recognition applications are applied on these images to give faster and better results than manual screening of slide. The dataset contained 500 total images (400 images are positive for TB and 100 images are negative for TB).

1.5 Significance of the Thesis

In this work the Mycobacterium Tuberculosis bacteria is identify, recognized and count to know the level of TB. The data set is acid fast stain sputum smear images are generated using light microscope and digital camera at Govt. TB hospital. Throughout this work these self developed data set is used for experiment. 60 AFB slides are used to prepare image library of 500 images which is to be analyzed using reposed system. During result process around 1000 objects are extracted from these images. Total 11000 features data set developed for 1000 objects.

The experimental work consists of image enhancement using two techniques by unsharp masking and contrast stretching the best one is contrast stretching which is used in proposed method. The M.TB cells are magenta in color on blue color sputum sample due to color property of ZN stain, color is first parameter for identification and separation from sputum. The segmentation of enhanced images using Euclidean distance and Mahalanobis distance the results from Mahalanobis distance are superior in visual perception. Hence Mahalanobis distance method used in RGB segmentation for separation of M.TB cells from sputum image. Shape is the second parameter of M.TB bacilli cells we can extract various shapes by outcomes of segmentation using edge detection. The pattern recognition using various types of classifier likes SVM, K-NN, Gaussian, PCA, and neural network techniques like perceptron network and back-propagation are used. The recognized cells are counts if the cells are overlapped that are separated and count to know TB level of patient. The results of the developed systems are verified by medical expertise the certificates of experts for analysis of results shown in Appendix “B”.

The main contribution of this work is image data set generation. The M.TB bacilli cells are accurately recognized and exactly counted. Cells, which are found overlapped, they are separated before counting to know the level of TB.
Another major contribution of this work is, 11 features for 1000 M.TB cells and Non-TB cells are designed i.e. total 11000 features are designed, library of design experiments has been shown in Appendix “A”.

1.6 Organization of the thesis

The thesis is organized in five chapters and aimed to discuss the research problem thoroughly and came up with best solution to the problem using propose system.

Chapter 1: Introduction - It gives introduction to Tuberculosis how it spread from person to person. Brief idea about the Mycobacterium Tuberculosis bacteria and its structure. This chapter presents the objective of the work and outline of the thesis.

Chapter 2 Background of Various Tuberculosis Detection Methods - It gives an introduction to backgrounds of various tuberculosis detection techniques. This chapter also present brief introduction of principles behind the manual TB detection technique and automatic TB detection techniques using literature surveys. The selected references which are used to compose this chapter and the list of related publications used during this study are given.

Chapter 3 Proposed Approach – It describes in detail the design of the proposed system. This chapter also describes details of our proposed approach, form data generation to M.TB bacilli recognition. Feature extraction for recognition and classification of TB cells and Non-TB objects.

Chapter 4 Experiments, Results and Discussion - It discusses the experiments, results and discussion. In this chapter, the discussion and interpretation of this research work is given. This chapter deals with the results of the experiments performed and the interpretation of them.

Chapter 5 Conclusion and Future work - The overall conclusion of this research work is presented along with some future directions to carry out this research work.

References


2. “Module for MPWs and other DOT providers”, Revised National Tuberculosis Control Program, Directorate general health and family welfare, Nirman bhavan new Delhi March 2006.


