

RESEARCH

Open Access

Novel coarse-to-fine dual scale technique for tuberculosis cavity detection in chest radiographs

Tao Xu¹, Irene Cheng², Richard Long³ and Mrinal Mandal^{1*}

Abstract

Although many lung disease diagnostic procedures can benefit from computer-aided detection (CAD), current CAD systems are mainly designed for lung nodule detection. In this article, we focus on tuberculosis (TB) cavity detection because of its highly infectious nature. Infectious TB, such as adult-type pulmonary TB (APT) and HIV-related TB, continues to be a public health problem of global proportion, especially in the developing countries. Cavities in the upper lung zone provide a useful cue to radiologists for potential infectious TB. However, the superimposed anatomical structures in the lung field hinder effective identification of these cavities. In order to address the deficiency of existing computer-aided TB cavity detection methods, we propose an efficient coarse-to-fine dual scale technique for cavity detection in chest radiographs. Gaussian-based matching, local binary pattern, and gradient orientation features are applied at the coarse scale, while circularity, gradient inverse coefficient of variation and Kullback–Leibler divergence measures are applied at the fine scale. Experimental results demonstrate that the proposed technique outperforms other existing techniques with respect to true cavity detection rate and segmentation accuracy.

Keywords: Classification, Segmentation, Computer-aided detection (CAD), Tuberculosis (TB)

1. Introduction

Chest radiographs or chest X-ray (CXR) images are widely used to diagnose lung diseases such as lung cancer, tuberculosis (TB), and pneumonia. Due to the superimposed anatomical structures in the human chest, the CXR images are generally noisy and the diagnosis requires careful examination by experienced radiologists. Computer-aided detection (CAD) systems in chest radiography have therefore been developed to reduce the workload of radiologists. Ginneken et al. reviewed the CAD technological development in 2001 [1] and 2009 [2]. Developing a single system that looks into all abnormalities on a chest radiograph is practically impossible due to the widely different characteristics of abnormalities, and specific focus of the image processing algorithms. Therefore, the current CAD systems often aim at a single aspect, e.g., detection of lung cancer nodules. This strategy has been proved to be successful, and

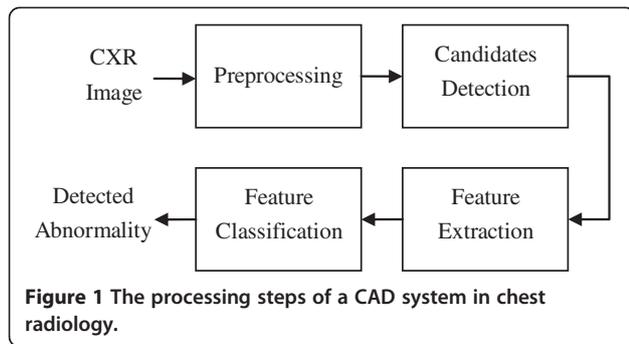
many effective algorithms have been developed for routine diagnostic procedures [2].

A general CAD system framework is shown in Figure 1. There are four modules in the system. First, a CXR image undergoes the preprocessing step, which generally includes image enhancement, noise removal, and lung field segmentation. In the next step, candidates that may contain abnormalities are coarsely detected using pattern recognition techniques. In the third step, features that can be used to identify abnormalities are identified from the candidates. Depending on the radiographic manifestation of the abnormalities, these features could be geometric, photometric, or textural. Finally, a classifier is applied to perform a high-level screening to reduce the false positive rate. An efficient CAD system relies on robust image processing, pattern recognition, and artificial intelligence techniques. For instance, a recent CAD system [3] designed for identifying lung nodules uses an active shape model for lung field segmentation, followed by a weighted multi-scale convergence-index filter for nodule candidates detection. To identify the nodules successfully, an adaptive distance-based threshold

* Correspondence: mmandal@ualberta.ca

¹Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada

Full list of author information is available at the end of the article



technique is applied to segment the contour of each candidate. The geometric, intensity, and gradient features are then extracted from the segmentation results. After the first level screening, a Fisher linear discriminant classifier is used on a subset of these features to perform the final detection.

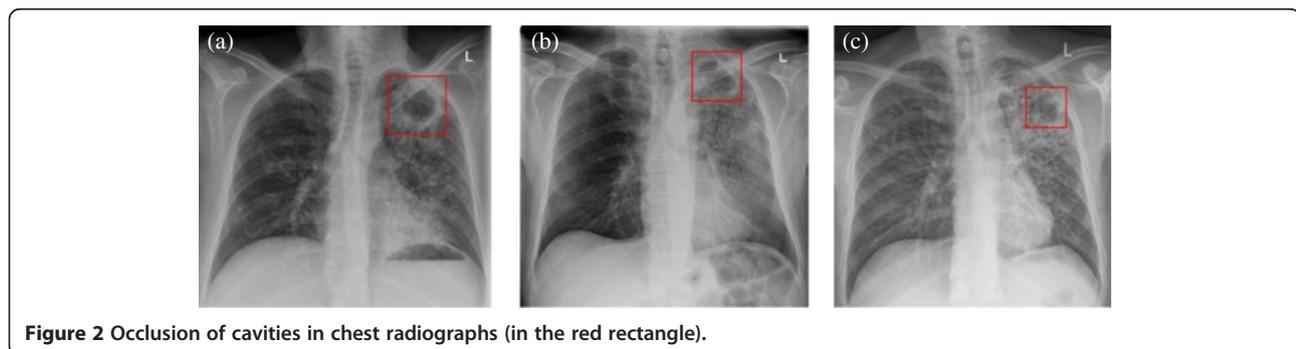
Nodule detection has been the main focus in current CXR CAD systems. However, as Ginneken et al. pointed out [2], there are other diseases, e.g. TB, that rely heavily on chest radiograph examination can benefit from the CAD systems. Infectious TB is still a public health problem in many countries [4]. Therefore, our research focus is on developing a CAD system for the diagnosis of infectious TB. The TB can be identified based on different radiographic patterns, such as cavity, airspace consolidation, and interstitial opacities [5]. A few existing CAD systems use texture analysis to detect interstitial changes [2]. However, the interstitial pattern is not a reliable radiographic cue for infectious TB. According to a recent research article on TB [6], cavitation in the upper lung zone (ULZ) is a typical radiographic feature of APTB. So far, insufficient research has been done for efficient detection of TB cavities. Shen et al. [7] recently proposed a hybrid knowledge-guided (HKG) framework for TB cavity detection, which contains three major steps. In Step 1, the cavity candidates are detected using adaptive thresholding on the mean-shift clustered CXRs. In Step 2, a segmentation technique is applied to the candidates to generate contours of important objects

present in the CXR image. In Step 3, the contour-based circularity and gradient inverse coefficient of variation (GICOV) features are extracted for the final cavity classification using a Bayesian classifier. Although, this technique provides a good performance, it has several limitations. First, due to cavity size variation and the occlusion from neighboring superimposed anatomical structures, the mean shift cluster result is sensitive to the parameter values used. Second, the adaptive threshold, which is a quadratic polynomial of GICOV score, does not perform well when the cavity boundary is weak. These two limitations lead to a high missing rate (MR) of true cavities. To overcome these problems, we propose a dual scale feature classification strategy for TB cavity detection in chest radiographs. First, a *coarse* feature classification step is performed to detect the cavity candidates by capturing the geometric, textural, and gradient features in the lung field. Second, a Hessian matrix-based technique is applied to enhance the cavity candidates, which leads to a more accurate contour segmentation. Finally, *fine* features based on the shape, edge, and region are extracted from the segmented contours for the final cavity classification. Experimental results show that the performance of the proposed candidates detection, segmentation, and cavity classification modules is superior compared to the results obtained using other related CAD systems.

The rest of this article is organized as follows. Section 2 explains the cavity pattern in CXRs. Section 3 describes our proposed method in detail. Section 4 reports and analyzes the performance of the proposed technique. Conclusion and future work are presented in Section 5.

2. Manifestation of cavity in chest radiographs

In chest radiography, a cavity is typically defined as a parenchymal cyst greater than 1 cm in diameter, containing either air or fluid or both [5]. Since the cavities are created by tissue necrosis within nodules or masses, their radiographic features are usually demonstrated as annular rings with variable wall thickness. Figure 2a shows a CXR image with a typical cavity (inside the



rectangle region), which manifests as a focal lucent area on the image and appears as a “hole” in the patient’s left upper lung zone. However, these holes might be blurred due to the overlapping projection of anatomical structures or some other abnormalities in the neighborhood, which makes the identification of cavities a difficult task for radiologists. Figure 2b is another example of a TB cavity obscured by the left clavicle. Figure 2c shows an example where the cavity is overlapped with interstitial opacities.

3. Proposed technique

Computer-aided feature identification in CXR images is comparatively more challenging than feature identification in medical images of other body parts because of the rib cage and other superimposed anatomical structures in the lung field as illustrated in Figure 2. After examining the geometric, textural, and photometric characteristics of TB cavities, we propose a coarse-to-fine feature classification technique for cavity detection. Figure 3 shows a schematic of the proposed technique. It is observed that there are three major steps: (i) coarse feature classification, (ii) contour segmentation, and (iii) fine feature classification. A CXR image is first divided into patches. In the first step, a coarse feature classification is performed on each image patch to identify candidates which are suspected to contain cavities. Two modules are used to capture the coarse features: Gaussian-model-based template matching (GTM), and local binary pattern (LBP) and histogram of oriented gradient (HOG)based feature classification (LHFC). In the second step, contours of the chosen candidates are segmented using two modules: Hessian-matrix-based image enhancement (HIE) and active contour-based segmentation (ACS). The HIE is used to boost the cavity edges. The edge-based ACS is then applied to segment the enhanced images. In the third step, a contour-based

feature classification (CFC) module is applied. Fine features including shape, edge, and region are extracted from the contours. Cavity classification is then performed based on these features. A detailed description of these five modules is presented in the following sections.

