

# Cavity contour segmentation in chest radiographs using supervised learning and dynamic programming

Pragnya Maduskar<sup>a)</sup> and Laurens Hogeweg

*Diagnostic Image Analysis Group, Radboud University Medical Center, Nijmegen, 6525 GA, The Netherlands*

Pim A. de Jong

*Department of Radiology, University Medical Center Utrecht, 3584 CX, The Netherlands*

Liesbeth Peters-Bax

*Department of Radiology, Radboud University Medical Center, Nijmegen, 6525 GA, The Netherlands*

Rodney Dawson

*University of Cape Town Lung Institute, Cape Town 7700, South Africa*

Helen Ayles

*Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, United Kingdom*

Clara I. Sánchez and Bram van Ginneken

*Diagnostic Image Analysis Group, Radboud University Medical Center, Nijmegen, 6525 GA, The Netherlands*

(Received 1 November 2013; revised 3 March 2014; accepted for publication 20 April 2014; published 24 June 2014)

**Purpose:** Efficacy of tuberculosis (TB) treatment is often monitored using chest radiography. Monitoring size of cavities in pulmonary tuberculosis is important as the size predicts severity of the disease and its persistence under therapy predicts relapse. The authors present a method for automatic cavity segmentation in chest radiographs.

**Methods:** A two stage method is proposed to segment the cavity borders, given a user defined seed point close to the center of the cavity. First, a supervised learning approach is employed to train a pixel classifier using texture and radial features to identify the border pixels of the cavity. A likelihood value of belonging to the cavity border is assigned to each pixel by the classifier. The authors experimented with four different classifiers: *k*-nearest neighbor (*k*NN), linear discriminant analysis (LDA), GentleBoost (GB), and random forest (RF). Next, the constructed likelihood map was used as an input cost image in the polar transformed image space for dynamic programming to trace the optimal maximum cost path. This constructed path corresponds to the segmented cavity contour in image space.

**Results:** The method was evaluated on 100 chest radiographs (CXRs) containing 126 cavities. The reference segmentation was manually delineated by an experienced chest radiologist. An independent observer (a chest radiologist) also delineated all cavities to estimate interobserver variability. Jaccard overlap measure  $\Omega$  was computed between the reference segmentation and the automatic segmentation; and between the reference segmentation and the independent observer's segmentation for all cavities. A median overlap  $\Omega$  of 0.81 ( $0.76 \pm 0.16$ ), and 0.85 ( $0.82 \pm 0.11$ ) was achieved between the reference segmentation and the automatic segmentation, and between the segmentations by the two radiologists, respectively. The best reported mean contour distance and Hausdorff distance between the reference and the automatic segmentation were, respectively,  $2.48 \pm 2.19$  and  $8.32 \pm 5.66$  mm, whereas these distances were  $1.66 \pm 1.29$  and  $5.75 \pm 4.88$  mm between the segmentations by the reference reader and the independent observer, respectively. The automatic segmentations were also visually assessed by two trained CXR readers as "excellent," "adequate," or "insufficient." The readers had good agreement in assessing the cavity outlines and 84% of the segmentations were rated as "excellent" or "adequate" by both readers.

**Conclusions:** The proposed cavity segmentation technique produced results with a good degree of overlap with manual expert segmentations. The evaluation measures demonstrated that the results approached the results of the experienced chest radiologists, in terms of overlap measure and contour distance measures. Automatic cavity segmentation can be employed in TB clinics for treatment monitoring, especially in resource limited settings where radiologists are not available. © 2014 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4881096>]

Key words: cavitation, tuberculosis, segmentation, chest radiography, automated

## 1. INTRODUCTION

Cavitation in the lung parenchyma is an important sign of the presence of pulmonary tuberculosis (TB), an infectious disease with high morbidity and mortality, especially in developing nations. A cavity is defined as a gas-filled space within a pulmonary consolidation, a mass, or a nodule, produced by the expulsion of the necrotic part of the lesion via the bronchial tree.<sup>1</sup> Cavities may also occur in diseases, such as primary bronchogenic carcinoma, lung cancer, pulmonary metastasis, pneumonia, and other infections. Cavitation is prominently visible and can be easily detected using computed tomography (CT) imaging [see Fig. 1(a)], but, in practice, the first-line imaging modality for TB diagnosis and screening is performed using chest radiography.<sup>2</sup> Cavities are often barely visible on chest radiographs (CXR) due to other superimposed 3D lung structures in the 2D projection image [Fig. 1(b)]. In CXRs, the appearance of cavities can be hazy [see Fig. 2(a)], and the cavity walls are often ill-defined or completely invisible [see Fig. 2(b)]. This poses a considerable challenge for the radiologists to detect and accurately measure cavities in chest radiographs.

Assessing cavity size and its variation between temporal scans is important for the disease diagnosis and for measuring the response to the TB treatment.<sup>3</sup> Cavities tend to increase in size during the initial treatment period and diminish in size toward the end of the treatment indicating the efficacy of drugs.<sup>4</sup> This can be used to decide on the treatment period. Hamilton *et al.*<sup>5</sup> showed that a persistent cavity after six months of TB treatment is a biomarker of TB relapse. Existence of cavitation is common in postprimary tuberculosis<sup>6</sup> and is higher in patients with TB and coexisting diabetes.<sup>7</sup> The presence of cavities is strongly associated with high TB mycobacterial load representing an advanced state of the disease.<sup>8,9</sup> The number and size of the cavities is an important category in scoring systems developed for TB diagnosis in CXRs.<sup>3,10,11</sup>

Automatic detection and segmentation of cavities is a relatively unexplored research area. Shen *et al.*<sup>12</sup> proposed a cavity detection system in CXRs for TB screening. Initial contours of suspected cavities were defined by adaptive thresholding integrated with the mean shift segmentation technique followed by an active contour model. Segmented candidates were classified as a cavity or a noncavity using

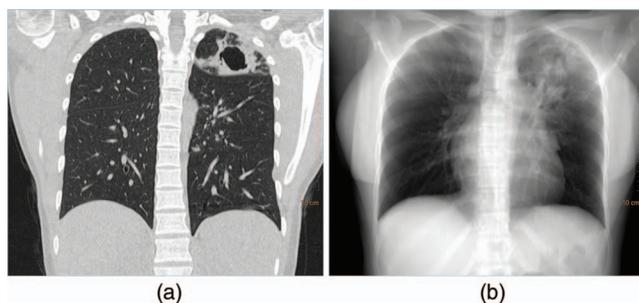


FIG. 1. Appearance of a cavity on a chest CT scan versus a chest radiograph: (a) Coronal CT section with a clearly visible cavity in upper left lobe. (b) A simulated chest radiograph by averaging intensity projection of coronal CT sections. Due to superimposed structures, the cavity is difficult to discern.

a Bayesian classifier. A second classification step was performed to detect the missed cavities in the upper lobes near clavicles. The technique was tested on only 16 CXRs with cavitation and a threshold on Tanimoto overlap measure (alternatively known as the Jaccard index)<sup>13</sup> was used to classify detected cavity regions as true or false positives. Shen *et al.* did not report the accuracy of their contour segmentation of the cavities. Another automatic cavity detection system was presented by Xu *et al.*<sup>14</sup> based on a coarse-to-fine dual scale methodology, where simpler features, such as Gaussian-based matching and local binary patterns, were applied at a coarse scale, while complex features, such as circularity and Kullback-Leibler divergence measures, were applied on a finer scale for the cavity classification. This method was validated on 35 CXRs containing 50 cavities and reported an average Tanimoto overlapping degree of 67.1%. The same author group proposed cavity segmentation techniques previously, which were based on an improved edge-based fluid vector flow snake model,<sup>15,16</sup> and were validated on 20 chest radiographs and resulted in a mean Jaccard overlapping degree of 77.2% (Ref. 15) and 68.8%.<sup>16</sup> Hence, previous studies were performed on small patient groups and were not extensively validated.

The focus of this work is cavity contour segmentation to assist cavity size and diameter measurement for therapy response. This task can be time consuming for a human reader and is prone to considerable interobserver variability. We use a database of 150 digital CXRs with cavities for training and evaluating the method.

Various contour segmentation methods are available in the literature which are based on energy minimization approaches, namely, active contour models<sup>17,18</sup> and active shape models.<sup>19</sup> These techniques deform the initial contour by minimizing the contour energy  $E$  being directed toward the object boundary. These techniques are highly dependent on contour initialization and the defined control points. All these methods are prone to local minima providing a suboptimal solution. In contrast, graph cuts<sup>20</sup> based segmentation solves a global optimization problem, but it assumes the foreground object to have a uniform structure which is different from background pixels. In case of cavities in CXRs, only the borders are distinctly visible, whereas the inner cavity region, often somewhat more radiolucent, shares similar texture with the remaining lung parenchyma. We therefore propose a supervised learning approach followed by a dynamic programming-based solution for automated segmentation of the cavity contour. Supervised learning approaches have been successfully used for edge detection of the target objects in various applications.<sup>21-24</sup> Likewise, dynamic programming has been employed in computer vision<sup>25</sup> and other medical image segmentation tasks.<sup>26,27</sup> In the proposed method, a seed point near the center of the cavity is required as a user input to define the region of interest.

This paper is organized as follows. Section 2 describes the data used. Section 3 describes the proposed method in detail. The experiments used to evaluate the method are described in Sec. 4. Section 5 presents results, which are discussed in Sec. 6. Section 7 concludes.

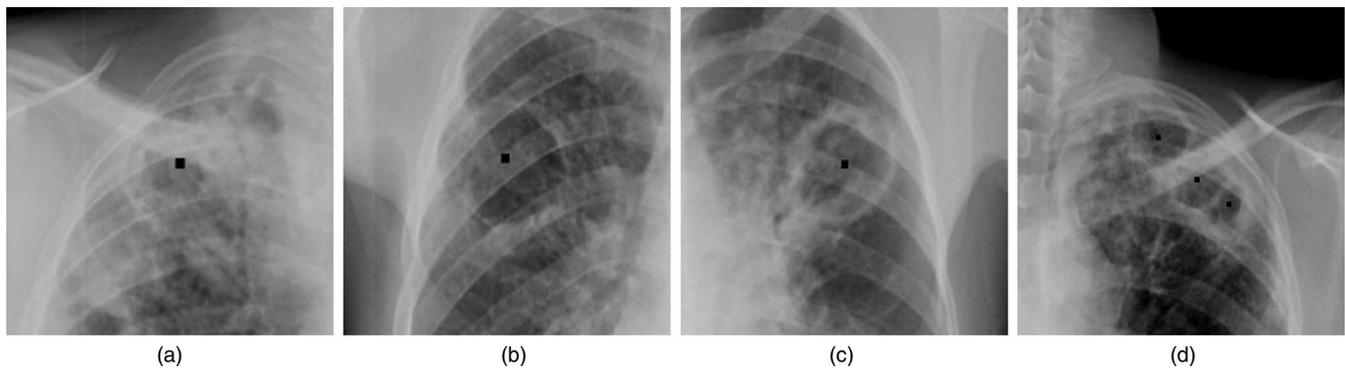


FIG. 2. Cavity example images on a chest radiograph, center of cavity is marked with a black pointer: (a) A subtle cavity. (b) Cavity with broken or incomplete borders. (c) An obvious cavity. (d) Multiple cavities.

