

Stratified Learning of Local Anatomical Context for Lung Nodules in CT Images

Dijia Wu^{1,2}, Le Lu¹, Jinbo Bi¹, Yoshihisa Shinagawa¹, Kim Boyer², Arun Krishnan¹, and Marcos Salganicoff¹

¹CAD and Knowledge Solutions, Siemens Medical Solutions, Malvern, PA 19355 USA,

²ECSE Department, Rensselaer Polytechnic Institute, Troy, NY 12180 USA,

Abstract

The automatic detection of lung nodules attached to other pulmonary structures is a useful yet challenging task in lung CAD systems. In this paper, we propose a stratified statistical learning approach to recognize whether a candidate nodule detected in CT images connects to any of three other major lung anatomies, namely vessel, fissure and lung wall, or is solitary with background parenchyma. First, we develop a fully automated voxel-by-voxel labeling/segmentation method of nodule, vessel, fissure, lung wall and parenchyma given a 3D lung image, via a unified feature set and classifier under conditional random field. Second, the generated Class Probability Response Maps (PRM) by voxel-level classifiers, are used to form the so-called pairwise Probability Co-occurrence Maps (PCM) which encode the spatial contextual correlations of the candidate nodule, in relation to other anatomical landmarks. Based on PCMs, higher level classifiers are trained to recognize whether the nodule touches other pulmonary structures, as a multi-label problem. We also present a new iterative fissure structure enhancement filter with superior performance.

For experimental validation, we create an annotated database of 784 subvolumes with nodules of various sizes, shapes, densities and contextual anatomies, and from 239 patients. High accuracy of multi-class voxel labeling is achieved 89.3% ~ 91.2%. The Area under ROC Curve (AUC) of vessel, fissure and lung wall connectivity classification reaches 0.8676, 0.8692 and 0.9275, respectively.

1. Introduction

Lung cancer is the number one cause of cancer death in US, and the early detection and diagnosis of cancer-relevant pulmonary nodules is vitally important. A pulmonary nodule is a small, often round lesion in the lungs as shown in Fig. 4. Recent studies reveal that the location of nodules, *i.e.*, purely intra-parenchymal (solitary), pleura-based (lung wall attached), juxtavascular (vessel attached) or fissure attached, provides information that can help distinguish be-

nign versus malignant lesions. For instance, vessel-feeding pulmonary nodules seem more likely to be lung cancers than solitary ones [22], while fissure attached or pleural-based nodules are typically benign, with volume-doubling period greater than 400 days [24]. Examples of nodules attached to the three pulmonary structures are shown in Fig.6. Despite its obvious clinical importance, automatic classification of solitary or various attached nodules has not been extensively studied before, manual labeling is typical. Wei *et al.* proposed a specific method of volume projection analysis to improve detection performance for nodules attached to vessels [22], but did not discuss the cases connecting to other pulmonary structures. In addition, this method is based on an ellipsoidal nodule model that does not always hold since nodules, especially malignant ones, may be spiculated, polylobulated or otherwise irregular in shape [21].

The spatial relation between lung nodule and its contextual anatomies will be easily obtained if we can parse the local anatomies correctly. It is similar to physicians' practices in labeling connectivity of nodules with other structures, by identifying and segmenting/coloring the nodule and its local contextual structures on voxel level and then determining whether they touch each other. Automated and robust segmentation of these pulmonary anatomies is highly desirable, but it has various challenges for each specific structure. For example, nodules are known to have a wide variety of morphologies, densities and sizes; fissures appear as low contrast surfaces with possibly blurred and incomplete boundaries; vessel segmentation is complicated by the weakness of contrast of small vessels and local ambiguities posed by crossing and branching points. Accordingly, most previous lung anatomy segmentation methods target only one type of structure. For fissure segmentation, various methods have been developed, such as fuzzy reasoning followed by graph search [26], curve-growing modeled by a Bayesian network [20], multi-stage supervised filters [19] and a computational geometry based method with surface Laplacian smoothing [13]; For vessels, Frangi, *et al.* proposed multiscale Hessian eigenvalue based filters to enhance vessels as a preprocessing step of segmentation [5]. Correlation based filters were used to enhance blood vessels and junctions while sup-

pressing nodules and other noise, followed by vessel tree reconstruction using fuzzy shape analysis [1]; The commonly used methods of pulmonary nodule segmentation include region growing [4], Gaussian intensity model fitting [12], and statistical learning based approach [17].

In this paper, we propose a generic approach which handles multiple structure (nodule, vessel, lung wall and fissure) voxel-level classification using a unified feature set and classification method. The aim at this stage is not to provide a perfect segmentation, but to use a fast and robust method for reasonable segmentation results that can serve as input to a higher level classification system to learn the spatial layout of nodule and its contextual anatomies. A similar voxel-level classification strategy for lung anatomy was presented in [11], but it differs from our approach in two main aspects. First, we propose an iterative Hessian eigenvalue based filter to enhance fissure structure with significantly fewer spurious responses as compared with traditional fissure enhancement filter used in [11]. Second, we use a conditional random field (CRF) model [9] incorporating gray-level, texture, shape, and edge cues to improve the segmentation of boundary voxels, especially of nodule boundary voxels, which is critical to the performance of the final anatomy connectivity classification system. At the higher level, instead of using simple n -connected component algorithm with hard segmentation results, we present a novel probability co-occurrence map (PCM) to measure the spatial interaction/attachment likelihood of nodule and other structures for connectivity classification.