3.1. GTM

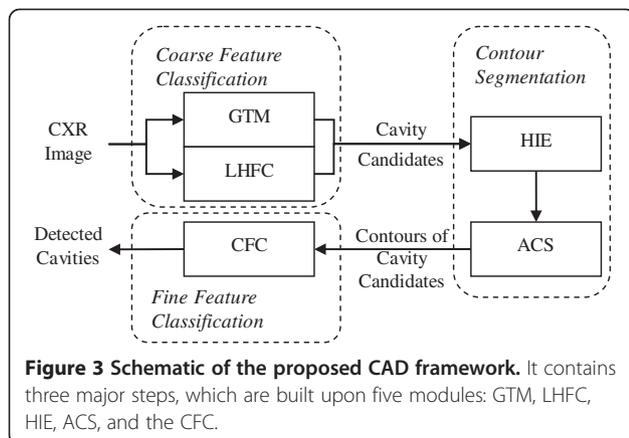
The template matching (TM) is a widely used technique in pattern recognition, where the presence of a pattern in an image is detected by comparing different parts of an image with a reference pattern known as template. In many TM techniques, instead of comparing a given template directly, a transformation of the template is matched with similar transformation of a candidate region using a similarity measure. Normalized cross correlation is often used to measure similarity because of its fast implementation using the fast Fourier transform. Since traditional TM is sensitive to rotation and scale, rotation and scale invariant transform such as Fourier–Mellin transform [8], or ring-projection transform [9] can be incorporated into TM. However, these transforms provide good results only when a cavity shape/size deviates very little from the template shape/size. To avoid missing true cavities, a solution is to use a large set of templates covering different cavity sizes and rotation angles.

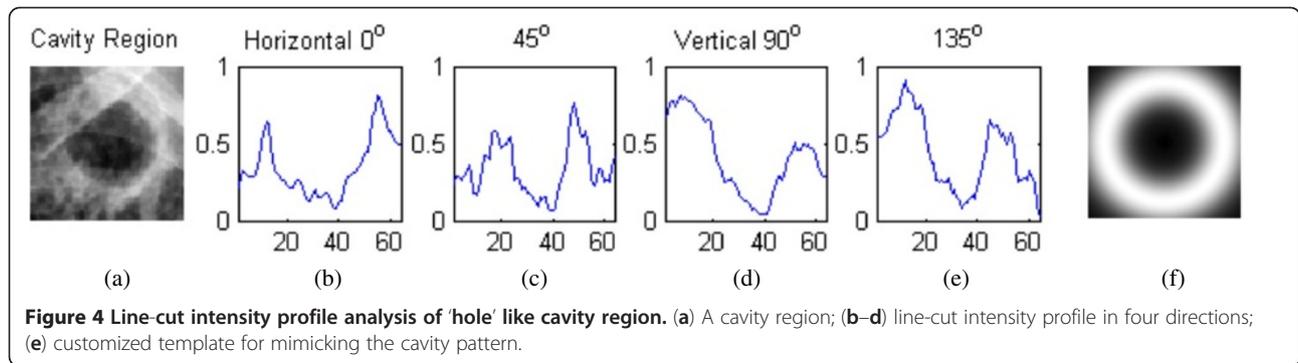
Using a large set of templates can be computationally expensive but still cannot guarantee to detect all cavities. Therefore, the proposed technique makes use of prior knowledge given by TB experts to generate a customized template database specific for TB cavities. Observe that in the “hole” like cavity shown in Figure 2a, line-cut intensity profiles in various directions of the cavity region appear similar. Figure 4a shows the magnified region of a cavity, and Figure 4b–e shows plots of the four intensity lines passing through the image center at 0°, 45°, 90°, 135°. Each line’s intensity profile appears as a bimodal Gaussian function. Based on the similarity of these intensity profiles, it is reasonable to mimic the cavity pattern using rotationally symmetric pattern such as 2D circular or elliptical Gaussian ring distribution (as shown in Figure 4f). Note that if a line-cut intensity profile of Figure 4f is calculated, a bimodal Gaussian distribution is obtained where the two major peaks correspond to the two sides of the ring.

A generic 2D Gaussian ring is defined as follows

$$I(x, y) = e^{-\frac{(1-w)^2(x^2+y^2)}{2\xi^2}} \quad (1)$$

where $w = \frac{ab}{\sqrt{a^2y^2 + b^2x^2}}$, a and b are the two radii (distance between the origin and the peaks on x, y axes), $I(x, y)$ is the image intensity function in the 2D domain, and ξ is





the standard deviation of the Gaussian distribution which determines the wall thickness of the ring. Noting that when $a = b = r$, Equation (1) represents a 2D circular Gaussian ring, where r is the inner radius. Rotated patterns can be generated by incorporating a rotation angle θ into the following coordinate transformation:

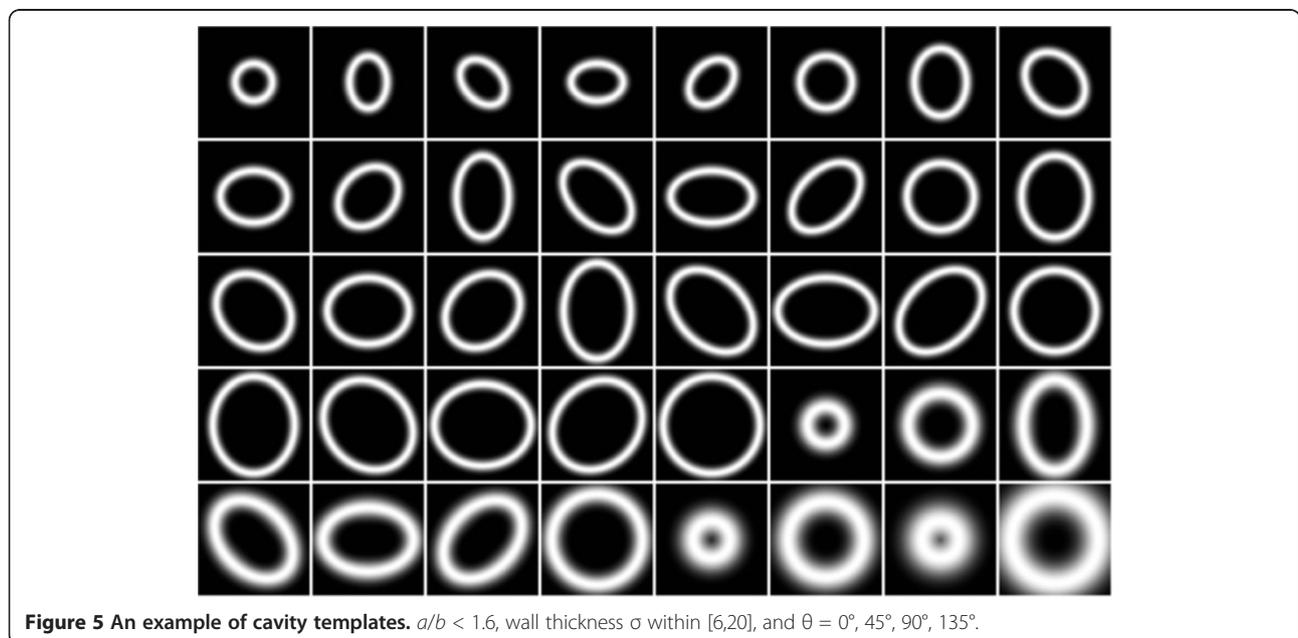
$$\begin{cases} x = x' \cos\theta + y' \sin\theta \\ y = y' \cos\theta - x' \sin\theta \end{cases} \quad (2)$$

where x', y' are the pixel's location before rotation. Using Equations (1) and (2), the template database can be built with various sizes, wall thicknesses, and rotation angles by changing the value of parameters a, b, ξ , and θ . For example, given a 512×512 CXR image with a pixel spacing $[0.8 \text{ mm}, 0.8 \text{ mm}]$, the physical size represented by the image is $40.96 \times 40.96 \text{ cm}^2$. Since the diameter of the largest cavity is usually less than 6 cm, we define the template size as 75×75 . While the wall thickness is

within the range of $[4 \text{ mm}, 16 \text{ mm}]$, parameter ξ is varied from 5 to 20 pixels. Figure 5 shows a set of templates, with various radii, rotation angle, and wall thickness, used in this article.

3.2. LHFC

Although the proposed GTM module works well for cavities of typical shape and intensity, it is difficult to detect cavities obscured by anatomical structures or some other abnormalities in the lung field. To address this issue, we combine the LBP and HOG features, which have been shown to be useful in human detection in handling partial occlusion [10]. The LBP [11] is a hybrid texture feature widely used in image processing. It combines the traditionally divergent statistical and structural models of texture analysis. The LBP feature has some key advantages, such as its invariance to monotonic gray level changes and computational efficiency. The HOG feature [12], similar to Lowe's scale-invariant feature



transform feature, is regarded as an excellent descriptor to capture the edge or local shape information. It has a great advantage of being robust to changes in illumination or shadowing. These two features are expected to complement well the GTM technique, especially in blurred regions containing cavities, to detect TB cavity candidates.

In the LHFC module, a feature vector, which combines the LBP and HOG features, is calculated for each candidate window. The feature vector is then fed to a classifier, which is trained offline using ground-truth (cavity and non-cavity) training data. The classifier will assess the windows as cavity candidates (positive samples) or not (negative samples). The candidate windows are generated using a sliding-window paradigm where an image is scanned from the top left to the bottom right with overlapping rectangular sliding windows. The windows are scanned row wise. The window size is consistent with the template size in GTM, i.e., each window has a size of 75×75 . The overlap between two consecutive windows is $2/3$ of the window size.

The computation of these two features and the classification using support vector machine (SVM) [13] are explained in the following sections.

3.2.1. Computation of the LBP feature

In this article, the LBP feature vector for a window is calculated in three steps. In Step 1, explained in Figure 6, the LBP values are calculated by applying the LBP labeling on each pixel. Here, each pixel in the window is compared to each of its eight neighbors. The LBP value for the pixel is then calculated as follows

$$LBP_{p,R} = \sum_{p=0}^{P-1} u(g_p - g_c) 2^p \left[\text{Note : } u(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \right] \quad (3)$$

where g_p, g_c are gray levels of the neighborhood pixels and center pixel, respectively, and $u(\cdot)$ is the unit-step function. For a window of 75×75 , there will be 5,625 LBP values, with dynamic range between 0 and 255. In Step 2, an LBP-histogram, with 256 bins, is generated for the window from the 5,625 computed LBP values.