## 2. DATA

The dataset collected for this study is part of a large database of over 20 000 CXRs originated from two sites in Africa (Cape Town, South Africa; Lusaka, Zambia) with a high TB incidence. Ethics approval was obtained from the University of Cape Town Health Research Ethics Committee and the University of Zambia Biomedical Research Ethics Committee that allowed usage of the data for research purposes. The digital CXRs of resolution in the range 1520–1812 pixels with an isotropic pixel size of 0.25 mm were acquired using a digital Odelca-DR system with a slotscan detector (Delft Imaging Systems, The Netherlands). An extensive cavity annotation procedure was followed to identify the CXRs containing cavities and to precisely annotate the cavity contours (Fig. 3). A medical student reviewed  $\approx 3000$  CXRs to identify images with possible presence of cavities. The images with identified cavities were reviewed by two experienced board-certified radiologists with varied years of experience in chest imaging. 150 CXRs were selected which were independently confirmed by both radiologists to have single

or multiple cavities. These 150 images were then split into a training and a test set with 50 images (50 cavities) and 100 images (126 cavities), respectively. The radiologist with  $>10$  years of experience precisely annotated cavity borders on the 100 test images to set the reference segmentation. To perform a detailed evaluation, the radiologist also assigned a subtlety rating to these cavities. The subtlety ratings were divided into five categories based on the level of visibility—“barely visible,” “subtle,” “moderately subtle,” “obvious,” and “prominent.” Cavity contours in the training images were annotated by a certified “B” reader trained to read chest radiographs according to the CRRS tuberculosis scoring system.<sup>10</sup> The other experienced chest radiologist with  $>5$  years of experience (independent observer) also outlined the borders of the cavities in the test set. The manual annotation procedure was as follows: the observer was shown a mark (Fig. 2) close to the center of the cavity and was requested to draw the corresponding cavity contour. For cavities with thick walls, the inner contours of the wall were considered to constitute the boundary of the cavity. Segmentation by the independent observer was employed to compare the interobserver variability

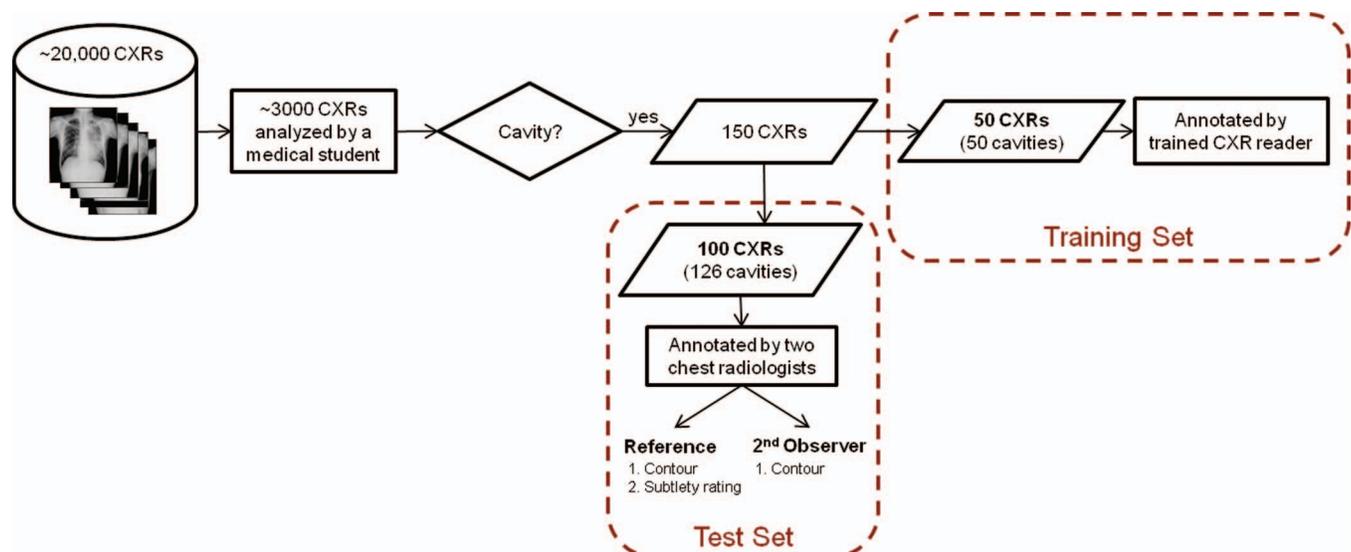


FIG. 3. Data selection and annotation procedure.

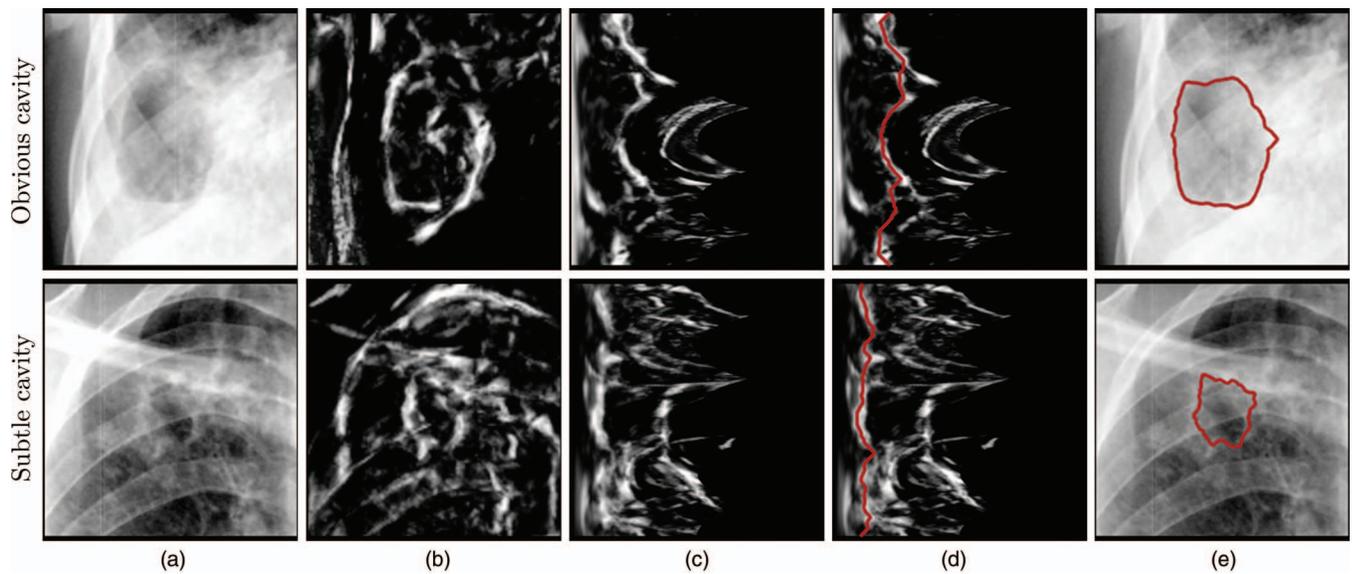


FIG. 4. Images depicting each step of the cavity contour segmentation technique for an obvious and a subtle cavity. (a) Cavity image. (b) Cavity border likelihood map. (c) Cost image after polar transform. (d) Maximum cost path overlaid on the cost image. (e) Segmented cavity contour overlaid on the cavity image.

between the manual segmentations with the reference and the automatic segmentation.

### 3. METHODS

We propose a method with two stages to segment the cavity borders. First, a supervised learning approach was used to train a pixel classifier to detect the border pixels of the cavity. The pixel classifier assigned each pixel a likelihood of belonging to the cavity border. This likelihood map was then used as an input cost image for dynamic programming to trace the optimal path in the polar transformed image space. This constructed path corresponds to the segmented cavity contour in image space. An illustration of all the steps is shown in Fig. 4 for an obvious and a subtle cavity. We discuss the details of each step in Subsections 3.A–3.C.

#### 3.A. Cavity border discriminative features

This step involves calculation of a set of descriptive features to allow discrimination between cavity border and the background. We calculated two sets of pixel-based local features to capture various properties associated with the cavity borders. The CXRs were resized to a pixel width of 1024 ( $\approx 0.4$  mm pixel size) and then subsampled by a factor of 2 to speed up the feature computation process.

##### 3.A.1. Texture features

The image  $I$  was filtered with a multiscale feature bank of Gaussian derivatives [Eq. (1)] to capture texture and strong edges of cavity borders<sup>28</sup>

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}. \quad (1)$$

Equation (1) shows the two-dimensional Gaussian function  $G(x, y)$ , where  $x$  and  $y$  are the pixel coordinates at a location in the image; and  $\sigma$  represents the spread of the function. Similarly, first order ( $G_x(x, y)$ ,  $G_y(x, y)$ ) and second order ( $G_{xx}(x, y)$ ,  $G_{xy}(x, y)$ , and  $G_{yy}(x, y)$ ) Gaussian derivatives were calculated and convoluted with image  $I$  to create Gaussian filtered images.

In CXRs, cavity borders give high response on higher order derivatives and this response is captured by Gaussian derivatives (Fig. 5). The image was filtered with Gaussian derivatives of orders 0, 1, and 2 at scales  $\sigma = 1, 2, 4, 8$ , and 16 pixels. In total, 31 features were extracted including the original intensity feature  $I(x, y)$ .

Cavity walls appear like broken line segments in CXRs (Fig. 2). This linelike structure can be captured using the eigenvalues of Hessian matrix  $H$  of the Gaussian filtered images.<sup>29</sup> If  $\lambda_1$  and  $\lambda_2$  ( $|\lambda_1| \geq |\lambda_2|$ ) are the two eigenvalues of  $H$ , then  $\lambda_1$  should be very large compared to  $\lambda_2$  for line-like structures. Two features at four different Gaussian scales ( $\sigma = 1, 2, 4, 8$  pixels) using the Hessian matrix were extracted, i.e., the largest absolute eigenvalue ( $|\lambda_1|$ ) and absolute difference of the two eigenvalues ( $\sqrt{|\lambda_1^2 - \lambda_2^2|}$ ). An example feature output of Hessian features at scale  $\sigma = 2$  pixels is demonstrated in Fig. 5. A total of 39 texture features were computed.

##### 3.A.2. Radial features

Radial features were calculated using the center of the cavity. Hence, a single user interaction was required, where the user of this segmentation system needed to click a single point close to the center of the cavity. A subimage of isotropic dimensions of 250 pixels (200 mm) was extracted around the input seed point. These dimensions were chosen based on the probable maximum size of a cavity which was estimated using the training dataset.

