

# DETECTION OF TUBERCULOSIS USING CHEST X RAY (CXR)

P.SAHANA<sup>1</sup>, T.SARAVANA KUMAR<sup>2</sup> and A.RISVIN<sup>3</sup>

<sup>1</sup>Department of Computer Science and Engineering, Dr.SivanthiAditanar College of Engineering, Tiruchendur.

<sup>2</sup>Department of Computer Science and Engineering, Dr.SivanthiAditanar College of Engineering, Tiruchendur.

<sup>3</sup>Department of Computer Science and Engineering, Dr.SivanthiAditanar College of Engineering, Tiruchendur.

## **ABSTRACT**

*Tuberculosis is an airborne disease that affects many organs in the body especially lungs. This disease is caused by a bacteria known as Mycobacterium Tuberculosis. When the bacteria becomes active it affects the body. If the disease is not treated properly a loss of life may occur. A robotic detection of tuberculosis is presented in this paper with the help of patient chest x ray(CXR).The input image is then filtered by Gaussian filter to remove noise and then the lung region gets segmented by using graph cut segmentation. The segmented lung region is partitioned into four lobes. The infected region is then segmented for that region the feature values are calculated. With these values it is classified as normal or abnormal by using Ada boost classifier.*

## **KEYWORDS**

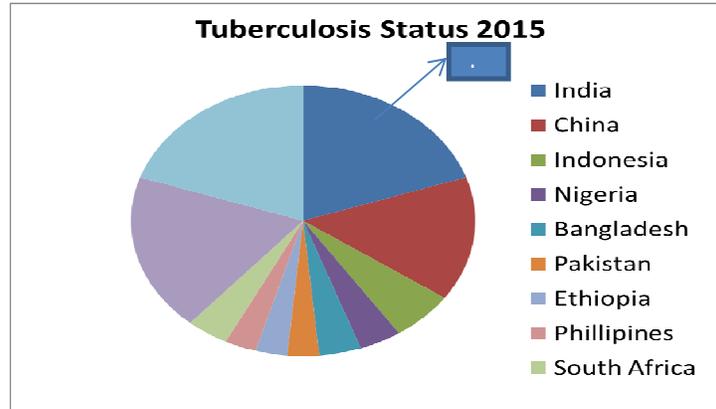
*CXR, Tuberculosis, AdaBoost, lobes, Classification, Feature extraction.*

## **1. INTRODUCTION**

### **1.1 TUBERCULOSIS**

Tuberculosis, a disease in many countries especially in India a burden one. A precise diagnostic method is necessary in order to identify the disease at an early stage to avoid spreading to others. Microscopic sputum culture is examined in many countries to identify the bacteria. But this is slow process to identify whether the bacteria is present in the sample. So a premature detection is necessary to avoid spreading of bacteria. There are two types of tuberculosis one is latent and the other is active. When the bacteria enters the body and starts fighting against with immune system and stop spreading bacteria inside the body but it cannot kill those bacteria which means latent. If immune system fails to stop preventing from spreading of those bacteria means it is active. It can take two to four weeks to start spreading after entering inside the body and affects mainly lungs and then to bone and nervous system. The symptoms are not known in latent TB in-case of active TB symptoms are clearly known to the infected patient. If the patient takes medicine for nearly six months if they are latent stage they can be prevented from loss of their life.

Figure 1: Chart that represent the Tuberculosis disease in world wide



Therefore it is necessary to eradicate the disease. For that spreading of the disease should be reduced. In order to avoid spreading it is essential to identify the presence of Tuberculosis disease in latent stage itself. Once the presence of disease is identified then proper medicine could save their life.

## 1.2 RELATED WORK

In order to detect Tuberculosis many research work is going on. But detecting tuberculosis with CXR is difficult because of the suppression of rib bones. Different techniques are used to segment the lung region by using combined lung masks, multiclass regulation parameters. Clavicle segmentation can be employed to remove clavicles from CXR images. So that identification of TB is easier. For the past 10 years there are many papers are published in CAD systems with CXR as input images. The developed CAD systems are successful for detecting cancer. With CXR images segmentation of lung region is difficult task due to the suppression of the clavicles. With many CXR images the regions are extracted with lung models. Segmentation masks are used for segmenting the lung regions. Average of three different masks is used. But combining the lung masks is a difficult task. With shape and texture the lung regions are segmented as normal and abnormal. Lung region gets segmented robustly with two stages one as by using average lung shape model and the other is done with lung boundary detection. Multiclass regulation parameters learning are used for graph cut image segmentation.

