LOCATION CLASSIFICATION OF LUNG NODULES WITH OPTIMIZED GRAPH CONSTRUCTION

Yang Song[1], *Weidong Cai*[1], *Yue Wang*[2], *David Dagan Feng*[1,3,4]

[1]Biomedical and Multimedia Information Technology (BMIT) Research Group, School of Information Technologies, University of Sydney, Australia
[2]Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, USA
[3]Center for Multimedia Signal Processing (CMSP), Department of Electronic & Information Engineering, Hong Kong Polytechnic University, Hong Kong
[4]Med-X Research Institute, Shanghai Jiao Tong University, China

ABSTRACT

The locations of lung nodules relative to the other lung anatomical structures are important hints of malignant cancers. In this paper, we propose a fully automatic method to identify if a lung nodule is well-circumscribed, juxta-vascular, juxta-pleural or pleural tail in computed tomography (CT) images. First, we design an optimized graph model, introducing new global and region-based energy terms, to label each voxel as background or foreground in a single graph cut algorithm. Then, the texture features of a lung nodule are extracted based on the voxel labeling outputs, and its location information is inferred. We evaluate the proposed method on low-dose CT images, and demonstrate highly effective nodule classification results comparatively.

Index Terms— CT, lung nodule, classification, graph cut

1. INTRODUCTION

Lung cancer is the most common cause of cancer-related death. About 20% of patient cases with lung nodules represent malignant cancers [1], and the ability to distinguish malignant nodules from the benign ones is critical for prognosis. Recent studies show interests in investigating the correlation with nodule locations and suggest that lung nodules that are intra-parenchymal are more likely to be malignant than those attached to vessels or pleura [2]. While such retrospective studies help to advance the knowledge in lung nodule diagnosis, annotating nodule locations is time consuming and potentially prevents research on larger populations.

A computerized system that can automatically label the nodule locations is thus beneficial to medical researchers, and can also be applied in content-based image retrieval to assist image interpretation and diagnosis. The aim of our work is therefore to classify the lung nodules into four types [3]: (i) *well-circumscribed* without any connection with other structures; (ii) *juxta-vascular* with connection between the nodule and a vessel; (iii) *juxta-pleural* with a large portion of nodule surface abutting the pleural wall; and (iv) *pleural-tail* with a thin connection between the nodule and the pleural wall. Sample images are shown in Fig. 1 (odd rows).

While lots have been reported in segmentation and detection of lung nodules [4], there are limited works in analyzing the location contexts. A stratified learning approach was proposed to detect the nodule locations by first classifying the various types of lung anatomies [5], with a similar voxel labeling technique as [6]. Such a method was motivated by the intuition that if each voxel can be labeled correctly as certain type of lung anatomies (e.g. nodule, vessel, and pleural wall), the location information of lung nodules can thereafter be inferred. However, this approach may actually be over complicated, since detecting location contexts does not necessarily require accurate labeling of each voxel especially for the bordering areas between nodules and the other structures. Furthermore, it has been shown that, without any voxel labeling but based solely on texture characterization, satisfactory results in location inference could be achieved [7]. However, it might not deliver optimal performance without first parsing any anatomical information.

Therefore, here we propose a new method to classify nodule locations in CT images exploiting the advantages of both voxel labeling and context characterization. Our main contributions are two folds. First, other than labeling each voxel into various types of lung anatomies, the voxels are classified into only two categories: background (i.e. parenchyma) and foreground (i.e. nodule, vessel, and pleural wall). Such a two-class labeling leads to simplified feature space, and hence becomes more generally applicable to other problem domains. To achieve effective labeling, we design an optimized graph model with global and region-based energy terms besides the standard unary and pairwise terms. Second, to infer the lo-

This work was supported in part by ARC and PolyU grants.

cation information, a learning-based context characterization based on the voxel labeling outputs is used, which can also accommodate intra-class variations of the nodule types.