Though mainly developed for medical imaging, our approach is well related to general computer vision problems. Recently, various contextual learning models [7, 14, 16] have been exploited to assist general object detection problems, mostly within the conditional random field (CRF) framework [9]. Our work also uses the context for voxel level segmentation, but focuses on more explicitly stratified modeling and learning the spatial relations of different objects (as nodules versus anatomical background). It is an integrated task of fusing detection, segmentation and semantic recognition for diagnostic quantification of lung nodules in 3D images. This work is also relevant with the multi-object interaction context modeling for activity recognition [6].

In summary, our contributions are *threefold*. First, we develop a fully automated voxel level parsing of multi-class lung anatomies using a single feature set and classifier. Second, partially inspired by [10], we present a two-level, hierarchical probabilistic framework of effectively inferencing 3D complex, multi-label nodule connectivity in lung images. Last, a new iterative Hessian eigenvalue based filter is proposed to enhance fissure structure. We show empirically and analytically that the new filter can significantly reduce spurious responses. Our experimental results are extensively evaluated using a large-scale medical image database of 784 various structured lung nodules from 239 subjects.

2. Voxel Level Classification of Lung Anatomy

At this stage, we aim to roughly segment the contextual pulmonary structures of lung nodules. The conditional random field (CRF) model [9, 16] is adopted that incorporates unified local gray-level, texture, shape and edge information to learn the conditional distribution of each voxel over class labeling. The conditional probability of class label c given a volume \mathbf{x} is defined below:

$$\log P(c|\mathbf{x}, \boldsymbol{\theta}) = \sum_i \psi(c_i, \mathbf{x}; \boldsymbol{\theta}_\psi) + \sum_{(i,j) \in N_\epsilon} \phi(c_i, c_j, \mathbf{x}; \boldsymbol{\theta}_\phi) - \log Z(\mathbf{x}, \boldsymbol{\theta}) \quad (1)$$

where i, j are indices of voxels in given volume, N_ϵ denotes the set of edges in 6-connected grid structure, $\boldsymbol{\theta} = \{\boldsymbol{\theta}_\psi, \boldsymbol{\theta}_\phi\}$ are the model parameters, $Z(\boldsymbol{\theta}, \mathbf{x})$ is the normalization function independent of label c . The unary potential $\psi(c_i, \mathbf{x}; \boldsymbol{\theta}_\psi)$ usually indicates individual label preference based on observed voxels and is defined as

$$\psi(c_i, \mathbf{x}; \boldsymbol{\theta}_\psi) = \log P(c_i|\mathbf{x}, i) \quad (2)$$

$P(c_i|\mathbf{x}, i)$ is the normalized distribution output from a classifier given local gray-level, texture and shape features of voxel i . These features and the classifier are detailed in Section 2.1 and 2.2, respectively. The pairwise potential $\phi(c_i, c_j, \mathbf{x}; \boldsymbol{\theta}_\phi)$ encourages spatial coherence by penalizing label discontinuities between neighboring voxels with similar gray-levels. Here it takes the form of a Potts Model:

$$\phi(c_i, c_j, \mathbf{x}; \boldsymbol{\theta}_\phi) = -\alpha e^{-\|x_i - x_j\|^2/\beta} \cdot \delta(c_i - c_j) \quad (3)$$

where α is the contrast parameter weighting unary and pairwise potentials, β denotes the normalization parameter dependent on volume \mathbf{x} and class label c_i, c_j . The $\delta(\cdot)$ is Dirac delta function. The use of intensity contrast based pairwise potential Eq.(3) is motivated by the observation that voxels within a single pulmonary structure of interest usually show similar intensities in lung CT images. For example, nodule intensities can vary greatly across population according to the nodule type (solid or ground glass), but they usually do not change much within one nodule.

Given the conditional random model and its learned parameters, the optimal labeling c^* can be found by maximizing Eq.(1) via graph-cut algorithm [3, 2].

2.1. Lower Level Feature Set

The features selected for detection of pulmonary structures can be categorized into three subsets.

Texture. We selected the texture features earlier used for ground glass nodule detection [27, 17] and parenchyma texture analysis [25]. These features consist of the well known gray-level co-occurrence matrix (GLCM) based features searching for gray-level pairs in 26 directions [25], intensity and rotation invariant local binary patterns (LBP) [27] and 3D Harr wavelets [17] capturing the local spatial and frequency information.

Gray-Level. In CT scans, radiation attenuation varies with body tissues of different atomic weight and density. Therefore the intensity can be a good indicator of which pulmonary structure the voxel comes from. The features in this subset comprise the original intensity value, and intensity statistics features including the minimum, maximum, mean, standard deviation, skewness and kurtosis calculated from a cubic subvolume of size $7 \times 7 \times 7$ voxels.

Shape. Lung nodules, vessels and fissures show quite different anatomical shapes which are usually modeled as faint or light blob-like, light tubular and faint plate-like structures, respectively. Therefore, a set of multiscale structure enhancement filters based on Hessian matrix eigenvalues is commonly used to differentiate the three structures [11, 17, 5, 19]. These filters are defined by the following or similar equations:

$$N(\boldsymbol{\lambda}) = (1 - \exp(-(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)/2\gamma^2))(1 - \exp(-\lambda_3^2/2\alpha^2|\lambda_1\lambda_2|))h(-\lambda_1)h(-\lambda_2)h(-\lambda_3) \quad (4)$$

$$V(\boldsymbol{\lambda}) = (1 - \exp(-(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)/2\gamma^2))(1 - \exp(-\lambda_2^2/2\alpha^2\lambda_1^2))\exp(-\lambda_3^2/2\beta^2|\lambda_1\lambda_2|)h(-\lambda_1)h(-\lambda_2) \quad (5)$$

$$F(\boldsymbol{\lambda}) = (1 - \exp(-(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)/2\gamma^2))\exp(-\lambda_2^2/2\beta^2\lambda_1^2)h(-\lambda_1) \quad (6)$$

where $\boldsymbol{\lambda} = \{\lambda_1, \lambda_2, \lambda_3\}$ are three eigenvalues of Hessian matrix with $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$. The thresholds α , β and γ control the sensitivity of filters to the eigenvalue measures, which are set as 0.5, 0.5 and 200 each. $h(\cdot)$ is the standard unit step function. $N(\boldsymbol{\lambda})$, $V(\boldsymbol{\lambda})$ and $F(\boldsymbol{\lambda})$ denote the enhancement filters of blob-like (nodule), tubular (vessel) and plate-like (fissure) structures, respectively. The theoretical basis for these filters is based on the different intensity curvatures predicted for each structure.

