LUNG SEGMENTATION IN CHEST RADIOGRAPHS USING DISTANCE REGULARIZED LEVEL SET AND DEEP-STRUCTURED LEARNING AND INFERENCE

Tuan Anh Ngo   *Gustavo Carneiro*

Australia Centre for Visual Technologies
The University of Adelaide, Australia

ABSTRACT

Computer-aided diagnosis of digital chest X-ray (CXR) images critically depends on the automated segmentation of the lungs, which is a challenging problem due to the presence of strong edges at the rib cage and clavicle, the lack of a consistent lung shape among different individuals, and the appearance of the lung apex. From recently published results in this area, hybrid methodologies based on a combination of different techniques (e.g., pixel classification and deformable models) are producing the most accurate lung segmentation results. In this paper, we propose a new methodology for lung segmentation in CXR using a hybrid method based on a combination of distance regularized level set and deep structured inference. This combination brings together the advantages of deep learning methods (robust training with few annotated samples and top-down segmentation with structured inference and learning) and level set methods (use of shape and appearance priors and efficient optimization techniques). Using the publicly available Japanese Society of Radiological Technology (JSRT) dataset, we show that our approach produces the most accurate lung segmentation results in the field. In particular, depending on the initialization used, our methodology produces an average accuracy on JSTR that varies from 94.8% to 98.5%.

Index Terms— Lung segmentation, Deep learning, Level set methods

1. INTRODUCTION

The automated segmentation of lung boundaries from digital chest X-ray (CXR) is one of the main stages in the computer-aided diagnosis (CAD) of lung health [1]. Lung boundaries can be used for computing lung volume or estimating shape irregularities [2], but it is also used as one of the stages in several CAD systems [6]. These CAD systems are particularly important for screening and detecting pulmonary pathologies, but with a major focus on tuberculosis, which is the second leading cause of death from infectious disease worldwide [3].
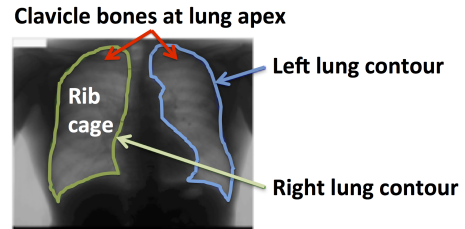
Fig. 1. Left and right lungs segmentation.

The highest incidences of these diseases occur in places of the world with inadequate health care infrastructure, so the deployment of such CAD systems in these places is important because they can help local clinicians in the screening and diagnosis processes mentioned above [1]. However, the automated segmentation of lung boundaries is a challenging task because of the following reasons (see Fig. 1) [1]: 1) the edges present at the rib cage and clavicle represent a challenge for optimization methods that can get stuck at local minima; 2) the appearance inconsistencies caused by the clavicle bone at the lung apex also represent an issue for most optimization approaches for the same reason above; and 3) the lack of a consistent lung shape among different individuals is a challenge for the use of shape priors.

There has been considerable effort applied in the development of automated lung segmentation methods [13], and the most successful approaches are usually based on hybrid methods that combine several techniques, such as methods that combine landmark learning with active shape and appearance models [11,12] or graph cuts with non-rigid registration [1]. Similarly, we propose a hybrid method based on a recent methodology that we developed for the problem of left ventricle segmentation from magnetic resonance image (MRI) [8], which has recently achieved the best results in the field. The extension of this methodology to this new problem requires some modifications to the original algorithm, but it is interesting to note that the core steps have remained almost unaltered, showing that this algorithm can be potentially applied to other similar problems.

The method being proposed in this paper for segmenting lungs from CXR images is based on the combination of distance regularized level set (DRLS) [7] and deep structured learning and inference using a deep belief network (DBN) [5]. Essentially, we use the DRLS [7] optimization with the usual shape and appearance terms, but with an additional term based on the top-down segmentation produced by a deep structured inference. This combination aims at exploring the advantages of both approaches, which are the efficient optimization and the prior shape and appearance terms from DRLS, and the robust statistical segmentation models produced by deep learning methods. We test our approach using the Japanese Society of Radiological Technology (JSRT) dataset, and our results show that, depending on the initialization used, our methodology can produce an average accuracy on JSTR that varies from 94.8% to 98.5%, which is significantly better than the current best approach in the field [1] (that reported an accuracy of 95.4%).