2. METHODS

2.1. Datasets

In this study, we use the publicly available Early Lung Cancer Action Program (ELCAP) database [8], which contains 50 sets of low-dose CT lung scans with 397 lung nodules annotated. The nodules are further classified into well-circumscribed (28.7%), juxta-vascular (14.2%), juxta-pleural (35%) and pleural-tail (22.1%) types. During preprocessing, a sub-window of 65×65 voxels is cropped from each image slice with the annotated nodule appearing in the center. Such preprocessing helps to restrict the problem scope to location classification of lung nodules only, and is similar to the handling in the related works [5, 7]. A set of randomly selected 40 images, with 10 for each nodule type, were used for training.

2.2. Graph-based Voxel Labeling

At the first stage, we aim to label each voxel as background or foreground. Denote an image (preprocessed) as $I = \{v_i : i = 1, ..., N\}$ with N voxels. To incorporate local spatial information and reduce spurious labeling due to noises, we employ the superpixel-based approach, by first performing over-segmentation using quick-shift clustering [9]. The image I is thus represented by $\{x_i : i = 1, ..., M\}$ with M superpixels ($M \ll N$). Each superpixel x_i is thus to be classified as B (background) or F (foreground) categories.

A new graph model for the labeling based on conditional random field (CRF) [10] is designed with the following energy function:

$$E(A|I) = \underbrace{\sum_i \phi(a_i)}_{\text{Unary}} + \underbrace{\sum_{i,j} \psi_L(a_i, a_j)}_{\text{Pairwise}} + \underbrace{(\phi_G(a_g) + \sum_i \psi_G(a_i, a_g))}_{\text{Global}} + \underbrace{(\sum_r \phi_R(a_r) + \sum_{i,r} \psi_R(a_i, a_r))}_{\text{Region-based}} \quad (1)$$

where A is the set of labels $a_i \in \{B, F\}$. The unary and pairwise terms follow the standard CRF formulations, and the global and region-based terms are newly introduced.

Graphical models have become popular for its ability to find a globally optimal labeling by integrating neighborhood-level interactions, and further enhancements to the basic model have recently been proposed for medical imaging to better describe the anatomy-specific features [11]. For our problem domain, as will be further described in the following, although CRF can assign a spatially smooth labeling with the pairwise term, accurate labeling for CT images requires features describing the intensity contrasts and region structures beyond the local neighborhood. Therefore, we choose to further optimize the basic CRF model by incorporating higher-order energy terms while maintaining the sub-modularity constraints for efficient solution using a single graph cut [12]:

$$A^* = \underset{A}{\operatorname{argmin}}\ E(A|I) \quad (2)$$

Unary term. The unary term $\phi(a_i)$ describes the cost of each node x_i in the graph labeled as B or F:

$$\phi(a_i) = 1 - P(x_i = a_i|f_i) \quad (3)$$

where f_i is the average intensity of x_i. A binary support vector machine (SVM) is trained to classify x_i based on its feature vector f_i, and the probability estimate is used as $P(.)$.

Pairwise term. The pairwise term $\psi_L(a_i, a_j)$ penalizes the differences in labeling of the neighboring superpixels x_i and x_j based on their feature distances:

$$\psi_L(a_i, a_j) = \delta(a_i - a_j) \cdot \exp(-\frac{\parallel f_i - f_j \parallel^2}{2\beta_L}) \quad (4)$$

where $\delta(.)$ is the Dirac delta function, and β_L is the normalization factor as the average of all L2 distances between neighboring feature vectors in image I.

