An Adaptive Region Growing Algorithm with Support Vector Machine Classifier for Tuberculosis Cavity Identification

[1]Meenakshi Sundaram, K. and [2]C.S. Ravichandran

[1]Department of Information Technology,
Dhanalakshmi Srinivasan College of Engineering, Coimbatore, India
[2]Department of Electrical and Electronics Engineering, Sri Ramakrishna Engineering College, Coimbatore, India

ABSTRACT

The major world health menace is Tuberculosis (TB) that has the effect on one-third of the global population and millions of new infections are occurring every year. The cavities in the upper lung zone are a strong indicator that the disease has developed into a highly contagious state. The study includes 52 Chest X-ray images with Tuberculosis and 43 Chest X-ray images without Tuberculosis. The identification of the TB cavities is mostly conducted by the clinicians by observing the chest radiographs. But the automatic screening has lot of advantages such as substantial reduction in the labor workload of clinicians, enhancing the sensitivity of the test and better precision in diagnosis by increasing the number of images that can be analyzed by the computer. Many researchers have proposed different techniques to improve the performance of automatic screening process. This paper improves the accuracy over the existing technique using the adaptive region growing property and SVM classifier. Initially, pre-processing is carried out for the input image using Gaussian filtering technique to make the image suitable for further processing. The contours of the image will be obtained using region growing technique. The SVM classifier is then used to confirm the suspected TB cavities. The classification will be carried out by the features which the study has taken from the segmented image. The proposed technique is implemented in MATLAB and the performance is compared with the existing technique. From the result, the study has achieved eighty-five percentage accuracy over the existing technique's seventy eight-percentage accuracy.

1. INTRODUCTION

In developing countries Tuberculosis (TB) lasting as a communal health problem of global proportions, although the effective therapies reduced the fatality rate from the infectious pulmonary Tuberculosis (TB) (WHO, 2009). Timely detection of TB and the treatment is essential because of high infectivity and fatality rate. Although cavity in the lungs is not very common in primary TB, cavity in the upper half of the lung is more common post primary TB. The cavity in the upper half of

the lung denotes that the disease has developed into high infectivity stage (Curvo-Semedo *et al.*, 2005). Therefore, the detection of such cavities is very important to prevent further transmission of disease. The chest radiography is generally used as the primary detection tool for TB during the immigration medical examination (Cain *et al.*, 2008; Klinkenberg *et al.*, 2009). Although there has the techniques such as skin test and blood test for the diagnosis of TB, the Chest Radiography (CXR) is usually used for the patients who have the pulmonary symptoms (Hopewell *et al.*, 2006) or the prevalence of the disease is high in the population

Corresponding Author: Meenakshi Sundaram, K., Department of Information Technology, Dhanalakshmi Srinivasan College of Engineering, Coimbatore, India

(Curvo-Semedo *et al.*, 2005). The combination of radiographic result and clinical data helps the physicians to decide the possibility of TB infection. This paper aimed to detect the TB cavities precisely from the CXR technique. Currently the detection of TB cavities from the CXR technique is mostly conducted by the radiologist visually based on their knowledge and experience. The objective of Computer Aided Detection (CAD) is to make this diagnostic procedure very efficient. The CAD system is used for the diagnoses of various other diseases such as breast cancer and lung cancer (Doi, 2007).

The development of CAD system is based on three basic ideas. First, the basic strategy for the development of techniques for detection and quantization of lesions in medical images has been based on the understanding of process that would be involved in image readings by radiologists (Doi, 1999; 2006). This strategy is quite logical and easier because the radiologists are doing very complex and hard tasks of image reading and radiological diagnosis. Therefore, computer algorithms should be developed based on the interpretation of image readings such as how radiologists can detect certain lesions and how they are distinguishing the benign and malignant lesions. The second idea was related to the way in which the achievement of the efforts could be evaluated if the study were successful in the development of CAD. It appeared that the great proof of the achievement would be the use of CAD in routine clinical work at many hospitals around the world. The third idea is to promote the wide acceptance of the CAD concept and to facilitate the global distribution of CAD research by many investigators at diverse institutions.

The primary challenge in the cavity detection from the CXRs is the complex texture and varied intensity distribution in the lung field caused by the TB infection and the superimposed anatomical structures which may blur the boundaries of the cavities or even partially occlude the cavities. The presence of the superimposed anatomical structures makes the discrimination of image features and the interpretation of a CXR more challenging even for the clinical experts (Ginneken *et al.*, 2001). To eradicate the normal or unrelated anatomical structures in CXRs, different subtraction techniques were discussed in the literature. The Dual energy subtraction and temporal subtraction are the methods which are proposed to categorize the abnormalities particularly for subtle lesions. Dual energy subtraction (MacMahon *et al.*, 2008) is based on the exploitation of differential attenuation of low-energy X-ray photons by calcium to develop separate images for bones and soft tissues. Dual energy subtraction radiography is not completely available in remote

communities for clinical purpose. Temporal subtraction (MacMahon *et al.*, 2008; Zhao *et al.*, 2001) based on the previous radiograph of the same patient is not applicable in this case because the proposed work is interested in detecting abnormalities in the first CXR of a patient. Analysis of the CXRs shows that the conditions such as texture and shape in the individual lung could differ significantly. Therefore the subtraction technique cannot be implemented directly.

In this study, the study has pre-processed the image, segmented the lung region from that image, segmented the cavity region in that lung region, extracted some features for training the classifier and used the SVM classifier to identify the tuberculosis affected lung. The pre-processing is done by using the Gaussian Filter to avoid the noise in the input image and to increase the image quality. The lung segmentation is done by comparing the region growing technique and the Local Gabor XOR Pattern (LGXP) based region growing technique. The cavity segmentation is done by evaluating the pixel range in the segmented lung region and setting a threshold value from that evaluated pixels and comparing every pixel with that threshold value. After the lung and cavity segmentation, the proposed work has chosen some parameters to train the classifier to identify whether an x-ray image is a normal or tuberculosis affected. The classifier used in the proposed technique is SVM classifier. The SVM classifier is then trained using the parameters that has chosen from the sample chest x-ray images to identify the normal lung and tuberculosis affected lung.