2. METHODOLOGY

In this section, assume that the annotated chest radiograph database is represented by $\mathcal{D} = \{(I, \mathbf{c}, q)_i\}_{i=1}^{|\mathcal{D}|}$, where $I : \Omega \to \mathbb{R}$ represents an image (with $\Omega \subseteq \mathbb{R}^2$ denoting the image lattice), $q \in \{\text{left lung,right lung}\}$ and $\mathbf{c} : [0,1] \to \Omega$ denotes the explicit contour representation of the segmentation. Also assume that the implicit contour representation is the zero level set of a signed distance function $\phi : \Omega \to \mathbb{R}$, and the lung segmentation map is represented by $\mathbf{y} : \Omega \to \{0,1\}$, where 1 represents the foreground (i.e., left or right lung) and 0 denotes the background. Below, we first explain the DRLS method, then we describe the DBN segmentation model, and finally explain the combined inference algorithm.

The main optimization method for producing the segmentation is based on the DRLS formulation [7], where the energy functional is represented by:

$$\mathcal{E}(\phi, \phi_{\text{DBN},q}) = \mu \mathcal{R}_p(\phi) + \mathcal{E}_{\text{ext}}(\phi, \phi_{\text{DBN},q}), \quad (1)$$

with the distance regularization $\mathcal{R}_p(\phi) = \int_\Omega 0.5(|\nabla\phi| - 1)^2 dx$ (this guarantees that $|\nabla\phi| \approx 1$), and

$$\mathcal{E}_{\text{ext}}(\phi, \phi_{\text{DBN},q}) = \lambda \mathcal{L}(\phi) + \alpha \mathcal{A}(\phi) + \gamma \mathcal{S}(\phi, \phi_{\text{DBN},q}), \quad (2)$$

where the length term $\mathcal{L}(\phi) = \int_\Omega g\delta(\phi)|\nabla\phi| dx$ (with $\delta(.)$ denoting the Dirac delta function and $g = \frac{1}{1+|\nabla G_\sigma * I|}$ representing the edge indicator function), the area $\mathcal{A}(\phi) = \int_\Omega gH(-\phi)dx$ (with $H(.)$ denoting the Heaviside step function), and $\mathcal{S}(\phi, \phi_{\text{DBN},q}) = \int_\Omega (\phi - \phi_{\text{DBN},q})^2 dx$ represents the shape term that drives the ϕ towards the shape $\phi_{\text{DBN},q}$, which is the distance function inferred from the deep belief network (DBN) structured inference described below (see Fig. 2-(a)). The minimization of the energy functional in (1) is achieved by finding the steady solution of the gradient flow equation $\frac{\partial\phi}{\partial t} = -\frac{\partial\mathcal{E}}{\partial\phi}$ [7].
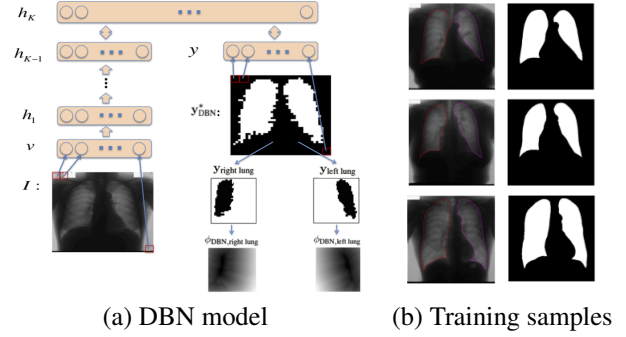

(a) DBN model   (b) Training samples

Fig. 2. DBN Model (a) and training samples for the DBN (b).

The DBN structured inference produces the following segmentation map (Fig, 2-(a)):

$$\mathbf{y}^*_{\text{DBN},q} = \arg\max_{\mathbf{y}} \sum_{\mathbf{h}_1}..\sum_{\mathbf{h}_K} P(\mathbf{v}, \mathbf{h}_1, ..., \mathbf{h}_K, \mathbf{y}; \Theta_q), \quad (3)$$

where $\mathbf{h}_k \in \{0,1\}^{|\mathbf{h}_k|}$ represents the $|\mathbf{h}_k|$ hidden nodes of layer $k \in \{1,..,K\}$ of the deep belief network, $\mathbf{v}$ is a vector representation of the input image I, and Θ_q denotes the DBN parameters (weights and biases). The probability term in (3) is computed as

$$P(\mathbf{v}, \mathbf{h}_1, ..., \mathbf{h}_K, \mathbf{y}) = P(\mathbf{h}_K, \mathbf{h}_{K-1}, \mathbf{y}) \left(\prod_{k=1}^{K-2} P(\mathbf{h}_{k+1}|\mathbf{h}_k)\right) P(\mathbf{h}_1|\mathbf{v}), \quad (4)$$