Global term. The discriminative power of f_i is limited due to the high variances between images. Especially in some cases, the B areas might appear lighter than usual, which could then be misclassified as F. Nevertheless, the B areas should always be darker than the F areas in the same image. Therefore, the intensity contrast is incorporated as a more discriminative feature, and we model it as a global energy term. An extra node g of label B is added into the graph, representing the background of I, which is obtained by Otsu thresholding on I. The unary cost $\phi_G(a_g)$ is defined as:

$$\phi_G(a_g) = \left\{ \begin{array}{ll} 0 & \text{if } a_g = B \\ \gamma_g & \text{otherwise} \end{array} \right\} \quad (5)$$

where γ_g is the size of the background, which effectively constraints the global node as B category with the high cost. Each x_i is then connected to g, with the edge costs $\psi_G(a_i, a_g)$ penalizing the labeling differences:

$$\psi_G(a_i, a_g) = \delta(a_i - a_g) \cdot \mu_G \cdot \exp(-\frac{\parallel f_i - f_g \parallel_0^2}{2\beta_G}) \quad (6)$$

where f_g is the average intensities of the segmented background, $\parallel f_i - f_g \parallel_0^2$ is the L2 distance square or 0 if $f_i < f_g$, μ_G is the weight allocated to the global term, and β_G is the normalization factor.

Region-based term. Unlike nodules and pleural wall, the vessels in CT images are sometimes quite obscure and difficult to distinguish from the background. However, they usually follow a tubular shape, and detecting such structures in

a separate step can then help to improve the labeling accuracies around these regions. Inspired by the recent proposal of incorporating object detectors into CRF inference [13], we create the region-based term by first adding Q region nodes $R_r : r = 0, ..., Q$ into the graph, each corresponding to a detected vessel structure, represented by a set of connected superpixels $X_r = \{x_i : i = 1, ..., N_r\}$. The unary cost $\phi_R(a_r)$ of node R_r is defined similarly to Eq. (5) to represent the F category. An edge is connected between R_r and each $x_i \in X_r$ with the edge cost $\psi_R(a_i, a_r)$ defined as:

$$\psi_G(a_i, a_r) = \delta(a_i - a_r) \cdot \mu_R \tag{7}$$

where μ_R is the weight of the region-based term. This region-based term thus encourages consistent labeling for X_r. However, unlike a standalone object detector, by integrating the detection hypothesis into the CRF, $x_i \in X_r$ is not necessarily labeled as F, but inferred with the globally optimal solution based on all energy terms. To detect the vessel structures, we use the Hessian matrix [5] together with SVM classification on a voxel level. Then the superpixels around each detected vessel are further classified into B or F according to the intensities; and those labeled as F are finally chosen as X_r.

2.3. Context Characterization

The second stage of our method is to extract the characteristics of the nodules based on the voxel labeling outputs. The feature vector is to describe the spatial textures surrounding the nodule; and be rotation invariant to accommodate appearance differences of a certain nodule type. We thus choose to employ the popular SIFT descriptor [14]. Similar to the recent approach [7], the SIFT descriptor is extracted by fixing the keypoint location at the nodule centroid, and with an empirically chosen scale parameter, the SIFT algorithm then automatically determines the principal orientation and a 128-dimensional feature vector is computed. A four-class SVM is then trained to classify the feature descriptors into four types of nodule locations.

Although the annotated nodule centroid can be directly used as the keypoint location [7], we choose to refine it since the annotated location often does not align with the actual geometric centroid. And since the misalignments are usually small, we employ an exhaustive search method within a 5×5 window around the annotated centroid for a region that best approximates a circle with the following score function:

$$o = \underset{o}{\operatorname{argmax}} \ (\omega_1 n_o - \omega_2 e_o) \tag{8}$$

where n_o is the percentage of F voxels in the local window around o, and e_o is the eccentricity of the F region around o. And ω_1 and ω_2 are the associated weights, which are trained using a linear-kernel SVM. The score corresponds to the classification margin, and the location o producing the highest score is chosen as final centroid.