- Lung nodules are expected to have three large negative eigenvalues $\lambda_1 \approx \lambda_2 \approx \lambda_3 \ll 0$ indicating the increased attenuation inside the nodule in any direction.
- Vessels should have two large negative eigenvalues $\lambda_1 \approx \lambda_2 \ll 0$ and one small eigenvalue $\lambda_3 \approx 0$ corresponding to the eigenvector in direction to the trajectory of the vessel.
- Fissures have one large negative eigenvalue $\lambda_1 \ll 0$ corresponding to the eigenvector normal to the fissure plane, and two small eigenvalues $\lambda_2 \approx \lambda_3 \approx 0$ corresponding to the eigenvectors parallel to the plane.

Fissure Enhancement: A typical problem of structure enhancement filter, especially the fissure enhancement filter is the spurious response rate as shown in Fig.2(b),2(f) which possibly originates from local planar properties of voxels. To remedy this, a two stage cascaded classification method was presented in [19]. It used a second supervised classifier with enhanced features extracted from the output of the first classifier to correct the previously misclassified voxels due to spurious responses. In this paper, we propose

a new and much simpler method which is completely unsupervised to reduce the spurious responses. As shown in Fig.2(b) and 2(f), despite spurious responses, the fissure is also effectively enhanced and a more clear plate-like structure is shown in the filtered image. Accordingly, it is expected that the same plate-like structure enhancement filter (Eq.(6)) can be applied on the filtered image to further distinguish the fissure from the background. To explain this, we define a simplified plate-like structure model as below:

$$I(x, y, z) = (h(x + w_x/2) - h(x - w_x/2)) \cdot (h(y + w_y/2) - h(y - w_y/2)) \quad (7)$$

where $h(\cdot)$ is the unit step function, w_x and w_y is the plate width in x and y direction with assumption $w_x \ll w_y$. The resulting Hessian matrix eigenvalues are

$$\begin{aligned} \lambda_x &= G'_\sigma(x + w_x/2) - G'_\sigma(x - w_x/2) & |y| \leq w_y/2 \\ \lambda_y &= G'_\sigma(y + w_y/2) - G'_\sigma(y - w_y/2) & |x| \leq w_x/2 \\ \lambda_z &= 0 \end{aligned} \quad (8)$$

where $G'_\sigma(\cdot)$ is the first derivative of Gaussian function with standard deviation σ and zero mean. Fig.1 shows the curves of eigenvalue function $\lambda(x) = G'_\sigma(x+w/2) - G'_\sigma(x-w/2)$ with various w/σ . Two special cases of interest regarding this function are considered: (1) when $w/2 \gg \sigma$, $\lambda(x)$ has two lobes of opposite signs on each side of the edges $x = \pm w/2$. It reaches the maximum absolute value at $|x \pm w/2| = \sigma$ and approaches to zero for $|x \pm w/2| \geq 4\sigma$; (2) when $w/2 \approx \sigma$, $\lambda(x)$ has a central negative lobe with width approximately equal to w , and it can be proved that $\lambda(x)$ reaches the minimum at $x = 0$ if $w/\sigma \leq 2\sqrt{3}$. Hence for plate-like structures with $w_x \ll w_y$, we select the scale $\sigma \approx w_x/2$ so that $|\lambda_x/\lambda_y|$ is maximized for $|x| < w_x/2$ and $|y| < w_y/2 - 4\sigma$ but drops significantly when $|y|$ approaches to $w_y/2 - \sigma$. As a result, the plate-like structure will keep plate width w_x and shrink around the edges $y = \pm w_y/2$ after applying the filter Eq.(6). If w_y is not significantly higher than w_x , the whole structure will be reduced after applying the filter for multiple times. Fig.2(c), 2(g) and 2(d), 2(h) show the results of filtering for 3 and 6 times respectively, which clearly reduces the spurious responses.

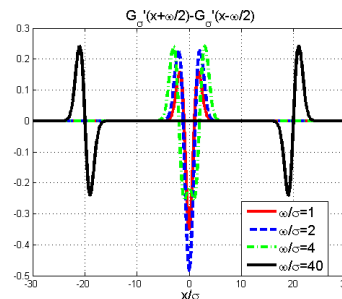


Figure 1. Curves of $\lambda = G'_\sigma(x + w/2) - G'_\sigma(x - w/2)$ as a function of x/σ with different w/σ .

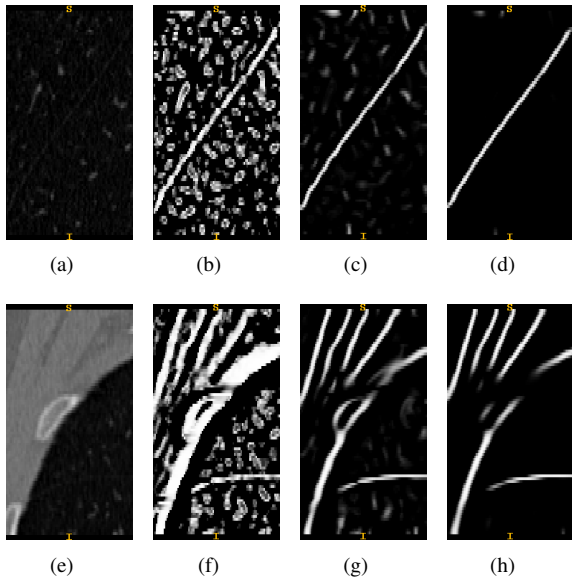


Figure 2. Example outputs from iterative fissure enhancement filter (a),(e) original slices; (b),(f) results of conventional filter; (c),(g) results of iterative filter (N=3); (d),(h) results of iterative filter (N=6). Filter scale σ is selected as 1 mm because typically the fissures are approximately 1-3 mm thick [13].

