This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2687939, IEEE Journal of Biomedical and Health Informatics

Lung Field Segmentation in Chest Radiographs from Boundary Maps by a Structured Edge Detector

Wei Yang, Yunbi Liu, Liyan Lin, Zhaoqiang Yun, Zhentai Lu,
Qianjin Feng*, *Member, IEEE*, and Wufan Chen, *Senior Member, IEEE*

***Abstract*—Lung field segmentation in chest radiographs (CXRs) is an essential preprocessing step in automatically analyzing such images. We present a method for lung field segmentation that is built on a high-quality boundary map detected by an efficient modern boundary detector, namely, a structured edge detector (SED). A SED is trained beforehand to detect lung boundaries in CXRs with manually outlined lung fields. Then, an ultrametric contour map (UCM) is transformed from the masked and marked boundary map. Finally, the contours with the highest confidence level in the UCM are extracted as lung contours. Our method is evaluated using the public JSRT database of scanned films. The average Jaccard index of our method is 95.2%, which is comparable with those of other state-of-the-art methods (95.4%). The computation time of our method is less than 0.1 s for a 256 × 256 CXR when executed on an ordinary laptop. Our method is also validated on CXRs acquired with different digital radiography units. The results demonstrate the generalization of the trained SED model and the usefulness of our method.**

***Index Terms*—chest radiography, lung field segmentation, boundary detection, structured edge detector**

I. Introduction

CHEST radiography (chest X-ray) is a diagnostic imaging technique widely used for lung diseases. The automatic segmentation of lung fields has received considerable attention from researchers as an essential preprocessing step in automatically analyzing chest radiographs (CXRs) [1-7]. An accurate automatic segmentation of lung fields can save physicians' efforts for manual identification of the lung anatomy. In addition, this process is a necessary component of a computer-aided diagnosis system for detecting lung nodules [8]. The segmentation of lung fields is also useful for the anatomic region-based processing of CXRs, such as contrast enhancement of lung regions and bone suppression [10].

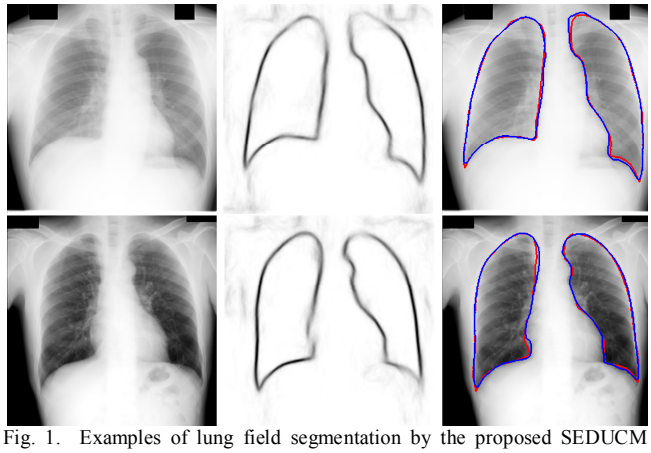


Fig. 1. Examples of lung field segmentation by the proposed SEDUCM method. From left to right: CXRs, corresponding boundary maps produced by a trained SED, and segmentation results. Red and blue contours indicate the ground truth and automatic segmentation results, respectively.

However, an accurate segmentation of lung fields in CXRs remains a challenge for several reasons. Lung fields exhibit large anatomical shape variations, including varying heart dimensions or other pathologies, across different patients in 2D radiographs. Lung fields in CXRs also contain several superimposed structures, such as lung vasculatures, clavicles, and ribs, which do not form the borders of lung fields. The strong edges at the rib and clavicle regions may result in inaccurate location of landmarks or inaccurate lung contours in some lung field segmentation approaches. In addition, segmenting the lung apex is difficult because of the varying intensities in the upper clavicle bone region.

Many lung field segmentation methods have been proposed for posterior–anterior (PA) CXRs to address these difficulties. These methods can be roughly divided into five categories: (1) rule-based methods, (2) pixel classification (PC)-based methods, (3) shape model-based methods, (4) hybrid methods, and (5) atlas-based methods. Rule-based segmentation methods [2, 3] contain sequences of steps and rules, such as thresholding or morphological operations. These methods have heuristic assumptions and compute approximate solutions below the global optimum. PC-based methods identify lung field segmentation as a classification problem and thus acquire a classifier to label each pixel as lung or background [4]. Most

This work was supported by the grants from National Natural Science Foundation of China (no. 61471187), Natural Science Foundation of Guangdong Province (no. 2015A030313280), and Guangdong Provincial Key Laboratory of Medical Image Processing (no. 2014B030301042).

W. Yang, Y. Liu, L. Lin, Z. Yun, Z. Lu, and W. Chen are with the Guangdong Provincial Key Laboratory of Medical Image Processing, School of Biomedical Engineering, Southern Medical University, Guangzhou, China (e-mail: weiyanggm@gmail.com; lly_neverchange@163.com; ybliu1994@gmail.com; yun1982@smu.edu.cn; luzhentai@163.com; wufanchen@gmail.com).

*Q. Feng is with the Guangdong Provincial Key Laboratory of Medical Image Processing, School of Biomedical Engineering, Southern Medical University, Guangzhou, China (correspondence e-mail: qianjinfeng08@gmail.com).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2687939, IEEE Journal of Biomedical and Health Informatics

classification errors on the pixels around the boundaries of lung fields lead to inaccurate locations. Active shape models (ASM) and active appearance models (AAM) can incorporate low-level appearance cues and high-level shape priors, and have been successfully applied to lung field segmentation [4, 6]. In general, shape model-based methods tend to produce average shapes and are ineffective with abnormal cases. The segmentation performance of shape models relies on the approximation accuracy of the initial model. Hybrid methods produce improved results by fusing several techniques, but the segmentation algorithm is sophisticated and time consuming [4]. A recent study has introduced an atlas-based method that exhibits state-of-the-art performance; in this method, the CXR database of pre-segmented lung fields is used as the anatomical atlas, and the SIFT Flow algorithm is employed to align the CXR with the atlas [1]. In general, atlas-based methods are very time consuming. Lung segmentation can be refined through post-processing typically by using graph cuts [1, 12]. Among the energy functions for graph cuts, the boundary term is critical to improve segmentation accuracy. An accurate detection of lung boundaries is crucial to realize an accurate and simple automatic segmentation of lung fields.