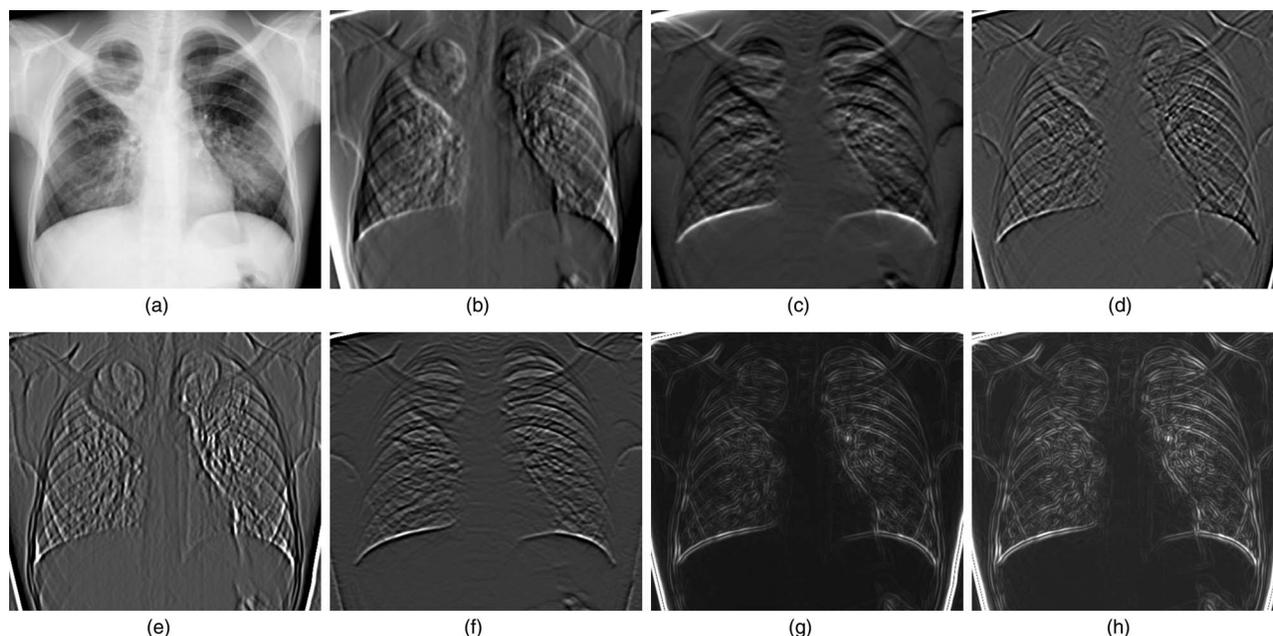


FIG. 5. Feature images of Gaussian derivatives and Hessian filter at a scale  $\sigma = 2$  pixels of a CXR image with a cavity in the right upper lobe. (a) Original CXR image. (b)  $G_x(x, y)$ . (c)  $G_y(x, y)$ . (d)  $G_{xy}(x, y)$ . (e)  $G_{xx}(x, y)$ . (f)  $G_{yy}(x, y)$ . (g) Largest eigenvalue. (h) Difference of the two eigenvalues.

Cavities are usually circular or elliptical in shape. Hence, gradient response for the cavity border pixels should be higher if gradients are calculated in the direction from center of the cavity (seed point) to the cavity border pixels. Other pixels have a low response on such a directional gradient filter. We refer to these features as radial features, as they were calculated in the direction from the center of the cavity to all the other pixels. These features were computed at scales  $\sigma = 2$  and 4 pixels for first- and second-order Gaussian derivatives. Also the distance to the center of the cavity was added to the feature set to implicitly encode knowledge about a typical cavity size. Radial feature images for an obvious cavity are shown in Fig. 6. In total, seven radial features were calculated and added to the feature set.

### 3.B. Pixel classification

Cavity border discriminative features were calculated at each pixel location in the training images to train the classi-

fier to discriminate between a cavity border pixel and a background pixel. For training purposes, we used CXRs with a single cavity and a seed point near the center of the cavity to calculate the radial features. All features were normalized to zero mean and unit standard deviation before classification.

A pixel classifier was trained to assign likelihood values of being a cavity border pixel to each location in a CXR. Cavity border pixels were sampled within 0.5 mm from manually annotated cavity contours to train the positive class of the pixel classifier. For the negative (background) class, random samples were chosen from the same CXRs. Half of the samples were randomly selected close to the cavity border but within 1.25–2.5 mm distance from the cavity border and the other half at a distance greater than 2.5 mm. The goal of this sample selection methodology for negative samples was to include texture diversity present in the lung parenchyma and nearby rib borders that can be mistaken as a cavity border pixel. Given a new test image, texture features were calculated and stored. When a user provided an input seed point,

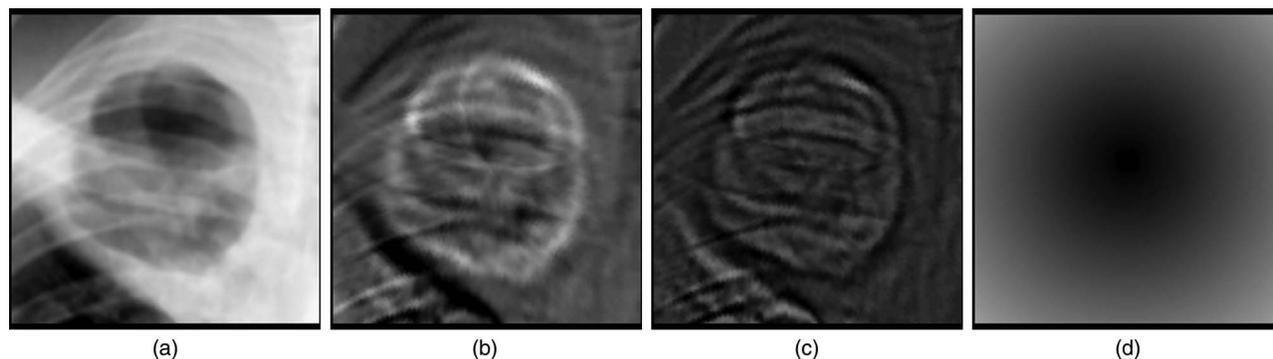


FIG. 6. Response images of radial features: (a) Cavity subimage (center represents the input), (b) and (c) First- and second-order directional gradient filter at scale  $\sigma = 2$ . Note how the sign of the response is positive for the inner wall border of the cavity where brightness changes from dark to bright, and negative for erroneous borders. (d) Distance to center of the cavity.

radial features were calculated on the extracted subimage which were then combined with the texture features. The subimage pixels were classified using the trained classifier which estimated the likelihood of belonging to the cavity border at each pixel location.

### 3.C. Contour segmentation

Given a cost image, dynamic programming can be utilized to find an optimal cost path between two pixels. As cavities are mostly elliptical in shape, optimal path calculation was done in the polar space. The polar image was constructed by extracting a circular region of interest (ROI) of radius  $R$  around the seed point on the cavity border likelihood subimage. The radius  $R$  was chosen sufficiently large so as to capture cavities of all sizes. In our application, we have chosen the maximum radius as 50 mm. The  $x$  axis in the polar image represents the radius from 0 to  $R$  and the  $y$  axis represents the angle from  $-\pi$  to  $\pi$  [see Fig. 4(c)]. Optimal cost path here refers to the maximum cost between two points as higher likelihood reflects more certainty of a pixel being part of a cavity border. We calculated the cost of the optimal path at various radii starting from  $R_{\min}$  to  $R_{\max}$  with a step size of  $\delta_R$  and chose the one with the maximum cost [Fig. 4(d)]. For our application, we chose  $R_{\min}$  as 7.5 mm,  $R_{\max}$  as 30 mm, and  $\delta_R$  as 0.5 mm. The start and the end points for the path calculation were set to the same location to ensure a closed contour when the maximum cost path was projected back to the original image space.

## 4. EXPERIMENTS

### 4.A. Cavity segmentation: Various classification configurations

The training set for training the pixel classifier consisted of 45 885 samples (11 605 cavity border pixels and 34 280 background pixels) sampled from the training images. We experimented with a number of classifiers:  $k$ -nearest neighbor ( $k$ NN),<sup>30,31</sup> linear discriminant analysis (LDA),<sup>30</sup> GentleBoost (GB),<sup>32</sup> and random forest (RF),<sup>33</sup> to analyze and compare cavity segmentations generated from different cavity likelihood maps. For  $k$ NN, we used  $k = 15$ , GB used 500 regression stumps as the weak classifiers, and the RF classifier used 100 decision trees with a maximum tree depth of 25. The number of nearest neighbors  $k$ , regression stumps and decision trees were chosen based on experimentation on an external training dataset outside this experiment. The external training dataset included CXRs used in training a detection system for TB. The effect of the parameters was evaluated using pixel-level area (AUC) under the receiver operating characteristics (ROC) curve in cross-validation on the external training set. The parameters with the maximum pixel-level AUC were the chosen parameters for each classifier.

We trained two variations of each pixel classifier based on the inclusion of different feature sets: (1) only texture features ( $F_T$ ), (2) texture and radial features ( $F_{All}$ ). This was done to evaluate the usefulness of including radial features. In this study, the seed point for each cavity was chosen only once by

a CRRS trained reader who was blinded to the segmentations by the radiologists.

### 4.B. Segmentation evaluation

The effect of different classification configurations was evaluated on the final cavity segmentation output as described in Subsections 4.B.1–4.B.2 Subsection 4.B.3 illustrates the visual assessment of the automatic cavity segmentation. The effect of seed point variation on the proposed segmentation method is explained in Subsection 4.C.

#### 4.B.1. Overlap measure

The Jaccard overlap measure  $\Omega$  [Eq. (2)]<sup>13</sup> was used to assess the segmentation accuracy of the proposed system. The measure produces a score of 1 for a perfect agreement and 0 for no agreement. The shape complexity and the size of the object can result in lower overlap values for small objects than for large objects.<sup>34</sup> In addition, the measure does not account for over or under-segmentation of the segmented object in comparison to the reference. Hence, we computed two additional measures—oversegmentation [Eq. (3)] and undersegmentation [Eq. (4)] to identify the false positive and false negative pixels with respect to the reference segmentation. To make these measures comparable for objects of different sizes, the amount is reported as a fraction of size of the reference object,

$$\Omega = \frac{|A \cap R|}{|A \cup R|}, \quad (2)$$

$$O = \frac{|\{v | R(v) = 0 \wedge A(v) = 1\}|}{|R|}, \quad (3)$$

$$U = \frac{|\{v | R(v) = 1 \wedge A(v) = 0\}|}{|R|}. \quad (4)$$

$A$  refers to the automatic segmentation and  $R$  to the reference segmentation. Oversegmentation is measured as the number of pixels  $v$ , for which  $R(v) = 0$  and  $A(v) = 1$ , and undersegmentation is the opposite.  $|R|$  is the total number of pixels in the reference segmentation.