2.2. Learning the CRF

Ideally, the model parameters in Eq.(1) is learned using maximum *a-posteriori* (MAP) algorithm that maximizes the conditional likelihood of labels given the training data. However, this is computationally expensive because it usually requires the evaluation of marginal probabilities over the class labels at each training pixel. Here, we adopted a more pragmatic stepwise training method as used in [16] which estimates the unary and pairwise potentials separately.

A variety of supervised classifiers can be used to learn the unary texture-intensity-shape potential parameters of Eq.(2). In this paper, we selected two different classifiers, the resilient subclass discriminant analysis (RSDA) [23] as a generative classifier and the relevance vector machine (RVM) [18, 15] as a discriminative classifier.

RSDA approximates the distribution of each class as a mixture of Gaussians, and estimates the probability model parameters using expectation-maximization integrated with linear discriminant analysis. This method is naturally phrased in a multiclass framework by calculating a *posteriori* probability of each class as follows.

$$P(c|x_i) = \frac{\sum_{j=1}^{N_c} \rho_{c_j} g(x_i; \mu_{c_j}, \Sigma_{c_j})}{\sum_{c=1}^C \sum_{j=1}^{N_c} \rho_{c_j} g(x_i; \mu_{c_j}, \Sigma_{c_j})} \quad (9)$$

where C is the number of classes and N_c is the number of subclasses in class c . ρ_{c_j} , μ_{c_j} and Σ_{c_j} is the prior, mean vector and covariance matrix of subclass j in class c obtained by RSDA, respectively.

RVM is a discriminative classifier using Bayesian inference to obtain sparse solutions for classification. It was originally derived for two-class problems and can be extended to multi-class setting by training and combining a series of n one-vs-rest binary classifiers. RVM has identical functional form to the popular support vector machine (SVM) but offers additional benefits. These include automatic estimation of regularization coefficients and probabilistic predictions so that no cross-validation based post optimization or heuristic calibration prior to combining multiple binary classifiers is required.

$$P(c|x_i) = \frac{\sigma(y(x_i, w_c))}{\sum_{c=1}^C \sigma(y(x_i, w_c))} \quad (10)$$

where $\sigma(x) = 1/(1 + e^{-x})$ is the logistic sigmoid function and $y(x_i, w_c) = w_c^T x_i$ including the offset or $y(x_i, w_c) = w_c^T \phi(x_i)$ if a nonlinear kernel is used. w_c is the weight vector for class c learned via RVM.

For pairwise edge potential Eq.(3), the normalization parameter β was volume dependent and set to $2\|\bar{x}_i - \bar{x}_j\|^2$ where \bar{x}_i denotes average pixel intensity of class i in the volume given *a posteriori* probabilities derived from the unary classifier described above. The contrast parameter α was manually tuned to minimize error on the validation data.

3. Nodule Connectivity Recognition

The hard segmentation results of lung anatomies obtained above can be used directly to determine whether the detected nodules are solitary or attached using simple n -connected component algorithm. However, this will be highly sensitive to the possible voxel-level segmentation errors. In stead, we train a second classifier using higher level features extracted from probability maps output by the lower level classifier. These features exploit the spatial co-occurrence statistics of pulmonary structures to make the system more robust to the results of voxel-level segmentation. They are detailed in the next section.

3.1. Probability Co-occurrence Map

If two objects are connected or very close to each other, both of their probability maps will have higher responses in the same area where two objects overlap after one of the maps is moved by a small offset. Based on this, we define the probability co-occurrence map (PCM) κ as follows:

$$\kappa_{n,m}(x, y, z) = \sum_{(i,j,k) \in W} P_n(i, j, k) P_m(i+x, j+y, k+z) \quad (11)$$

where P_n and P_m are the probability maps of nodule and another class of interest obtained from the classifier Eq.(9) or (10). W is a local window around the detected nodule, and (x, y, z) is the offset. The probability co-occurrence map gives a quantitative measure of the possibility that a pair of objects co-occur in the volume with different spatial offsets. Eq.(11) is essentially the correlation of two

class probability maps where the nodule probability map $P_n(x, y, z)$ serves as a filter with kernel roughly in the shape and size of this nodule, which smooths the output to make the upper level classifier more robust to the voxel-level segmentation noise.

3.2. Higher Level Feature Set

As mentioned above, if the nodule is close to another structure, the probability co-occurrence map is expected to have higher responses for some small offset. Accordingly, we define a set of basic features which simply sums up all $\kappa(x, y, z)$ on the spherical surfaces with different radius $r = \sqrt{x^2 + y^2 + z^2}$ from 1 to 10. This is equivalent to measuring the correlation at discrete distances. Therefore a total of 10 basic and rotation invariant features are generated for each of the three nodule connectivity cases.