However, lung boundaries are not always located on well-defined edges, where the gradient magnitude is maximum along the gradient direction. Simple gradients or derivatives of CXRs are insufficient for handling many anatomical structures and textures. Hence, an accurate detection of lung boundaries in CXRs is traditionally considered to be highly difficult. The classical Canny edge detector [13] and other edge detection methods based on image derivatives from CXRs can detect the edges not only along the borders of the lung but also along the borders of other anatomical structures such as the ribs and clavicles, which are almost not close contours. From the edges detected by the Canny edge detector, the candidate segments of lung boundaries are selected by the sophisticated rule-based reasoning method as in Ref. [3]. Tsujii et al. [14] developed a supervised lung boundary detector that uses 1D convolution neural networks trained to classify the pixels in CXRs into lung boundaries or otherwise. However, the detected lung boundaries in [14] are still not continuous and cannot be transformed directly to the lung segmentation.

Modern boundary detectors, such as Pb [15], structured edge detector (SED) [16], DeepEdge [17], and HED [18], are different from the classical Canny edge detector because these detectors emphasize the importance of suppressing false edge responses through an explicit oriented analysis of higher-order statistics. These statistics are obtained in various ways, including supervised learning. Such boundary detectors can benefit from global normalization provided by graph-spectral analysis or ultrametric consistency. These analyses enforce closure, thereby boosting the contrast of contours that completely enclose salient regions [19]. Most modern boundary detectors can be trained and provide a feasible way to detect the boundaries of particular objects. Among these modern boundary detectors, SED emerges as a distinguished system for edge detection because of its state-of-the-art performance and high speed [16].

In the present work, we aim to develop an accurate method for the real-time segmentation of lung fields in standard PA CXRs for practical applications. Unlike previous methods that use PC or shape models, we initially detect lung boundaries and then produce segmentation results from the detected boundary map. We select SED, which can be trained on samples of manually outlined lung fields, to detect lung boundaries efficiently. Lung contours are then extracted from an ultrametric contour map (UCM) [20], which is transformed from the boundary map detected by a trained SED and marker-controlled watershed transform (MWT) [21]. Our proposed method for lung field segmentation is called SEDUCM. In the SEDUCM segmentation pipeline, PC and initialization of shape models are not necessary. Fig. 1 shows two examples of the boundary maps detected by the trained SED and the segmentation results of lung fields through SEDUCM.

The remainder of this paper is organized as follows. The framework and details of our method are described in Section 2. Experimental results are provided in Section 3. The summary and discussion of results are shown in Section 4.

II. Methods

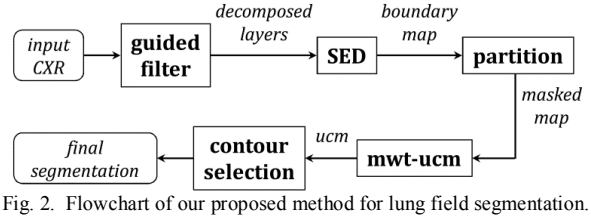


Fig. 2. Flowchart of our proposed method for lung field segmentation.

A. Overview

This work aims to develop a practical and useful method for automatically segmenting lung fields in CXRs. The core of our proposed method is the effective use of the lung boundary map produced by SED. As shown in Fig. 2, an input CXR was first normalized into the intensity range [0, 1] and decomposed as the input of SED to the base and detail layers by a guided filter [22]. Next, a boundary map was produced by the SED model trained for detecting the boundaries of lung fields. From the boundary map and the input CXR, the ribcage and spinal centerline were extracted. These segments were used to partition the CXR into the right and left thorax areas as well as clean the boundary map for further processing. Subsequently, the candidate lung regions and contours were generated by using MWT and UCM transforms (*mwt-ucm*). Finally, the contours with the highest confidence level were selected as the right and left lung contours. To effectively perform segmentation, each step of the proposed method employed highly efficient algorithms for executing the corresponding functions, including guided filter [22], dynamic programming, and watershed transform (WT).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2687939, IEEE Journal of Biomedical and Health Informatics

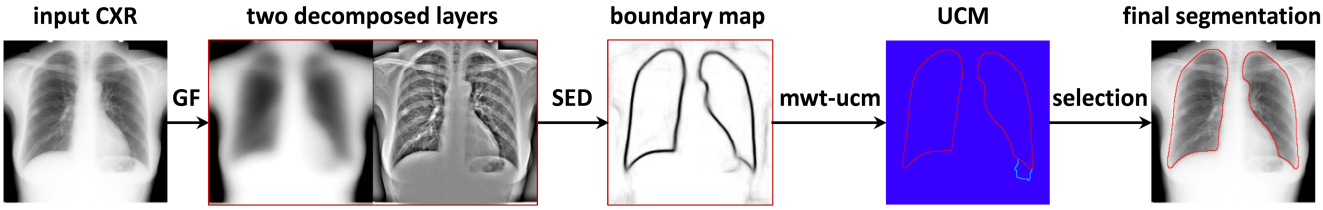


Fig. 3. Intermediate results of an input CXR in SEDUCM pipeline.

B. SED for Detecting Lung Boundaries

We first reviewed the SED proposed by Dollár and Zitnic [16]. Dollár and Zitnic formulated the edge detection task in a general structured learning framework where a random decision forest [23, 24] is exploited to general structured output spaces. SED has the advantage of the inherent structure in edge patches and can be computed efficiently.

A decision tree $f_t(x)$ classifies an input $x \in \mathcal{X}$ by splitting the data between the left and right sub-trees according to a binary split function $h(x, \theta_j)$ with parameter θ_j at each node j. Given a node j and a training set $\mathbf{S} \subset \mathcal{X} \times \mathcal{Y}$, the training goal of the decision tree is to find parameter θ_j that maximizes the information gain criterion I_j defined by

$$I_j = I(\mathbf{S}_j, \mathbf{S}_{jL}, \mathbf{S}_{jR}), \quad (1)$$

At each internal node of the tree, a feature is chosen to split the incoming training samples to maximize some criteria. A random decision forest comprises multiple independent decision trees [23, 24]. Given a sample, the predictions from the set of decision trees are combined into a single output by using an ensemble model.

Dollár et al. [16] extended random forests to structured random forests for predicting structured outputs. Given an image patch $x \in \mathcal{X}$, the output $y \in \mathcal{Y}$ stores the corresponding segmentation mask or binary edge map. A segmentation mask is denoted by $y \in \mathcal{Y} = \mathbb{Z}^{d \times d}$ and a binary edge map is represented by $y' \in \mathcal{Y} = \{0, 1\}^{d \times d}$, where d is the patch width. The main goal of structured random forests is to map all structured labels to a discrete set $c \in \mathcal{C}$. Dollár et al. solved this problem by first mapping the structured output space $\mathcal{Y}$ to an intermediate space $\mathcal{Z}$. $z = \Pi(y)$ is defined as a long binary vector that encodes whether each pair of pixels in y belongs to the same or different segments. The problem with the high dimensionality of the structured output space $\mathcal{Y}$ is alleviated by sampling a few dimensions of $\mathcal{Z}$ followed by conducting principal component analysis (PCA). Then, the intermediate space $\mathcal{Z}$ is mapped to the discrete label space $\mathcal{C}$ by PCA quantization. Thus, the standard information gain criteria based on Gini impurity can be adopted to train structured random forests. In a trained structured random forest, the learned edge masks y' are averaged as a soft edge response and stored at each left node. Although SED is originally designed to detect general edges on natural images, it can be adopted to detect the boundaries of particular objects, such as lung boundaries in CXRs.