#### 4.B.2. Contour distance measures

Contour distance measures,<sup>34</sup> namely, Hausdorff and mean absolute contour distance, were applied to evaluate the distance between the segmented contours. Let  $A$  be the automatic segmentation and  $R$  be the reference segmentation, mean absolute contour distance and Hausdorff distance can be written as Eqs. (6) and (7), respectively. For each point on contour  $A$ , the closest point on contour  $R$  was computed using  $d(a, R)$  [Eq. (5)]. The distance measurements were repeated with contours  $A$  and  $R$  interchanged to make the measure symmetric. The distances are measured in millimeters; the pixel spacing being 0.25 mm,

$$d(a, R) = \min_{r \in R} |r - A|, \quad (5)$$

$$\text{mean}_{a \in A} d(a, R), \quad (6)$$

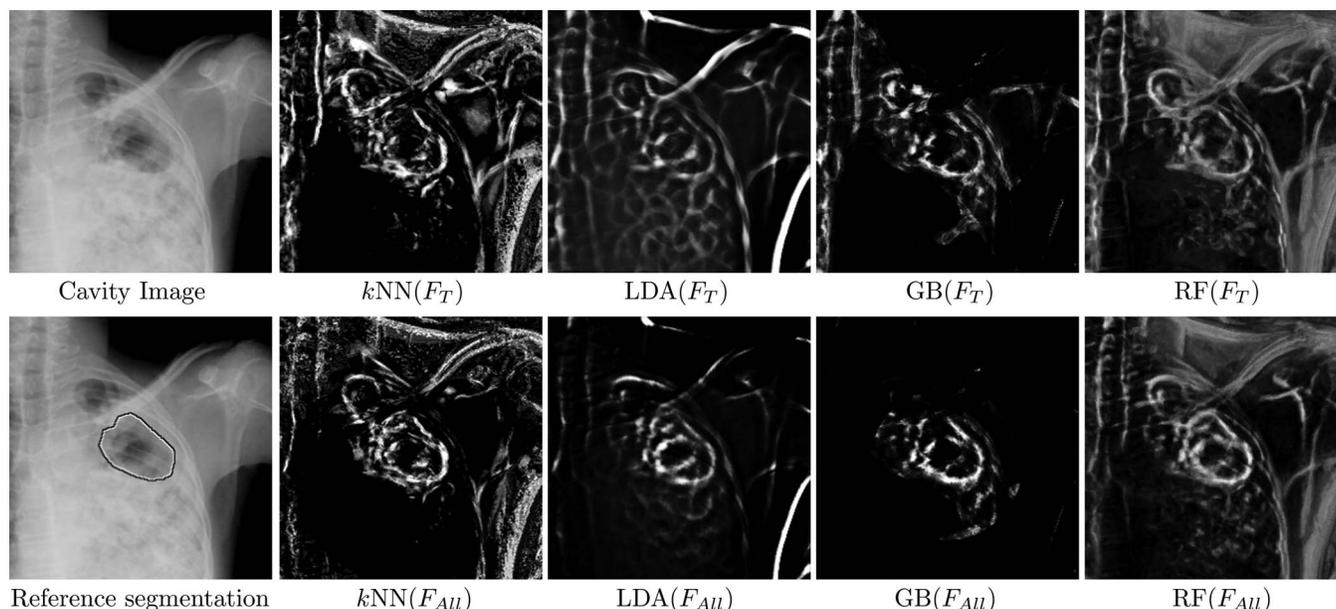


FIG. 7. Cavity border likelihood map generated using different classifier models with inclusion of only texture features ( $F_T$ ) and all features ( $F_{All}$ ).

$$\max_{a \in A} d(a, R). \quad (7)$$

All the measures were calculated between the reference segmentation and the automatic segmentation; and between the reference segmentation and the segmentation provided by the independent observer, for all the cavities. Additionally, the overlap measure was calculated individually for each subtlety category. The usefulness of inclusion of radial features was measured by pairwise comparison of measurement values between the classifier trained with only texture features ( $F_T$ ), and all features ( $F_{All}$ ). Wilcoxon signed rank test was performed using statistical software package R (v2.15.0) to compute  $p$ -values for this comparison on the test set. A  $p$ -value of less than 0.05 was considered to indicate a statistically significant difference.

#### 4.B.3. Visual assessment

The automatically generated cavity contours were shown to two CRRS certified readers independently and they were asked to rate the quality of the segmentation. They could provide one of the following ratings to each cavity: “excellent,” “adequate,” or “insufficient.” We report their rating results to judge the acceptance of the automatically generated contours. Following the visual assessment, the segmentations rated as “insufficient” by any of the readers were reviewed again and the cause of the failure of the method was characterized and reported.

#### 4.C. Cavity segmentation: Seed point variation and reproducibility

Cavity segmentation accuracy can be affected by the choice of the seed point by the user. In the real world applications of the method, the user can be an experienced radiologist, a clinical officer or a radiographer. Our method assumes

that the provided seed point is close to the center of the cavity. Experiments were performed by varying the seed point location to determine its implication on the segmentation output. This was executed as follows:

1. Consider a seed point  $s$  defined at a location  $(x, y)$ . Concentric circles were assumed around  $s$  at varying radii  $r$ .
2. Two diametrically opposite points were chosen as seed points at radius  $r$ .
3. The proposed method was run on these simulated seed points.
4. Segmentation evaluation measures as mentioned in Sec. 4.B.1–4.B.2 were computed for quantitative analysis of the effect of seed point variation on the segmentation algorithm.

We chose six radii on a logarithmic scale, i.e.,  $r = 2^n$  pixels,  $n = 1, 2, 3, 4, 5, 6$ .

## 5. RESULTS

### 5.A. Cavity segmentation: Various classification configurations

Cavity border likelihood maps generated using different classifier configurations are shown in Fig. 7. The response likelihood map varied with the classifier model and with inclusion of only texture features ( $F_T$ ) or all features ( $F_{All}$ ). The  $kNN$  classifier gave a high response on the cavity borders but also on other anatomical structures such as clavicle contours. GB classifier produced a high response only on the cavity borders but missed subtle cavity borders. With the other two classifiers, LDA and RF, cavity borders showed a high response in comparison to other anatomical structures. In all classifier models, inclusion of the radial features reduced false responses of the classifier on bony structures and other

TABLE I. Jaccard overlap measure  $\Omega$ , oversegmentation  $O$ , and undersegmentation  $U$  between the reference segmentation and the independent observer, and between the reference and the automatic segmentation, for all the cavities. Results are shown with inclusion of only texture features ( $F_T$ ) and all features ( $F_{All}$ ) for training different classifier models.  $p$ -values were computed with a Wilcoxon signed rank test between  $F_T$  and  $F_{All}$  for each classification configuration. \* indicates a significant result with  $p$ -value  $<0.05$  calculated for Jaccard overlap measure.

Evaluation	Feature set	Jaccard overlap ( $\Omega$ )		Oversegmentation ( $O$ )		Undersegmentation ( $U$ )	
		$\mu \pm \sigma$	Median	$\mu \pm \sigma$	Median	$\mu \pm \sigma$	Median
Independent observer		$0.82 \pm 0.11^*$	0.85	$0.07 \pm 0.10$	0.03	$0.12 \pm 0.12$	0.08
AS <sub>kNN</sub>	$F_T$	$0.71 \pm 0.18$	0.76	$0.38 \pm 1.22$	0.10	$0.14 \pm 0.17$	0.08
	$F_{All}$	$0.74 \pm 0.16$	0.78	$0.31 \pm 1.13$	0.08	$0.15 \pm 0.15$	0.10
AS <sub>LDA</sub>	$F_T$	$0.74 \pm 0.16$	0.78	$0.24 \pm 0.31$	0.15	$0.11 \pm 0.18$	0.03
	$F_{All}$	$0.76 \pm 0.16^*$	0.81	$0.13 \pm 0.20$	0.06	$0.15 \pm 0.18$	0.08
AS <sub>GB</sub>	$F_T$	$0.70 \pm 0.19$	0.76	$0.31 \pm 1.12$	0.09	$0.17 \pm 0.20$	0.08
	$F_{All}$	$0.74 \pm 0.17$	0.79	$0.22 \pm 1.00$	0.04	$0.19 \pm 0.17$	0.14
AS <sub>RF</sub>	$F_T$	$0.74 \pm 0.15$	0.77	$0.30 \pm 1.09$	0.12	$0.13 \pm 0.14$	0.08
	$F_{All}$	$0.76 \pm 0.16^*$	0.81	$0.18 \pm 0.50$	0.06	$0.15 \pm 0.15$	0.11

normal anatomy. All the classifiers occasionally respond to other structures in the radiograph but, overall, the likelihood values of the cavity border pixels were higher in comparison to the surrounding tissues. This makes it possible to use this likelihood map as a cost function for dynamic programming to segment the cavity border.

## 5.B. Segmentation evaluation

### 5.B.1. Overlap and contour distance measures

Table I shows the comparison of segmentation overlap ( $\Omega$ ) statistics between the reference segmentation and the independent observer; and the reference segmentation and the automatic segmentation for the four classifiers (AS<sub>kNN</sub>, AS<sub>LDA</sub>, AS<sub>GB</sub>, AS<sub>RF</sub>) with inclusion of only texture features ( $F_T$ ) and all features ( $F_{All}$ ). The independent observer achieved an overlap  $\Omega$  of  $0.82 \pm 0.11$  for all cavities.  $\Omega$  between the reference segmentation and various classifier configurations were in range of 0.70–0.76 where AS<sub>LDA</sub>– $F_{All}$  and AS<sub>RF</sub>– $F_{All}$  achieved the best overlap of  $0.76 \pm 0.16$ .

The inclusion of all the features for cavity border likelihood map computation performed significantly better ( $p < 0.05$ ) than using only texture features for LDA and RF classifiers. Mean undersegmentation scores for the independent observer (0.12) and the classifier configurations (0.11–0.19) were similar, although the automatic segmentation had higher mean oversegmentation scores (0.13–0.38) than the independent observer (0.07).

The results with contour based measures for the evaluation of automatic segmentation are shown in Table II. Mean contour distance and Hausdorff distance between the reference and various classifier configurations were on average between 2.48–3.47 mm, and 8.32–9.80 mm, respectively, whereas the contour distance values for the independent observer were smaller with  $1.66 \pm 1.29$  mm mean contour distance and  $5.75 \pm 4.88$  mm Hausdorff distance. Mean contour distances were significantly lower for the classifier models trained with all the features than only texture features.

Overall, it can be seen in the box plots (Fig. 8) of different evaluation measures that the classifiers trained with all features have improved median scores (maximum overlap and

TABLE II. Mean contour distance (in mm) and Hausdorff distance (in mm) between the reference segmentation and the independent observer, and between the reference and the automatic segmentation, for all the cavities. Results are shown with inclusion of only texture features ( $F_T$ ) and all features ( $F_{All}$ ) for training different classifier models.  $p$ -values were computed with a Wilcoxon signed rank test between  $F_T$  and  $F_{All}$  for each classification configuration. \* indicates a significant result with  $p$ -value  $<0.05$ .