If the nodule is attached to fissure or vessel, the probability co-occurrence map κ will have higher responses in a plate-like or tubular area crossing the origin because κ is essentially the correlation of two probability maps and nodule is usually much smaller than fissure and vessel. Hence we can extract enhanced features from κ exploiting these specific structure shapes in a similar manner as the Hessian filters used in segmentation. This can make up for the voxel-level segmentation errors to some extent and is particularly useful for fissure connectivity classification. Here we define the enhanced feature similar to Eq.(6) for fissure as follows:

$$F = (1 - \exp(-m_\kappa^2/2\gamma^2)) \exp(-\lambda_3^2/2\beta^2\lambda_2^2) \exp(-d^2/2\alpha^2r_n^2) \quad (12)$$

where γ , β and α are the threshold parameters. m_κ is the average response of probability co-occurrence map in the local window W , r_n is the nodule size approximated as:

$$r_n = (3/4\pi \sum_{(x,y,z) \in W} P_n(x, y, z))^{1/3} \quad (13)$$

λ_1 , λ_2 and λ_3 are three eigenvalues of the weighted covariance matrix of all offset vectors $(x, y, z) \in W$ with normalized weights $\kappa_{n,m}(x, y, z)/(m_\kappa|W|)$, in descending order of absolute value $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$. d is the distance between the origin to the planar surface which is defined as $(\bar{x}, \bar{y}, \bar{z})^T v_3$ where $(\bar{x}, \bar{y}, \bar{z})$ denotes the weighted mean of the offset vectors, and v_3 is the eigenvector corresponding to the smallest eigenvalue λ_3 . If there exists a plate-like structure crossing the origin in the probability co-occurrence map κ , $|\lambda_3|$ will be much smaller than $|\lambda_2|$ with d smaller than the nodule size r_n .

Because a single nodule can be attached to multiple structures as shown in Fig.6, the nodule connectivity classification is actually a multi-label problem. Therefore, three binary RVM classifiers are trained and tested independently for each of the three connectivity cases on 10, 10 and 11 features respectively.

4. Experimental Results

In the experiment, we collected 784 lung CT subvolumes from 239 subjects. These subvolumes are $83 \times 83 \times 83$

voxels in size and each subvolume has a nodule in the center with a variety of appearances and contextual anatomy connectivity as shown in Fig.4. The labeled voxels of pulmonary structures were obtained from 34 randomly selected subvolumes of different patients by two radiologists based on visual assessment and consensus review. The radiologists were requested to sample evenly different parts of each structure, including major and minor lobar fissures, veins and arteries from hilum to peripheral level, parenchyma in the open area and close to the structures such as vessel and lung wall, solid and ground glass opacity nodules. In total, more than 200,000 sample voxels were marked with roughly 6%, 20%, 25%, 11% and 38% samples labeled as nodule, vessel, lung wall, fissure and parenchyma, respectively. Then a radiologist annotated the nodule connectivity of the 784 subvolumes, of which 254 cases are vessel attached, 150 nodules are pleural-based or lung wall connected, and 81 cases are fissure attached. Data were collected from multiple hospitals.

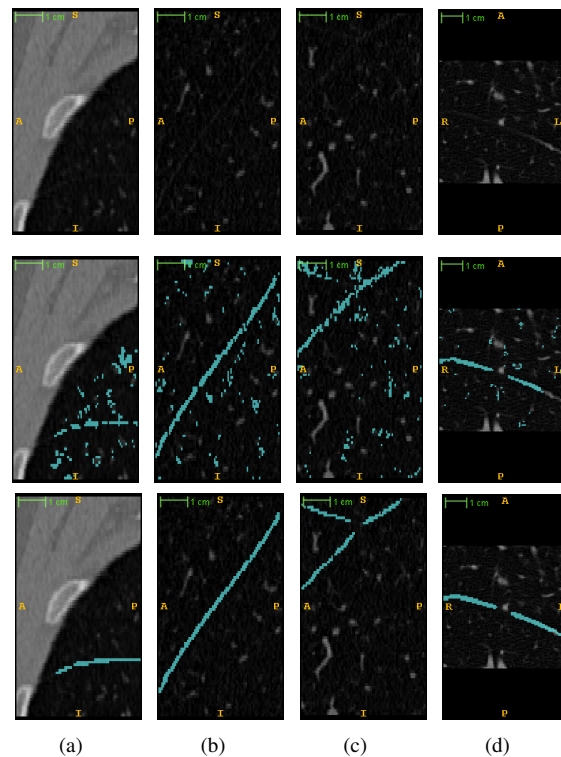


Figure 3. Example fissure segmentation results. Top row: original slices. Middle row: segmentation results with original fissure enhancement filter. Bottom row: segmentation results with iterative fissure enhancement filter.

4.1. Voxel-Level Lung Anatomy Segmentation

We first report the accuracy of voxel-level lung anatomy segmentation by unary classifier. The “voxel painted” 34 subvolumes were randomly split into two folds as the training and testing data. First of all, we compare the performance of two classifiers, RVM and RSDA, using the

Table 1. The pairwise confusion matrix of five pulmonary structure classes without CRF model. (a) RSDA (b) RVM

		(a)				
GROUND TRUTH	PREDICTION (%)					
	N	V	W	F	P	
NODULE	77.9	8.6	8.5	1.2	3.8	
VESSEL	3.3	90.0	3.9	1.4	1.4	
WALL	0.7	1.1	97.2	0	1.0	
FISSURE	1.0	3.3	0.1	83.6	12.0	
PARENCHYMA	0.5	0.6	0.6	1.4	96.9	

		(b)				
GROUND TRUTH	PREDICTION (%)					
	N	V	W	F	P	
NODULE	81.8	8.8	5.7	0.5	3.2	
VESSEL	6.8	87.2	3.0	0.9	2.0	
WALL	0.2	0.6	95.7	0.0	3.5	
FISSURE	0.8	2.4	0.0	84.7	12.1	
PARENCHYMA	0.1	0.4	0.4	0.5	98.6	

texture, intensity and shape features as described in Section 2.1. As shown in Table 1, the classification performance of two classifiers is very close, reaching **90.1%** and **89.3%** overall recognition rate by RSDA and RVM, respectively. In addition, both confusion matrices show similar pairs of classes where misclassification is more likely to occur. Fissures can be faint and incomplete thus they are more easily confused with parenchyma. Vessels and nodules are confused because vessel bifurcation or branch points may appear locally more blob-like as nodules and less tubular. These observations are consistent with [11] as the state-of-the-art and our classification results are better in fissure, nodule and parenchyma classes (by margins of 1.6% ~ 18.9%) and comparable in vessel¹. Without significant difference, we select RSDA as the classifier in all of the following experiments.