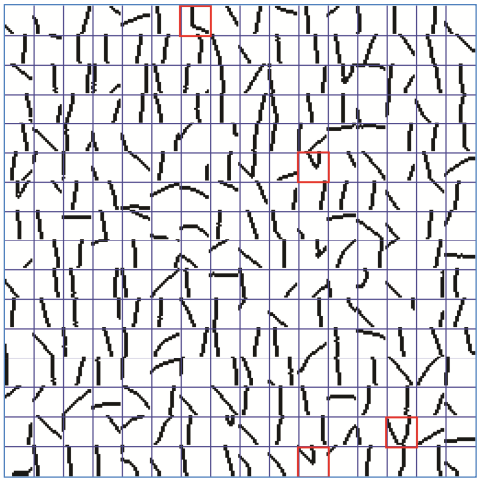
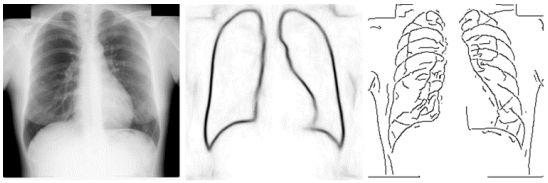


Fig. 4. Examples of boundary patches in the leaf nodes of the trained SED. The red rectangles indicate some representative boundary fragments which locate along the lower lobes of lung.

We began by decomposing an input CXR into its base and detail layers using a guided filter as shown in Fig. 3. The guided filter performs edge-preserving smoothing on an image like the bilateral filter, using the content of a guidance image, to influence the filtering. One advantage of the guided filter is that it can be implemented efficiently through integral image technique. We used the input CXR itself as the guidance image, and the kernel radius and the regularization coefficient of the guided filter were set to 8 and 0.1, respectively. The base layer is for extracting coarse-scale features of the input CXR, whereas the detail layer is for extracting fine-scale features. These two layers were normalized to zero mean and one variance as the input feature maps of SED. The feature extraction and the SED training approach presented in Ref. [16] were adopted in this work. The SED predicted the centered 16 × 16 lung boundary response from a 32 × 32 × 2 image patch. Each image patch was augmented to obtain 12 channels, including 2 input channels, 2 gradient magnitude channels, and 8 gradient orientation channels. A total of 6,672 candidate features were efficiently extracted from a 32 × 32 × 12 patch, similar to that in Ref. [16].

Fig. 4 shows 256 randomly selected patches of the lung boundary response in the leaf nodes of a learned structured random forest. Each patch of the boundary response exhibits the shape characteristics of particular locations along the lung boundaries. Compared with a global shape model, the patches of the boundary response can be viewed as local shape fragments and are more flexible to composite the lung contours.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2687939, IEEE Journal of Biomedical and Health Informatics



(a) CXR (b) boundary map by SED (c) edge map by Canny detector

Fig. 5. Example of boundary map produced with the trained SED and edge map detected by the Canny edge detector.

ALGORITHM I. PREDICTING BOUNDARY MAP

Input: $\boldsymbol{I}$, a CXR. T, the total tree number of the trained SED model. T_{eval}, the evaluation number of trees.

Output: $\boldsymbol{E}$, the soft boundary map.

1: Decompose $\boldsymbol{I}$ into the base layer $\boldsymbol{B}$ and the detail layer $\boldsymbol{I}_b$.
2: Sample densely the patches $\{\boldsymbol{p}_i\}$ of size 32×32×2 from $\boldsymbol{B}$ and $\boldsymbol{I}_b$ with stride 2.
3: **for** each patch $\boldsymbol{p}_i$ at location (r_i, c_i) **do**
4: compute the indices of trees $\{k$=mod(mod(r_i+c_i, 2×T_{eval})+t, T), t=0, 1, 2, ..., T_{eval}−1$\}$.
5: apply the subset $\{k\}$ of decision trees in the trained SED model to the patch $\boldsymbol{p}_i$ and predict the corresponding patch $\boldsymbol{b}_i$ of boundary response.
6: **end for**
7: aggregate and average the overlapped patches $\{\boldsymbol{b}_i\}$ of boundary response to yield the soft boundary map $\boldsymbol{E}$.

During the prediction stage, the image patches are sampled densely from the base and detail layers of an input CXR, and SED predicts their corresponding patches of boundary response. Then, the overlapped patches of boundary response were averaged and aggregated to yield a map of soft boundary response. The efficiency of SED can be further improved by reducing the number of densely sampled image patches and the number of decision trees evaluated during the prediction stage since both the inputs and outputs of each decision tree overlap. Similar to the settings in [16], the stride of image patch sampling was set to 2 pixels and an alternating subset of decision trees was evaluated on the sampled image patches at each adjacent location. The procedure of predicting boundary map for a CXR is described in Algorithm I.

Dollár et al. [16] proposed a sharpening procedure on the boundary map using local color or intensity cues. However, sharpening the boundary map would lead to less smooth lung contours and degrade the segmentation performance. Therefore, the lung boundary maps predicted by SED were not sharpened but rather slightly smoothed using a triangle filter with a kernel size of 1 pixel in our work. Examples of boundary maps detected by the trained SED are shown in Figs. 1 and 5. An edge map detected by the Canny edge detector for a CXR is shown in Fig. 5(c). The borders of the ribs and body were identified as the edges by the Canny edge detector. The trained SED can effectively distinguish the lung boundaries from the other structures. The response values along the lung contours are significantly larger than other regions in the boundary map produced by SED.

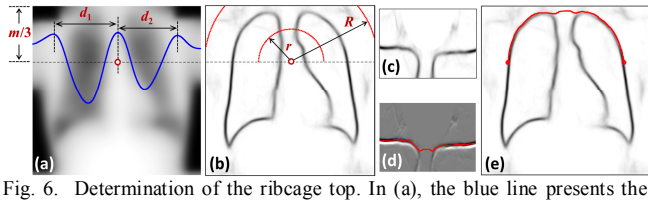


Fig. 6. Determination of the ribcage top. In (a), the blue line presents the horizontal intensity profile at 1/3 of image height of the base layer of a CXR, and the small red circle indicates the center point for polar transformation. The two dashed red semicircles in (b) delimit the sampling range of polar transform. (c) Polar transform result from (b). The optimal path overlaid on the energy function is displayed in (d). The red line is the detected boundary of lung top in Cartesian coordinates in (e).