2. MATERIALS AND METHODS

Lot of researches has been performed for the segmentation of normal and abnormal lung in chest x-ray image. Some of the modern related works regarding the identification of tuberculosis is as follows: Shen *et al.* (2010) have proposed a Hybrid Knowledge-Guided identification method for examining of contagious Pulmonary Tuberculosis from Chest Radiographs. The proposed computerized segmentation model, which takes a hybrid knowledge related Bayesian categorization approach to detect TB cavities automatically. They have implemented gradient inverse coefficient of disparity and circularity measures to classify detected features and confirm true TB cavities. By matching with non hybrid approaches and the traditional active contour techniques for feature extraction in medical images, experimental outcomes demonstrate that the proposed

approach achieves high precision with a low false positive rate in detecting TB cavities.

Polzehl *et al.* (2010) have presented Structural adaptive segmentation for numerical parametric mapping. The proposed work includes a novel structural adaptive Segmentation Algorithm (AS) that logically combines the signal detection with noise reduction in one procedure. Moreover, the novel method is very much related to a lately proposed structural adaptive smoothing algorithm and preserves shape and spatial extent of activation areas without blurring their borders.

Li *et al.* (2011) have presented an enhanced identification of Subtle Lung Nodules by utilizing the Chest Radiographs with Bone Suppression Imaging: Receiver Operating Characteristic examination. The reason of this article is to estimate radiologists' ability to identify subtle nodules by use of standard chest radiographs alone contrasted with bone suppression imaging used mutually with standard radiographs. The two image sets were examined by three skilled radiologists, with an interval of more than 2 weeks between the sessions. Receiver Operating Characteristic (ROC) curves, with and without localization, was gained to estimate the observers' performance. The mean value of the area under the ROC curve for the three observers was considerably enhanced, from 0.840 with standard radiographs alone to 0.863 with extra bone suppression images (p = 0.01). The area under the localization ROC curve was too enhanced with bone suppression imaging. The use of bone suppression images enhanced radiologists' presentation in the detection of subtle nodules on chest radiographs.

Dawson *et al.* (2010) have presented Chest radiograph analysis and recording system: estimation for tuberculosis screening in patients with highly developed HIV. This study provides proof of the good inter-observer pact using the CRRS regular reporting technique when used among patients with highly developed HIV-associated immunodeficiency and a high occurrence of culture-proven pulmonary TB. The usefulness of radiology as a diagnosis for TB in this patient group, however, remains restricted.

Iakovidis *et al.* (2009) have introduced Robust model-based identification of the lung field boundaries in portable chest radiographs assisted by selective threshold. Iakovidis *et al.* (2009) presented a new technique for the detection of the lung field boundaries in moveable chest radiographs of patients with bacterial pulmonary infections. Such infections are radio

graphically manifested as foci of consolidations that can lead to vague or unseen lung field limitations, difficult to distinguish even by experienced physicians. Conventional and advanced approaches address mostly stationary radiographs, whereas only some of them cope with pulmonary infections. The suggested technology is based on an active shape model including shape prior information about the lung fields. The model is initialized by a new technique using a set of most important points discovered on the peripheral anatomic structures of the lungs. A selective thresholding algorithm related on a spinal cord sampling procedure supports both the initialization and the development of the model for the detection of the lung field boundaries. The experiments show that the suggested technique outperforms state-of-the-art approaches.

Xu *et al.* (2010) have presented an enhanced fluid vector flow for cavity segmentation in chest radiographs. The proposed work enhances the technique from two aspects: edge leakage and control point selection. Experimental results of cavity segmentation in chest radiographs show that the suggested technique gives at least 8% progress over the original FVF method.

Reid and Shah (2009) have presented Approaches to tuberculosis screening and diagnosis in people with HIV in resource-limited settings. Tuberculosis is a major reason of morbidity and mortality in people who are affected by HIV/AIDS worldwide. Early diagnosis and treatment is crucial to addressing the dual epidemic of tuberculosis and HIV. Increasing recognition of the significance of combining tuberculosis services-together with screening-into HIV care has led to global policies and the beginnings of implementation of joint actions at the national level. However, debate remains about the most excellent type of screening for pulmonary tuberculosis amid people affected by HIV/AIDS in resource-limited settings. Mycobacterial culture, the gold standard for tuberculosis diagnosis, is very slow and difficult to use for screening test in such settings. More extensively available techniques, such as symptom screening, sputum smear microscopy, chest radiography and tuberculin skin testing have considerable shortcomings, particularly in people living with HIV/AIDS. However, until simpler, cheaper and more responsive diagnostics for tuberculosis are existing in peripheral healthcare settings, an approach must be developed that uses current evidence to combine available screening tools.

2.1. Proposed Technique for the Identification of Cavity

The process of this work is precisely explained in the **Fig. 1** as block diagram. In this figure, some sample chest x-ray images with tuberculosis and without tuberculosis. The sample images are then preprocessed and then send for segmenting the lung and cavity regions. After the lung and cavity regions are segmented from the sample images, some parameters are chosen to train the classifier. Similarly, the study has to preprocess the input image which is in need to find whether it is tuberculosis affected or not. After the preprocessing process, the study requires to segment the lung and the cavity region. Thereafter, the same parameters has to chosen which has been chosen in the sample images and give it to the classifier. The classifier used in the technique is SVM classifier. The SVM classifier then identify whether the input chest x-ray image is tuberculosis affected or not by comparing the parameters from the sample images and from the input image.

The main contributions of the proposed technique are:

- Using the Local Gabor XOR Pattern (LGXP) in normal region growing technique for Lung Segmentation
- Computed the pixel range in the lung region to get a threshold value to find the cavities in the lung region after comparing the threshold value with all the pixels in the lung region
- Extracted features such as number of cavities in the lung region, minimum area of the cavity region, maximum area of the cavity region, total number of pixels in each cavity, maximum repeated pixel intensity in the cavity region and maximum repeated pixel in the lung region to train the classifier

2.2. Pre-Processing

X-ray image cannot be fed directly as input. The input image is subjected to a set of pre-processing steps to make the image suitable for further process. The pre-process is used to load the input image to the MATLAB environment and it will remove the noise present in the input image. Here the Gaussian filter is used as pre-processing technique. The image is passed through the Gaussian filter to lower the noise and to get a better image. The Gaussian filter will also increase the image quality.