## 2. METHOD

### 2.1 OVERALL ARCHITECTURE

The diagram presents the implementation methods for preprocessing, segmentation, partitioning, lesion extraction and then feature extraction and classification by using Ada boost which results the image as abnormal and also it indicates the infected region is present in which lobe too.

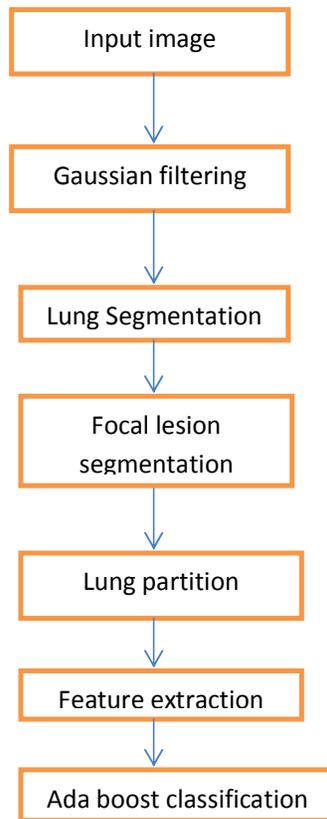


Figure2: shows the architecture of the system processing at each steps

## 2.2 ALGORITHM

- 1)The input for diagnosing tuberculosis is Chest X Ray(CXR) images.
- 2)Noise may present in the images so preprocessing is necessary to remove noises. Filtering is the preprocessing step which removes noise from the image. Gaussian filter, a filter used here.
- 3)Graph cut algorithm is used for segmenting the lungs from CXR images.
- 4)Then the segmented lungs are partition into upper, middle, lower and center parts.
- 5)Focal lesion is an infected area of tuberculosis mostly occur in the upper lobe of CXR images and it is detected
- 6)For that lesion feature values are calculated. Various features are skew, kurtosis, standard deviation.
- 7)Ada boost classifier is used to classify the image whether it is normal or abnormal. If it is abnormal then the abnormal lobe is specified.

## 2.3 GAUSSIAN FILTERING

Gaussian filter blur the image and then remove the noise from it. It is used in numerous research areas. The Gaussian filter works by using the 2D distribution as a point-spread function.

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad (2.3)$$

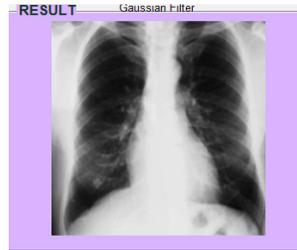


Figure 3: Gaussian Filter

## 2.4 GRAPH CUT SEGMENTATION

Segmentation is a process of separating individual objects from the whole image. There are various segmentation algorithms. The most relevant method for CXR image is graph cut based segmentation. In graph cut segmentation a graph contains vertex and edges. Assign a constant value for an image and also calculate the modulus value of an image. With these values construct a height and weight. By connecting four points in the graphs height and weight edges are calculated. Vertices are calculated from the modulus and edge value. Select one or more foreground and background pixel. For each pixel the pixel values are calculated. Then pixel values are again converted into pixel location once the segmentation is completed.

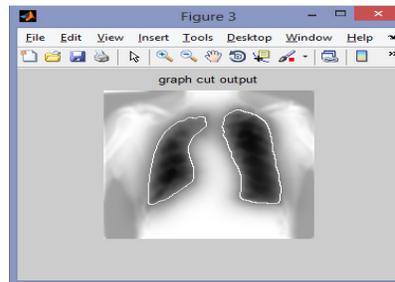


Figure 4: Graph cut output

## 2.5 LUNG PARTITION

The lungs are partitioned into four lobes as upper, middle, lower and central lobes. Tuberculosis mainly affects in the upper and lower lobes. The Lungs are partitioned by distance measures of rows and columns. So this partition helps in calculating the presence of disease exact manner