where $-\log P(\mathbf{h}_K, \mathbf{h}_{K-1}, \mathbf{y}) \propto \mathcal{E}_{\text{RBM}}(\mathbf{h}_K, \mathbf{h}_{K-1}, \mathbf{y})$ with

$$\mathcal{E}_{\text{RBM}}(\mathbf{h}_K, \mathbf{h}_{K-1}, \mathbf{y}) = -\mathbf{b}_K^\top \mathbf{h}_K - \mathbf{a}_{K-1}^\top \mathbf{h}_{K-1} - \mathbf{a}_y^\top \mathbf{y} - (\mathbf{h}_K)^\top \mathbf{W}_K \mathbf{h}_{K-1} - (\mathbf{h}_K)^\top \mathbf{W}_y \mathbf{y} \quad (5)$$

representing the energy function of a restricted Boltzmann machine (RBM) [5], where $\mathbf{b}_K, \mathbf{a}_{K-1}, \mathbf{a}_y$ denote the bias vectors and $\mathbf{W}_K, \mathbf{W}_y$ are the weight matrices. Also in (4), we have

$$P(\mathbf{h}_{k+1}|\mathbf{h}_k) = \prod_j P(\mathbf{h}_{k+1}(j) = 1|\mathbf{h}_k), \quad (6)$$

with $P(\mathbf{h}_{k+1}(j) = 1|\mathbf{h}_k) = \sigma(\mathbf{b}_{k+1}(j) + \mathbf{h}_k^\top \mathbf{W}_{k+1}(:,j))$, $P(\mathbf{h}_1(j) = 1|\mathbf{v}) = \sigma(\mathbf{b}_1(j) + \frac{\mathbf{v}^\top \mathbf{W}_1(:,j)}{\sigma^2})$ [1], where $\sigma(x) = \frac{1}{1+e^{-x}}$, the operator (j) returns the jth vector value, and $(:,j)$ returns the jth matrix column.

The DBN in (3) is trained with a dataset containing training image I and respective segmentation map $\mathbf{y}$, as shown in Fig. 2-(b). The training process is based on the unsupervised bottom-up training of each pair of layers, where the weights

---

[1]That is, we assume zero-mean Gaussian visible units for the DBN.

and biases of the network are learned to build an auto-encoder for the values at the bottom layer, and a top RBM is trained with the segmentation map $\mathbf{y}$ [5]. The structured inference process consists of taking the input image and performing bottom-up inferences, until reaching the top two layers, which form an RBM, and then initialize the layer $\mathbf{y} = \mathbf{0}$ and perform Gibbs sampling on the layers $\mathbf{h}_K$ and $\mathbf{h}_{K-1}, \mathbf{y}$ until convergence [5].

The combination of DRLS and DBN is explained in the Alg. 1, where essentially, we iteratively run the DRLS method until convergence using the segmentation result from the DBN as one of the optimization terms.

Algorithm 1 Combined Level Set and DBN Segmentation

- INPUT: test image I and initial segmentation ϕ_0
- Infer $\mathbf{y}^*_{\text{DBN},q}$ from I using (3) for $q \in \{$left lung,right lung$\}$
- Compute distance function $\phi_{\text{DBN},q}$ from map $\mathbf{y}^*_{\text{DBN},q}$ (Fig. 2-(a))

for t = 1:T **do**
  - Run DRLS using $\phi_{t-1}, \phi_{\text{DBN},q}$ to produce ϕ_t

end for
- Segmentation is the zero level set $\mathcal{C} = \{\mathbf{x} \in \Omega | \phi_T(\mathbf{x}) = 0\}$