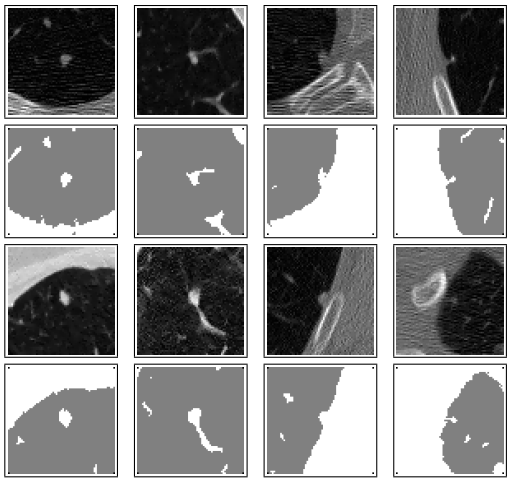

Fig. 1. Voxel labeling outputs of two example images for each nodule type: (a) well-circumscribed; (b) juxta-vascular; (c) juxta-pleural; and (d) pleural-tail. The first two rows are the first set of examples, and the last two rows are the second set, where the upper row shows the CT slices and the lower row shows the labeling results with white for foreground and gray for background.

3. RESULTS

Examples of voxel labeling on the four types of lung nodules are shown in Fig. 1, with nodules at the center of the preprocessed slices. The overall results suggest good labeling performance based on our visual inspection, and the contours of nodules are well depicted, which is especially beneficial for differentiating between juxta-pleural and pleural-tail nodules. Mislabelings arise in some areas with very faint vessel structures or nodules appearing less distinguishable from the surrounding parenchyma.

Fig. 2 demonstrates the advantages of the proposed global and region-based terms in the graph formulation. As shown in Fig. 2a, with the basic CRF, the areas around the nodules tend to be misclassified as foreground; and the global term helps to refine the labeling based on the average intensities of the foreground. Such improvements result in better delineation of the nodule boundaries, and are particularly useful for classifying the pleural attached nodules. Fig. 2b shows two examples of juxta-vascular nodules, where parts of vessels are labeled as background initially because of their darker appearances; and the region-based term helps to improve the labeling by incorporating the detector prior. As seen from the third image, although there is still a small gap between the nodule and vessel (due to imperfect vessel detector performance), the gap size is largely reduced, and hence becomes more likely to render correct nodule location with the SIFT descriptor.

The nodule classification rates are illustrated in Fig. 3,

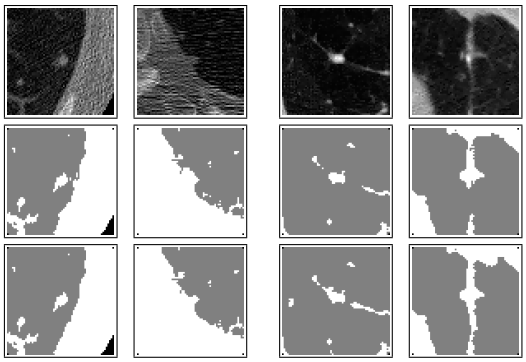

Fig. 2. Examples of labeling improvements due to the proposed (a) global and (b) region-based energy terms. Top row is the CT slices, middle row is the labeling results with basic CRF, and bottom row is the final labeling results with all four energy terms.

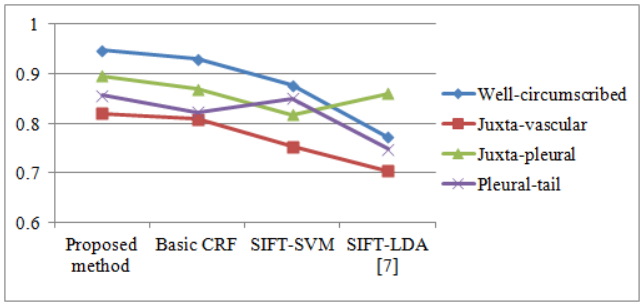

Fig. 3. The classification rates of the four types of nodules comparing various methods.

comparing the following four approaches: (i) our proposed method; (ii) basic CRF for voxel labeling; (iii) SIFT descriptor extracted based on raw voxel intensities without voxel labeling; and (iv) SIFT descriptor with linear discriminant analysis (LDA) as proposed in [7]. All approaches are trained using 50% of the data for the nodule classification step, with (i)-(iii) using SVM and (iv) using LDA. It is apparent that our proposed method achieves the highest classification rates; and by comparing results of (i) and (iii), an average 6% improvement introduced by the voxel labeling stage can be seen.