2.3. Gaussian Filter

A Gaussian filter Haddad and Akansu (1991) is a filter which has the impulse response as Gaussian function. Gaussian filters are created to exclude the overshoot of step function input while reducing the rise and fall time. The group delay in the Gaussian filter is less. In mathematical terms, a Gaussian filter converts the input signal by convolution with a Gaussian function. This process of changing the input is also called Weierstrass transform. The Gaussian function will not get the value as zero for the condition $x \in [-\infty, \infty]$ and would require an infinite window length. The Gaussian kernel is continuous and is not discrete. The cut-off frequency F_c of the filter is defined as the ratio between the sample rate R_s and the standard deviation σ:

$$F_c = \frac{R_s}{\sigma}$$

2.4. Lung Segmentation

Lung segmentation is a process of segmenting the lungs from the chest x-ray image. The normal process of region growing technique for segmenting the lungs is as follows. To segment the lungs, the proposed work has to choose a pixel from the chest x-ray image as default. Thereafter, a threshold value is set for comparison to find the pixel intensity for the lung area in the chest x-ray. The default pixel chosen already is then compared with the adjacent pixel values. If the difference between the default pixel and the adjacent pixel is greater than the threshold value, the study has to exclude that adjacent pixel. If the difference between the default pixel and the adjacent pixel is less than the threshold value, the adjacent pixel has to be included for region growing. Similarly, the proposed work has to compare all the pixels except the left pixels with its adjacent pixels by keeping one pixel as default. The process of normal region growing technique is shown in the **Fig. 2** as block diagram.

In this proposed work, the normal region growing technique is compared with the Local Gabor XOR Pattern (LGXP) based region growing technique to segment the lungs from the chest x-ray image. The LGXP technique is used to find the texture image.
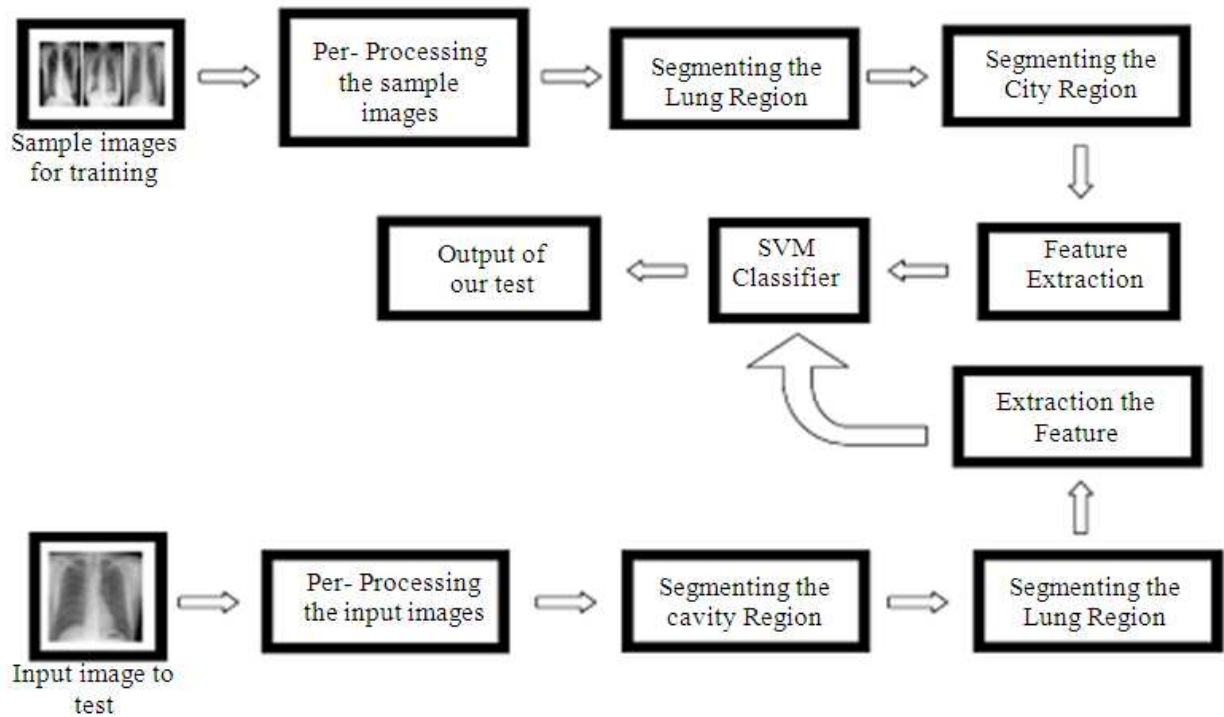
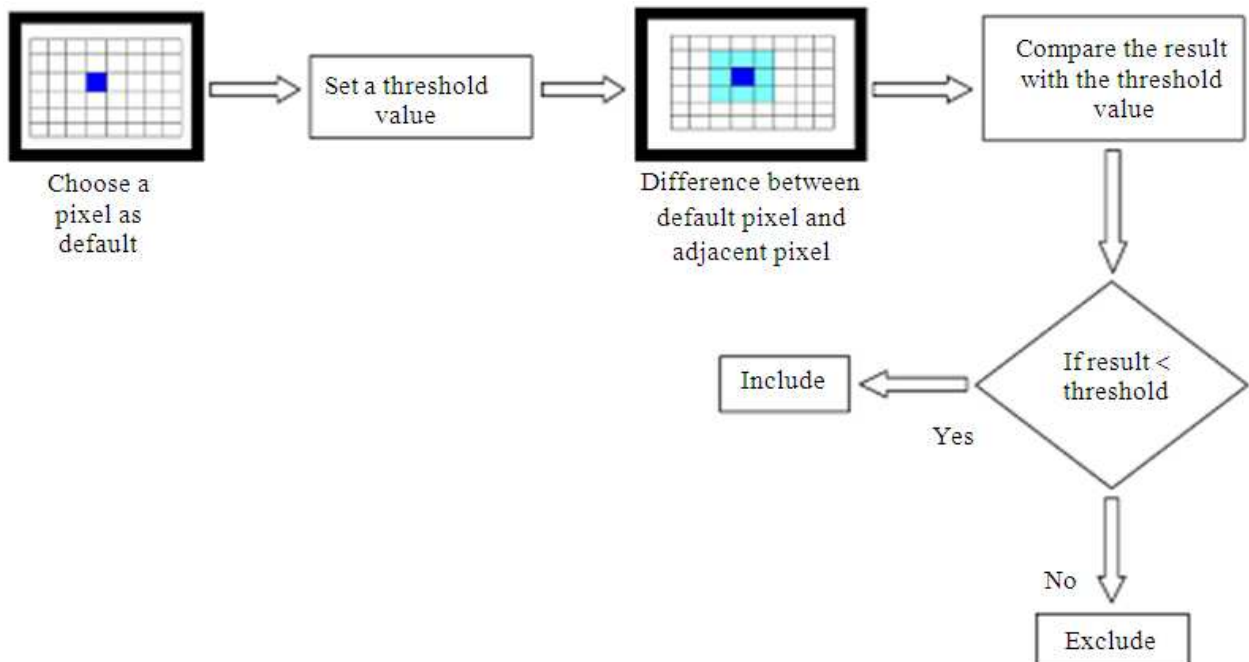**Fig. 1.** Block diagram of proposed technique