---

3. EXPERIMENTS

The evaluation of the accuracy of our methodology uses the publicly available Japanese Society of Radiological Technology (JSRT) dataset [10], which contains manual segmentations of lung fields, heart and clavicles [12]. The JSRT database contains 247 chest radiographs, where 154 contain lung nodules (100 malignant, 54 benign) and 93 have no nodules, and each sample is represented by 12-bit gray scale image with size 2048×2048 pixels and $0.175mm$ pixel resolution. This database is randomly split into three sets: training (84 images), validation (40 images) , and test (123 images), and the assessment is based on following three measures: Jaccard Similarity Coefficient (Ω), Dice's Coefficient (DSC), and Average Contour Distance (ACD) [1]. We use the training set for the estimation of the DBN and DRLS parameters and the validation set for the DBN model selection (e.g. select the number of layers and number of nodes per layer in the network). The model selection estimated the following configuration for DBN: each hidden layer has 1000 nodes, with the input and segmentation layers with 1600 nodes. The initial guess ϕ_0 in Alg. 1 used by our approach is not automatically produced, so we show how the performance of our approach is affected by initial guesses of different accuracies, which are generated by random perturbations from the manual annotation. We denote the different initial guesses by the index $k \in \{1, 2, 3\}$, where $k = 1$ indicates the highest precision and $k = 3$ means the lowest precision initial guess. The estimation of the level set parameters is performed separately for each type of initial guess, and we achieve the following result: $\mu = 0.12, \lambda = 2, \alpha = -3, \gamma = 0.0005$ for $k = 1$; $\mu = 0.12, \lambda = 2, \alpha = -10, \gamma = 0.003$ for $k = 2$; and
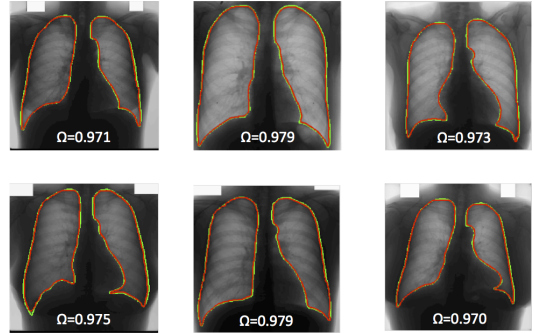

Fig. 3. Lung segmentation results with initial guess $k = 2$. The green contour shows expert annotation and the red illustrates the final result.

$\mu = 0.12, \lambda = 2, \alpha = -15, \gamma = 0.007$ for $k = 3$.

3.1. Results

Table 1 shows the results of our proposed methodology for lung segmentation with the different types of initial guesses. In this table, we also show the results when $\gamma = 0$, which is denoted by "Model without DBN" (this shows the influence of the DBN in the proposed methodology); and we also show the results for the initial guess, represented by "Initial guess only". Table 2 compares our results with the ones produced by the current state of the art on the JSRT database. Finally, Fig. 3.1 shows a few lung segmentation results using initial guess $k = 2$ on images of the test set. Using a standard computer (Intel(R) Core(TM) i5-2500k 3.30GHz CPU with 8GB RAM), and processing an input image with size 256×256 pixels, our method runs on average in 20.68 sec/image, which is comparable to the result by Candemir et al. [1], who report a running time of between 20 and 25 sec/image using the same input resolution and similar computer configuration.

4. DISCUSSIONS AND CONCLUSIONS

The results show that our proposed method using the initial guesses $k \in \{1, 2\}$ produces the best results in the field and comparable running times to the current state of the art. The main step that is missing from our approach is an automated initial guess, and we plan to address this issue by using the initial guess proposed by Candemir et al. [1]. We also plan to extend this method to other lung segmentation databases [1] and other segmentation problems (i.e., other anatomies) from different imaging techniques.

5. REFERENCES

[1] S. Candemir, et al. Lung segmentation in chest radiographs using anatomical atlases with non-rigid registration. 2014.

[2] F. Carrascal, et al. Automatic calculation of total lung capacity from automatically traced lung boundaries in

Table 1. Quantitative experiments on the JSRT database [10] showing the performance of the proposed method as a function of the initial guess used, where each cell is formatted as "mean (standard deviation) [min value - max value]".

Initial guess	Method	Ω	DSC	ACD
$k = 1$	Proposed model	$0.985(0.003)[0.972 - 0.991]$	$0.992(0.002)[0.986 - 0.996]$	$1.075(0.065)[0.825 - 1.267]$
	Model without DBN	$0.984(0.003)[0.969 - 0.990]$	$0.992(0.002)[0.984 - 0.995]$	$1.376(0.221)[1.234 - 6.184]$
	Initial guess only	$0.955(0.006)[0.919 - 0.968]$	$0.977(0.003)[0.958 - 0.984]$	$1.392(0.006)[1.372 - 1.404]$
$k = 2$	Proposed model	$0.973(0.007)[0.944 - 0.985]$	$0.986(0.004)[0.971 - 0.993]$	$1.120(0.165)[0.628 - 1.916]$
	Model without DBN	$0.946(0.007)[0.910 - 0.961]$	$0.972(0.004)[0.953 - 0.980]$	$2.408(0.232)[0.021 - 7.232]$
	Initial guess only	$0.912(0.013)[0.844 - 0.935]$	$0.954(0.007)[0.916 - 0.967]$	$2.519(0.041)[2.369 - 2.621]$
$k = 3$	Proposed model	$0.948(0.012)[0.893 - 0.970]$	$0.973(0.006)[0.943 - 0.985]$	$1.852(0.286)[1.120 - 3.708]$
	Model without DBN	$0.866(0.018)[0.790 - 0.900]$	$0.928(0.010)[0.883 - 0.947]$	$4.695(0.276)[3.792 - 9.112]$
	Initial guess only	$0.828(0.024)[0.712 - 0.873]$	$0.906(0.014)[0.832 - 0.932]$	$4.936(0.105)[4.391 - 5.200]$

Table 2. Quantitative experiments on the JSRT database [10] comparing our results with the state of the art on the same database, sorted from best (top) to worst (bottom). The symbol '?' indicates that the result is not available.