C. Finding Thorax Centerline and Ribcage for Partition

Although high-quality lung boundary maps can be produced using the trained SED, false responses for lung boundary still exist, as shown in Figs. 1 and 5. Many irrelevant regions are evident among the regions generated by WT directly from the original boundary map. In particular, false strong boundary responses complicate the subsequent selection of correct regions as lung fields. We determined the thorax centerline and the ribcage boundary to filter out irrelevant boundary responses and delimit the search area for the right/left lung field. Consequently, the effect of false boundary responses was reduced. We defined specific energy functions as the thorax centerlines and the ribcage boundaries to search for minimum cost paths using dynamic programming.

We determined the thorax centerline in a CXR based on high intensity values and low boundary responses in spinal regions. This finding led to the following simple energy function of an input CXR $\boldsymbol{I}$:

$$e(\boldsymbol{I}) = 1 - \boldsymbol{I}_b + \boldsymbol{B}, \tag{2}$$

where $\boldsymbol{I}_b$ is the base layer of $\boldsymbol{I}$, and $\boldsymbol{B}$ is the boundary map of CXR $\boldsymbol{I}$. We defined the cost of an 8-connected path s as:

$$C(s) = \sum_{k=1}^{m} e(s_k) + v\,\ell(s), \tag{3}$$

where $\ell(s)$ denotes the length of s, v is a tunable parameter, and m is the row number of CXR $\boldsymbol{I}$. We searched for the optimal path s^* that minimizes this cost: $s^* = \min_s C(s)$. The second term $v\ell(s)$ in the cost $C(s)$ was used to render the optimal path less zigzag. v was set to 0.3 in the experiments. The optimal path can be found using dynamic programming. Traversing the energy map from the second row to the last row, we calculated the cumulative minimum energy C for all possible 8-connected paths for each entry (i, j):

$$C(i,j) = e(i,j) + \min\left(C(i-1,j-1) + \sqrt{2}v, C(i-1,j) + v, C(i,j-1) + v\right). \tag{4}$$

The minimum value of the last row in C indicates the end of minimum cost path. Then, we traced back from the minimum entry on C to find the optimal path as the thorax centerline. To reduce computation time and avoid finding unreasonable thorax centerlines, the search range of the optimal paths was limited in the area from $n/3$ to $2n/3$ columns in a CXR of $m \times n$ pixels similar to [25].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2687939, IEEE Journal of Biomedical and Health Informatics

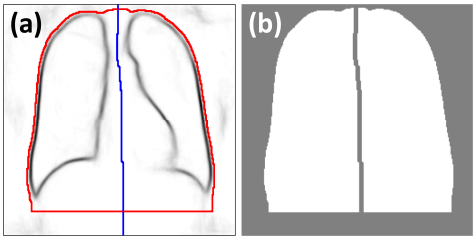


Fig. 7. Ribcage contour and thorax centerline (a), as well as the corresponding mask (b) for subsequent processing.

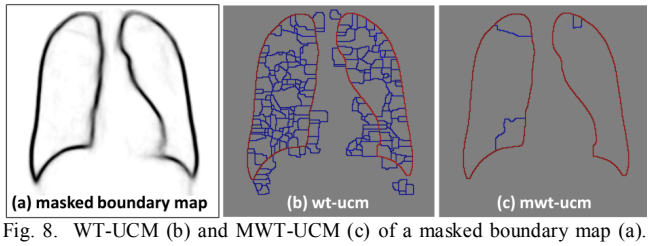


Fig. 8. WT-UCM (b) and MWT-UCM (c) of a masked boundary map (a). Red/blue color indicates high/low UCM value.

The produced lung boundary map provided sufficient cues to identify the ribcage. We first determined the top of the ribcage similar to [2]. Considering that the top of the ribcage was more or less circular, we realized the polar transformation of a hemicircle in the boundary map to 128 × 180 pixels for the CXRs of 256 × 256 pixels and then applied dynamic programming to find the optimal path. Fig. 6 shows the computation of the center point and the radius for polar transformation. The radius of the hemicircle was estimated from the intensity profile in the base layer at 1/3 height. We denoted the point with the maximum value on the intensity profile as the center point. Two peaks of the intensity profile on the sides of the center point indicated the rough positions of the right and left ribcages at 1/3 height. The maximum distances (d_1 and d_2 in Fig. 6(a)) from these two peak locations to the center point were computed as the hemicircle radius. The minimum radius r and the maximum radius R (Fig. 6(b)) were set as 0.5 and 1.4 times the radius for polar transformation, respectively. Fig. 6(c) shows the corresponding polar transformation of the hemicircle region in the boundary map. We regarded the first derivative along the circle radius of the polar transformation image as the energy function for the ribcage top. Similar to finding the thorax centerline, the optimal path was determined by traversing the energy map from the second column to the last column using dynamic programming (Fig. 6(d)).

To determine the left ribcage boundary, we used the end points of the top ribcage boundary as the starting points of the optimal paths. The optimal paths were searched in the rectangular regions. We used the first derivatives along the vertical direction of the boundary map to define the cost functions. For the left/right ribcage boundary, the cost function was regarded as the first forward/backward derivative of the boundary map. The minimum cost path was determined through dynamic programming. We also located the minimum peak on the bottom of vertical projection of boundary map to delimit the area for selection of the lung contours.

We divided each CXR into right, left, and irrelevant areas using the detected ribcage contour and thorax centerline, as shown in Fig. 7(b). To assure that the lung fields are included in the detected ribcage, we enlarged the ribcage using the morphologic dilatation operation as shown in Fig. 7(a). The responses in the irrelevant areas, i.e., gray regions in Fig. 7(b), of the boundary map were masked out and reset to 0.

D. mwt-ucm: From Boundary Map to Segmentation of Lung Fields

Once obtained, a boundary map can be utilized for segmentation through several means. However, lung boundary maps predicted by SED are diffused in certain areas. Retrieving segmentation masks from boundary maps is not a straightforward process. The edges produced using ordinary techniques, such as non-maximum suppression from a boundary map, may not be closed. Thus, the edges do not separate the image into regions. Another approach is to integrate a boundary map into the pixel or region probability for labeling in a maximum-a-posteriori framework. However, the trade-off between the region and boundary constraints produces smooth contours, which may not appear in real object boundaries. We opted to create candidate regions or segmentation masks directly from a boundary map by using the over-segmentation or superpixel approaches. WT can be implemented efficiently and effectively; thus, we applied WT to generate over-segmentation from the boundary map. The traditional WT typically produces an excessive number of small irrelevant regions. To reduce the number of such regions, we use MWT to produce over-segmentation regions from the boundary map. We set the pixels of boundary response values less than 0.01 as the markers. Each marker indicates a specific location within the boundary map which is modified by using the minima imposition technique [21]. The modified boundary map only has regional minima in the locations of the markers. Then, the traditional WT is applied to the modified boundary map to obtain segmentation in the belt between the markers. An example of the regions produced from a masked boundary map by WT and MWT is presented in Fig. 8.