Fig. 2. Block diagram of normal region growing technique

The LGXP based region growing technique is as follows. In LGXP technique, the proposed work applies the Gabor Phase Technique on every pixel in the chest x-ray image. The Gabor Phase Technique will convert all the pixel values to phase values (0 to 360). After converting all the pixel values to phase values, the study needs to find these phase values comes under which quadrant. Each quadrant has certain values. For the first quadrant the value is zero and for the second quadrant the value is one and for the third quadrant the value is two and for the fourth quadrant the value is three. Thereafter, a default phase value of a pixel is chosen and checked under which quadrant this phase value comes and the study has to assign respective quadrant value to that pixel. After assigning respective quadrant value to the default pixel, the studies checks the adjacent pixel's phase values and assign the respective quadrant values to those adjacent pixels. Thereafter, the proposed work converts the adjacent pixel's value as zero which has the same quadrant value of the default pixel. If the adjacent pixels value does not have the same quadrant value of the default pixel, convert the adjacent pixel's value as one. Now the pixel values would be like binary values as zeros and ones. After converting the pixel values as binary format, then make that binary format as decimal value and apply that decimal value to the default pixel. The process of taking the binary value is shown in the figure. Similarly, apply this LGXP process for all the pixels in the chest x-ray by keeping one pixel as default. The sample process of LGXP technique is shown in the **Fig. 3** as block diagram.

After applying the LGXP technique in all the pixels, the work implements the region growing technique for segmenting the lungs using the phase value of the pixels that is got from the LGXP process. Thereafter, the normal region growing technique and LGXP based region growing technique is compared. While comparing both the techniques, the study checks for the same pixel as default. During this process, if the difference between the adjacent pixel and the default pixel got the value as less than the threshold value on both the techniques separately, the work includes that adjacent pixel for region growing or else the work needs to exclude that adjacent pixel. But the adjacent pixel and the default pixel which has been chosen to compare should be same on both the techniques.

2.5. Local Gabor XOR Pattern (LGXP)

The fundamental idea of the proposed technique (Xie *et al*., 2010) is to ease the sensitivity of Gabor phase to

the differing positions, whether two phases reflect same local feature must be determined in a "looser" way. Specifically, if two phases belongs to the same interval (for instance: 0^0, 90^0), they are believed to have similar local features or else they reflect different local features. In the section, the LGXP descriptor is depicted using the images.

The **Fig. 4** shows an instance for the LGXP encoding method where the phase is quantized into 4 ranges. In the proposed LGXP technique, phases are first quantized into disparate ranges and the LGXP operator is applied to the quantized pixels of the central pixel and each of its neighboring pixels and eventually the result of the binary labels are concatenated together as a local pattern of the central pixel. In the **Fig. 3** is the matrix with initial phase of the pixels after applying the Gabor filter and (b) is the result after quantization and (c) is the result after XOR comparison with the center quantized value. From the matrix which the study has got after XOR comparison, then deduce the binary value obtained is 01011101 and its equivalent decimal value is 93. The pattern of LGXP in binary and decimal form is as follows:

$$LGXP^i_{\mu v} = (p_c)\left[LGXP^N_{\mu v}, LGXP^{N-1}_{\mu v}, ..., LGXP^1_{\mu v}\right]_{binary}$$

$$= \left[\sum_{i=1}^N 2^{i-1}.LGXP^i_{\mu v}\right]_{decimal}$$

where, P_c denotes the central pixel in the Gabor phase map with scale v and orientation, μ, N is the size of the neighborhood and $LGXP^i_{\mu v}$ $\tau = 1, 2, ..., N$ denotes the pattern calculated between P_c and its neighbor P_i, which is computed as follows:

$$LGXP^i_{\mu v} = q(\Phi_{\mu v}(P_c)) \otimes q(\Phi_{\mu v}(P_i)),$$
$$i = 1, 2, ..., N$$

where, $\Phi_{\mu v}$ denotes the phase, $\otimes$denotes the LXP operator, which is based on XOR operator, q denotes the quantization operator which calculates the quantized code of the phase according to the number of phase ranges:

$$a \otimes b = \begin{cases} 0, \text{if } a = b \\ 1, \text{else} \end{cases}$$
$$q(\varnothing_{\mu v}(\bullet)) = i, \text{if}$$
$$\frac{360 * i}{e} < \varnothing_{\mu v}(\bullet) < \frac{360 * (i+1)}{e},$$
$$i = 0, 1, ..., b-1$$
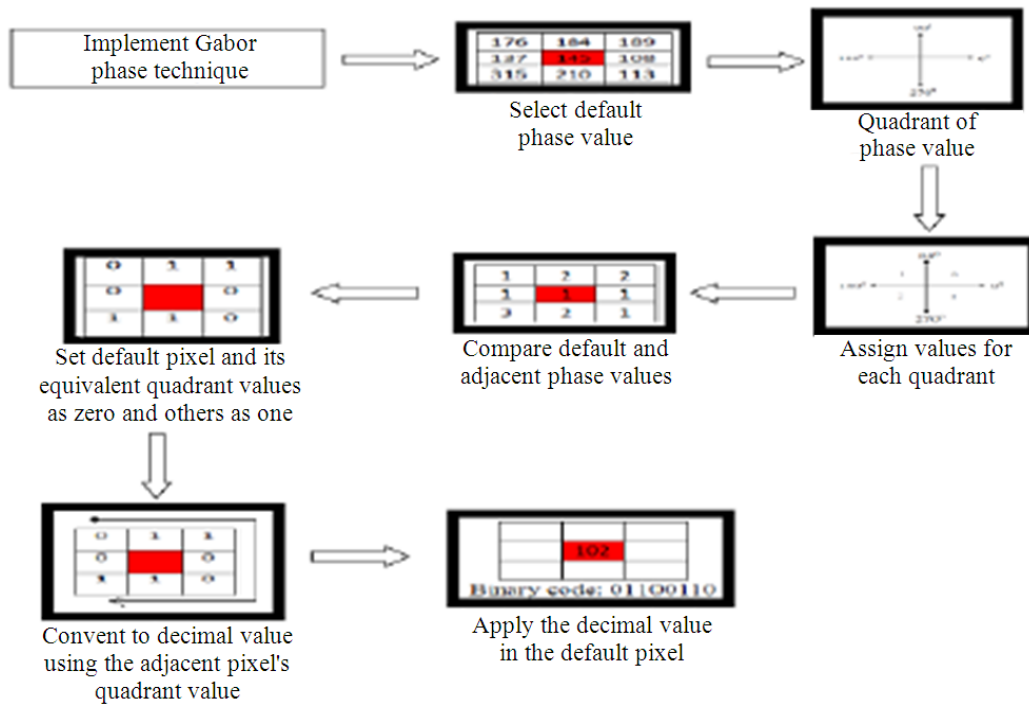$$q(\varnothing_{\mu v(\bullet)}) = i \text{ if}$$
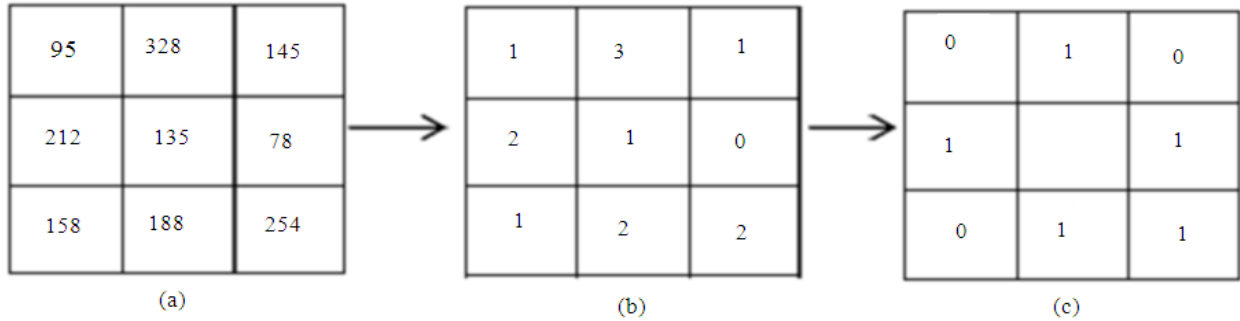
Fig. 3. Block diagram of LGXP Technique

Fig. 4. Example of LGXP method where the phase is quantized into 4 ranges

where, e denotes the number of phase ranges. With the pattern explained above, one pattern map is computed for each Gabor kernel. Thereafter, each pattern map is split into m non-overlapping sub blocks and the histograms of all the sub blocks of scales and the orientations are concatenated to form the proposed LGXP descriptor of the input face image:

$$H = \left\lfloor H_{\mu_0 v_0 1}, \ldots\ldots H_{\mu_0 v_0 m 1}; \ldots\ldots; H_{\mu_{0-1} v_{s-1} 1}, \ldots\ldots; H_{\mu_{0-1} v_{s-1} m} \right\rfloor$$

where, $H_{\mu v}$ (i = 1,2,….m) denotes the histogram of the ith sub block of the LGXP map with scale v and orientation μ.

2.6. Cavity Segmentation

After the lung segmentation, the study identifies the cavities in the lung. The cavities present in the lung region are an essential thing to identify the tuberculosis affected lung. To identify the cavity in the lung, an adaptive threshold value is set. The threshold value is chosen by calculating the pixel range in the lung region and dividing that pixel range by two. Thereafter, the comparison is done with the threshold of all the pixels. While the comparison is done for the pixels to the threshold value, if the pixel value is

greater than the threshold value then it would be the cavity region and if the pixel value is less than the threshold value then it would be the lung region. The **Fig. 5** shows the block diagram for segmenting the cavity region from the lung region.

2.7. Feature Extraction

After finding the regions, the work extracts some features to diagnose the disease in the lung. To discover the disease in the lung, the extracted feature has to be fed as the input to the classifier, because the extracted features will give vital information about the region which is used to train the classifier. The classifier which used here is SVM classifier. The features extracted are number of cavities in the lung region, minimum area of cavity region, maximum area of cavity region, total number of pixels in each cavity, maximum repeated pixel intensity in the cavity region and maximum repeated pixel in the lung region.

The total number of cavities in the lung region is identified in this study. Because the normal lung would also have some cavities present in its region. So in order to distinguish the normal lung image and the tuberculosis affected lung, the total numbers of cavities present in the lung region is found and then the result is given to the SVM classifier:

$$f1 = \{c_1, c_2, c_3, \ldots\ldots, c_n\}$$

Where:
f_1 = No. of cavities in the lung region
c_1 = First cavity
c_2 = Second cavity
c_3 = Third cavity
c_n = n^{th} number of cavity

Thereafter, the work carries out computation of the area of each cavity in the lung region. After computing the area of each cavity in the lung region, the cavity is selected with maximum area and minimum area. The cavity with maximum area and the cavity with minimum area are essential feature to distinguish the normal lung and the lung with tuberculosis. Because the value of the maximum cavity area of the tuberculosis affected lung would have larger value when compare to the value of the maximum cavity area of a normal lung. The result of maximum area and minimum area of the cavity of lung after computation is give to the SVM classifier:

$$f_2 = \max \{ac_1, ac_2, ac_3, \ldots\ldots, ac_n\}$$

Where:
f_2 = Maximum area of cavity region
ac_1 = of first cavity
ac_2 = Area of second cavity
ac_3 = Area of third cavity
ca_n = Area of nth cavity

$$f_3 = \min \{ac_1, ac_2, ac_3, ac_n, \ldots\ldots,,,\}$$

where, f_3 = Minimum area of cavity region.