Evaluation	Feature set	Mean contour distance (in mm)		Hausdorff distance (in mm)	
		$\mu \pm \sigma$	Median	$\mu \pm \sigma$	Median
Independent observer		$1.66 \pm 1.29^*$	1.23	$5.75 \pm 4.88^*$	3.86
AS <sub>kNN</sub>	$F_T$	$3.34 \pm 3.46$	2.18	$9.66 \pm 6.05$	7.79
	$F_{All}$	$2.93 \pm 2.89^*$	2.09	$8.96 \pm 5.68$	7.23
AS <sub>LDA</sub>	$F_T$	$2.92 \pm 3.03$	2.09	$8.81 \pm 5.62$	7.48
	$F_{All}$	$2.48 \pm 2.19^*$	1.76	$8.44 \pm 5.85$	7.07
AS <sub>GB</sub>	$F_T$	$3.47 \pm 3.44$	2.34	$9.80 \pm 5.91$	7.93
	$F_{All}$	$2.85 \pm 3.01^*$	1.85	$8.80 \pm 5.93$	7.26
AS <sub>RF</sub>	$F_T$	$2.79 \pm 2.55$	2.14	$8.80 \pm 4.92$	7.38
	$F_{All}$	$2.55 \pm 2.39^*$	1.75	$8.32 \pm 5.66$	6.75

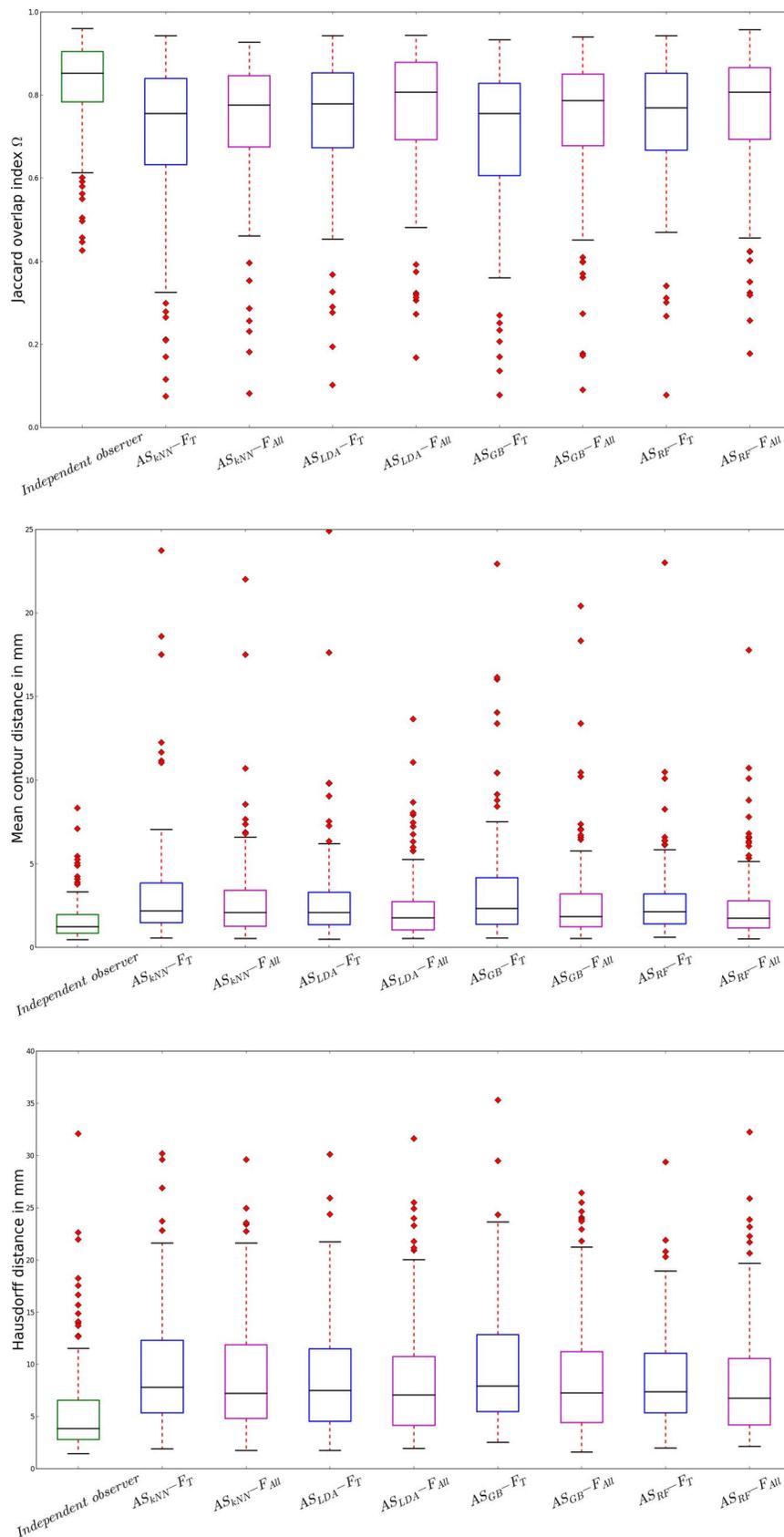


FIG. 8. Box plots of Jaccard overlap measure ( $\Omega$ ), mean contour distance (in mm), and Hausdorff distance (in mm) calculated between the reference segmentation and segmentation by the independent observer; and between the reference and the automatic segmentations generated using different classifier configurations for 126 cavities. The corresponding values are listed in Tables I and II. The central line of the box indicates the median, the box edges the 25th and 75th percentiles and the whiskers the extremes of the data excluding outliers. Points plotted as diamonds are the outliers lying beyond 1.5 interquartile range from the mean.

TABLE III. Jaccard overlap measure  $\Omega$  between the reference segmentation and the segmentation by the independent observer, and between the reference and the automatic segmentation, for various subtlety categories. Results are shown with inclusion of only texture features ( $F_T$ ) and both texture and radial features ( $F_{All}$ ) for different classifier models.  $\tilde{\Omega}$  represents the median value of overlap  $\Omega$  for the cavities in the test set.

$\Omega$	Feature set	Barely visible (22 cavities)		Subtle (26 cavities)		Moderate (39 cavities)		Obvious (29 cavities)		Prominent (10 cavities)	
		$\mu \pm \sigma$	$\tilde{\Omega}$	$\mu \pm \sigma$	$\tilde{\Omega}$	$\mu \pm \sigma$	$\tilde{\Omega}$	$\mu \pm \sigma$	$\tilde{\Omega}$	$\mu \pm \sigma$	$\tilde{\Omega}$
Independent observer		$0.69 \pm 0.15$	0.74	$0.81 \pm 0.10$	0.82	$0.85 \pm 0.09$	0.87	$0.87 \pm 0.04$	0.89	$0.91 \pm 0.06$	0.93
AS <sub>kNN</sub>	$F_T$	$0.66 \pm 0.15$	0.71	$0.63 \pm 0.23$	0.72	$0.73 \pm 0.18$	0.79	$0.75 \pm 0.14$	0.78	$0.85 \pm 0.09$	0.86
	$F_{All}$	$0.68 \pm 0.16$	0.72	$0.70 \pm 0.20$	0.75	$0.74 \pm 0.17$	0.81	$0.78 \pm 0.12$	0.81	$0.84 \pm 0.09$	0.85
AS <sub>LDA</sub>	$F_T$	$0.65 \pm 0.17$	0.73	$0.68 \pm 0.21$	0.77	$0.76 \pm 0.12$	0.78	$0.80 \pm 0.13$	0.84	$0.83 \pm 0.08$	0.86
	$F_{All}$	$0.68 \pm 0.17$	0.74	$0.73 \pm 0.18$	0.78	$0.78 \pm 0.15$	0.82	$0.79 \pm 0.16$	0.83	$0.85 \pm 0.09$	0.86
AS <sub>GB</sub>	$F_T$	$0.66 \pm 0.13$	0.65	$0.70 \pm 0.18$	0.75	$0.68 \pm 0.21$	0.75	$0.71 \pm 0.20$	0.78	$0.83 \pm 0.09$	0.85
	$F_{All}$	$0.68 \pm 0.16$	0.69	$0.68 \pm 0.23$	0.77	$0.75 \pm 0.17$	0.80	$0.79 \pm 0.12$	0.80	$0.83 \pm 0.11$	0.87
AS <sub>RF</sub>	$F_T$	$0.67 \pm 0.14$	0.70	$0.68 \pm 0.19$	0.72	$0.76 \pm 0.13$	0.77	$0.79 \pm 0.14$	0.82	$0.83 \pm 0.09$	0.86
	$F_{All}$	$0.70 \pm 0.14$	0.74	$0.73 \pm 0.17$	0.78	$0.75 \pm 0.18$	0.82	$0.81 \pm 0.11$	0.84	$0.84 \pm 0.09$	0.85

minimum contour distances) and lower standard deviation than classifiers trained with only texture features. We also performed overlap analysis separately for each subtlety category as rated in the reference standard (Table III). The independent observer had significantly better  $\Omega$  than the automatic segmentation variations for all the subtlety categories except for “barely visible” cavities where all overlap values were similar. Reported results show increasing trend for  $\Omega$  from most subtle cavities (“barely visible”) to the most obvious cavities (“prominent”). Example images for each subtlety category consisting of the reference segmentation, the segmentation by the independent observer and the automatic segmentation are shown in Fig. 9. We show all the result images for automatic segmentation with the likelihood map constructed using  $RF-F_{All}$  classifier.

### 5.B.2. Visual assessment

Cavity segmentations constructed using the RF classifier trained with all the features were shown independently to two CRRS readers for visual assessment. Out of the 126 automatic cavity segmentations, 72 (57%), 42 (33%), and 12 (10%) were rated by CRRS Reader 1 as “excellent,” “adequate,” and “insufficient,” respectively. According to CRRS Reader 2, 93 (74%), 17 (13%), and 16 (13%) had “excellent,” “adequate,” and “insufficient” segmentations, respectively. Table IV reports agreement between the ratings given by the two CRRS readers. They demonstrated good agreement in assessing the cavities and rated 84% of the cavities as “excellent” or “adequate.” Remaining 16% were labeled as “insufficient” by at least one of the CRRS readers. The inaccurately segmented cavities were further reviewed and classified based on their cause of failure (Table V). The reason of failure can be divided into four major categories; 35% of the inaccurately segmented cavities were attracted to clavicle or rib borders, 25% were very subtle lesions, 20% were located in severely diseased region, 15% were huge cavities with  $>7$  cm diameter, and 5% were overlapping with other cavities.