We adopted the contour-based hierarchical segmentation method proposed by Arbelaez et al. [19, 20] to generate candidate segments of lung field from the MWT or WT of the masked boundary map. The result of this hierarchical segmentation method is a weighted contour image called UCM, the values of which reflect contour strength and the contrast between neighboring regions [19]. This method generally preserves the global contours of objects while providing hierarchical segments. Such segments are obtained using a greedy graph-based region merging algorithm. Let $G(P_0, K_0, W(K_0))$ denote an initial graph, where the nodes are the regions P_0 generated by WT or MWT, the links are the arcs K_0 separating adjacent regions, and the weights $W(K_0)$ are a measure of dissimilarity between two adjacent regions which is defined as the average boundary response of SED of their common boundary in K_0. The algorithm proceeds by iteratively merging the most similar regions, and produces a tree of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2687939, IEEE Journal of Biomedical and Health Informatics

regions, where the leaves are the initial elements of P_0, the root is the entire image, and the regions are ordered by the inclusion relation [19]. A real-valued image is obtained by weighting each boundary by its scale of disappearance as the UCM, which has the remarkable property of producing a set of closed curves for any threshold. Hierarchical segmentations can be created by setting the UCM thresholds. Fig. 8 presents two examples of UCM with different over-segmentation regions P_0.

SED was trained to detect lung boundaries; hence, lung contours are expected to have high UCM values. On the right/left partition of a CXR generated from the detected ribcage and thorax centerline, the contour with the highest UCM value was selected as the right/left lung contour.

E. Evaluation Metrics

We used three widely used metrics to evaluate our SEDUCM method quantitatively and compare it with other lung segmentation methods, namely, the Jaccard index (Ω), the Dice similarity coefficient (DSC) [26], and the mean boundary distance (MBD). These metrics were defined and computed as follows.

Let us denote S as the estimated segmentation mask and T as the ground truth mask. Jaccard index was computed as:

$$\Omega = \frac{|S \cap T|}{|S \cup T|},$$

where $|\cdot|$ is the cardinality of the set, $S \cap T$ is the intersection of S and T, and $S \cup T$ is the union of S and T. DSC is the overlap ratio between the ground truth mask T and the estimated segmentation mask S:

$$DSC = \frac{2 \times |S \cap T|}{|S| + |T|}.$$

MBD is the average distance between the estimated segmentation boundary S and the ground truth boundary T. Let s_i and t_j be the points on boundaries S and T, respectively. The minimum distance of point s_i on S to boundary T was computed as:

$$d(a_i, T) = \min_j \left\| a_i - t_j \right\|.$$

For MBD computation, the minimum distance for each point on boundary S to boundary T was calculated, and vice versa. These minimum distances were averaged as MBD:

$$MBD(S,T) = \frac{1}{2}\left(\frac{\sum_i d(s_i, T)}{|\{s_i\}|} + \frac{\sum_j d(t_j, S)}{|\{t_j\}|} \right).$$

III. EXPERIMENTS

A. Experimental Datasets and Settings

The proposed SEDUCM method was evaluated on three CXR datasets. The first dataset is the publicly available Japanese Society of Radiological Technology (JSRT) dataset [27]. The JSRT dataset consists of 247 standard PA CXRs that are scanned from plain film radiographs to a size of 2048 × 2048 pixels, with a spatial resolution of 0.175 mm and 12 bit gray levels. The manual segmentation of lung fields for CXR in the JSRT dataset is available at http://www.isi.uu.nl/Research/Databases/SCR/. The JSRT dataset is divided into two folds: fold 1 (124 images) and fold 2 (123 images). We used the two-fold cross-validation method to evaluate the segmentation performance. The second dataset is the Chest Radiograph Anatomical Structure Segmentation (CRASS) dataset (http://crass.grand-challenge.org/). The CRASS dataset consists of 548 PA CXRs without manual segmentation of lung fields from a database containing images acquired at two sites in sub Saharan Africa with high tuberculosis incidence. Images from digital radiography (DR) units were used (Delft Imaging Systems, The Netherlands) with a typical size of 1800 × 2000 pixels and a spatial resolution of 0.25 mm. The third dataset includes 650 PA CXRs acquired using three types of DR systems (Discovery XR656, GE Healthcare; FD-X, Siemens Healthcare; T-D3000, SONTU Medical Imaging) in Guangdong, China. 18 manual segmentation of the lung fields was performed in this dataset; such segmentation was used to evaluate the cross-data set generalization of the proposed method qualitatively. Similar to most studies [1, 6, 9], we downscaled the original radiographs in the JSRT database to 256 × 256 pixels in the experiments. Thus, the pixel size became 1.4 mm. Image downsampling to a lower resolution prior to segmentation significantly speeds up runtime without compromising accuracy.

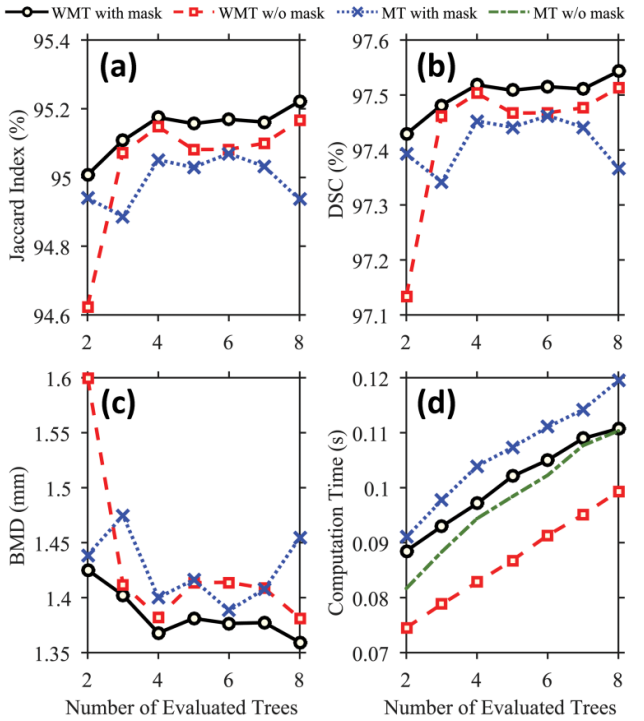


Fig. 9. Segmentation performance of SEDUCM with different settings on the JSRT dataset.