Finally, in Step 3, to reduce the dimensional numbers of features, we adopt a popular approach used in texture analysis, e.g., [14], by calculating the six statistical features (mean, standard deviation, smoothness, skewness, uniformity, and entropy) based on the LBP histogram. Figure 7b shows the six LBP features calculated from the image window shown in Figure 7a.

3.2.2. Computation of HOG feature

For computational convenience, we first resize each 75×75 image window into a 64×64 window using bicubic interpolation. The HOG feature for each resized window is then calculated as follows.

- Step 1. *Gradient computation*: The gradient of each pixel in the window is calculated using two filter kernels: $[-1, 0, 1]$ and $[-1, 0, 1]^T$. Let the magnitude and orientation of the gradient of the i th pixel ($1 \leq i \leq 4096$) be denoted by m_i and ϕ_i , respectively.
- Step 2. *Orientation histogram*: Each window is first divided into non-overlapping cells of equal dimension, e.g., a rectangular cell of 8×8 . The orientation histogram is then generated by quantizing ϕ_i into one of the nine major orientations: $\frac{(2k-1)\pi}{9} \pm \frac{\pi}{9}$, $1 \leq k \leq 9$. The vote of the pixel is weighted by its gradient magnitude m_i . Thus, a cell orientation histogram H_c is a vector with dimension of 1×9 .
- Step 3. *Block normalization*: In order to account for changes in illumination and contrast, the cell histogram must locally be normalized, which requires grouping the cells together into larger, spatially connected blocks. The block size we use is 2×2 cells (i.e., 16×16 pixels), and the overlap between two neighboring blocks is $1/2$ of the block size. Therefore, a whole window contains 49 blocks. The block divisions for a window image are shown in Figure 8. The feature vector of one block H_b is concatenated by four cell histograms: $H_b = [H_{c1} H_{c2} H_{c3} H_{c4}]$. Note that the orientation histogram of a block H_b is a vector with a dimension of 1×36 . The

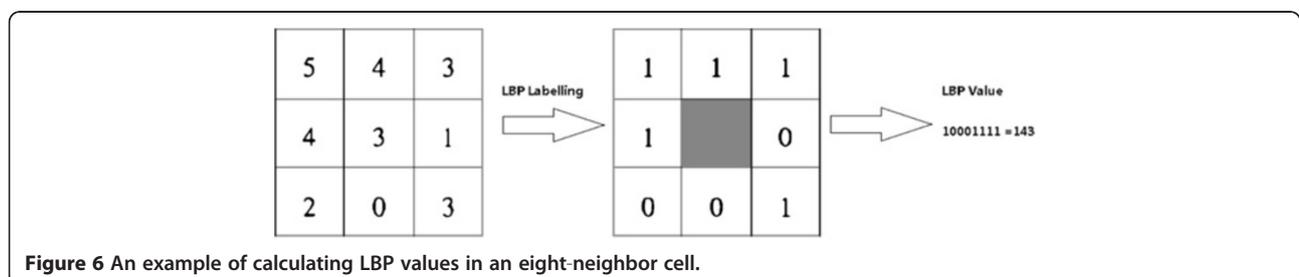
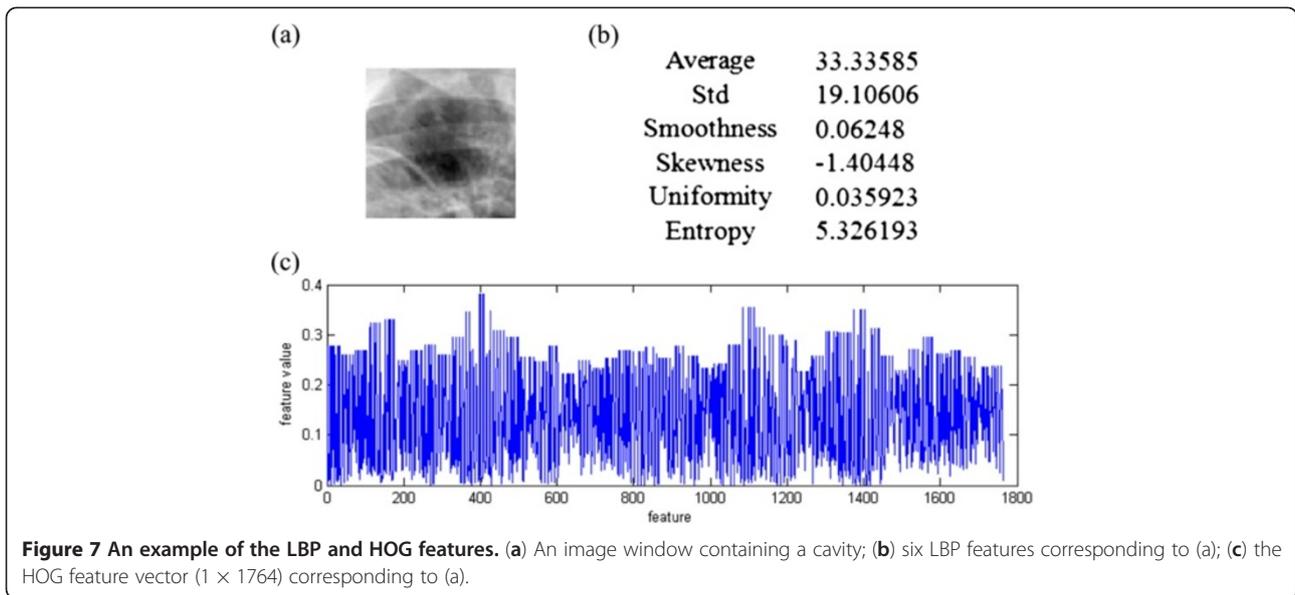


Figure 6 An example of calculating LBP values in an eight-neighbor cell.



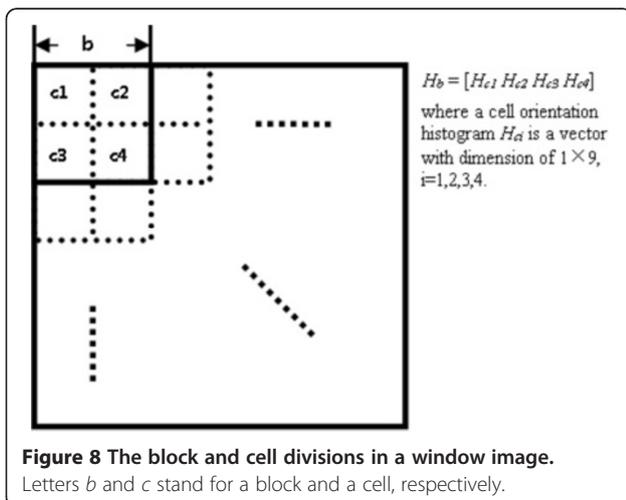
normalized HOG vector is then calculated as follows [12].

$$\hat{H}_b = \frac{H_b}{\|H_b\|} \quad (4)$$

where $\|\cdot\|$ represents the L^2 norm.

The HOG feature vector of an image window (with 49 blocks) is a concatenated vector of all 49 normalized block orientation histogram (\hat{H}_b), and will have a dimension of 1×1764 in our case. Figure 7c shows the plot of the HOG feature vector of the image window shown in Figure 7a.

Combining the LBP and HOG features, a feature vector of size 1×1770 is obtained for each image window. These features vectors are fed to the SVM classifier, explained in the following section, for cavity candidates detection.



3.2.3. Classification using SVM

Although SVM can perform both linear and nonlinear classifications, the basic SVM is a non-probabilistic binary linear classifier [13]. It is commonly used in machine learning as a supervised learning technique for recognizing patterns. Our goal is to use a pattern's feature vectors to identify which class it belongs to. The classification decision is based on the value of a linear combination of these feature vectors. Researchers use SVM classifiers in applications because of its efficiency in handling both linear and nonlinear classification problems. Once the separating hyperplane is obtained after the training step and the classification accuracy is satisfied, the given task (data) could linearly be separated in a high-dimensional feature space using this hyperplane.

For two-class classification, the optimal separating hyperplane in SVM to separate two sets of data in a feature vector space is defined by $\vec{w} \cdot \vec{x} + b = 0$, where \vec{x} is the feature vector space, \vec{w} is the normal vector to the hyperplane, and b is the offset of the hyperplane from the origin. Given M training feature vectors $\{\vec{x}_k, 1 \leq k \leq M\}$, and the corresponding ground-truth classification result $\{y_k \in [1, -1], 1 \leq k \leq M\}$, the optimal hyperplane coefficients vector \vec{w} is generated as follows

$$\min \frac{1}{2} \|\vec{w}\|^2, s.t. y_k [\Gamma(\vec{w}, \vec{x}_k) + b] \geq 1, 1 \leq k \leq M \quad (5)$$

where $\Gamma(\cdot)$ denotes a kernel function [13]. Linear, polynomial, radial basis function (RBF), and sigmoid are widely used as SVM kernels. In our tasks, we use the RBF kernel function which performs better than other kernels.

The SVM training builds a model that is able to distinguish the belonging class of any future data based on the

support vectors obtained by the training dataset. Any new feature vector \vec{x}_i is classified according to the output of the decision function:

$$f(\vec{x}_i) = \sum_{k=1}^M \alpha_k y_k \Gamma(\vec{x}_k, \vec{x}_i) + b \quad (6)$$

where α_k is the Lagrange multiplier. If $f(\vec{x}_i) \geq 0$, it means \vec{x}_i belongs to class $y = 1$, and if $f(\vec{x}_i) < 0$, it means \vec{x}_i belongs to class $y = -1$.

An example of cavity candidate detection using GTM + LHFC (note: LHFC includes the LBP and HOG features) is shown in Figure 9. Figure 9a shows the original CXR

image, and Figure 9b shows three detected TB cavity candidates, C1, C2, C3. The magnified images of these candidates are also shown in Figure 9c. To eliminate the false positive candidates (C1 and C3), further contour segmentation and fine feature classification are necessary.