### 5.C. Seed point variation and reproducibility

The effect of the seed point variation on the performance of the automatic segmentation using  $RF-F_{All}$  classifier model is illustrated in Fig. 10. The seed point was varied in a circular neighborhood from 2 pixels (0.5 mm) to 64 pixels (16 mm). It is evident from the graphs that the performance is stable till a displacement of 8 pixels but begins to drop from 16 pixels with worse performance at a displacement of 64 pixels. This trend is observed in all the three evaluation measures: Jaccard overlap, mean contour distance, and Hausdorff distance. Jaccard overlap remains  $\approx 0.76$  till a distance of 8 pixels and starts diminishing at a distance of 16 pixels ( $\Omega = 0.73$ ) and performs poorly at a distance of 64 pixels ( $\Omega = 0.30$ ). Similarly, the mean contour distance and Hausdorff distance, respectively, increases from 11.5 to 37.9 mm, and 37.6 to 105.7 mm when the distance increases from 16 to 64 pixels.

## 6. DISCUSSION

A novel technique was presented to automatically segment cavities on chest radiographs using dynamic programming by computing an optimal path in an input cost image. The cost image was defined as the cavity border likelihood map in the polar space. The best automatic segmentation technique performed well for most of the cavities with a mean overlap of  $0.76(\pm 0.16)$  as shown in Table I. Additional contour distance measures, namely, mean contour distance and Hausdorff distance demonstrated low average distance values of  $2.55 \text{ mm}(\pm 2.39)$  and  $8.32(\pm 5.66)$ , respectively. The automatic segmentation was compared with a reference segmentation set by an experienced chest radiologist on a large dataset which included prominent as well as subtle cavities. Analyzing the results per subtlety category (see Table III) show that the automatic segmentation technique performs as good as the independent human expert for “barely visible” cavities but there is slight room for improvement for the other subtlety categories. In general, the automatic annotations look very similar to the manual annotations by the experienced chest

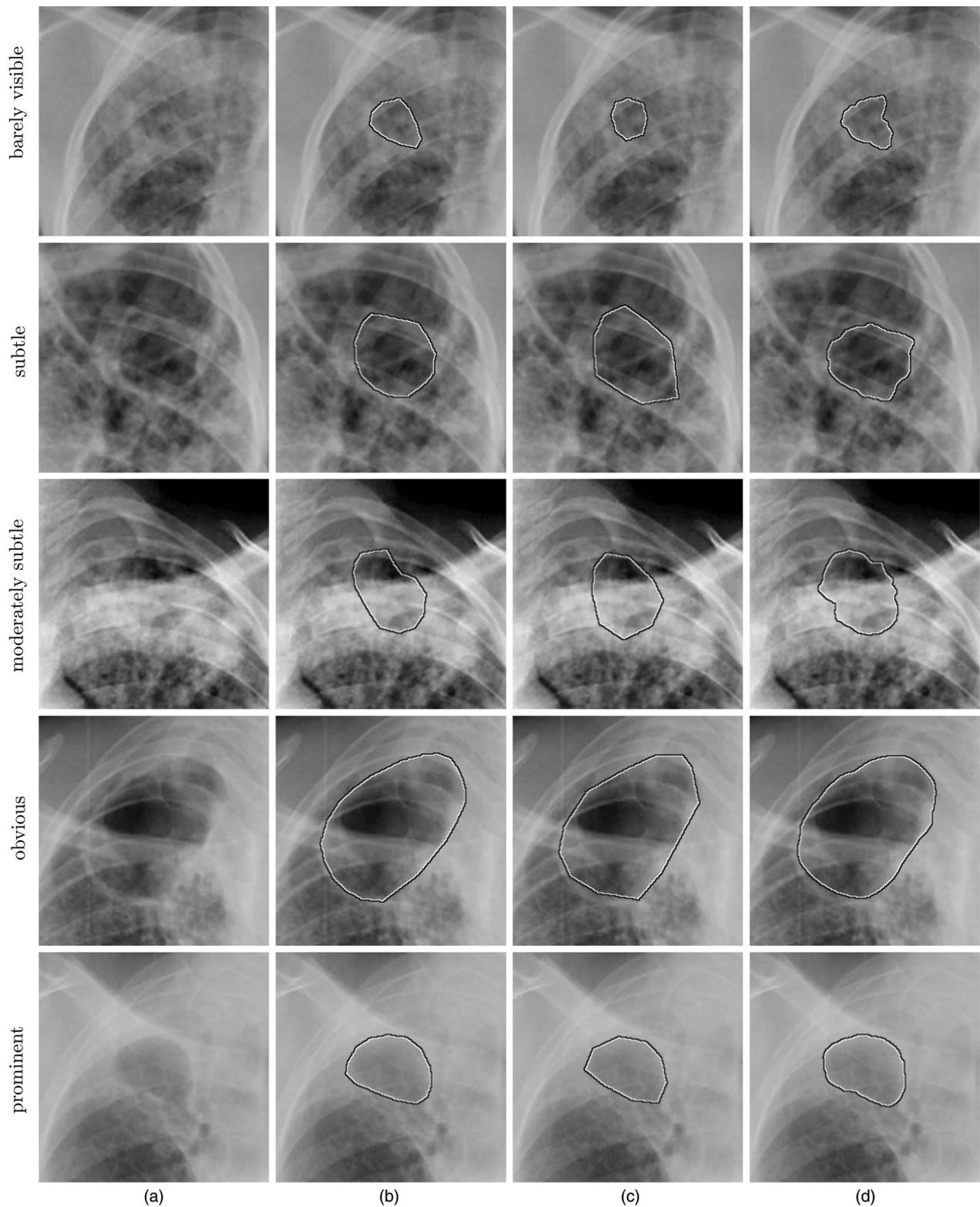


FIG. 9. Cavity segmentation results for various subtlety categories. (a) Cavity image. (b) Reference segmentation. (c) Segmentation by the independent observer. (d) Automatic segmentation using  $RF-F_{All}$  classifier.

radiologist and the independent observer, as shown in Fig. 9. The results were visually validated by two trained readers who accepted the automatic segmentation in 84% of the cavities (Table IV). The segmentation accuracy has a few outliers which is the cause of lower overlap statistics when

compared with the overlap between the two radiologists. Figure 9 shows this variation in segmentation for “barely visible” and “subtle” cavities. In the “barely visible” cavity example, the cavity borders are obscured by overlapping diseased lung parenchyma which makes delineation of the cavity

TABLE IV. Agreement between the ratings by two CRRS readers for automatically segmented cavities constructed using  $RF-F_{All}$  classifier.

CRRS Reader1	CRRS Reader2			All
	Excellent	Adequate	Insufficient	
Excellent	67	5	0	72
Adequate	23	11	8	42
Insufficient	3	1	8	12
All	93	17	16	126

very difficult. The example subimage of “subtle” cavity misses part of the cavity contour and as a result, there is no response in the cavity border likelihood map for that contour segment. Hence, the reference, the independent observer and the automatic segmentation have different segmentations in that part of the cavity (Fig. 9, “subtle” cavity). For “obvious” and “prominent” cavities, the segmentations were visually very similar to the radiologists’ annotations.

The proposed method requires a seed point as a user input to be close to the center of the cavity. Hence, the accuracy of the method might be affected by placement of the seed point based on the user’s judgment. We tested the implication of the location of the seed point by varying it in a circular neighborhood of the predefined seed point. Results in Fig. 10 show that the variation in distance up to 8 pixels (2 mm) has no effect on the performance and the performance begins to degrade when the distance was increased to 16 pixels (4 mm) and onwards. This can be explained by the fact that the seed point starts moving away from the center at larger displacements and gets closer to the cavity boundary which would result in wrong cavity subimage and subsequently wrong cost image in the polar space. Hence, minor displacement of the seed point would have no implications on the segmentation accuracy. Although the allowed freedom of displacement would largely depend on the size of the cavity, for example, larger displacement of seed point from the center for big cavities can still yield the same performance.

If we compare our results with the state of the art,<sup>14–16</sup> we achieved higher overlap on a substantially larger number of CXRs (100) containing 126 cavities than what has been reported in the literature (on 50 and 20 cavities). Analysis of the manual segmentations revealed that there were multiple cases where the human readers exhibited large disagreement. A few examples can be seen in Fig. 11. This indicates that the delineation of cavities is a very difficult task. The appearance of the cavities differs with the disease type and its progression.

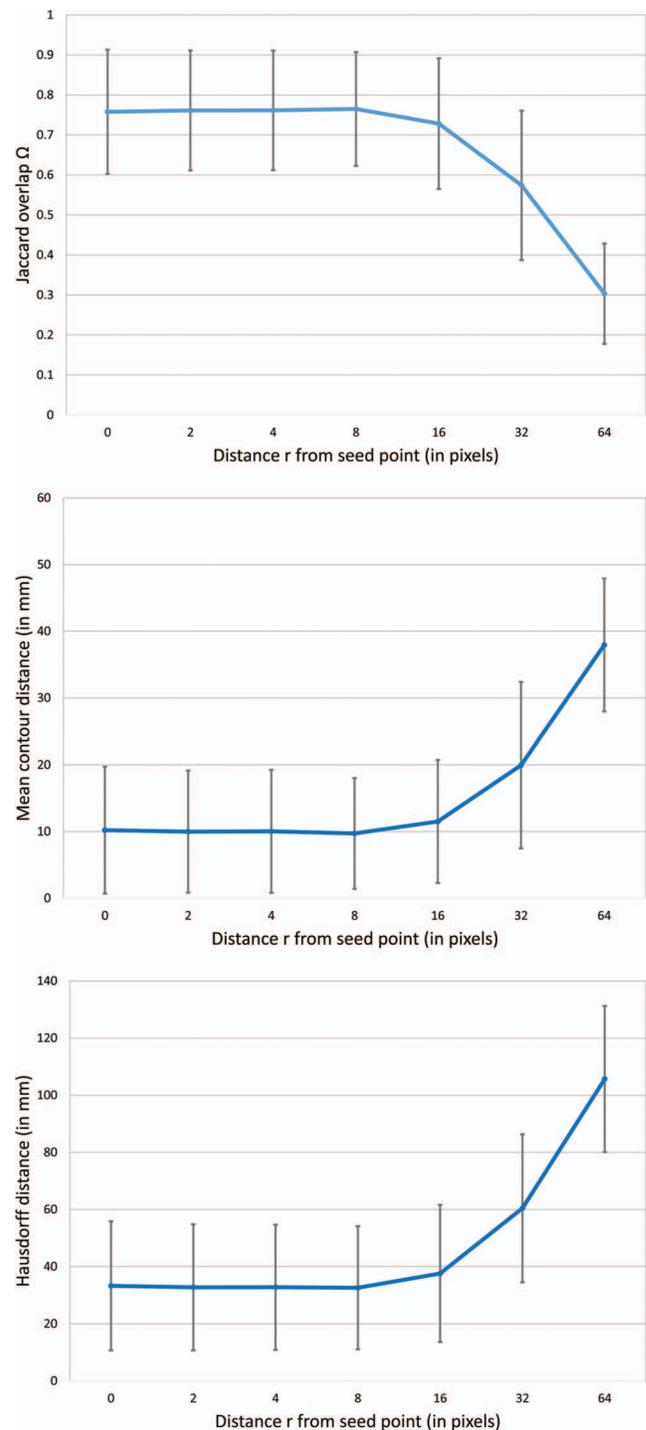


FIG. 10. Effect of seed point variation on various performance measures. The radius  $r$  is in pixels with a pixel size of 0.25 mm. The automatic segmentation was constructed using  $RF-F_{All}$  classifier for the cavity border likelihood map.