We then validate the fissure segmentation improvement by the iterative fissure enhancement filters in section 2.1. As shown in the middle row of Fig.3, the classification results using the original fissure enhancement filter features [11, 19] output many pepper-and-salt labeling noises (as short strokes in 3D). With our enhanced features (applying the filter iteratively for 3 and 6 times), the misclassification errors are significantly reduced so that they can be easily eliminated by a simple small connected component removal algorithm, as in the bottom row of Fig.3. Moreover, the proposed iterative filter is computationally efficient and it takes about 2 seconds to complete 6 iterations of filtering on one $83 \times 83 \times 83$ subvolume, implemented with Insight Toolkit (ITK) [8] on Intel Xeon platform. Nevertheless, the segmentation results break up where the fissure connects the

¹We have no access to the data used in [11] and comparison is numerical only.

Table 2. The classification performance of nodule connectivity using hard anatomy segmentation results and n -connected ($n = 6, 18, 26$) component algorithm. SN stands for sensitivity (%) and SP for specificity (%).

	VESSEL		WALL		FISSURE	
	SN	SP	SN	SP	SN	SP
N = 6	79.1	82.9	72.7	96.9	18.5	97.6
N = 18	79.9	81.0	78.0	96.1	21.0	96.8
N = 26	80.3	81.0	79.3	95.9	23.5	96.4

Table 3. The area under ROC curve (AUC) of nodule connectivity classification based on original probability maps and probability maps masked by CRF segmentation results.

	VESSEL	WALL	FISSURE
ORIGINAL	0.7793	0.9184	0.7555
MASKED	0.8676	0.9275	0.8318

Table 4. The performance of nodule connectivity classification. SN stands for sensitivity (%) and SP for specificity (%).

	VESSEL	WALL	FISSURE
AUC	0.8676	0.9275	0.8692
SN	82.0	93.3	83.8
SP	81.9	93.9	83.5

lung wall (Fig.3(a)), two or more fissures join (Fig.3(c)) or the fissure is attached to other structures like nodule or vessel (Fig.3(d)). This is reasonable because the assumption of locally plate-like structure does not hold in these places. A post-processing algorithm like fissure extension in [13] can be applied to fill up the discontinuities, if more accurate fissure segmentation is required.

Finally we segment the subvolumes by incorporating the pairwise constraints from edge cues and optimizing the conditional random field model via graph-cut [9, 3, 2]. The overall classification rate increases from 90.1% to 91.2%. While the improvement is numerically small, the contour of the structure, especially the nodule, is more accurately delineated which gives a considerable increase in perceived accuracy as shown in Fig.4. It is clear that a more precise segmentation of nodule boundaries is critical for the upper contextual classification system to classify nodule connectivity, which will be shown in the next Section.

4.2. Nodule Connectivity Classification

In this experiment, the 784 subvolumes are split into training and testing data via ten-fold cross validation which is repeated for 20 times and all results presented here are the averages over 20 runs. The subvolumes from the same patient are always kept within the same fold. We first test the performance using simple n -connected component algorithm with hard anatomy segmentation results. No training is required for this method. The experimental results in Table 2 show that sensitivity and specificity are quite unbalanced especially for the case of fissure connectivity, the

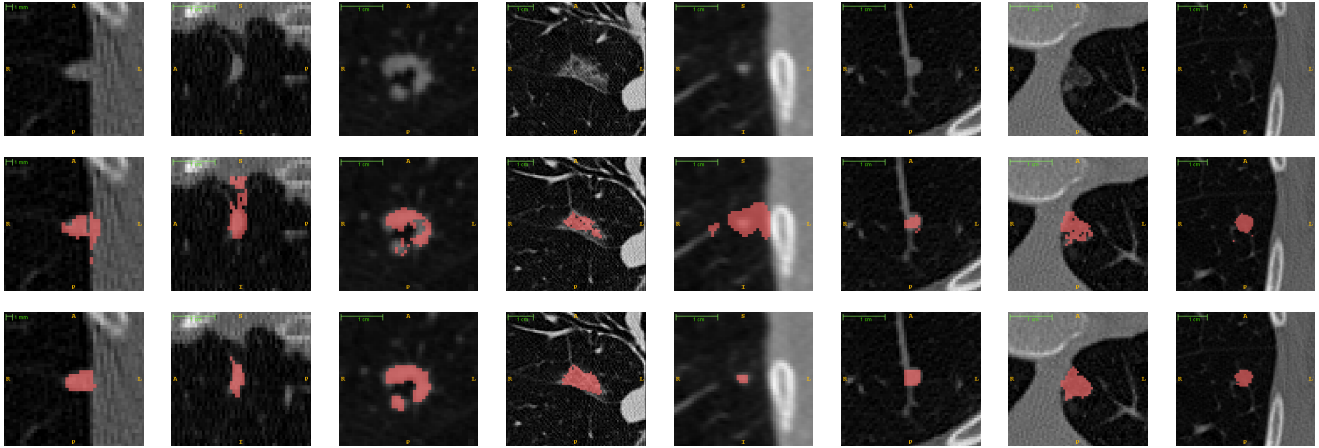


Figure 4. Example nodule segmentation results. Top row: original slices. Middle row: segmentation results without CRF model. Bottom row: segmentation results with CRF model.