3.3. HIE

As shown in Figure 9b, the GTM + LHFC detects a large number of cavity candidates some of which may be false positives (e.g., C1 and C3 shown in Figure 9b). In this section, we present a technique to enhance the cavity feature in a candidate, which will help in reducing the number of false positives. In order to reduce the effect

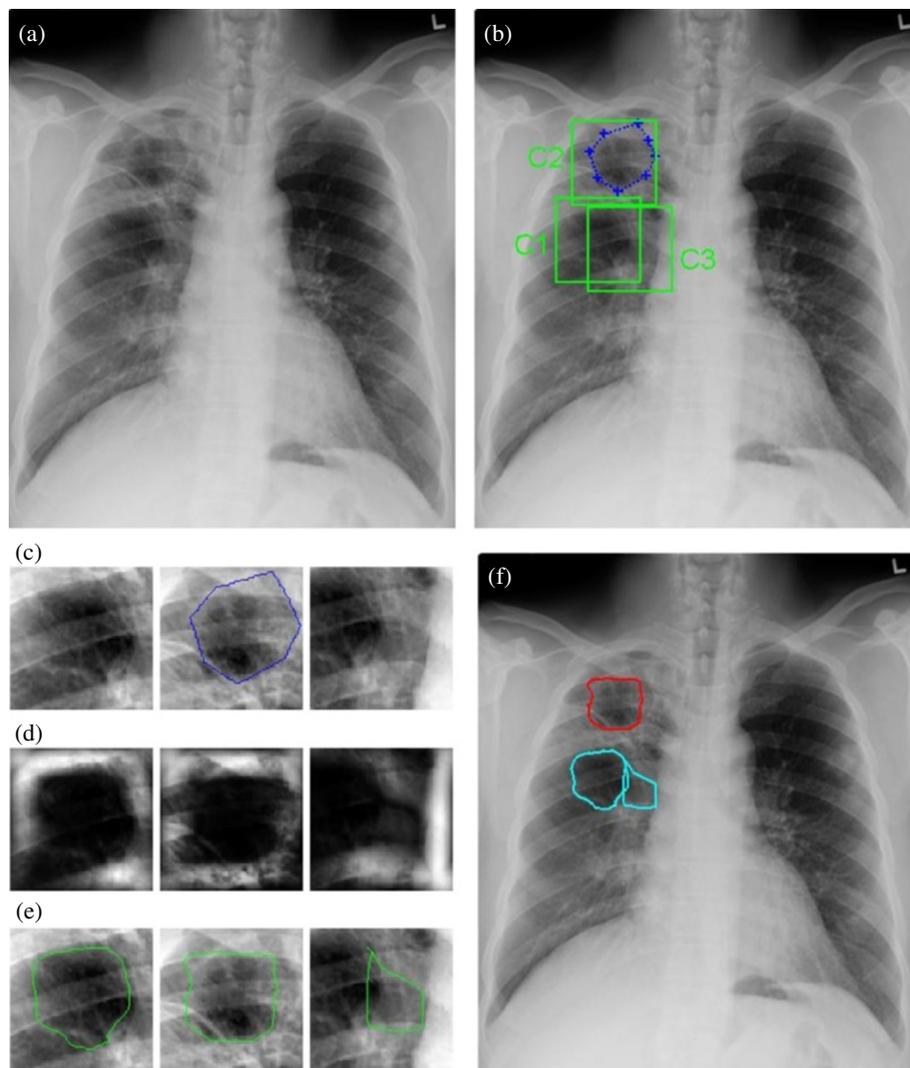


Figure 9 An example of cavity candidates detection using the proposed technique. (a) Original CXR image; (b) candidate detection results in ULZ obtained using GTM + LHFC where the green rectangular windows (C1, C2, C3) represent the candidates, and the blue dotted contour is the true cavity annotated by radiologists; (c) magnified candidate windows: C1–C3 (left to right); (d) HIE results of C1–C3; (e) improved fluid vector flow (IFVF) results of C1–C3 with the help of HIE; (f) final cavity detection results using fine feature classification. Red contour is the detected cavity, while the cyan ones are the non-cavity contours.

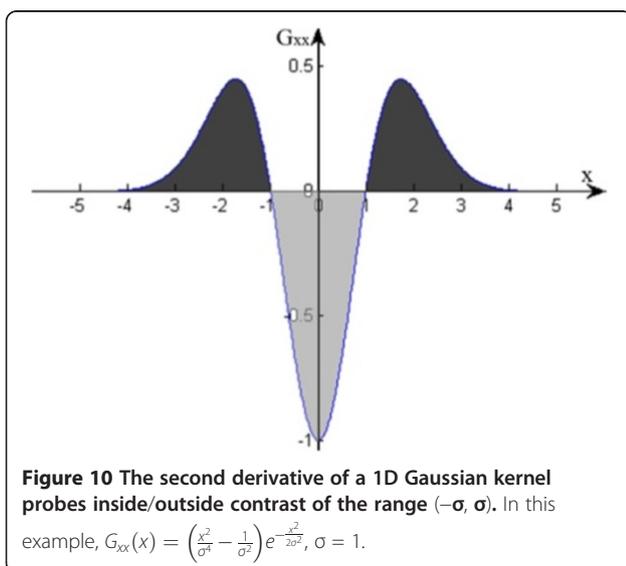
of noise and irrelevant anatomical structures or abnormalities, we apply the HIE to enhance the candidates. Note that the Hessian matrix has been applied in the literature to enhance local patterns such as plate-like, line-like, or blob-like structures [15]. The proposed HIE has three steps, which are described below.

Step 1. Laplacian of Gaussian smoothed image: In this step, three Laplacians (in three directions) of a Gaussian smoothed image, at scale σ , are obtained by convolving a cavity candidate with the second derivative of Gaussians as follows. where $I(x, y)$ is the candidate and G is the Gaussian kernel. Note that for a candidate of size 75×75 , each of the three L matrices in Equation (7) will have a size of 75×75 . Figure 10 shows the second derivative of a 1D Gaussian kernel. The intrinsic characteristic of this analysis is that the second derivative of the Gaussian kernel at scale σ generates a probe kernel that measures the contrast between the regions inside and outside the range $(-\sigma, \sigma)$ in the direction of the derivative.

$$\begin{cases} L_{xx}(x, y, \sigma) = \sigma^2 I(x, y) * G_{xx}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) = \sigma^2 I(x, y) * G_{xy}(x, y, \sigma) \\ L_{yy}(x, y, \sigma) = \sigma^2 I(x, y) * G_{yy}(x, y, \sigma) \end{cases} \quad (7)$$

Step 2. Hessian matrix calculation: For a given σ value, the Hessian matrix corresponding to pixel (x_i, y_i) in the candidate is calculated as follows

$$H_\sigma(x_i, y_i) = \begin{bmatrix} L_{xx}(x_i, y_i, \sigma) & L_{xy}(x_i, y_i, \sigma) \\ L_{yx}(x_i, y_i, \sigma) & L_{yy}(x_i, y_i, \sigma) \end{bmatrix} \quad (8)$$



where $L_{xy}(x_i, y_i, \sigma) = L_{yx}(x_i, y_i, \sigma)$. A known problem of multi-scale analysis using Hessian matrix is that over-blurring can occur during the multi-scale smoothing, which may increase false detections [16]. Therefore, in this article, we set the σ value equal to the object scale calculated using the method described in [17]. The object scale at every pixel is defined as the radius of the largest hyperball centered at the pixel such that all pixels within the ball satisfied a predefined image intensity homogeneity criterion. Object scale represents the geometric information (size) of the local structure. Object scale at the center of a blob-like structure is approximately equal to the radius of the blob in pixel size.

Step 3. Image enhancement using eigenvalues of Hessian matrix: The pixel (x_i, y_i) in the candidate with intensity $I(x_i, y_i)$ is enhanced using the following equation:

$$I_E(x_i, y_i) = |\lambda_1| I(x_i, y_i) \quad (9)$$

where λ_1 and λ_2 are eigenvalues of $H_\sigma(x_i, y_i)$, and $|\lambda_1| \geq |\lambda_2|$. The intuition in Equation (9) of using only the largest eigenvalue for cavity enhancement is based on the fact that the Hessian matrix has a strong edge effect (for those strong edge points, $|\lambda_1| \gg |\lambda_2| \approx 0$) [18]. Although cavities are usually embedded in noisy surroundings due to the neighboring necrosis caused by cavitation, the inside of a cavity (filled with air or fluid or both) still has lower intensity than the background. Thus, the strong edge between the inside and outside of a cavity gives a good clue to identify the contour of cavity. Different techniques of edge enhancement were evaluated in this study, such as contrast-limited adaptive histogram equalization [19], fuzzy C means [20], and speckle reducing anisotropic diffusion technique [21], and the proposed HIE technique achieves the best performance.

The enhanced window candidates C1–C3 are shown in Figure 9d. It is observed that the annular ring-like structure is greatly enhanced.

3.4. ACS

Active contours or deformable models are generally divided into two types: parametric active contours (typically known as snakes) and geometric active contours (level set). The snake-based techniques are often faster than level sets in virtue of efficient numerical methods. In addition, the level sets produce more false detections due to its multiple objects capturing ability. Therefore, in this article, we use a snake-based technique known as

IFVF [22]. In this technique, a snake contour represented by v evolves through the candidate window to reach a force balance equation $F_{\text{int}}(v) + F_{\text{ext}}(v) = 0$, where $F_{\text{int}}(v)$ is the internal force constraining contour's smoothness, and $F_{\text{ext}}(v)$ is the external force attracting the contour toward image features.

The IFVF is a fast and accurate edge-based snake technique, because of the introduction of both static and dynamic terms in the external force.

$$F_{\text{ext}}(v) = F_{\text{static}}(v) + F_{\text{dynamic}}(v) \quad (10)$$

The F_{static} could be a static external force which overcomes the edge leakage problem, e.g., we use boundary vector flow (BVF) proposed in [23] as the F_{static} . The BVF extends the capture range further to the entire image based on simpler interpolation. Four potential functions Ψ_x , Ψ_y , Ψ_{xy} , and Ψ_{yx} are computed using line-by-line interpolations in the horizontal, vertical, and two diagonal directions. The F_{static} is calculated as follows

$$\begin{aligned} F_{\text{static}} &= \Phi_1 = (\nabla\Psi_x, \nabla\Psi_y) \text{ or} \\ F_{\text{static}} &= \Phi_2 = \left(\frac{\sqrt{2}}{2}(\nabla\Psi_{xy} + \nabla\Psi_{yx}), \frac{\sqrt{2}}{2}(\nabla\Psi_{xy} - \nabla\Psi_{yx}) \right) \end{aligned} \quad (11)$$

The F_{dynamic} is achieved in three steps.