TABLE V. Classification of the cavities with inaccurate cavity segmentation categorized on the cause of failure.

Failed cases	Clavicle/rib border	Very subtle	Diseased region	Big cavity (>7 cm)	Overlapping cavities
Insufficient by one	4	2	1	1	0
Insufficient by both	3	3	3	2	1
All	7(35%)	5(25%)	4(20%)	3(15%)	1(5%)

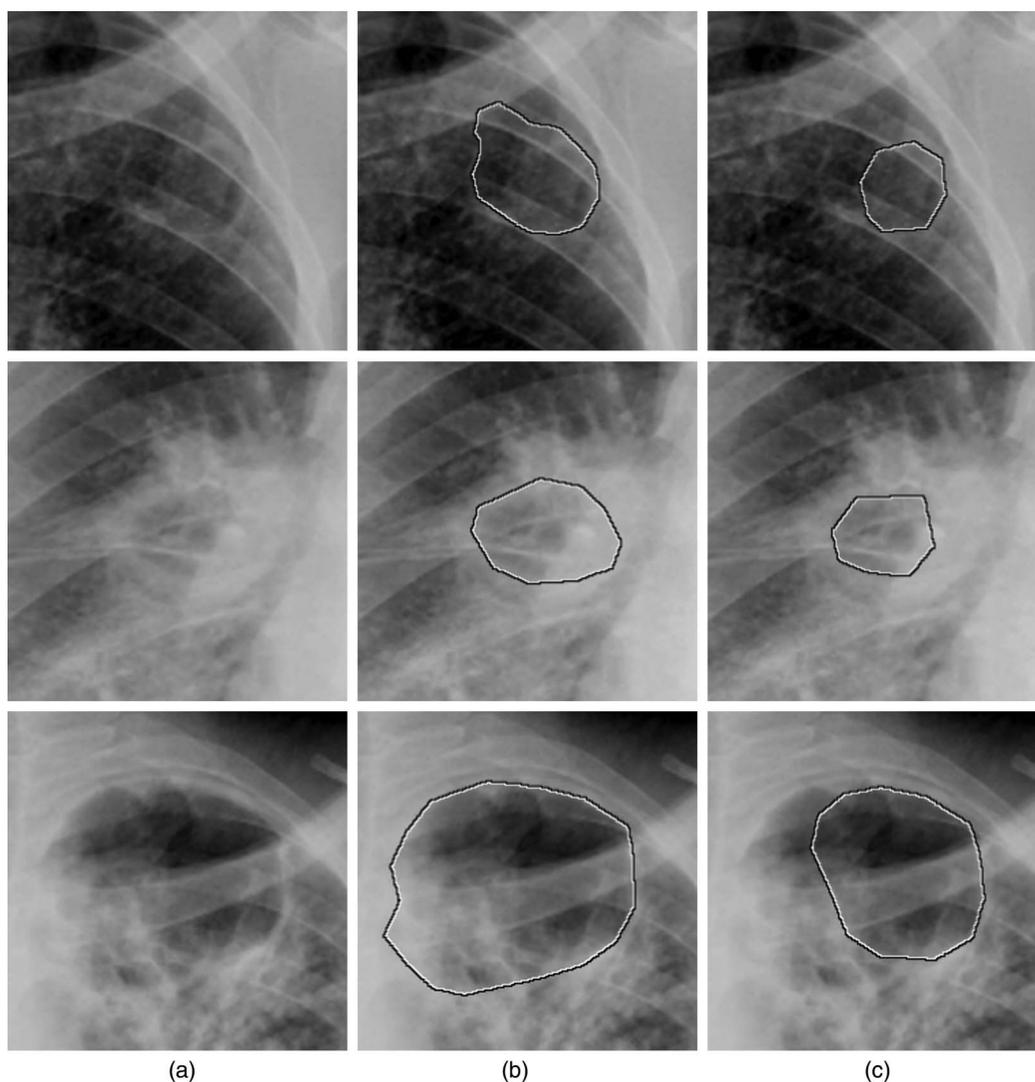


FIG. 11. Example cavities with disagreement in segmentation between the manual segmentations. (a) Cavity image. (b) Reference segmentation. (c) Segmentation by the independent observer.

In case of severe disease, part of the cavity borders might be missing or overlapping with the diseased lung tissue as shown in Figs. 9, 11, and 12.

The automatic technique for cavity segmentation can be a very useful tool at a clinic for treatment monitoring of TB. The prevalence of TB is high in resource constrained countries and there such a tool can be beneficial for a clinical officer analyzing the CXRs. This argument can be testified based on the assessment of the segmentations by two CRRS readers who were both satisfied with the automatic segmentation in 84% of the cavities. An advantage of our proposed automated technique is that it provides precise locations for all points of the border, whereas a human reader only annotates a few points on the cavity border. Dynamic programming determines all the connected points in the cost image to find the maximum cost path, making the automatic technique more accurate and robust.

The proposed automatic segmentation method has a few limitations. Detailed review of the 16% of the cavities rated as “insufficient” by any of the trained readers revealed the cause

of low performance. It can be concluded from Table V that the major reasons for the failure were: (1) false contour on the rib/clavicle borders, (2) very subtle cavities, (3) diseased regions, and (4) big cavities with a diameter  $>7$  cm. Example images with incorrect cavity segmentation are shown in Fig. 12. Rib and clavicle borders close to the cavity might get a high response after applying the cavity border classifier attracting dynamic programming to rib borders [Fig. 12(a)]. This shortcoming could potentially be alleviated by introducing a bony structure suppression technique,<sup>26,35–39</sup> prior to the execution of the cavity segmentation method. When cavities are present in the diseased region, it is very hard to define and identify the correct cavity contour [Fig. 12(b)]. The technique also sometimes undersegments cavities which are elliptical and very large in size [Fig. 12(c)]. The dynamic programming path can be calculated more precisely if a few reference points on the contour are clicked and the path is forced to pass through those points. Providing more than one reference point can be useful for subtle cavities for precise boundary segmentation.

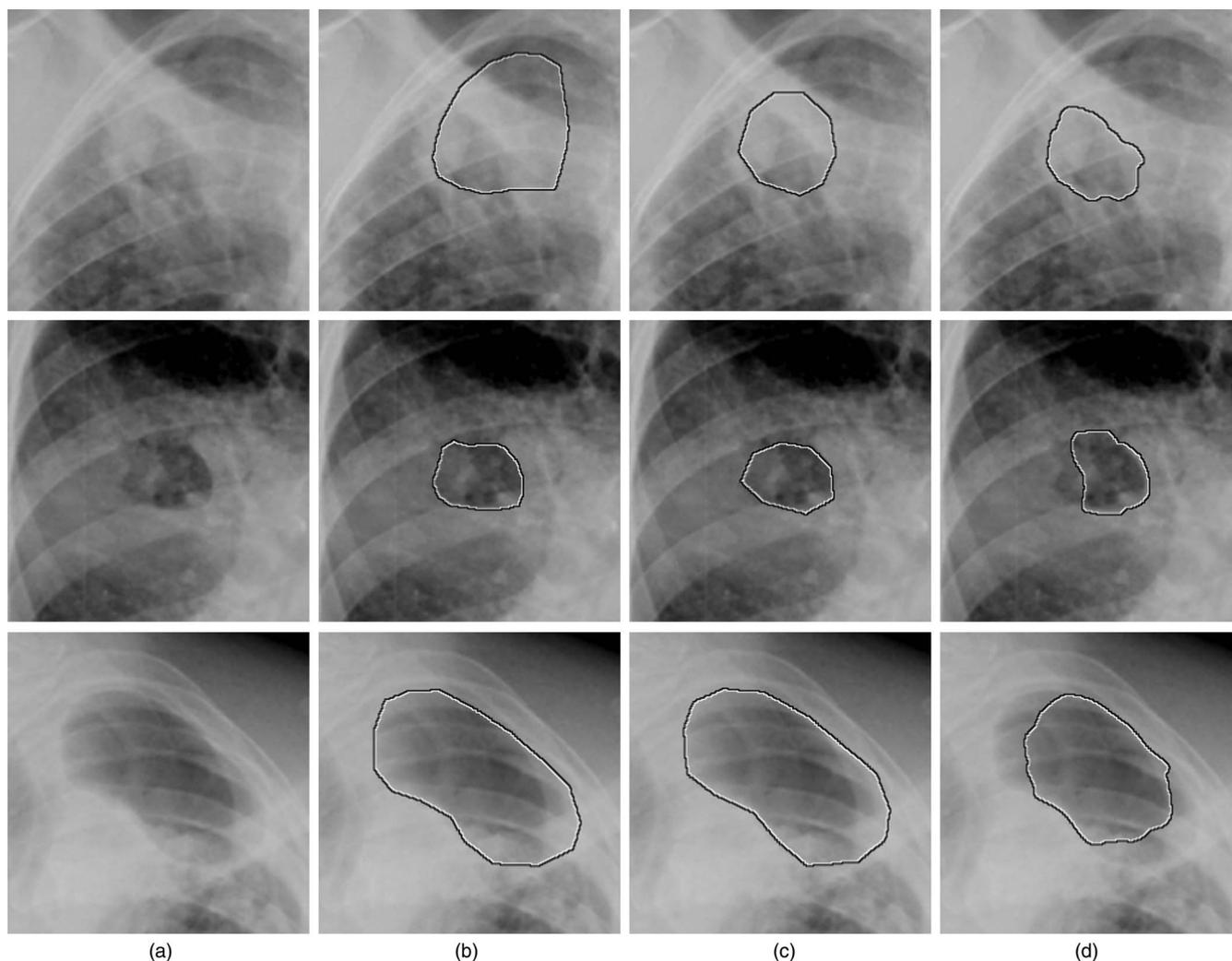


FIG. 12. Examples of inaccurate segmentations calculated by the proposed automatic cavity segmentation method: (a) Cavity image. (b) Reference segmentation. (c) Segmentation by the independent observer. (d) Automatic segmentation using  $RF-F_{All}$  classifier.