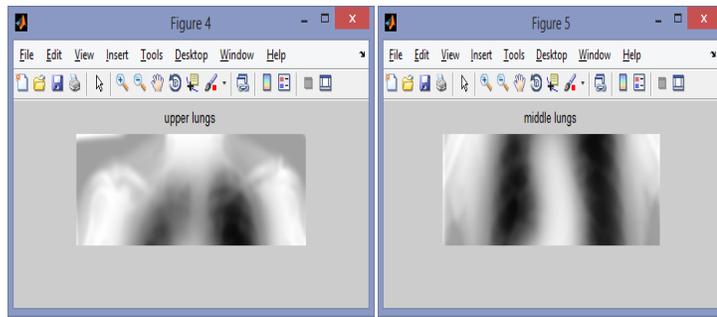


Figure 5: Upper and middle lobes

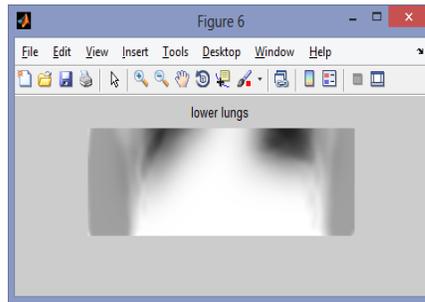


Figure 6: Lower lobes

## 2.6 FOCAL LESION

Focal lesion is an infected area. In CXR images focal lesion is the tuberculosis infected region. Tissues are more abnormal in this region. This region needs to be detected in order to detect the presence of tuberculosis in an accurate manner. This is detected by fixing a threshold value.



Figure 7: Infected region

## 2.7 FEATURE EXTRACTION

Feature extraction is necessary in image processing. The goal of feature selection is to classify the image as normal or abnormal. By eliminating the redundant features the classifier performance can be improved. The benefits of feature selection are to understanding the data, effectiveness and prediction. The classifier function is to determine whether or not a patient affected by Tuberculosis disease. When the feature generation is success the classifier performance can be improved much and so the disease presence can be predicted accurately. The various features

calculated for identifying the presence of this disease is Skewness, kurtosis, standard deviation, area and perimeter.

$$\text{Skewness} = \sum \left( \frac{x - \bar{x}}{\sigma} \right)^3 \dots\dots\dots 2.7.1$$

$$\text{Kurtosis} = \frac{\sum \left( \frac{x - \bar{x}}{\sigma} \right)^4}{n} - 3 \dots\dots\dots 2.7.2$$

$$\text{Standard Deviation} = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} \dots\dots\dots 2.7.3$$

$$\text{Mean} = \frac{\sum \bar{x}}{n} \dots\dots\dots 2.7.4$$

$$\text{area} = \text{bw area}(x) \dots\dots\dots 2.7.5$$

$$\text{perimeter} = \text{sum}(\text{sum}(\text{bwperim}(x))) \dots\dots\dots 2.7.6$$

$$\text{maxlen} = \text{regionprops}(x, \text{MajorAxisLength}) \dots\dots\dots 2.7.7$$

FEATURE EXTRACTION		
	FOCAL LESION	
	Left Lung	Right Lung
STANDARD DEVIATION	3.613975e+01	3.432031e+01
MEAN	2.310192e+01	2.204232e+01
SKEWNESS	4.309697e+00	4.307547e+00
KURTOSIS	4.345737e+00	4.371145e+00
SHAPE	151	498

Figure 2.7.1 shows feature extraction values

$$\text{minlen} = \text{regionprops}(x, \text{MinorAxisLength}) \dots\dots\dots 2.7.8$$

$$\text{elongation} = \left( \frac{\text{maxlen.MajAxisLen}}{\text{minlen.MinAxisLen}} \right) * 100 \dots\dots\dots 2.7.9$$

$$\text{majorax} = \text{maxlen}(x). \text{MajorAxisLength} \dots\dots\dots 2.7.10$$

$$\text{minax} = \text{minlen}(x). \text{MinorAxisLength} \dots\dots\dots 2.7.11$$

## 2.8 CLASSIFICATION

Ada boost is a device learning meta algorithm. In order to boost the performance of Ada boost algorithm it can be used with some other machine learning algorithm. If these boosting algorithms are employed then the weighted sum is the output of the algorithm. Ada boost is sensitive to pierce data and outliers. Boosting can be seen as minimization of a convex loss

function over a convex set of functions. Ada boost algorithm correctly predicts misclassified tuples as a correct one. It is simple to implement.