1. Given an HIE-enhanced candidate image, a binary edge map B is generated using smoothing technique speckle reducing anisotropic diffusion [21] and the Canny edge detector [24].
2. By comparing the edge map points to the current snake contour points (snaxels), a new control point (x_c, y_c) is selected by considering the point which contributes more to the distance between snake contour and object boundary [22]. We use the Hausdorff distance to find such a point. Assuming two sets of points S and O , the Hausdorff distance is then defined as $h(S, O) = \max_{o \in O} \{ \min_{s \in S} \{ d(s, o) \} \}$ where $d(s, o)$ is the Euclidean distance between a snaxel s and a object boundary point o . So, the control point is chosen as the point on the object boundary which has the Hausdorff distance value.
3. For any pixel (x, y) on the contour v , its $F_{\text{dynamic}}(x, y)$ is then calculated as follows

$$F_{\text{dynamic}}(x, y) = (1 - B)\delta \frac{\nabla d'(x, y)}{\|\nabla d'(x, y)\|} \quad (12)$$

where $\delta = \pm 1$ controls the outward or inward direction. In this article, we use $\delta = 1$, as the initial contour is automatically set as a small circle in the center of the window image with radius of 3 pixels. $d'(x, y)$ is the Euclidean distance between points (x, y)

and (x_c, y_c) . Note that the term $(1 - B)$ makes the F_{dynamic} zero for those points which already reach edges. Based on the edge map generated from the enhanced candidates images using HIE, the IFVF segmentation result of these candidates C1–C3 are shown in Figure 9e. The stopping criterion of the evolution is determined by computing the difference in locations (defined by the x and y coordinates) of the corresponding contour points between two consecutive iterations. If it is less than a convergence threshold t , the active contour evolution will be stopped. In our experiments, t is empirically set to 0.05. Based on our tests, there is no significant improvement even if t is smaller than 0.05.

3.5 CFC

The last module in our proposed technique is the CFC, which performs the fine scale feature classification. Three types of contour-based features, shape, edge and region, are extracted for the final cavity detection. These features include circularity measure [25], GICOV [26], and Kullback–Leibler divergence (KLD) [27] between the pixel intensity distributions inside and outside the contour. The computations of these three features are explained below.

1. Assuming a contour has one centroid, L points are selected from the contour in L cardinal directions. The circularity of the contour is then calculated as scaled variance as follows

$$C = \frac{\text{var}(d(x_i, y_i))}{\max(d(x_i, y_i))}, i = 1, 2, \dots, L \quad (13)$$

where $d(x_i, y_i)$ is the distance from the centroid to the contour point (x_i, y_i) in the i th direction. In this article, we use $L = 16$. The circularity feature is a feature which could effectively reduce the false positives.

2. Based on the observation that the inner boundary of a cavity often has dark-to-bright transition, the GICOV value of L points on the contour is calculated as follows

- (a) For the contour point (x_i, y_i) in the i th direction, its gradient in normal direction $g_n(x_i, y_i)$ is calculated as $g_n(x_i, y_i) = \nabla I(x_i, y_i) \cdot \vec{n}(x_i, y_i)$, where $\vec{n}(x_i, y_i)$ is the unit outward normal vector at this point.
- (b) The mean and standard deviation of g_n denoted by

$$\begin{aligned} m \text{ and } s, \text{ are then calculated as } m &= \frac{1}{L} \sum_{i=1}^L g_n(x_i, y_i) \\ \text{and } s^2 &= \frac{1}{L-1} \sum_{i=1}^L (g_n(x_i, y_i) - m)^2 \end{aligned}$$

Table 1 Fine feature values of three contours in Figure 9e

	Circularity	GICOV	KLD
Contour-1	0.11	15.33	1.49
Contour-2	0.15	13.68	2.28
Contour-3	0.69	15.26	0.28

(c) The GICOV value of the contour is finally achieved using following equation:

$$GICOV = \frac{m}{s/\sqrt{L}} \quad (14)$$

3. Given the probability distributions, P and Q , of the pixel intensity values inside and outside the cavity, respectively, the KLD for a candidate window is calculated as follows

$$KLD = \sum_{i=1}^B P(i) \ln \frac{P(i)}{Q(i)} \quad (15)$$

where B is the number of bins in the histogram span by P and Q . The KLD compares the difference in gray level distribution between the pixels inside and outside the contour.

Table 1 shows the above feature values corresponding to three contours shown in Figure 9e. As in the coarse feature classification step, we select the SVM as the fine feature classifier in this step. Based on the feature values (Table 1), the trained SVM classifier identifies the Contour-2 as a positive and Contour-1 and Contour-3 as negatives. The final detected cavity (corresponding to Contour-2) in the CXR image is shown in Figure 9f as the red contour. The result matches with the ground truth.

4. Performance evaluation

In this section, we evaluate our proposed coarse-to-fine dual scale technique with respect to three aspects: the effectiveness of candidate selection, the accuracy of contour segmentation, and the accuracy of final cavity detection.

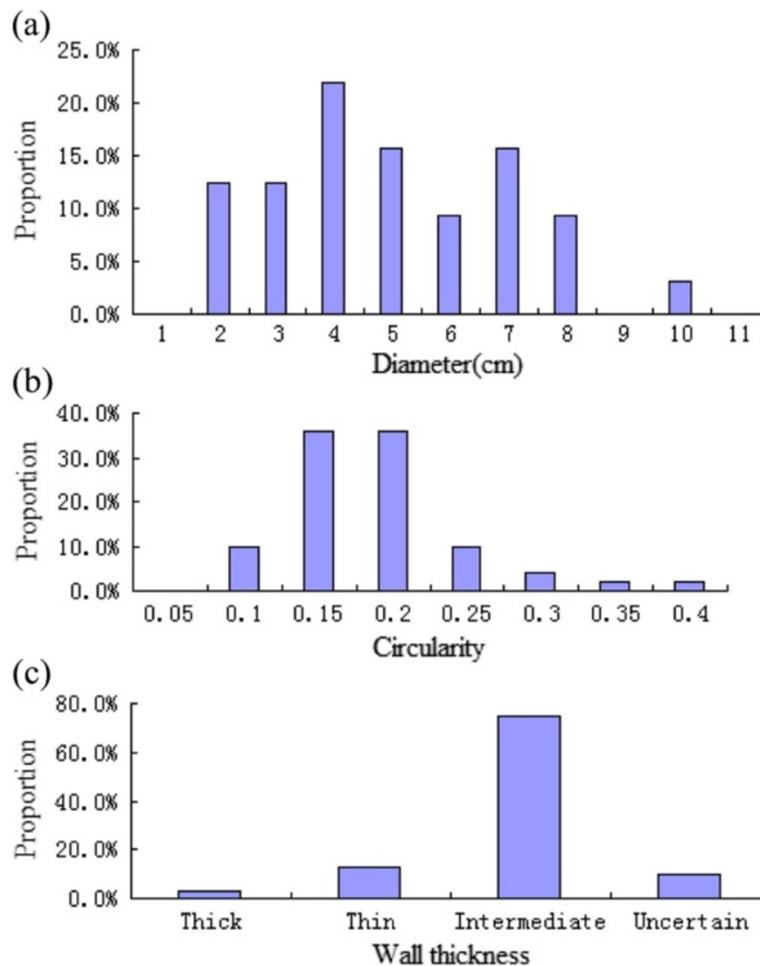


Figure 11 Sample histograms of cavity properties. (a) Histogram of diameter; (b) histogram of circularity; (c) histogram of wall thickness of four categories: "Thick" (≥ 16 mm), "Intermediate" (4–15 mm), "Thin" (< 4 mm), and "Uncertain" (wall not discernible).

4.1. Experimental dataset and parameters configuration

A cavity dataset of 35 CXR images containing 50 cavities is obtained from the University of Alberta Hospital. All the images were independently read by three experienced chest radiologists who are specialized in TB diagnosis. The presence of TB cavities was confirmed by the agreement of at least two radiologists. The sample histograms of cavity properties such as diameter, circularity, and wall thickness are shown in Figure 11. From the histograms, it can be seen that the cavities vary in diameters while their circularities range mainly from 0.15 to 0.2 and most of them have intermediate thickness. For computational efficiency, the original CXR images are resized as 512×512 (or close to this size) with a fixed pixel spacing [0.8 mm, 0.8 mm]. Since all the cavities are located in the ULZ, a similar preprocessing procedure as described in [7] was applied to segment the target lung region, which reduces the processing area to a smaller rectangular bounding box. Figure 12 shows an example of the target area.

The proposed cavity detection technique is implemented in MATLAB 2007b on an Intel Pentium 4 CPU 2.8 GHz with 2 GB RAM computer. All the parameters in the proposed technique are listed in Table 2. The SVM classifiers in both coarse and fine feature classification are built using LIBSVM software [28]. To train the SVM classifiers, we applied the ‘leave-one-out’ method [29] since the size of samples with cavities is small. For example in LHFC, to detect the candidate regions in one of the 30 CXR images, we use the remaining 29 CXR images for the training. The training set contains the



Figure 12 An example of the target area. The enhanced subimage inside the green rectangle is the result of the preprocessing procedure.

Table 2 Parameters configuration in the proposed technique

Modules	Parameters names	Parameters values
GTM	Template size	75×75 pixels
	Wall thickness σ	[6,20]
	Aspect ratio a/b	[1, 1.6]
	Rotation angle θ	{ $0^\circ, 45^\circ, 90^\circ, 135^\circ$ }
LHFC	Window size	75×75 pixels
	Cell size	8×8 pixels
	Block size	2×2 cells
	Block overlap	0.5
	SVM parameters	Default values in LIBSVM software [28]
ACS	Snake evolution direction δ	1
	Convergence threshold t	0.05
CFC	SVM parameters	Default values in LIBSVM software [28]

LBP and HOG feature vectors extracted from windows with and without cavities (positive and negative samples) in these 29 CXR images. Note that the negative samples for training were selected from the contralateral position of the positive samples based on the approximate symmetry of the lung field. The SVM classifier in CFC is trained in a similar way.