Discovering total number of pixels in each cavity is also a crucial feature for distinguishing a normal lung and a lung with tuberculosis. Because the tuberculosis affected lung would have more pixels in its cavities when compared to the pixels in the cavities of a normal lung. Therefore, the number of pixels is found in each cavity of a lung and gives the result to the classifier:

$$f_4 = \sum_{i=1}^{N} ac_i$$

where, f_4 = Total pixels in each cavity.

Thereafter, the work has to find the maximum repeated pixel intensity in the cavity regions of a lung. To discover the maximum repeated pixel, the intensities of all the pixels has to be found in each cavity of a lung by implementing histogram and thereafter the comparison of all the pixels of every cavity with each other is needed. After discovering the maximum repeated pixel in the cavities of a lung, the result is given to the classifier. Similarly, the maximum repeated pixel in the whole lung region is identified and the result is given to the classifier. The SVM classifier which will then compare all the features and it will find out whether the lung is affected by tuberculosis or not by comparing all the features.

2.8. Training and Testing Using SVM

To train the SVM classifier, some data features are needed to identify the normal lung region and tuberculosis affected lung region. The data features will then train the classifier and the classifier will find whether the given x-ray image is normal or abnormal. The data features that are chosen for training the SVM classifier are number of cavities in the lung region, maximum area of cavity in the lung region, minimum area of cavity in the lung region, total number of pixels in each cavity, maximum repeated pixel in the cavity regions together and maximum repeated pixel in the lung region. After computing all the data features, the values are given to the classifier.
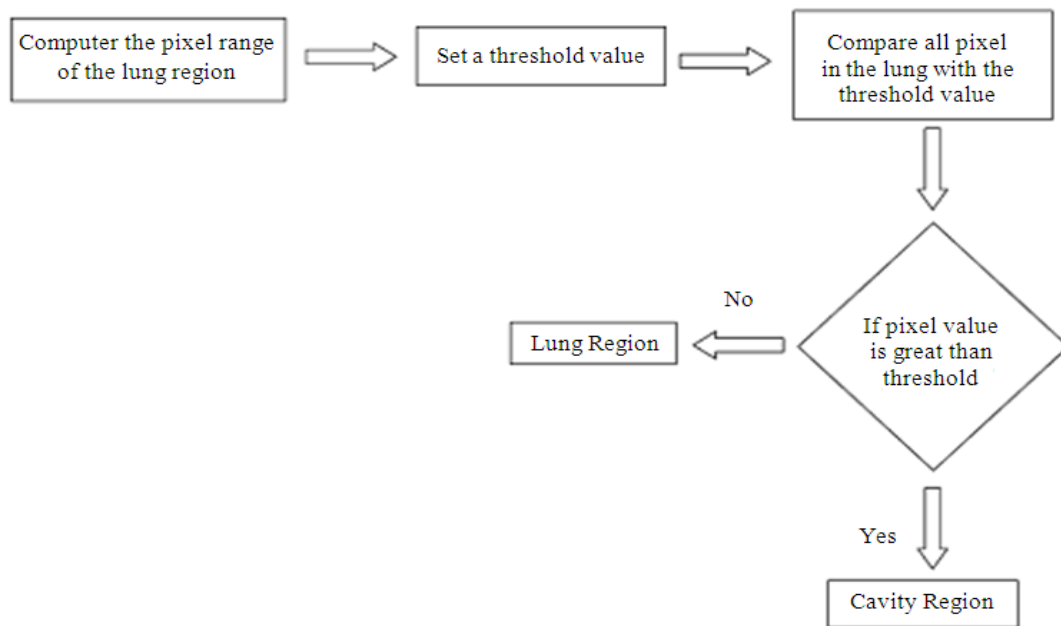
Fig. 5. Block diagram of Cavity Segmentation

For instance, if there is five normal x-ray images and five abnormal x-ray images, calculation is done for all the six data features separately for the x-ray images. After calculating all the six data features for every chosen x-ray images, the result is given to the SVM classifier. Using those results the classifier is trained to identify the normal and abnormal lung from the given x-ray image. After the SVM classifier is trained, a new x-ray image is given to find whether it has tuberculosis or not. Thereafter, the six data features such as number of cavities in the lung region, maximum area of the cavity region, minimum area of the cavity region, total number of pixels in each cavity, maximum repeated pixel in the cavity region and maximum repeated pixel in the lung region are computed for the new x-ray image. The computed values of all the six data features are then give to the SVM classifier.

The SVM classifier is then compare the values of all the six data features with the stored values of normal and abnormal x-ray images. Because during training all the six data features of the five normal x-ray images and five abnormal x-ray images are stored. After comparison, the SVM classifier will identify whether the given x-ray image comes under normal category or abnormal category and give the result to us.

2.9. Support Vector Machine (SVM)

In most cases, an object is assigned to one of several categories based on some of its characteristics in the real life situation. For instance, based on the outcome of several medical tests the diagnosis is done whether the patient has a particular disease or not. In computer science such situations are explained as classification issue.

The Support Vector Machine (SVM) which was derived from the statistical theory is a powerful supervised classifier and is an accurate learning technique. The SVM was introduced in 1995. It gives successful classification outcomes in different application domains such as medical diagnosis (Guyon *et al*., 2002; Zhang and Liu, 2004). SVM works under the principle of structural risk reduction from the statistical learning theory. To maximize the margin between the classes and to minimize the true cost (Zhang *et al*., 2006) its kernel is used to control the empirical risk and categorization capacity. A support vector machine can search an optimal separating hyper plane amid the members and non-members of a given class in a high dimension feature space (Kim and Park, 2003). There are many general kernel functions such as linear, polynomial of degree

and Radial Basis Function (RBF). Among these kernel functions, a radial basis function proves to be useful because of the fact the vectors are mapped nonlinearly to a very high dimension feature space.

3. RESULTS AND DISCUSSION

This section explains the experimental results of the adaptive proposed segmentation technique using the x-ray images with and without tuberculosis. The proposed technique is implemented in MAT LAB. Here, the proposed segmentation technique is tested using the x-ray images that are taken from the medical hospitals.

3.1. Chest X-Ray Image Dataset Description

The x-ray image used for the image segmentation technique is taken from the resources that are available publicly. The study includes 52 Chest X-ray images with Tuberculosis and 43 Chest X-ray images without Tuberculosis. The **Fig. 6** shows the sample chest x-ray images which are not affected by tuberculosis (normal) and the sample images which are affected by tuberculosis (abnormal).