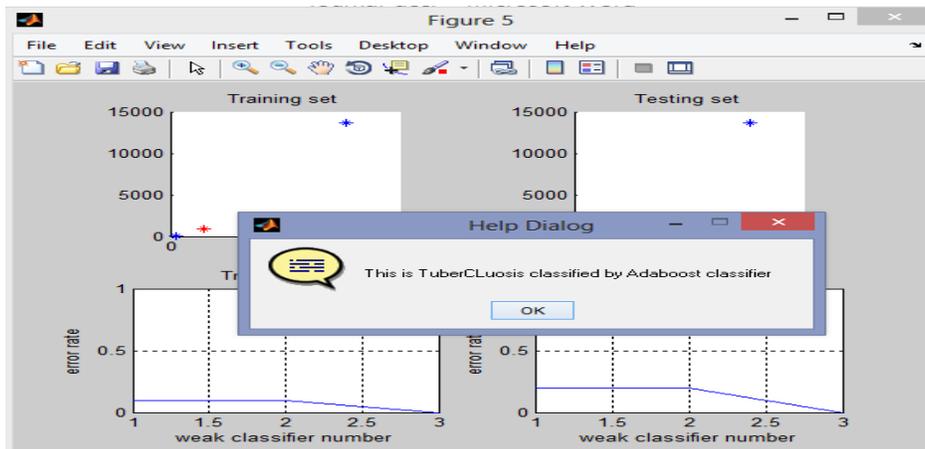
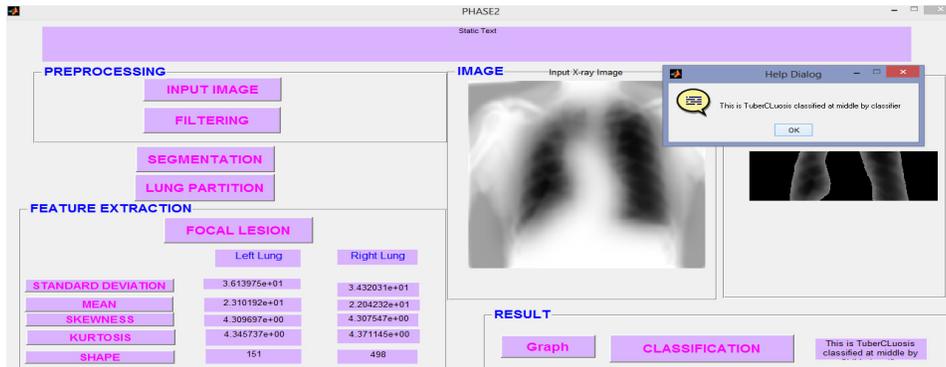


Figure 2.8.1 shows output of the classifier

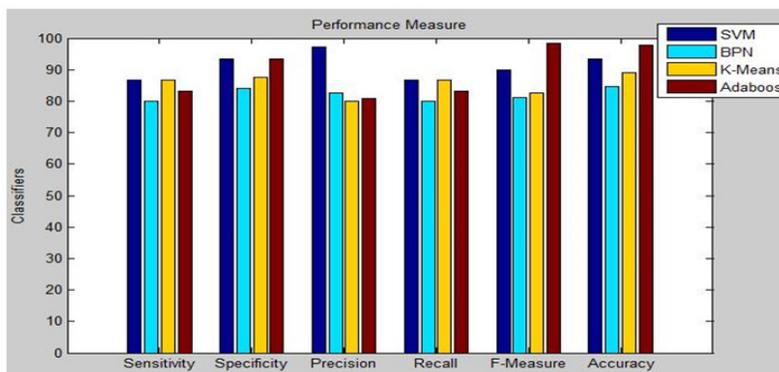


Figure: 2.8.2 shows sample comparative analysis of a disease classification with many classifiers in which Ada boost is the best one

### 3. CONCLUSION

Tuberculosis disease is a disease and it spreads in rapid manner by air when someone is infected by this disease. Therefore an early diagnosis is necessary in order to avoid from spreading of the disease. An automated system that diagnosis the presence of tuberculosis is necessary. By using Gaussian filter the noise is removed which makes the image more clear. The graph cut approach is applied to segment the lung portion and then by feature extraction the features are extracted. By using Ada Boost classifier the normal or abnormal portion of lung along with lung lobes are classified. Thus the presence of tuberculosis is detected using this automated system. This system is more useful in country where the presence of radiologist is rare.