4.2. Effectiveness of candidate selection

The proposed coarse feature classification technique for candidate detection is evaluated by the MR, which is calculated as follows

$$MR = \frac{\text{of Cavities Excluded from Candidates}}{\text{Total of True Cavities}} \times 100\% \quad (16)$$

A preliminary experiment using only GTM for candidate detection has already been reported in [30]. We anticipate that by integrating with other novel techniques, a better result can be obtained. Thus, we used different combinations of LBP and/or HOG features together with GTM, and checked whether the MR could be reduced. Table 3 shows our test results.

From the results, we observe that the HKG framework for TB cavity detection [7] missed more cavities than our proposed approach. HKG is based on an adaptive thresholding on the mean-shifted clustered image for candidate detection. Its high MR is due to two reasons. First, the mean-shift clustering approximates nearest neighbors intensities and space information but neglects the texture. Second, the adaptive threshold, which is a quadratic polynomial of the GICOV feature, is not suitable for modeling all

Table 3 Candidates detection results

	HKG [7]	GTM [30]	GTM + LBP	GTM + HOG	GTM + LBP + HOG
Number of cavities	50	50	50	50	50
Number of candidates	170	164	315	229	160
Number of missing cavities	22	18	10	17	7
MR (%)	44	36	20	34	14

shapes, especially when the boundary of a cavity is weak. Figure 13 compares the detection results of HKG and our technique. The green boxes represent cavity regions reported by the classifier. In Figure 13a, HKG cannot identify both cavities due to the failure of mean-shift clustering in the noisy ULZ. Our technique is able to identify the two cavities (Figure 13b). Figure 13c is yet another example showing the adaptive threshold value used in HKG unable to identify the cavity. However, our technique is able to detect all cavities correctly (Figure 13d).

Using the same parameter values for LBP and HOG as in the literature, we found that a combination of LBP and HOG together with GTM achieved better

performance. Our finding is consistent with the results in human detection using LBP and HOG features [10]. HOG performs poorly when the background is cluttered with noises. LBP is able to alleviate this deficiency. It can filter out noises following the uniform pattern estimation. However, if LBP is used alone without HOG, the entire ULZ will be extracted if some other abnormalities are also present in the area. In that case, the HOG helps to reduce the false positives based on the available edge information. Figure 14 illustrates the complementary effect of LBP and HOG. The window reported by the classifier should contain a complete cavity in order to be qualified as a positive candidate. Note that in the first row second column

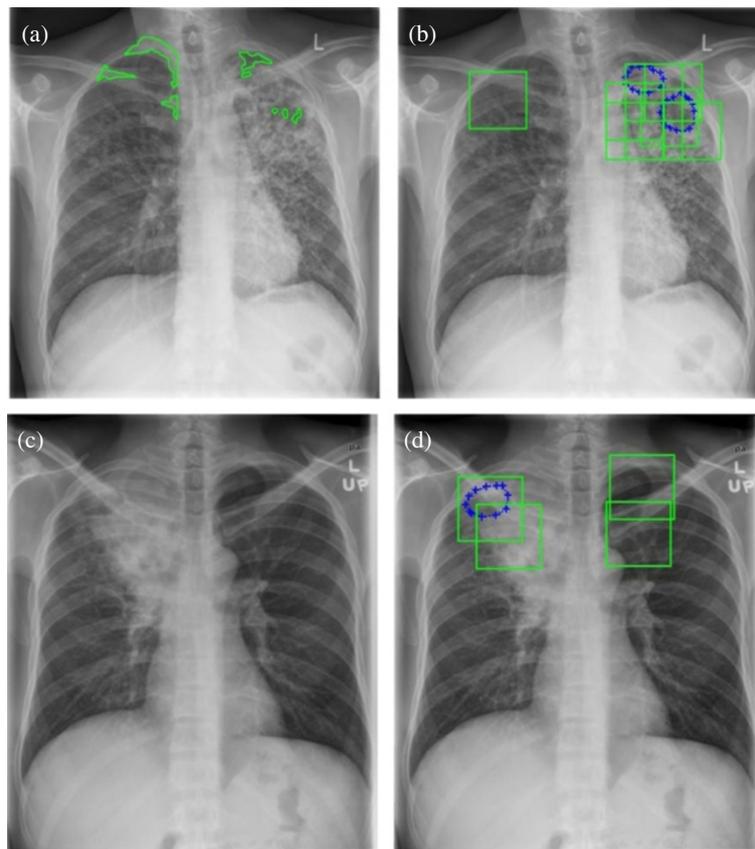


Figure 13 Comparison of candidates detection between HKG [7] and the proposed technique; (a, c) the results of HKG, (b, d) generated from the proposed technique. Green regions in the images are cavity candidates regions reported by the classifier, and blue dotted contours are the true cavities annotated by radiologists.

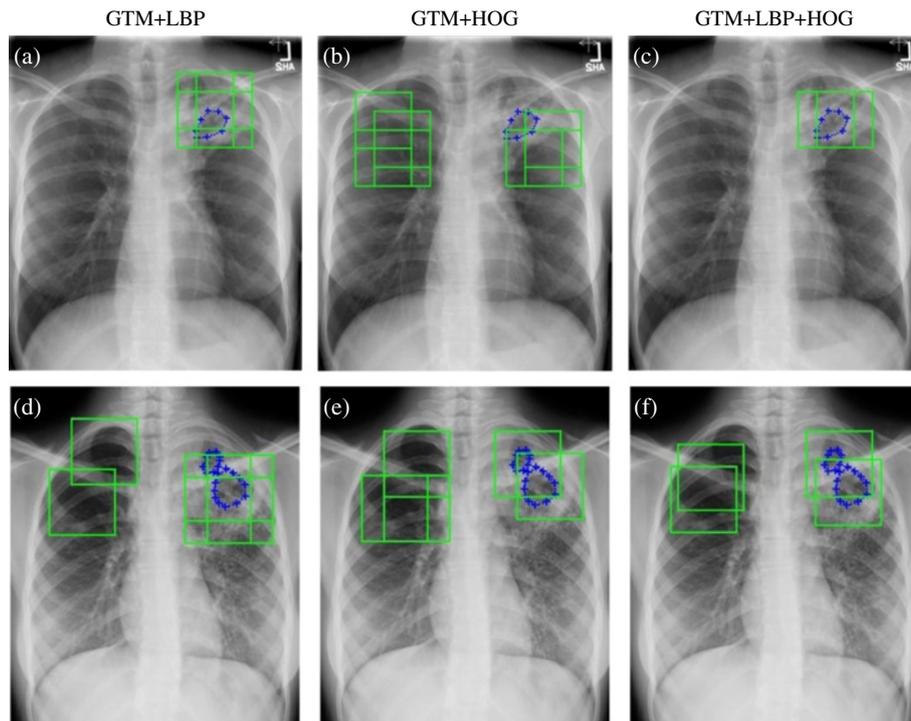


Figure 14 Comparison of candidate detection in the coarse feature classification step using (a, d) GTM + LBP, (b, e) GTM + HOG, (c, f) GTM + LBP + HOG. Note that in the first row HOG misses the cavity but LBP is able to detect it. In the second row, LBP misses the small cavity but HOG can detect it. In both rows our technique is able to detect all the cavities.

when using only HOG, no reported window contains a complete cavity. The HOG performs poorly when the background is cluttered with noises, and the edge information is no longer reliable. Similarly, in the second row first column, when using only the LBP, the small cavity is missing because no reported window contains the complete small cavity, and only the larger cavity is fully contained in a reported window.

The above test results show that combining the LBP and HOG features for capturing the texture and gradient information around the cavity region, and using the GTM for shape recognition, contributes to the low MR of the proposed coarse feature classification technique.

4.3. Accurate contour segmentation

We evaluate segmentation accuracy using the following Tanimoto measure (TMM) [7]:

$$TMM = \frac{\|R_c \cap R_g\|}{\|R_c \cup R_g\|} \quad (17)$$

where R_c denotes the region enclosed by the contour generated by the segmentation techniques, such as DBC-GVF [7] and our IFVF [22]; R_g denotes the region

of a TB cavity that is enclosed by the ground-truth contour manually drawn by radiologists; and $\|\cdot\|$ denotes the cardinality (number of pixels). $TMM = 0$ indicates that the segmented contour has no intersection with the ground truth, while $TMM = 1$ indicates that the segmented contour is identical to the exact cavity. To improve the segmentation accuracy, we apply the HIE on the candidates before segmentation.

The performance of the DBC-GVF and the IFVF techniques with and without the HIE is shown in Table 4. Note that around 10% accuracy improvement is achieved for both DBC-GVF and IFVF when HIE is incorporated. The results are also more robust as demonstrated by the lower standard deviations of the TMM. Figure 15 presents subjective comparison of different segmentation techniques. With the HIE, the segmented contours are closer to the ground truth compared to the same techniques without the HIE.

Note that image patterns, even without cavities, may generate close to ring-like shape after the HIE step. Figure 16 shows some of these cases. For example, the image in the bottom row contains a pattern similar to a cavity. To eliminate this type of candidates, the fine scale feature classification step in our approach is necessary. The accuracy of our final cavity detection is evaluated in the next section.