3.2. Experimental Results

The sample images that have been taken from the available resources are filtered using Gaussian Filtering technique. The Gaussian filtering technique is used remove the noises which the user can't identify with his/her naked eye. The Gaussian filter thus improves the quality of the images.

The sample images after applied the filtering technique are given to the process of lung segmentation. The lung segmentation process only segments the lung region from the sample x-ray images. The **Fig. 7** shows a sample image of segmented lungs with tuberculosis and without tuberculosis.

After the lung is segmented from the sample images, the cavities are segmented from the lung region. Using the cavities in the lung region, the lung image is diagnonised whether it is affected with tuberculosis or not. The **Fig. 8** shows a sample image of segmented cavities and segmented cavities with x-ray image for the tuberculosis affected lung.

3.3. Performance Analysis using Evaluation Metrics

The evaluation of the tuberculosis identification in different images is carried out using the following metrics:

$$\text{Sensitivity} = \text{TP} (\text{TP} + \text{FN})$$
$$\text{Specificity} = \text{TN} (\text{TN} + \text{FP})$$
$$\text{Accuracy} = (\text{TN} + \text{TP})/(\text{TN} + \text{TP} + \text{FN} + \text{FP})$$

Where:
TP = True Positive
TN = True Negative
FN = False Negative
FP = False Positive

As indicated by the above equations.

Sensitivity is the proportion of true positives that are correctly identified by a diagnostic test. It shows how good the test is at detecting a disease.

Specificity is the proportion of the true negatives correctly identified by a diagnostic test. It suggests how good the test is at identifying normal (negative) condition.

Accuracy is the proportion of true results, either true positive or true negative, in a population. It measures the degree of veracity of a diagnostic test on a condition. Either true positive or true negative, in a population. It measures the degree of veracity of a diagnostic test on a condition.

The **Table 1** shows the tabular column for the accuracy comparison of training dataset and testing dataset in normal region growing technique and LGXP based region growing technique. The tabular column is plotted after the calculations of sensitivity, specificity and accuracy. In this tabular column, the accuracy of the training dataset got is eighty three percentages using normal region growing technique and hundred percentages using the LGXP based region growing technique. Also, the accuracy of the testing dataset is sixty two percentages using normal region growing technique and seventy five percentages using LGXP based region growing technique.

3.4. Comparative Analysis

The proposed region growing technique is compared with the existing technique. The **Table 2** shows the block diagram of accuracy comparison between the proposed technique and the existing technique. The tabular column shows that the accuracy is eighty five percentages for the proposed technique and seventy eight percentages for the existing technique. From this reading, it can be inferred that the proposed technique shows better performance than the existing technique.

Lungs without tuberculosis (Normal)

(A)
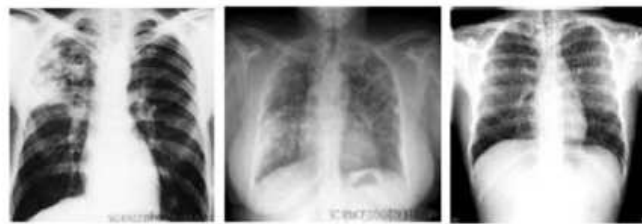
Lungs without tuberculosis (Abnormal)

(B)

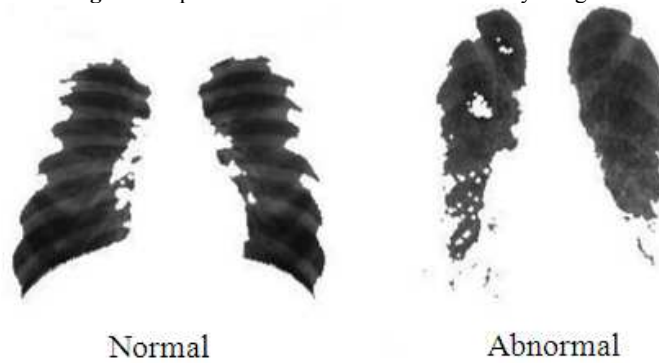**Fig. 6.** Sample normal and abnormal chest x-ray images

Normal   Abnormal

Fig. 7. Sample images of segmented lung with tuberculosis and without tuberculosis

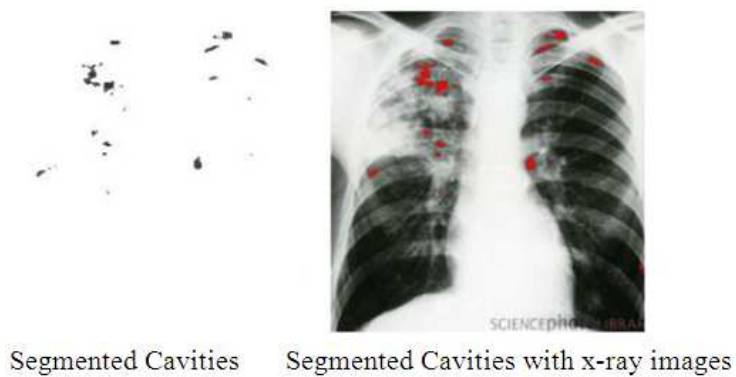Segmented Cavities   Segmented Cavities with x-ray images

Fig. 8. Segmented cavities with x-ray image and without x-ray image for the tuberculosis affected lung

Table 1. Accuracy detection of the proposed method in training and testing dataset

		Training dataset		Testing dataset	
	Region Evaluation metrics	LGXP based growing	Region region growing	LGXP based growing	Region growing
Input x-ray image dataset	True positive	3.000	3	3.000	3.00
	True negative	2.000	3	2.000	3.00
	False positive	1.000	0	2.000	1.00
	False negative	0.000	0	1.000	1.00
	Sensitivity	1.000	1	0.750	0.75
	Specificity	0.667	1	0.500	0.75
	Accuracy	0.830	1	0.625	0.75

Table 2. Comparative analysis of existing technique with the proposed technique

	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy
Existing technique Shen *et al*. (2010)	4	7	3	0	1.00	0.70	0.786
Proposed technique	3	9	1	1	0.75	0.90	0.857

4. CONCLUSION

In the study, an effective adaptive technique is proposed for the identification of tuberculosis in the lung region. The proposed technique contains pre-processing, lung segmentation, cavity segmentation, feature extraction, training and testing using SVM. The chest x-ray image dataset used for the proposed technique is taken from the publically available resources. The performance of the proposed technique and the existing technique is analyzed using evaluation metrics. To evaluate these metrics, some terms like True Positive, True Negative, False Positive and False Negative are needed. After evaluating these metrics, the performance of the proposed technique is better when compared to the existing technique in terms of accuracy. The result shown that the accuracy of the proposed technique is eighty-five percentages over the existing technique's seventy-eight percentage accuracy.