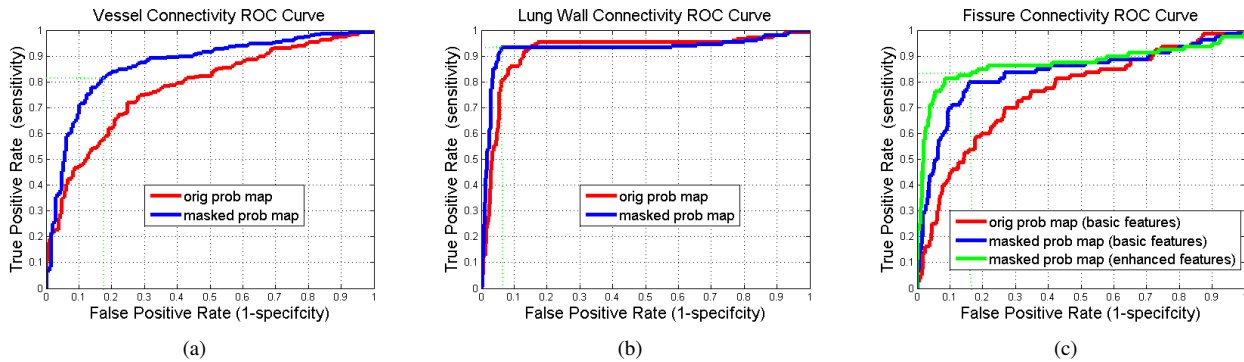


Figure 5. The ROC curves of lung nodule connectivity classification based on different probability maps. For fissure connectivity, enhanced feature of Eq.(12) is used and compared.

sensitivity or detection rate is very low. This can be explained by the above observation that the fissure segmentation misses where it is attached to lung nodules (Fig.3(d)). The trade-off between specificity and sensitivity can be achieved by changing n in connected component algorithm, but it is quite limited.

To make the upper-level system more robust to voxel-level segmentation errors and obtain complete ROC curves, we use the supervised classification method [15, 18] on 10 basic features extracted from soft Probability Co-occurrence Maps as described in Section 3.2. The improvement of nodule connectivity classification by introducing CRF model with pairwise edge cues in voxel-level anatomy segmentation is validated in Fig.5. The red ROC curves are based on original probability maps directly derived from unary classifier, and blue ROC curves use the probability maps masked by the hard segmentation results from CRF model. The area under ROC curves (AUC) increases in all three cases, by 0.0091 for wall and especially 0.0883, 0.0763 for vessel and fissure connectivity classification respectively, as given in Table 3.

Next for fissure connectivity classification, we experiment with the enhanced feature as described in Section 3.2.

The comparison is given in Fig.5(c) with AUC further increased from 0.8318 to 0.8692. The connectivity classification performance of all three structures is summarized in Table 4 with threshold values selected for balanced sensitivity and specificity as given by green dotted lines in Fig.5(a), 5(b) and 5(c). Finally, Fig.6 show examples of attached nodules which are correctly identified by the system with corresponding segmentation results. It clearly demonstrates that the upper level classifier can recognize the connectivity regardless of the missed fissure segments surrounding the nodule as described above, where n -connected component algorithm will fail.

5. Conclusion & Discussion

In this paper, we proposed a two level classification system to identify the nodule contextual anatomy connectivity in CT images. The use of local feature descriptors and single statistical classifier provides a simple and robust voxel-level segmentation (i.e., PRM) of pulmonary structures, which can be used as the input for post-processing algorithms to further improve the segmentation accuracy or higher learning systems, for example, the nodule connec-

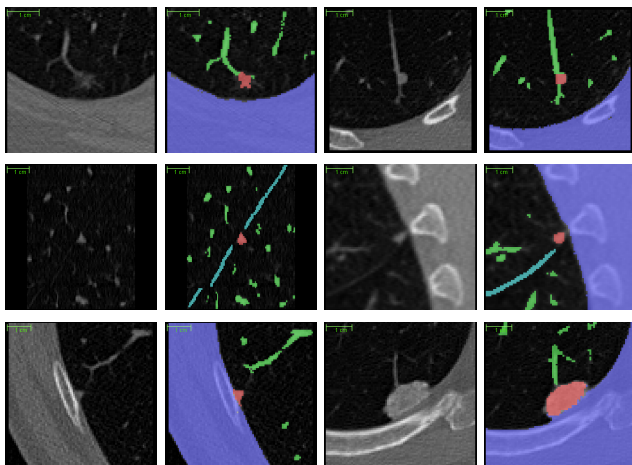


Figure 6. Examples of different nodule connectivity cases and segmentation results. Top Row: vessel attached. Middle Row: fissure attached, the second case is also wall attached. Bottom Row: lung wall attached, the second case is also vessel attached.

tivity classification as addressed in the paper. At this level, the iterative fissure enhancement filter and CRF model show success in the improvement of segmentation performance. Based on the segmentation results, a probability co-occurrence map (PCM) was presented to provide likelihood measures of two objects co-occurring in the same 3D volume with different spatial offsets. Higher level statistics features can be extracted from this map revealing the spatial relations between nodule and other anatomical structures.

Ongoing work will include extension of this hierarchical framework to lung airway which can be confused with surrounding parenchyma, especially for higher generation airway segments, due to local appearance resemblance of two structures, as with a thin and blurred layer of airway wall in the between. The combination with top-down process such as seeded region growing might be needed for better segmentation results. In addition, it is of clinical value to further classify those vessel attached nodules to vessel feeding or touching sub-categories (e.g., touching the vessel borders because of nodule growth), which demands more accurate and specific analysis of the connectivity. Also, accurate segmentation of pulmonary structures from abnormal parenchyma, caused by diffused lung diseases such as emphysema, is another challenging problem.