We trained one SED on each fold of the JSRT dataset. The maximum number of trees in the SED was set to 8, and the maximum depth of trees was set to 16. The minimum sample number of leaf nodes was set to 8. A total of 0.2 million positive patches (in which lung boundaries existed) and 0.2 million

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2687939, IEEE Journal of Biomedical and Health Informatics

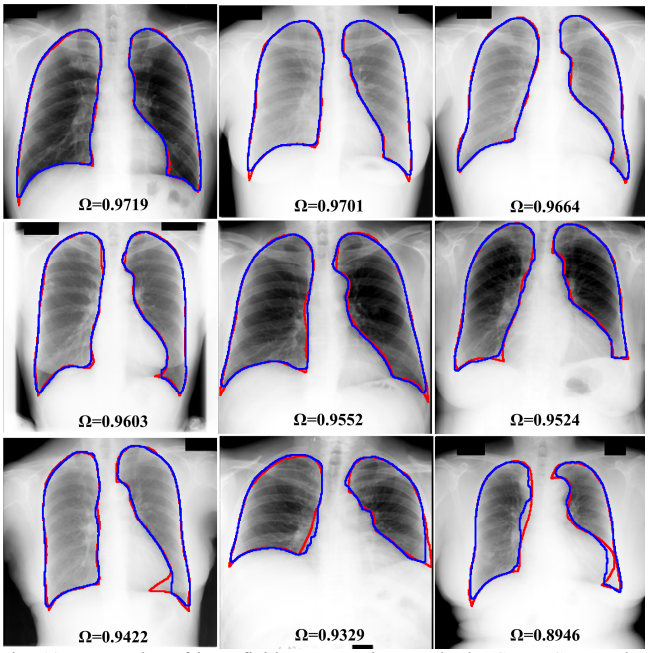


Fig. 10. Examples of lung field segmentation results by SEDUCM on the JSRT dataset. Red and blue contours indicate the ground truth and automatic segmentation results, respectively.

negative patches were randomly selected as training samples for each SED. Experiments were conducted on a laptop with Intel Core i7 CPU (2.3 GHz) and 8 GB RAM. Executions were performed using MATLAB 2014b with the toolboxes provided by Piotr Dollár (https://pdollar.github.io/toolbox/; https://github.com/pdollar/edges). The training time of each tree was approximately 5 min. Decision trees were evaluated in parallel in the prediction stage. The codes for our SEDUCM method and the trained SED models for lung field segmentation are available online (https://github.com/SMU-MedicalVision/SEDUCM).

B. *Performance of Our Proposed Method*

In Fig. 9, we identified the effect of the numbers of evaluated trees and the variants of UCM on the segmentation performance. The numbers of evaluated trees varied from 2 to 8. Applying MWT or WT on a boundary map masked/unmasked by the ribcage mask produced four variants of UCM: MWT with mask, MWT without mask, WT with mask, and WT without mask. Fig. 9 shows how varying parameters and UCM variants affect the Jaccard index, DSC, BMD, and computation time. High Jaccard indices and small MBD values were achieved by evaluating numerous trees when UCM was generated from MWT on the masked boundary map. DSC highly correlated with the Jaccard index. The effect of false strong boundary response on the constant merging procedure of UCM was effectively reduced by MWT and masking out the irrelevant areas. Thus, the use of MWT with mask produced the best segmentation results for different numbers of evaluated trees.

When the number of evaluated trees was greater than 4, all three segmentation evaluation metrics were marginally improved; however, additional computation time was needed. The segmentation results from the UCM through WT on the unmasked boundary maps were the worst. The values of corresponding three evaluation metrics were out of the displayed ranges shown in Figs. 9(a)–9(c). MWT led to a more efficient segmentation than WT on the same boundary maps. MWT without the step to find the ribcage was the most efficient strategy as shown in Fig. 9(d). When four trees were evaluated, the average numbers of segments per CXR produced by MWT and WT (with/without mask) were 6.5/31.5 and 134.3/336.2, respectively. The computation time for UCM generation was reduced when fewer segments were merged and processed.

The majority of the computation time of the SEDUCM segmentation procedure was spent during CXR preprocessing and the generation of the ribcage mask and UCM from the boundary map. Increasing the number of evaluated trees marginally increased the computation time of SEDUCM segmentation (from 0.075 s to 0.12 s) as shown in Fig. 9(d). Reliable segmentation results and reasonable performance can be achieved with four evaluated trees as compared with eight evaluated trees. Consequently the results of setting the number of evaluated trees to four and using UCMs generated by MWT from the masked boundary maps are reported in the rest of this paper. The average Jaccard index and DSC are 95.2% and 97.5%, respectively. The Jaccard indices for most cases are approximately 95%, whereas those for a few cases are less than 90%. Fig. 10 illustrates the visual quality of the extracted lung contours from one test set of JSRT dataset.

TABLE I

PERFORMANCE OF LUNG FIELD SEGMENTATION IN TERMS OF JACCARD INDEX (Ω), DSC, MBD, AND COMPUTATION TIME ON CPU ON THE JSRT DATASET. VALUES (BESIDES COMPUTATION TIME) ARE REPORTED AS MEAN ± STANDARD DEVIATION.

Method	Ω (%)	DSC (%)	MBD (mm)	time (s)
SEDUCM	95.2±1.8	**97.5±1.0**	1.37±0.67	**<0.1**
SIFT-Flow [1]	**95.4±1.5**	96.7±0.8	1.32±0.32	20~25
MISCP [5]	95.1±1.8	/	1.49±0.66	13~28
Hybrid voting [9]	94.9±2.0	/	1.62±0.66	>34
Local SSC [6]	94.6±1.9	97.2±1.0	1.67±0.76	35.2
Human observer [9]	94.6±1.8	/	1.64±0.69	/
GTF [11]	94.6±2.2	/	1.59±0.68	38
InvertedNet [28]	94.6	97.2	**0.73**	7.1
PC post-processed [9]	94.5±2.2	/	1.61±0.80	30
ASM tuned [9]	92.7±3.2	/	2.30±1.03	1
ASM_SIFT [9]	92.0±3.1	/	2.49±1.09	75
AAM whiskers [9]	91.3±3.2	/	2.70±1.10	3

C. *Comparison With Other Methods for Lung Field Segmentation*

Van Ginneken et al. [4] reported the quantitative results of several segmentation methods, including PC, ASM, AAM, hybrid voting, and human observation. Recently, the SIFT Flow algorithm for dense correspondence has been applied to lung segmentation in CXRs with excellent performance [1]. Shao et al. [6] proposed another method for lung field

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2687939, IEEE Journal of Biomedical and Health Informatics