Method	Ω	DSC	ACD
Proposed model, $k = 1$	$0.985(0.003)[0.972 - 0.991]$	$0.992(0.002)[0.986 - 0.996]$	$1.075(0.065)[0.825 - 1.267]$
Proposed model, $k = 2$	$0.973(0.007)[0.944 - 0.985]$	$0.986(0.004)[0.971 - 0.993]$	$1.120(0.165)[0.628 - 1.916]$
Candemir, Hybrid [1]	$0.954(0.015)[?-?]$	$0.967(0.008)[?-?]$	$1.321(0.316)[?-?]$
Ginneken, Hybrid voting [12]	$0.949(0.020)[0.818 - 0.978]$	$?(?)[?-?]$	$1.62(0.66)[0.95 - 7.72]$
Proposed model, $k = 3$	$0.948(0.012)[0.893 - 0.970]$	$0.973(0.006)[0.943 - 0.985]$	$1.852(0.286)[1.120 - 3.708]$
Ginneken, PC post-processed [12]	$0.945(0.022)[0.823 - 0.972]$	$?(?)[?-?]$	$1.61(0.80)[0.83 - 8.34]$
Dawoud [4]	$0.940(0.053)[?-?]$	$?(?)[?-?]$	$2.46(2.06)[?-?]$
Ginneken, PC [12]	$0.938(0.027)[0.823 - 0.968]$	$?(?)[?-?]$	$3.25(2.65)[0.93 - 15.59]$
Ginneken, Hybrid ASM/PC [12]	$0.934(0.037)[0.706 - 0.968]$	$?(?)[?-?]$	$2.08(1.40)[0.91 - 11.57]$
Seghers [9]	$0.930(?)[?-?]$	$?(?)[?-?]$	$?(?)[?-?]$
Ginneken, AAM whiskers BFGS [12]	$0.922(0.029)[0.718 - 0.961]$	$?(?))[?-?]$	$2.39(1.07)[1.15 - 12.09]$
Yu [14]	$0.907(0.033)[?-?]$	$?(?)[?-?]$	$?(?)[?-?]$
Ginneken, Mean shape [12]	$0.713(0.075)[0.460 - 0.891]$	$?(?)[?-?]$	$10.06(3.18)[3.50 - 23.77]$

postero-anterior and lateral digital chest radiographs. *Medical physics*, 25(7):1118–1131, 1998.

[3] Charles L Daley, M Gotway, and R Jasmer. Radiographic manifestation of tuberculosis. *A primer for clinicians*, 1, 2003.

[4] A Dawoud. Lung segmentation in chest radiographs by fusing shape information in iterative thresholding. *IET Computer Vision*, 5(3):185–190, 2011.

[5] G. Hinton and R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.

[6] S. Jaeger, et al. Automatic tuberculosis screening using chest radiographs. 2013.

[7] C. Li, et al. Distance regularized level set evolution and its application to image segmentation. *IEEE TIP*, 19(12):3243–3254, 2010.

[8] T. Ngo and G. Carneiro. Fully automated non-rigid segmentation with distance regularized level set evolution initialized and cosntrained by deep-structured inference. In *CVPR*. 2014.

[9] D. Seghers, et al. Minimal shape and intensity cost path segmentation. *IEEE TMI*, 26(8):1115–1129, 2007.

[10] J. Shiraishi, et al. 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*, 174(1):71–74, 2000.

[11] B. Ginneken et al. Active shape model segmentation with optimal features. *IEEE TMI*, 21(8):924–933, 2002.

[12] B. Ginneken, M. Stegmann, and M. Loog. Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. *MIA*, 10(1):19–40, 2006.

[13] B. Ginneken, Bart M ter Haar Romeny, and Max A Viergever. Computer-aided diagnosis in chest radiography: a survey. *IEEE TMI*, 20(12):1228–1241, 2001.

[14] T. Yu, J. Luo, and N. Ahuja. Shape regularized active contour using iterative global search and local optimization. In *CVPR*, 2005.