Table 4 Segmentation accuracy evaluation

	DBC-GVF without HIE (%)	DBC-GVF with HIE (%)	IFVF without HIE (%)	IFVF with HIE (%)
Average of TMM	55.1	64.6	56.8	67.1
Standard division of TMM	15.8	12.6	12.2	9.3
Mean of TMM	58.2	64.9	59.3	66.1

4.4. Accuracy of final cavity detection

Before performing the final cavity detection, 160 candidate contours are divided into cavity and non-cavity contours. Candidate region reported by the classifier as highlighted by the green windows in Figure 12 may not contain true cavities. Also, even if a reported window contains the entire cavity, its segmented contour may not be the same as the ground truth. To evaluate

the accuracy of the final contour classification, we need to impose a value $TMM > 0.7$ (based on the segmentation accuracy of 67.1% reported in Table 4), in order to qualify a candidate to be a true cavity; otherwise it is considered as non-cavity. Three contour-based features (circularity, GICOV, and KLD) are extracted from the candidate contours for the final cavity classification. To evaluate the performance of

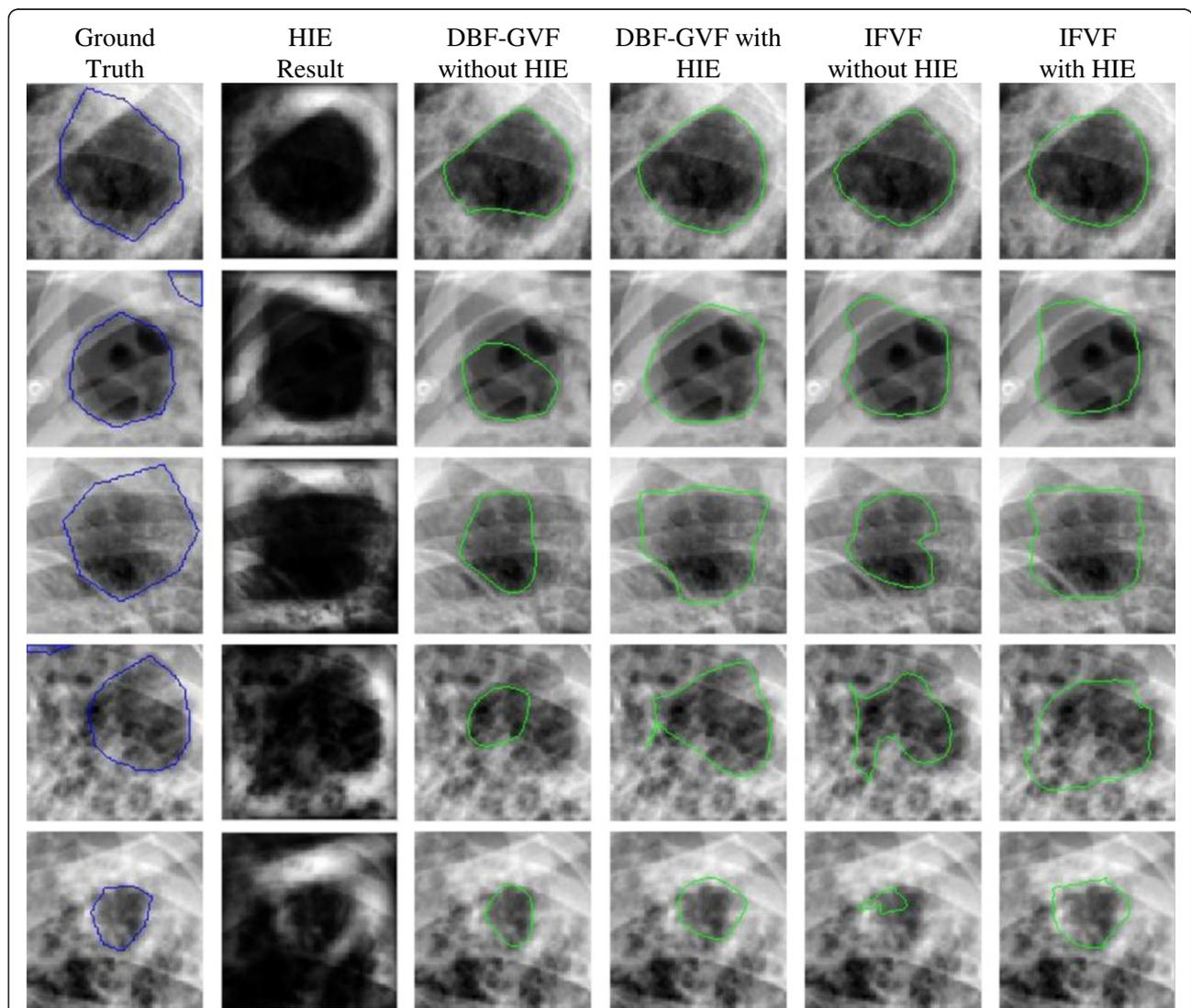


Figure 15 Cavity segmentation result comparison using different edge-based snakes with and without HIE. From top to bottom, the cavity is more and more difficult to identify. Blue contours are the true cavities annotated by radiologists. Green contours are the computer segmentation results.

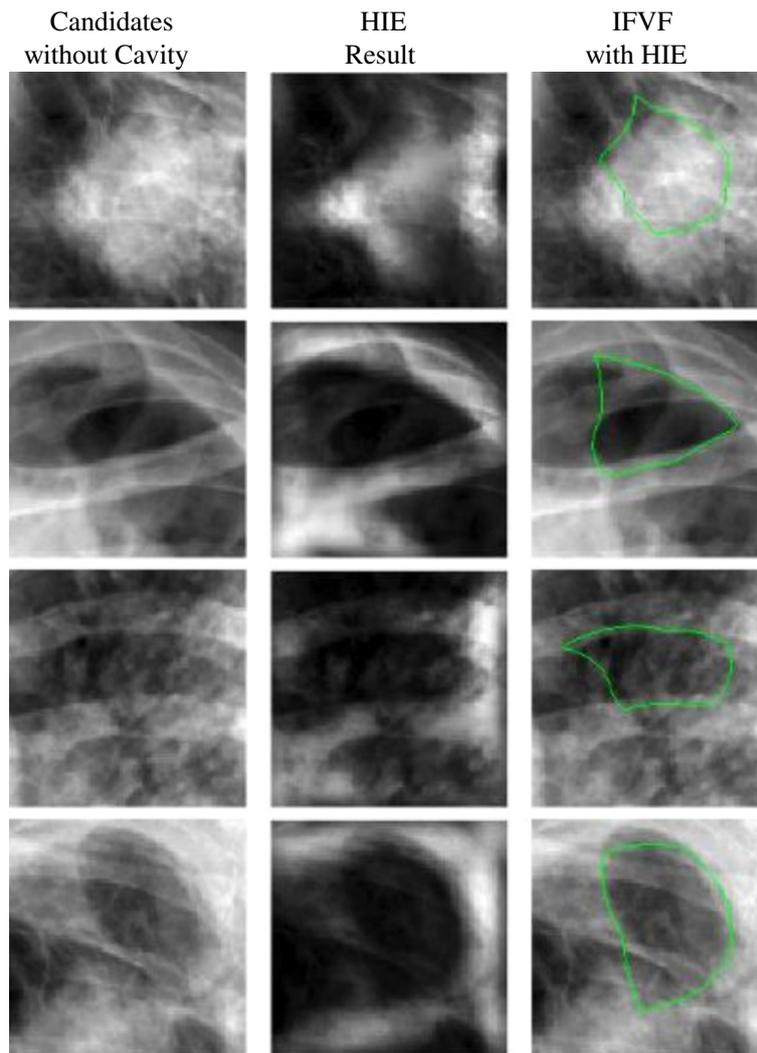


Figure 16 Segmentation results of candidates without cavity.

classification, sensitivity, specificity, and accuracy are calculated as follows.

$$\text{Sensitivity} = \frac{\text{Number of Correctly-Detected Cavity Contours}}{\text{Total Number of Cavity Contours}} \times 100\% \quad (18)$$

$$\text{Specificity} = \frac{\text{Number of Correctly-Detected Non-Cavity Contours}}{\text{Total Number of Non-Cavity Contours}} \times 100\% \quad (19)$$

$$\text{Accuracy} = \frac{\text{Number of Correctly-Detected Contours}}{\text{Total Number of Candidates Contours}} \times 100\% \quad (20)$$

Table 5 Cavity detection evaluation

	Sensitivity (%)	Specificity (%)	Accuracy (%)
Circularity + GICOV [7]	62	46	54
Circularity + GICOV + KLD	70	60	65

For our sample size, we use cross-validation method [29] for the SVM classification. The classification result for the 160 candidate contours is shown in Table 5. It can be observed that the detection accuracy is increased by more than 8% in our approach after adding KLD feature. Figure 17 shows cavity detection results of HKG [7] and the proposed technique, which demonstrate that our technique can detect more true cavities and detect fewer false cavities. As illustrated in

Figure 17, the proposed cavity detection system identifies all cavities annotated by the radiologists and there is only one false alarm. The presence of cavities in the upper half of the lungs, especially when there are multiple or bilateral cavities, should raise suspicion of TB in the appropriate epidemiologic and/or clinical context. Unfortunately, in practice, a lot of these findings are not mentioned in the radiologist's report, because the epidemiologic or clinical information, necessary to raise suspicion, is not provided by the ordering physician on the requisition. This is often the case in geographic regions where TB rate is low. Based on the clinician's perspective, a relatively higher false positive rate is better than false negatives because the latter can cause an infectious TB to spread. Even with false positives, clinicians find automatic cavity detection system helpful in reducing