## REFERENCES

- [1] [www.tbfacts.org/tb-statistics-india](http://www.tbfacts.org/tb-statistics-india)
- [2] [http://www.who.int/tobacco/resources/publications/factsheet\\_tub\\_tob.pdf?ua=1](http://www.who.int/tobacco/resources/publications/factsheet_tub_tob.pdf?ua=1)
- [3] S.Jaeger,A.Karargyris,S.Candemir,J.Siegelman, L. Folio, S.Antani,and G. Thom “Automatic screening for tuberculosis in chest radiographs. A survey,” *Quant. Imag. Med. Surg.*, vol. 3, no. 2, pp. 89–99,2013.
- [4] S.Sakai, H.Soeda, N.Takahashi radiography: Validation test on consecutive T1 cases of resectable lung cancer,” *J. Digit. Imag.*,vol 19, no. 4, pp. 376–382, 2000.
- [5] S.Candemir, K.Palaniappan, and Y.Akgul, “Multi-class regularization parameter learning for graph cut image segmentation,” in *Proc.Int. Symp. Biomed. Imag.*, 2013, pp. 1473–1476.
- [6] S.Candemir, S.Jaeger, K.Palaniappan, S. Antani, and G. Thoma,“Graph-cut based automatic lung boundary detection in chest radiographs,”in*Proc. IEEE Healthcare Technol. Conf.: Translat. Eng.Health Med.*, 2012, pp. 31–34.
- [7] P. Maduskar, L. Hogeweg, H. Ayles, and B. van Ginneken, “Performance evaluation of automatic chest radiograph reading for detection of tuberculosis (TB): A comparative study with clinical officers and certified readers on TB suspects in sub-Saharan Africa,” in *Eur. CongrRadiol.*, 2013.
- [8] G. Zhao and M. Pietikainen, “Dynamic texture recognition using local binary patterns with an application to facial expressions,” *IEEE Trans.Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915–928, Jun. 2007.
- [9] G. Srinivasan and G. Shobha, “Statistical texture analysis,” *Proc. World Acad. Sci., Eng. Technol.*, vol. 36, pp. 1264–1269, 2008.
- [10] J. Shiraishi, H. Abe, F. Li, R. Engelmann, H. MacMahon, and K. Doi,“Computer-aided diagnosis for the detection and classification of lung cancers on chest radiographs: ROC analysis of radiologists’ performance,” *Acad. Radiol.*, vol. 13, no. 8, pp. 995–1003, 2006.
- [11] K. Doi, “Current status and future potential of computer-aided diagnosis inmedical imaging, ”*Br. J.Radiol.*, vol. 78, no. 1, pp. 3–19, 2005. Ryszard S. Chora’s “Image Feature Extraction Techniques and Their Applications for CBI and Biometrics Systems”
- [12] S. Jaeger, A. Karargyris, S. Antani, and G. Thoma, “Detecting tubercu-losis in radiographs using combined lung masks,” in *Proc. Annu. Int.Conf. IEEE Eng. Med. Biol. Soc.*, 2012, pp. 4978–4981.
- [13] S. Chatzichristofis and Y. Boutalis, “FCTH: Fuzzy color and texture histogram—A low level feature for accurate image retrieval,” in*Proc. Int. Workshop Image Anal. Multimedia Interactive Services*, 2008, pp.191-196.
- [14] K. Palaniappan, F.Bunyak,P.Kumar,I.Ersoy,S. Jaeger, K.Ganguli, Haridas, J. Fraser, R. Rao, and G. Seetharaman, “Efficient feature extraction and likelihood fusion for vehicle tracking in low frame rate airborne video,” in*Proc. Int. Conf. Inf. Fusion*, 2010, pp. 1–8.
- [15] B. Van Ginneken, M. Stegmann, and M. Loog, “Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database,”*Med. Image Anal.*, vol.10, no. 1, pp. 19–40, 2006.
- [16] B. van Ginneken and B. terHaarRomeny, “Automatic segmentation of lungfields in chest radiographs,”*Med. Phys.*, vol. 27, no. 10, pp. 2445–2455, 2000.
- [17] Laurens Hogeweg\*, Clara I. Sánchez, PragnyaMaduskar, Rick Philipsen, Alistair Story, Rodney Dawson, Grant Theron, KeertanDhedha, Liesbeth Peters-Bax, and Bram van Ginneken “Automatic

Detection of Tuberculosis in Chest Radiographs Using a Combination of Textural, Focal, and Shape Abnormality Analysis”, IEEE transactions on medical imaging, vol34, no12, dec2015, pgn02429-2442.

- [18] Rusdah Department of Computer Science and Electronic, Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada Yogyakarta, Indonesia Edi Winarko ; Retantyo Wardoyo “Preliminary diagnosis of pulmonary tuberculosis using ensemble method” 2015 conferences.