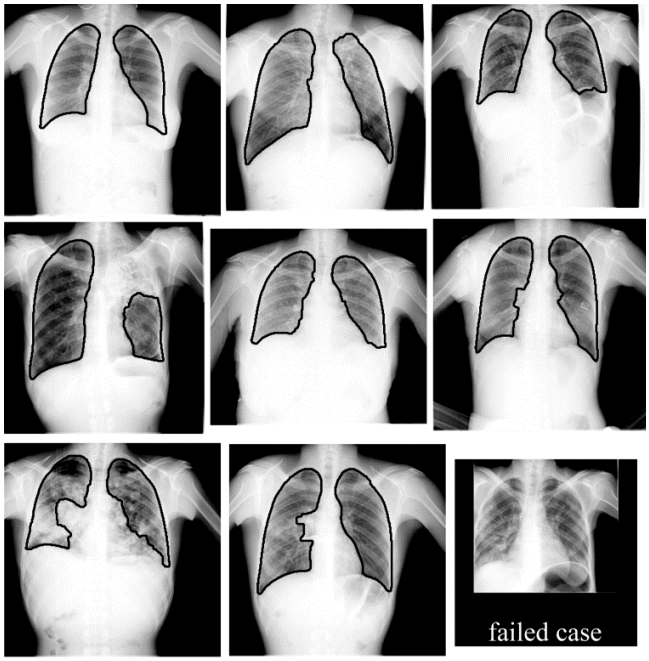


Fig. 11. Segmentation results of SEDUCM on the CRASS dataset.

segmentation by learning local sparse shape and appearance models; this method outperformed conventional shape and appearance models. However, the two latter state-of-the-art methods are time consuming (over 20 s for a CXR with 256 × 256 pixels). Recently, Novikov et al. [28] trained the convolutional neural networks to segment the anatomical structures in CXRs. A model called InvertedNet in [28] achieved the average Jaccard index of 94.6% for lung field segmentation on the JSRT dataset. The computation times of the InvertedNet for each CXR were 7.1 s on CPU and 0.06 s on GPU. As listed in Table I, the average Jaccard index and DSC of our proposed SEDUCM on the JSRT dataset are 95.2% (±1.8%) and 97.5% (±1.0%), respectively. Our SEDUCM method is comparable with the atlas-based method [1] in terms of the aforementioned evaluation metrics. In addition, SEDUCM outperforms all the methods based on shape model, such as ASM, ASM-SIFT [4], and MISCP [5]. The average computation time of SEDUCM pipeline including all steps in Fig. 2 for an input CXR with 256 × 256 pixels is less than 0.1 s. Our SEDUCM method is faster than all other methods according to the reported computation time on CPU in the related literatures without considering of running environments and computers for different methods. With regard to segmentation accuracy, only SEDUCM fulfills the practical requirement of real time.

D. Cross-dataset Generalization

To validate the generalization of SEDUCM, we used the SED that was trained on fold 1 of the JSRT database. There were ten cases of raw data in the third dataset which had not been processed by the enhancement algorithms. For these images of raw data, negative logarithm transform was applied on intensity values to compress the dynamic range for

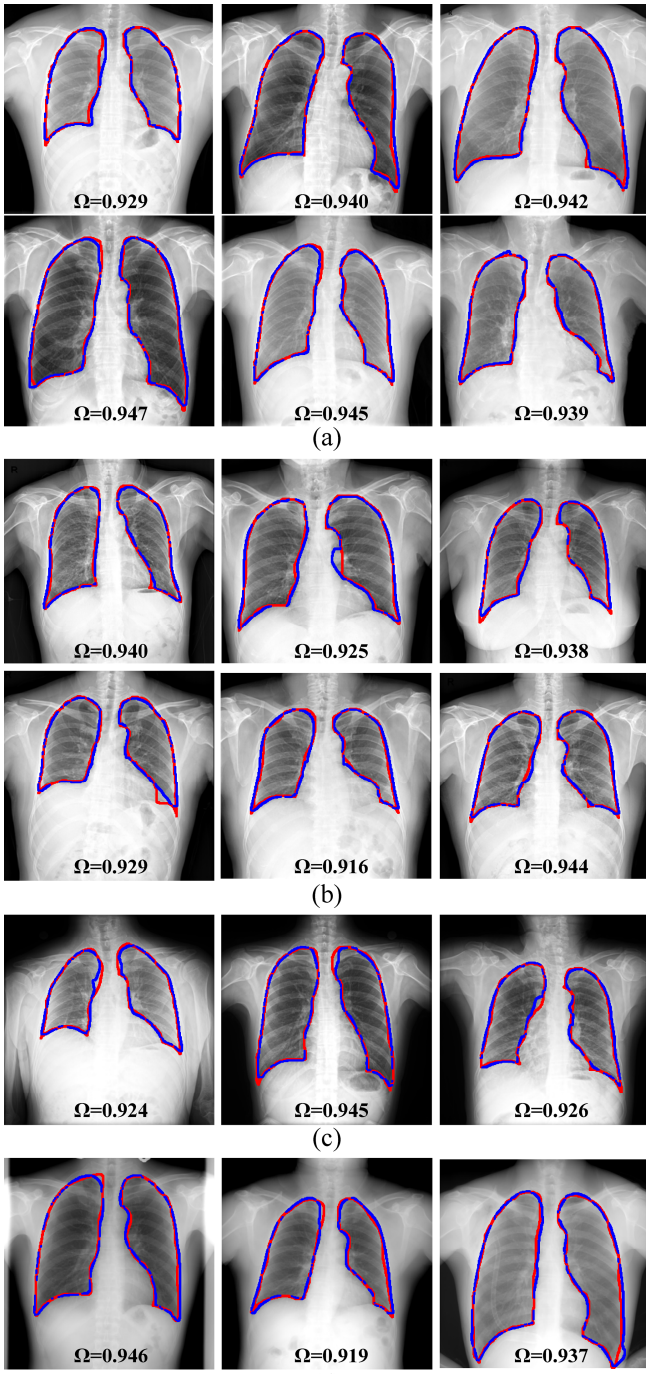


Fig. 12. Segmentation results of SEDUCM on the CXRs acquired by different DR units. (a) Discovery GE XR656, (b) Siemens FD-X, (c) SONTU T-D3000, and (d) SONTU T-D3000 (raw data). Red and blue contours indicate the lung contours outlined by a radiologist and the automatic segmentation results, respectively.

subsequent processing. The DR radiographs were downscaled, and pixel size became 1.4 mm. The radiographs were then normalized and fed to the aforementioned SED. Figs. 11 and 12 illustrate the segmentation examples from the two datasets of digital CXRs. The scanned film radiographs in the JSRT database and the DR radiographs for cross-dataset validation are considerably different. The DR radiographs we collected

have a significantly higher contrast than the CXRs in the JSRT database. In specific, many CXRs in the CRASS dataset were cases of abnormal lungs, and the imaging conditions were not thoroughly controlled. Although our SEDUCM method can provide reasonable segmentation results on many cases of the CRASS dataset, the method failed on some abnormal cases. One failed case is shown in Fig. 11. Extra post-processing rules should be developed to deal with these abnormal cases. SEDUCM can produce segmentation results effectively even in cases with abnormal lungs and raw data, as shown Fig. 12. These results provide evidence that the trained SED demonstrates good generalization capability and that the proposed SEDUCM method is effective and robust.