This work focuses only on the task of cavity segmentation and not on detection, which is clinically relevant for TB diagnosis. Future research could extend the presented technique to a fully automatic cavity detection system which can be added as a subsystem to an automatic tuberculosis detection system.<sup>40</sup>

## 7. CONCLUSION

An automatic cavity segmentation on chest radiographs was presented which displayed promising results, approaching the agreement between the segmentation assessed by the reference radiologist and the independent observer. Segmented cavities with a low agreement with the reference segmentation often contained subtle cavities or cavities in diseased regions. The proposed automatic cavity segmentation method requires minimal user input and can be potentially used in TB clinics for TB treatment monitoring and/or assessment of disease prognosis, especially in a resource limited setting where TB is highly prevalent but radiologists are generally not available.

## ACKNOWLEDGMENTS

This study was supported by the European and Developing Countries Clinical Trials Partnership (EDCTP), the Evaluation of multiple novel and emerging technologies for TB diagnosis, in smear-negative and HIV-infected persons, in high burden countries (TB-NEAT) project.

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: pragnya.maduskar@radboudumc.nl; Telephone: +31 24 3652272; Fax: +31 24 3540866.

<sup>1</sup> R. R. Gill, S. Matsusoka, and H. Hatabu, "Cavities in the lung in oncology patients: Imaging overview and differential diagnoses," *Appl. Radiol.* **39**(6), 10–21 (2010).

<sup>2</sup> World Health Organization, "Systematic screening for active tuberculosis: Principles and recommendations," World Health Organization, 2013.

<sup>3</sup> A. P. Ralph, M. Ardian, A. Wiguna, G. P. Maguire, N. G. Becker, G. Drogumuller, M. J. Wilks, G. Waramori, E. Tjitra, Sandjaja, E. Kenagalem, G. J. Pontororing, N. M. Anstey, and P. M. Kelly, "A simple, valid, numerical score for grading chest x-ray severity in adult smear-positive pulmonary tuberculosis," *Thorax* **65**, 863–869 (2010).

<sup>4</sup> M. Roy and S. Ellis, "Radiological diagnosis and follow-up of pulmonary tuberculosis," *Postgrad. Med. J.* **86**, 663–674 (2010).

- <sup>5</sup>C. D. Hamilton, J. E. Stout, P. C. Goodman, A. Mosher, R. Menzies, N. W. Schluger, A. Khan, J. L. Johnson, A. N. Vernon, and Tuberculosis Trials Consortium, "The value of end-of-treatment chest radiograph in predicting pulmonary tuberculosis relapse," *Intl. J. Tuberc. Lung Dis.* **12**, 1059–1064 (2008).
- <sup>6</sup>W. Koh, Y. J. Jeong, J. Kwon, H. J. Kim, E. H. Cho, W. J. Lew, and K. S. Lee, "Chest radiographic findings in primary pulmonary tuberculosis: Observations from high school outbreaks," *Korean J. Radiol.* **11**, 612–617 (2010).
- <sup>7</sup>D. Tatar, G. Senol, S. Alptekin, C. Karakurum, M. Aydin, and I. Coskunol, "Tuberculosis in diabetics: Features in an endemic area," *Jpn. J. Infect. Dis.* **62**, 423–427 (2009).
- <sup>8</sup>F. M. R. Perrin, N. Woodward, P. P. J. Phillips, T. D. McHugh, A. J. Nunn, M. C. I. Lipman, and S. H. Gillespie, "Radiological cavitation, sputum mycobacterial load and treatment response in pulmonary tuberculosis," *Intl. J. Tuberc. Lung Dis.* **14**, 1596–1602 (2010).
- <sup>9</sup>M. Palaci, R. Dietze, D. J. Hadad, F. K. C. Ribeiro, R. L. Peres, S. A. Vinhas, E. L. N. Maciel, V. do Valle Dettoni, L. Horter, W. H. Boom, J. L. Johnson, and K. D. Eisenach, "Cavitary disease and quantitative sputum bacillary load in cases of pulmonary tuberculosis," *J. Clin. Microbiol.* **45**, 4064–4066 (2007).
- <sup>10</sup>R. Dawson, P. Masuka, D. J. Edwards, E. D. Bateman, L.-G. Bekker, R. Wood, and S. D. Lawn, "Chest radiograph reading and recording system: Evaluation for tuberculosis screening in patients with advanced HIV," *Intl. J. Tuberc. Lung Dis.* **14**, 52–58 (2010).
- <sup>11</sup>L. M. Pinto, K. Dheda, G. Theron, B. Allwood, G. Calligaro, R. van Zyl-Smit, J. Peter, K. Schwartzman, D. Menzies, E. Bateman, M. Pai, and R. Dawson, "Development of a simple reliable radiographic scoring system to aid the diagnosis of pulmonary tuberculosis," *PLoS One* **8**, e54235 (2013).
- <sup>12</sup>R. Shen, I. Cheng, and A. Basu, "A hybrid knowledge-guided detection technique for screening of infectious pulmonary tuberculosis from chest radiographs," *IEEE Trans. Biomed. Eng.* **57**, 2646–2656 (2010).
- <sup>13</sup>W. R. Crum, O. Camara, and D. L. G. Hill, "Generalized overlap measures for evaluation and validation in medical image analysis," *IEEE Trans. Med. Imaging* **25**, 1451–1461 (2006).
- <sup>14</sup>T. Xu, I. Cheng, R. Long, and M. Mandal, "Novel coarse-to-fine dual scale technique for tuberculosis cavity detection in chest radiographs," *EURASIP J. Image Video Process.* **3**, 1–18 (2013).
- <sup>15</sup>T. Xu, I. Cheng, and M. Mandal, "Automated cavity detection of infectious pulmonary tuberculosis in chest radiographs," in *Proceedings of the Annual International Conference of the IEEE on Engineering in Medicine and Biology Society, EMBC, 2011* (IEEE, New York, 2011), pp. 5178–5181.
- <sup>16</sup>T. Xu, I. Cheng, and M. Mandal, "An improved fluid vector flow for cavity segmentation in chest radiographs," in *Proceedings of the International Conference on Pattern Recognition* (IEEE, New York, 2010), pp. 3376–3379.
- <sup>17</sup>M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Intl. J. Comput. Vis.* **1**, 321–331 (1988).
- <sup>18</sup>W. J. Niessen, B. M. ter Haar Romeny, and M. A. Viergever, "Geodesic deformable models for medical image analysis," *IEEE Trans. Med. Imaging* **17**, 634–641 (1998).
- <sup>19</sup>T. F. Cootes, C. J. Taylor, D. Cooper, and J. Graham, "Active shape models—Their training and application," *Comput. Vision Image Understanding* **61**, 38–59 (1995).
- <sup>20</sup>Y. Boykov and G. Funka-Lea, "Graph cuts and efficient N-D image segmentation," *Intl. J. Comput. Vision* **70**, 109–131 (2006).
- <sup>21</sup>K. Suzuki, "A review of computer-aided diagnosis in thoracic and colonic imaging," *Quant. Imaging Med. Surg.* **2**, 163–176 (2012).
- <sup>22</sup>J. Ma, L. Lu, Y. Zhan, X. Zhou, M. Salganicoff, and A. Krishnan, "Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine deformable model," *Med. Image Comput. Comput.-Assist. Interv.* **6361**, 19–27 (2010).
- <sup>23</sup>P. Dollár, T. Zhuowen, and S. Belongie, "Supervised learning of edges and object boundaries," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006* (IEEE, New York, 2006), Vol. 2, pp. 1964–1971.
- <sup>24</sup>K. Suzuki, I. Horiba, N. Sugie, and M. Nanki, "Extraction of left ventricular contours from left ventriculograms by means of a neural edge detector," *IEEE Trans. Med. Imaging* **23**, 330–339 (2004).
- <sup>25</sup>A. A. Amiri, T. E. Weymouth, and R. C. Jain, "Using dynamic programming for solving variational problems in vision," *IEEE Trans. Pattern Anal. Mach. Intel.* **12**, 855–867 (1990).
- <sup>26</sup>L. Hogeweg, C. I. Sánchez, P. A. de Jong, P. Maduskar, and B. van Ginneken, "Clavicle segmentation in chest radiographs," *Med. Image Anal.* **16**, 1490–1502 (2012).
- <sup>27</sup>X. Wu and M. Merckel, "Simultaneous border segmentation of doughnut-shaped objects in medical images," *J. Graph Algor. Appl.* **11**, 215–237 (2007).
- <sup>28</sup>L. M. J. Florack, B. M. ter Haar Romeny, M. A. Viergever, and J. J. Koenderink, "The Gaussian scale-space paradigm and the multiscale local jet," *Intl. J. Comput. Vis.* **18**, 61–75 (1996).
- <sup>29</sup>A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," *Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science Vol. 1496* (Springer Berlin Heidelberg, Berlin, 1998), pp. 130–137.
- <sup>30</sup>R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. (Wiley, New York, 2001).
- <sup>31</sup>S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching in fixed dimensions," *J. ACM* **45**, 891–923 (1998).
- <sup>32</sup>J. Friedman, T. Hastie, and R. Tibshirani, "Special invited paper. Additive logistic regression: A statistical view of boosting," *Ann. Statist.* **28**, 337–374 (2000).
- <sup>33</sup>L. Breiman, "Random forests," *Mach. Learn.* **45**, 5–32 (2001).
- <sup>34</sup>G. Gerig, M. Jomier, and M. Chakos, "Valmet: A new validation tool for assessing and improving 3D object segmentation," *Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science Vol. 2208* (Springer Berlin Heidelberg, Berlin, 2001), pp. 516–523.
- <sup>35</sup>S. Chen and K. Suzuki, "Separation of bones from chest radiographs by means of anatomically specific multiple massive-training ANNs combined with total variation minimization smoothing," *IEEE Trans. Med. Imaging* **33**(2), 246–257 (2014).
- <sup>36</sup>L. Hogeweg, C. I. Sánchez, and B. van Ginneken, "Suppression of translucent elongated structures: Applications in chest radiography," *IEEE Trans. Med. Imaging* **32**, 2099–2113 (2013).
- <sup>37</sup>S. Chen and K. Suzuki, "Computerized detection of lung nodules by means of "virtual dual-energy" radiography," *IEEE Trans. Biomed. Eng.* **60**, 369–378 (2013).
- <sup>38</sup>P. Maduskar, L. Hogeweg, R. Philipsen, S. Schalekamp, and B. van Ginneken, "Improved texture analysis for automatic detection of tuberculosis (TB) on chest radiographs with bone suppression images," *Proc. SPIE* **8670**, 86700H (2013).
- <sup>39</sup>K. Suzuki, H. Abe, H. MacMahon, and K. Doi, "Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN)," *IEEE Trans. Med. Imaging* **25**, 406–416 (2006).
- <sup>40</sup>L. Hogeweg, C. Mol, P. A. de Jong, R. Dawson, H. Ayles, and B. van Ginneken, "Fusion of local and global detection systems to detect tuberculosis in chest radiographs," *Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science Vol. 6363* (Springer Berlin Heidelberg, Berlin, 2010), pp. 650–657.