References

- [1] G. Agam, S. G. A. III, and C. Wu. Vessel Tree Reconstruction in Thoracic CT Scans With Application to Nodule Detection. *IEEE Trans. Med. Imaging*, 24(4):486–499, April 2005.
- [2] Y. Boykov and G. Funka-Lea. Graph Cuts and Efficient N-D Image Segmentation. *IJCV*, 70(2):109–131, 2006.
- [3] Y. Boykov, O. Veksler, and R. Zabih. Fast Approximate Energy Minimization via Graph Cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, November 2001.
- [4] J. Dehmshki, H. Amin, M. Valdivieso, and X. Ye. Segmentation of Pulmonary Nodules in Thoracic CT Scans: A Region Growing Approach. *IEEE Trans. Med. Imaging*, 27(4):467–480, April 2008.

- [5] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale Vessel Enhancement Filtering. *Proc. MICCAI*, 1496:130–137, 1998.
- [6] A. Gupta, A. Kembhavi, and L. Davis. Observing human-object interactions: Using spatial and functional compatibility for recognition. *IEEE Trans. PAMI*, 31(10), 2009.
- [7] G. Heitz and D. Koller. Learning spatial context: Using stuff to find things. *Proc. ECCV*, 2008.
- [8] Insight Segmentation and Registration Toolkit (ITK). <http://www.itk.org/>.
- [9] J. Lafferty, A. McCallum, and F. Pereira. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. *Proc. ICML*, pages 282–289, 2001.
- [10] L. Lu, K. Toyama, and G. D. Hager. A Two Level Approach for Scene Recognition. *Proc. CVPR*, 1:688–695, 2005.
- [11] R. A. Ochs, J. G. Goldin, and et al. Automated Classification of Lung Bronchovascular Anatomy in CT Using AdaBoost. *Med. Image Anal.*, 11(3):315–324, June 2007.
- [12] K. Okada, D. Comaniciu, and A. Krishnan. Robust Anisotropic Gaussian Filtering for Volumetric Characterization of Pulmonary Nodules in Multislice CT. *IEEE Trans. Med. Imaging*, 24(3):409–423, March 2005.
- [13] J. Pu, J. K. Leader, and et al. A Computational Geometry Approach to Automated Pulmonary Fissure Segmentation in CT Examinations. *IEEE Trans. Med. Imaging*, 28(5):710–719, May 2009.
- [14] A. Rabinovich, A. Vedaldi, C. Galleguillos, E. Wiewiora, and S. Belongie. Objects in context. *Proc. ICCV*, 2007.
- [15] V. Raykar, B. Krishnapuram, J. Bi, M. Dundar, and R. B. Rao. Bayesian multiple instance learning: automatic feature selection and inductive transfer. *Proc. ICML*, 808–815, 2008.
- [16] J. Shotton, J. Winn, C. Rother, and A. Criminisi. TextonBoost for Image Understanding: Multi-Class Object Recognition and Segmentation by Jointly Modeling Texture, Layout, and Context. *IJCV*, 81:2–23, 2009.
- [17] Y. Tao, L. Lu, M. Dewan, and et al. Multi-level Ground Glass Opacity Detection and Segmentation in CT Lung Images. *Proc. MICCAI*, 2009.
- [18] M. E. Tipping. Sparse Bayesian Learning and the Relevance Vector Machine. *JMLR*, 1:211–244, 2001.
- [19] E. M. van Rikxoort, B. van Ginneken, M. Klik, and M. Prokop. Supervised Enhancement Filters: Application to Fissure Detection in Chest CT Scans. *IEEE Trans. Med. Imaging*, 27(1):1–10, January 2008.
- [20] J. Wang, M. Betke, and J. P. Ko. Pulmonary Fissure Segmentation on CT. *Med. Image Anal.*, 10:530–547, August 2006.
- [21] Y. Wang, R. J. van Klaveren, and et al. Effect of Nodule Characteristics on Variability of Semiautomated Volume Measurements in Pulmonary Nodules Detected in a Lung Cancer Screening Program. *Radiology*, 248(2):625–631, August 2008.
- [22] G. Wei, L. Fan, and J. Qian. Automatic Detection of Nodules Attached to Vessels in Lung CT by Volume Projection Analysis. *Proc. MICCAI*, 2488/2002:746–752, 2002.
- [23] D. Wu and K. L. Boyer. Resilient Mixture Discriminant Analysis. *Proc. ICCV*, 2009.
- [24] D. M. Xu, H. J. van der Zaag-Loonen, and et al. Smooth or Attached Solid Indeterminate Nodules Detected at Baseline CT Screening in the Nelson Study: Cancer Risk during 1 Year of Follow-up. *Radiology*, 250(1):264–272, January 2009.
- [25] Y. Xu, M. Sonka, G. McLennan, J. Guo, and E. A. Hoffman. MDCT-Based 3-D Texture Classification of Emphysema and Early Smoking Related Lung Pathologies. *IEEE Trans. Med. Imaging*, 25(4):464–475, April 2006.
- [26] L. Zhang, E. A. Hoffman, and J. M. Reinhardt. Atlas-driven Lung Lobe Segmentation in Volumetric X-Ray CT Images. *IEEE Trans. Med. Imaging*, 25(1):1–16, January 2006.
- [27] J. Zhou, S. Chang, D. N. Metaxas, and et al. Automatic Detection and Segmentation of Ground Glass Opacity Nodules. *Proc. MICCAI*, 4190:784–791, 2006.