TB BACILLI DETECTION FROM ZN STAINED SPUTUM IMAGES USING CONNECTED COMPONENT TECHNIQUE

S.Srividhya and R.Kalpana

Department of Biomedical Engineering, Rajalakshmi Engineering College, Chennai, India
E-mail: srividhyavalase@gmail.com, kalpana.r@rajalakshmi.edu.in

Abstract— Tuberculosis (TB) is one of the contagious diseases that pose major threat to public health. It mainly affects the lungs but can also spread to other organs such as brain, intestine, kidney and spine. TB bacilli are normally present in the sputum of infected subject which presently is detected by manual observation through microscopes. Any incorrect diagnosis will lead to serious health related issues. Therefore presented here is an approach to detect the bacterium from Ziehl - Neelsen stained sputum smear. Microscopic images of sputum smear are collected from infected subjects. These images contain TB bacilli and also other objects. In order to extract the bacilli, other objects that are smaller in size when compared to bacilli is removed by checking for connected components and fixing an area (as threshold) that is less than the area of bacilli. Then by performing morphological operations, the boundary of the bacilli is identified. They are superimposed on the original image, so that boundaries of bacilli are highlighted. However better classification tools like neural network, could still improve the accuracy and sensitivity of detection.

Index Terms— Tuberculosis, Sputum smear, ZN-stain, connected components, Morphological operations.

I. INTRODUCTION

Tuberculosis (TB) is a contagious disease which is caused by a bacterium called mycobacterium tuberculosis. These bacteria usually affect the lungs; it will also attack any part of the body such as the kidney, spine, and brain. In the pulmonary TB, bacteria are found only in lungs where as in extra pulmonary TB bacteria affect outside the lungs. The extra pulmonary tuberculosis is often found in human Immunodeficiency virus infected people. If not identified and treated properly, TB disease can be fatal. Major symptoms may include chest pain and a prolonged cough producing sputum. It can be cured if identified sand treated with the combination of drugs minimum for the period of six months. Even though TB is easily curable, its rate is high in India. It is estimated that about 40% of the Indian population is infected with TB. According to the World Health Organization (WHO) statistics for 2015 report, globally the TB mortality rate in 2015 was 47% lower than in1990 [1]. The conventional method used for diagnosing TB is from sputum smear microscopy as it is simple, cost-efficient and accessible tool for pulmonary TB diagnosis and treatment monitoring method. Mycobacterium tuberculosis also known as Acid fast bacilli is small, nonmotile bacillus which divides very 16 to 20 hours. A fluorescence microscope is used to examine auramine stained sputum smear while a light or bright-field microscopy is used to examine ZN stained sputum images. Ziehl-Neelsen (Z-N) staining method is widely used in many developing countries, to stain mycobacterium species including mycobacterium tuberculosis [2]. Color is the most important feature that is utilized in detecting the TB bacilli. During the staining process, carbol fuchsin dye is used to color the TB bacilli red, while the methylene blue turns the tissues and backgrounds to the blue color.

Manual observation is done by microbiologist to identify bacteria in microscope which will take 2-3 hours for a single slide. Some of the bacilli appear stained in deep red and in some cases they appear in pale-red making a challenging task for an observer. However, the intensity distribution of TB bacilli and background are often varying from image to image, due to manually ZN staining method.

Although several computation methods for TB detection have been developed, accurate bacteria and non-bacteria classification still remains a big challenge. A technique of image segmentation by colour thresholding method was done to identify the coloured bacilli in a stained sputum image. Further, image enhancement like contrast stretching was done to extract bacilli [3].

Khuutlang et al. [4] used pixel classifier which uses linear mapping to segment bacteria from other objects. Using Euclidean distance it is differentiated from other images. Various classifiers (such as Bayes classifier, SVM, etc.) are compared. Further Manual segmentation is compared with the result of Pixel classifier using Haudorff Distance.

Rachna et al. [5] has explained the limitation of pixel classification approaches that they cannot incorporate spatial or regional information about the object of interest. Hence they have used tubeness filtering to detect rod like bacteria. The structures which are curvilinear or tube like in shape are
detected easily by multi-scale line filter. Forero et al. [6] uses an adaptive color thresholding followed by morphological operations to detect bacteria based on their shape and size which was captured using a fluorescence microscope. Different feature descriptors were evaluated for bacilli characterization.

In this work, TB bacilli which are found in digital microscopic image was segmented by ignoring the other objects lesser than that of the bacilli using a connected component labeling algorithm. Thresholding is done to find bacilli in a varying intensity background.

II. MATERIALS AND METHODS

Sputum was collected from a TB affected subjects and ZN staining method is followed. The images are acquired from slides using optical microscope interfaced with computer. The slides are examined using Leica DM 2700M Optical Microscope with 100X magnification.

This particular microscope can be interfaced with the computer to acquire the images. The images are of 32 bit color images in size of 1024 X 768. The images were stored in PNG file format. The steps involved in processing of input images for segmentation of mycobacterium tuberculosis using the connected component labeling algorithm is shown in Figure 2.