4. CONCLUSIONS

We present a new method for classifying lung nodule in CT images into four location types. The voxels are firstly labeled as foreground or background using a CRF construct with additional global and region-based terms, incorporating features describing the intensity contrasts and region structures. The SIFT features are then extracted originated from the refined nodule centroid for the nodule classification. Our evaluation on the ELCAP database shows higher classification performance with the voxel labeling approach.

5. REFERENCES

[1] J.J. Erasmus, J.E. Connolly, H.P. McAdams, and V.L. Roggli, "Solitary pulmonary nodules: part i. Morphologic evaluation for differentiation of benign and malignant lesions," *RadioGraphics*, vol. 20, pp. 43–58, 2000.

[2] D.M. Xu, H.J. Zaag-Loonen, M. Oudkerk, Y. Wang, R. Vliegenthart, E.T. Scholten, J. Verschakelen, M. Prokop, H.J. Koning, and R.J. Klaveren, "Smooth of attached solid indeterminate nodules detected at baseline CT screening in the nelson study: cancer risk during 1 year of follow-up," *Radiology*, vol. 250, no. 1, pp. 264–272, 2009.

[3] S. Diciotti, G. Picozzi, M. Falchini, M. Mascalchi, N. Villari, and G. Valli, "3-D segmentation algorithm of small lung nodules in spiral CT images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 1, pp. 7–19, 2008.

[4] S.L.A. Lee, A.Z. Kouzani, and E.J. Hu, "Automated detection of lung nodules in computed tomography images: a review," *Mach. Vis. Appl.*, pp. 1–13, 2010.

[5] D. Wu, L. Lu, J. Bi, Y. Shinagawa, K. Boyer, A. Krishnan, and M. Salganicoff, "Stratified learning of local anatomical context for lung nodules in CT images," *in Proc. CVPR*, pp. 2791–2798, 2010.

[6] R.A. Ochs, J.G. Goldin, F. Abtin, H.J. Kim, K. Brown, P. Batra, D. Roback, M.F. McNitt-Gray, and M.S. Brown, "Automated classification of lung bronchovascular anatomy in CT using adaboost," *Med. Image Anal.*, vol. 11, no. 3, pp. 315–324, 2007.

[7] A. Farag, S. Elhabian, J. Graham, A. Farag, and R. Falk, "Towards precise pulmonary nodule descriptors for nodule type classification," *in MICCAI 2010, LNCS*, vol. 6363, pp. 626–633, 2010.

[8] "ELCAP public lung image database," url-http://www.via.cornell.edu/databases/lungdb.html.

[9] A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking," *in ECCV 2008, LNCS*, vol. 5305, pp. 705–718, 2008.

[10] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: probabilistic models for segmenting and labeling sequence data," *in Proc. ICML*, pp. 282–289, 2001.

[11] V. Pamulapati, B.J. Wood, and M.G. Linguraru, "Intra-hepatic vessel segmentation and classification in multi-phase CT using optimized graph cuts," *in Proc. ISBI*, pp. 1982–1985, 2011.

[12] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, 2004.

[13] L. Ladicky, P. Sturgess, K. Alahari, C. Russell, and P.H.S. Torr, "What, where and how many? Combining object detectors and CRFs," *in ECCV 2010, LNCS*, vol. 6314, pp. 424–437, 2010.

[14] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV*, vol. 60, no. 2, pp. 91–110, 2004.