IV. Discussion and Conclusion

Our method for lung field segmentation employed structured random forests to detect lung boundaries. In principle, modern boundary detectors, such as DeepEdge [17], Oriented Edge Forests [29], and HED [18], can be trained to detect these particular boundaries of lung fields. Among modern boundary detectors, SED exhibits high efficiency, which promotes a fast and practical procedure of lung field segmentation.

The segmentation performance of our method can be further improved. One technique is to combine pixel classification results and the boundary map detected by SED. Another approach is to combine shape models with the boundary map. However, computation time and algorithm complexity increase when these methods are used. A direct approach is to reduce the false boundary responses of SED. In general, a large number of training samples can lead to a relatively good performance of prediction models. We can collect many CXRs with the manual segmentation ground truth to train a SED. Variations in lung field boundaries can be effectively identified using the trained SED.

The segmentation of abnormal lungs is typically difficult. We should develop appropriate rules to address abnormal cases and improve the robustness of SEDUCM. As shown in Fig. 11, our SEDUCM method produces some unreasonable lung contours. The resulting notches by the discontinuity of the detected lung contours can be linked by the smooth curves. Alternatively, the active contour model [30] with a few iterations can be applied to refine the lung contours but with a long computation time.

In summary, we present an effective and efficient lung field segmentation method that can achieve state-of-the-art segmentation accuracy and fulfill the practical requirement of real time. Our method uses a SED to detect lung boundaries. The results demonstrate that effective detection of lung contours using SED and mwt-ucm transform is feasible. Our method can be adopted to simplify approaches for analyzing CXRs.

References

[1] S. Candemir, S. Jaeger, K. Palaniappan, J. P. Musco, R. K. Singh, X. Zhiyun, A. Karargyris, S. Antani, G. Thoma, and C. J. McDonald, "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," IEEE Trans. Med. Imaging, vol. 33, no. 2, pp. 577-90, Feb, 2014.

[2] B. v. Ginneken, and B. M. t. H. Romeny, "Automatic segmentation of lung fields in chest radiographs," Med. Phys., vol. 27, no. 10, pp. 2445-2455, 2000.

[3] M. S. Brown, L. S. Wilson, B. D. Doust, R. W. Gill, and C. Sun, "Knowledge-based method for segmentation and analysis of lung boundaries in chest X-ray images," Computerized Medical Imaging and Graphics, vol. 22, no. 6, pp. 463-477, 1998.

[4] B. v. Ginneken, M. B. Stegmann, and M. Loog, "Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database," Med. Image Anal., vol. 10, pp. 19-40, 2006.

[5] D. Seghers, D. Loeckx, F. Maes, D. Vandermeulen, and P. Suetens, "Minimal shape and intensity cost path segmentation," IEEE Trans. Med. Imaging, vol. 26, no. 8, pp. 1115-1129, 2007.

[6] Y. Shao, Y. Gao, Y. Guo, Y. Shi, X. Yang, and D. Shen, "Hierarchical lung field segmentation with joint shape and appearance sparse learning," IEEE Trans. Med. Imaging, vol. 33, no. 9, pp. 1761-1780, 2014.

[7] T. Xu, M. Mandal, R. Long, I. Cheng, and A. Basu, "An edge-region force guided active shape approach for automatic lung field detection in chest radiographs," Computerized Medical Imaging and Graphics, vol. 36, no. 6, pp. 452-463, 2012.

[8] A. M. R. Schilham, B. van Ginneken, and M. Loog, "A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database," Med. Image Anal., vol. 10, no. 2, pp. 247-258, 2006.

[9] B. van Ginneken, M. B. Stegmann, and M. Loog, "Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database," Med. Image Anal., vol. 10, no. 1, pp. 19-40, 2006.

[10] 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, vol. 33, no. 2, pp. 246-257, 2014.

[11] B. Ibragimov, B. Likar, F. Pernus, and T. Vrtovec, "A Game-Theoretic Framework for Landmark-Based Image Segmentation," IEEE Trans. Med. Imaging, vol. 31, no. 9, pp. 1761-1776, 2012.

[12] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 11, pp. 1222-1239, 2001.

[13] J. Canny, "A Computational Approach to Edge Detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 8, no. 6, pp. 679-698, 1986.

[14] O. Tsujii, M. T. Freedman, and S. K. Mun, "Lung contour detection in chest radiographs using 1-D convolution neural networks," Journal of Electronic Imaging, vol. 8, no. 1, pp. 46-53, 1999.

[15] D. R. Martin, C. C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 5, pp. 530-549, 2004.

[16] P. Dollár, and C. L. Zitnick, "Fast Edge Detection Using Structured Forests," IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 8, pp. 1558-1570, 2015.

[17] G. Bertasius, S. Jianbo, and L. Torresani, "DeepEdge: A multi-scale bifurcated deep network for top-down contour detection," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 4380-4389.

[18] S. Xie, and Z. Tu, "Holistically-Nested Edge Detection," in 2015 IEEE Conference on Computer Vision (ICCV), 2015, pp. 1395-1403.

[19] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour Detection and Hierarchical Image Segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 5, pp. 898-916, 2011.

[20] P. Arbelaez, "Boundary Extraction in Natural Images Using Ultrametric Contour Maps," in IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW '06) 2006.

[21] P. Soille, Morphological Image Analysis: Principles and Applications, 2nd ed.: Springer Berlin Heidelberg, 2004.

[22] K. He, J. Sun, and X. Tang, "Guided image filtering," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 6, pp. 1397-1409, 2013.

[23] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.

[24] A. Criminisi, and J. Shotton, Decision Forests for Computer Vision and Medical Image Analysis: Springer, 2013.

[25] Y. Zhanjun, A. Goshtasby, and L. V. Ackerman, "Automatic detection of rib borders in chest radiographs," IEEE Trans. Med. Imaging, vol. 14, no. 3, pp. 525-536, 1995.
[26] L. R. Dice, "Measures of the amount of ecologic association betweenspecies," Ecology, vol. 26, no. 3, pp. 297-302, 1945.
[27] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, "Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules," American Journal of Roentgenology, vol. 174, no. 1, pp. 71-74, 2000.
[28] A. A. Novikov, D. Major, D. Lenis, J. Hladůvka, M. Wimmer, and K. Bühler. "Fully Convolutional Architectures for Multi-Class Segmentation in Chest Radiographs," http://arxiv.org/abs/1701.08816, Feb. 2017.
[29] S. Hallman, and C. C. Fowlkes, "Oriented edge forests for boundary detection," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1732-1740.
[30] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," International Journal of Computer Vision, vol. 1, no. 4, pp. 321-331, 1988.