A. Pre-processing

Resolution is one of most commonly used ways to describe the image quantity of digital camera. The image is of varying intensity, to avoid complications due to this it is converted to a gray scale image. Gray levels represent the interval number of quantization in gray scale image processing. At present, the most commonly used storage method is 8-bit storage. There are 256 gray levels in an 8 bit gray scale image, and the intensity of each pixel can have from 0 to 255, with 0 being black and 255 being white. Colour component are also known as colour channels or colour planes (components). In the RGB colour model, a colour image can be represented by the intensity function.

\[ I_{RGB} = (F_R, F_G, F_B) \]  

Where \( F_R(x,y) \) is the intensity of the pixel \((x,y)\) in the red channel, \( F_G(x,y) \) is the intensity of pixel \((x,y)\) in the Green channel, and \( F_B(x,y) \) is the intensity of pixel \((x,y)\) in the blue channel. The intensity of each colour channel is usually stored using eight bits, which indicates that the quantization level is 256. During the end of this step gray image is obtained. Figure 3. Show the original image in the gray scale.

B. Binarization and Thresholding

A binary image is a digital image that has only two possible values for each pixel. A binary image can be processed well than a gray-scale image. The basic idea in thresholding is to select a threshold \(T\) to extract an object or several objects with the same value from background. Otsu’s method selects the threshold by minimizing the within-class variance of the two groups of pixels separated by the thresholding operator [11]. Thresholding is used for separation of light and dark regions in a given image. The basic global threshold, \(T\) is calculated as follows:

1. Calculate the average gray level in the image and assign it as \(T\).
2. Separate the image with respect to \(T\) to obtain two groups of pixels: \(G_1\) consisting of pixels with grey levels > \(T\) and \(G_2\) consisting pixels with grey levels \(\leq T\).
3. Compute the average grey levels of pixels in \(G_1\) to give \(\mu_1\) and \(G_2\) to give \(\mu_2\).
4. Calculate a net threshold value as follows,

\[ T = \frac{\mu_1 + \mu_2}{2} \]  

5. Repeat steps 2-4 until the difference in \(T\) in successive iterations is less than a predefined limit \(T_{max}\).

In global thresholding [11], the threshold value is held constant throughout the image:

\[ g(x, y) = \begin{cases} 0 & f(x, y) < T \\ 1 & f(x, y) \geq T \end{cases} \]  

Equation (3) will be the binary image containing the value 0 and 1. The pixels of the binary image B are 0’s and 1’s; the 1’s will be used to denote foreground pixels and the 0’s background pixels.
C. Connected Components

Connected Components Labeling (CCL) is a very important tool in Image Processing, Engineering, Physics, and others; therefore there have been many proposed algorithms and implementations. Each maximal region of connected pixels is called a connected component. Two pixels are connected if they are neighbors that share a common property that defines a component. The property may be color, brightness, range of brightness values, or anything else of interest. Pixels may be 4-connected or 8-connected [16] as shown in Figure 5.

- Pixels p and q are 4-connected if p and q both have the required property and q is in the 4-neighborhood of p.
- Pixels p and q are 8-connected if p and q both have the required property and q is in the 8-neighborhood of p.

Fig. 5. Connected component Labeling (a)4-connected neighbor (b)8-connected neighbor

A set (object) S of pixels is a connected component if there is at least one path in S that joins every pair \{p, q\} of pixels in S. The path must contain only pixels in S. During the end of this process, grouping of the pixel of same characters separately is done.

Morphology process is someway used in the image processing technique to extract or modify the information of shape and structure of the object which contained in the image[7]. Dilation is one of the basic operations in mathematical morphology. In binary morphology, dilation is a shift-invariant. 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 [17]. The dilation of the binary image A by the structuring element B is defined by:

\[ A \oplus B = \bigcap_{b \in B} A_b \]

where \( A_b \) is the translation of A by b

Erosion is one of two fundamental operations in morphological operations. In the erosion the basic effect of the operator on a binary image is to erode away the boundaries of regions of foreground pixels (i.e. white pixels, typically). Thus areas of foreground pixels shrink in size, and holes within those areas become larger [17].

\[ A \ominus B = \bigcap_{b \in B} A_{-b} \]

After applying morphological operators in the image, unwanted dots in the background is removed.

D. Edge Detection

The purpose of edge detection in general is to significantly reduce the amount of data in an image, while preserving the structural properties to be used for further image processing. The probability of detecting real edge points should be maximized while the probability of falsely detecting non-edge points should be minimized. This corresponds to maximizing the signal-to-noise ratio. The algorithm runs in 5 separate steps:

1. Smoothing: Blurring of the image to remove noise.
2. Finding gradients: The edges should be marked where the gradients of the image has large magnitudes.
3. Non-maximum suppression: Only local maxima should be marked as edges.
4. Double thresholding: Potential edges are determined by thresholding.
5. Edge tracking by hysteresis: Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

The result obtained is shown in Figure 6.

It shows the boundaries of the bacilli and other objects are also ignored.

III. RESULTS AND CONCLUSION

The bacilli detection was tested and implemented using MATLAB R2012b version for few images. Throughout the visual inspection and comparison with the manual segmentation image, it can be seen that the algorithms is able to segment TB bacilli. As the main objective is to detect the presence of bacilli, the higher TB segmentation rate is important to avoid missed segmentation of TB bacilli during performing the task.

Manual segmentation images are prepared through manually tracing the TB bacilli and removing the background. Bacilli is highlighted in the original image as shown in Figure 7.

Fig. 6. Edge Detection

For quantitative evaluation of the method, a pixel area comparison between the resultant image and manual segmentation is performed. Finally, the bacilli are detected accurately in a short period of time. Further, the work may be extended using neural network to find the characteristics of...
the bacteria for classification. By this system, it could be possible to perform the bacilli detection with still better accuracy even on different kinds of staining.

REFERENCES