5. REFERENCES

Cain, K.P., S.R. Benoit, C.A. Winston and W.R. Mac Kenzie, 2008. Tuberculosis among foreign-born persons in the United States. J. Am. Med. Assoc., 300: 405-412. DOI: 10.1001/jama.300.4.405

Curvo-Semedo, L., L. Teixeira and F. Caseiro-Alves, 2005. Tuberculosis of the chest. Eur. J. Radiol., 55: 158-172. DOI: /10.1016/j.ejrad.2005.04.014

Dawson, R., P. Masuka, D.J. Edwards, E.D. Bateman and L.G. Bekker *et al*., 2010. Chest radiograph reading and recording system: Evaluation for tuberculosis screening in patients with advanced HIV. Int. J. Tuberculosis, Lung, Dis., 14: 52-58. PMID: 20003695

Doi, K., 1999. Computer-Aided Diagnosis in Medical Imaging. 1st Edn., Elsevier Science Limited, Amsterdam, ISBN-10: 0444500588, pp: 563.

Doi, K., 2006. Diagnostic imaging over the last 50 years: Research and development in medical imaging science and technology. Phys. Med. Biol., 51: 5-27. DOI: 10.1088/0031-9155/51/13/R02

Doi, K., 2007. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Comput. Med. Imag. Graph, 31: 198-211. DOI: 10.1016/j.compmedimag.2007.02.002

Ginneken, V., B.M. Ter Haar Romeny and M.A. Viergever, 2001. Computer-aided diagnosis in chest radiography: A survey. IEEE Trans. Med. Imag., 20: 1228-1241. DOI: 10.1109/42.974918

Guyon, I., J. Weston, S. Barnhill and V. Vapnik, 2002. Gene selection for cancer classification using support vector machines. Mach. Learn., 46: 389-422. DOI: 10.1023/A:1012487302797

Haddad, R.A. and A.N. Akansu, 1991. A class of fast gaussian binomial filters for speech and image processing. IEEE Trans. Acoustics, Speech Sig. Proces., 39: 723-727. DOI: 10.1109/78.80892

Hopewell, P.C., M. Pai, D. Maher, M. Uplekar and M.C. Raviglione, 2006. International standards for tuberculosis care. Lancet Infect. Dis., 6: 710-725. PMID: 17067920

Iakovidis, D.K., M.A. Savelonas and G. Papamichalis, 2009. Robust model-based detection of the lung field boundaries in portable chest radiographs supported by selective thresholding. Measurement Sci. Technol., 20: 104019-104019.

Kim, D. and J. Park, 2003. Network-based intrusion detection with support vector machines. Lecture Notes Comput. Sci., 2662: 747-756. DOI: 10.1007/978-3-540-45235-5_73

Klinkenberg, E., D. Manissero, J.C. Semenza and S. Verver, 2009. Migrant tuberculosis screening in the EU/EEA: Yield, coverage and limitations. Eur. Respir. J., 34: 1180-1189. DOI: 10.1183/09031936.00038009

Li, F., T. Hara1, J. Shiraishi, R. Engelmann, H. MacMahon and K. Doi, 2011. Improved detection of subtle lung nodules by use of chest radiographs with bone suppression imaging: Receiver operating characteristic analysis. Am. J. Roentgenol., 196: 535-541. DOI: 10.2214/AJR.10.4816

MacMahon, H., F. Feng, R. Engelmann, R. Roberts and S. Armato, 2008. Dual energy subtraction and temporal subtraction chest radiography. J. Thorac. Imag., 23: 77-85. DOI: 10.1097/RTI.0b013e318173dd38.

Polzehl, J., U.H. Voss and K. Tabelow, 2010. Structural adaptive segmentation for statistical parametric mapping. Neuroimage, 52: 515-523. DOI: 10.1016/j.neuroimage.2010.04.241

Reid, M.J.A. and N.S. Shah, 2009. Approaches to tuberculosis screening and diagnosis in people with HIV in resource-limited settings. Lancet Infect. Dis., 9: 173-184. DOI: 10.1016/S1473-3099(09)70043-X

Shen, R., I. Cheng and A. Basu, 2010. A hybrid knowledge-guided detection technique for screening of infectious pulmonary tuberculosis from chest radiographs. IEEE Trans. Biomed. Eng., 57: 2646-2656. DOI: 10.1109/TBME.2010.2057509

WHO, 2009. Epidemiology. In: Global Tuberculosis Control: Epidemiology, Strategy, Financing, WHO (Eds.), WHO Press, Geneva, ISBN-10: 924156380X, pp: 6-33.

Xie, S., S. Shan, X. Chen and J. Chen, 2010. Fusing local patterns of gabor magnitude and phase for face recognition. IEEE Trans. Imag. Proces., 19: 1349-1361. DOI: 10.1109/TIP.2010.2041397

Xu, T., I. Cheng and M. Mandal, 2010. An improved fluid vector flow for cavity segmentation in chest radiographs. Proceedings of the 20th International Conference Pattern Recognition, Aug. 23-26, IEEE Xplore Press, Istanbul, pp: 3376-3379. DOI: 10.1109/ICPR.2010.824

Zhang, J. and Y. Liu, 2004. Cervical cancer detection using SVM based feature screening. Proceeding of the 7th International Conference, Sep. 26-29, Springer Berlin Heidelberg, Saint-Malo, France, 2, pp: 873-880. DOI: 10.1007/978-3-540-30136-3_106

Zhang, K., H.X. CAO and H. Yan, 2006. Application of support vector machines on network abnormal intrusion detection. Applic. Res. Comput., 5: 98-100. DOI: 10.1016/j.cor.2004.03.019

Zhao, H., S.C.B. Lo, M.T. Freedman and S.K. Mun, 2001. Automatic temporal subtraction of chest radiographs and its enhancement for lung cancers. Proc. SPIE, 4322: 1867-1872. DOI: 10.1117/12.431078