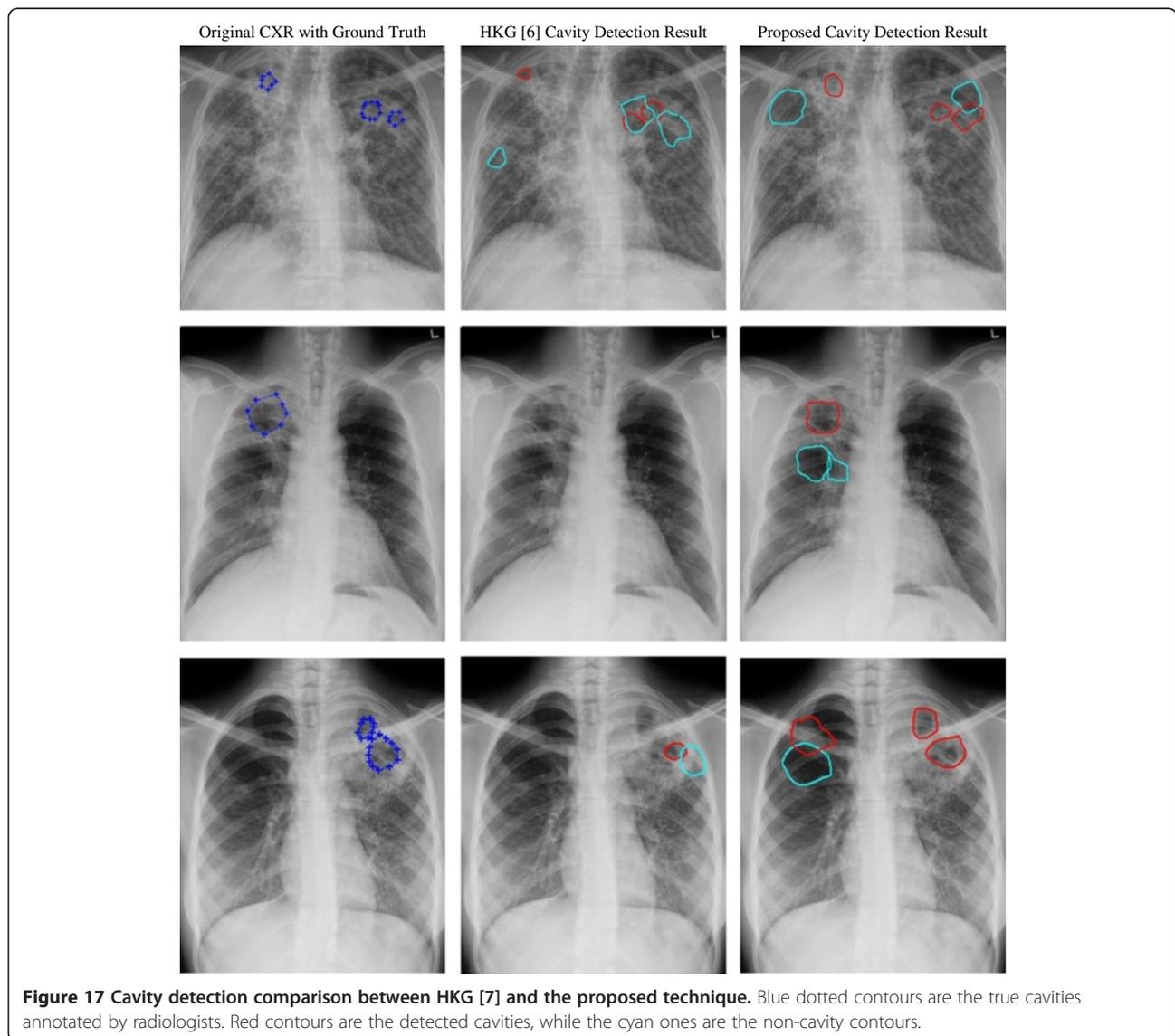


Table 6 Cavity detection evaluation of E-Group

	Sensitivity (%)	Specificity (%)	Accuracy (%)
Circularity + GICOV [7]	65	78.2	71.6
Circularity + GICOV + KLD	78.8	86.8	82.8

a large number of true negatives and radiograph examinations. This is beneficial given the limited radiologists available particularly in remote communities and developing countries.

The radiologists also classified the true cavity contours into two categories: E-Group and D-Group, containing cavities which are ‘easy’ or ‘difficult’ to identify, respectively. The D-Group contains cavities even radiologists found them difficult to identify without other demographic or additional information. False cavity contours were then combined with each of these two groups. The cross-validation SVM classification results of these groups are shown in Tables 6 and 7. Observe that on average the classification accuracy in each group is higher than the result reported in Table 5. The performance of the E-Group is significantly improved by adding the KLD feature. In the D-Group, although the intensity variation inside and outside a cavity changes only slightly making it very difficult to identify the contour even for radiologists, there is still improvement in the detection result. This shows that the classifier can perform better if trained using more specific knowledge.

5. Conclusions

In this article, we proposed an efficient coarse-to-fine dual scale feature classification technique for TB cavity detection in chest radiographs. Experimental results demonstrate that the proposed technique outperforms existing methods in three aspects. First, a lower MR is achieved because in the proposed method local cavity region-related coarse features, such as geometric, textural, and gradient features, are taken into consideration. Second, edge-based segmentation becomes more accurate by incorporating HIE to enhance the contours. Third, the final cavity detection accuracy is greatly increased by introducing the fine scale feature classification using three types of contour-related features, which includes shape, edge, and region. This study contributes in the development of CAD systems for infectious TB diagnosis, because of the higher detection rate and lower MR compared to other techniques. Future work will

Table 7 Cavity detection evaluation of D-Group

	Sensitivity (%)	Specificity (%)	Accuracy (%)
Circularity + GICOV [7]	57.6	88	72.8
Circularity + GICOV + KLD	69.4	81.6	75.5

focus on exploring novel algorithms to model other characteristics of infectious TB.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada. ²Department of Computing Science, University of Alberta, Edmonton, AB T6G 2E8, Canada. ³Division of Pulmonary Medicine, Department of Medicine, University of Alberta, Edmonton, AB T6G 2V2, Canada.

Received: 14 December 2011 Accepted: 28 November 2012

Published: 8 January 2013

References

- BV Ginneken, BMT Hromeny, MA Viergever, Computer-aided diagnosis in chest radiography: a survey. *IEEE Trans. Med. Imag.* **20**(12), 1228–1241 (2001)
- BV Ginneken, L Hogeweg, M Prokop, Computer-aided diagnosis in chest radiography: beyond nodules. *Eur. J. Radiol.* **72**(2), 26–30 (2009)
- RC Hardie, SK Rogers, T Wilson, A Rogers, Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs. *Med. Image Anal.* **12**(3), 240–258 (2008)
- World Health Organization, *Epidemiology, in Global Tuberculosis Control: Epidemiology, Strategy, Financing*, WHO Press, Geneva, 2009, pp. 6–33
- R Long, E Ellis et al., *Canadian Tuberculosis Standards*, 6th edn. (Public Health Agency of Canada, 2007)
- A Lau, R Long et al., The public health consequences of smear positive pulmonary tuberculosis in patients with typical and atypical chest radiographs, in *15th Annual International Union Against Tuberculosis and Lung Disease*, North American Region Conference, Vancouver, BC, 2011
- R Shen, I Cheng, A Basu, A hybrid knowledge-guided detection technique for screening of infectious pulmonary tuberculosis from chest radiographs. *IEEE Trans. Biomed. Eng.* **57**(11), 2646–2656 (2010)
- BS Reddy, BN Chatterji, An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans. Image Process.* **5**(8), 1266–1271 (1996)
- YH Lin, CH Chen, Template matching using the parametric template vector with translation, rotation and scale invariance. *Pattern Recognit.* **41**(7), 2413–2421 (2008)
- X Wang, TX Han, S Yan, *An HOG-LBP Human Detector with Partial Occlusion Handling* (in Proceedings of the ICCV, Kyoto, 2009), pp. 32–39
- T Ojala, M Pietikainen, D Harwood, A comparative study of texture measures with classification based on feature distributions. *Pattern Recognit.* **29**(1), 51–59 (1996)
- N Dalal, B Triggs, Histograms of oriented gradients for human detection, in *Proceedings of the IEEE CVPR*, ed. by, Vol. 1st, San Diego, 2005, pp. 886–893
- C Cortes, V Vapnik, Support-vector networks. *Mach. Learn.* **20**(3), 273–297 (1995)
- B Li, M Meng, Computer-aided detection of bleeding regions for capsule endoscopy images. *IEEE Trans. Biomed. Eng.* **56**(4), 1032–1039 (2009)
- Y Sato, S Nakajima et al., 3D multiscale line filter for segmentation and visualization of curvilinear structure in medical images. *Med. Image Anal.* **2**, 143–168 (1998)
- J Liu, JM White, RM Summers, *Automated Detection of Blob Structures by Hessian Analysis and Object Scale* (Proceedings of the ICIP, Hong Kong, 2010), pp. 841–844
- PK Saha, JK Udupa, Scale-based image filtering preserving boundary sharpness and fine structure. *IEEE Trans. Med. Imag.* **20**, 1140–1155 (2001)
- H Zhang, M Wan, Z Bian, Complementary tensor-driven image coherence diffusion for oriented structure enhancement. *EURASIP J. Adv. Signal Process.* **70**(2011) (2011)
- K Zuiderveld, Contrast limited adaptive histogram equalization, in *Chapter VIII.5, Graphics Gems IV*, Academic Press, Cambridge, MA, 1994, pp. 474–485
- M Gil, EG Sarabia et al., Fuzzy c-means clustering for noise reduction, enhancement and reconstruction of 3D ultrasonic images, in *Proceedings of the ETFA*, Barcelona, 1999, pp. 465–472
- Y Yu, ST Acton, Speckle reducing anisotropic diffusion. *IEEE Trans. Image Process.* **11**(11), 1260–1270 (2002)

22. T Xu, I Cheng, M Mandal, An Improved fluid vector flow for cavity segmentation in chest radiographs, in *Proceedings of the ICPR*, Istanbul, 2010, pp. 3376–3379
23. KW Sum, PYS Cheung, Boundary vector field for parametric active contours. *Pattern Recognit.* **40**(6), 1635–1645 (2007)
24. J Canny, A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**(6), 679–714 (1986)
25. C Di Ruberto, A Dempster, Circularity measures based on mathematical morphology. *Electron. Lett.* **36**(20), 1691–1693 (2000)
26. G Dong, N Ray, S Acton, Intravital leukocyte detection using the gradient inverse coefficient of variation. *IEEE Trans. Med. Imag.* **24**(7), 910–924 (2005)
27. C Bishop, *Pattern Recognition and Machine Learning, Chap. 1* (Springer, New York, 2006)
28. CC Chang, CJ Lin, LIBSVM: A Library for Support Vector Machines. (2011). available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
29. T Hastie, R Tibshirani, J Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Chap. 7*, 2nd edn. (Springer, New York, 2009)
30. T Xu, I Cheng, M Mandal, Automated cavity detection of infectious pulmonary tuberculosis in chest radiographs, in *Proceedings of the IEEE EMBC*, Boston, 2011, pp. 5178–5181

doi:10.1186/1687-5281-2013-3

Cite this article as: Xu et al.: Novel coarse-to-fine dual scale technique for tuberculosis cavity detection in chest radiographs. *EURASIP Journal on Image and Video Processing* 2013 **2013**:3.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
