FALSE POSITIVE REDUCTION IN LUNG GGO NODULE DETECTION WITH 3D VOLUME SHAPE DESCRIPTOR

Ming Yang[‡]*, Senthil Periaswamy*[†]*, Ying Wu*[‡]

[†]Siemens Medical Solution
CAD Group
Malvern, PA 19355

[‡]Northwestern University
EECS Dept.
Evanston, IL 60208

ABSTRACT

Lung nodule detection, especially ground glass opacity (GGO) detection, in helical computed tomography (CT) images is a challenging Computer-Aided Detection (CAD) task due to the enormous variances in nodules' volumes, shapes, appearances, and the structures nearby. Most of the detection algorithms employ some efficient candidate generation (CG) algorithms to spot the suspicious volumes with high sensitivity at the cost of low specificity, e.g. tens even hundreds of false positives per volume. This paper proposes a learning based method to reduce the number of false positives given by CG based on a new general 3D volume shape descriptor. The 3D volume shape descriptor is constructed by concatenating spatial histograms of gradient orientations, which is robust to large variabilities in intensity levels, shapes, and appearances. The proposed method achieves promising performance on a difficult mixture lung nodule dataset with average 81% detection rate and 4.3 false positives per volume.

Index Terms— Medical imaging, computer aided analysis, computer vision, lung nodule detection, shape descriptor.

1. INTRODUCTION

Rapid growth in volume of lung nodule, i.e. small mass of tissues in lung, often reveals an early stage of lung cancer which is one of the leading fatal cancers in western countries. With early detection, the mean 5 year survival rate of lung cancer can dramatically increases from 14% to 49% [1], however, the vast amount of chest radiography to be interpreted manually are too burdensome for radiologists, which calls for an automation way to assist the diagnosis with CAD approach.

Although computer-aided lung nodule detection has been studied since early of 90s, it has not been taken into common clinical practice. The fundamental challenge is the lack of a clear definition and an in-depth understanding of nodules from the image perspective. Generally speaking, for object detection in vision, we need to know the model of the target or in simple words what we are looking for, i.e. what image attributes or features can reveal the presence of the targets and distinguish them from similar objects. Usually, the model comes from either the empirical knowledge or the training data, which largely determines the detection performance. Unfortunately, for lung nodules, we could only rely on very rough assumptions: the nodules often appear as small round shadows with spherical or ellipsoidal shapes in CT images. However, in practice, the detection system faces much more complicated cases: no nodules are perfect spherical but exhibit very large variations in intensities, sizes, shapes, appearances, and the surrounding structures, as shown in Fig. 1.

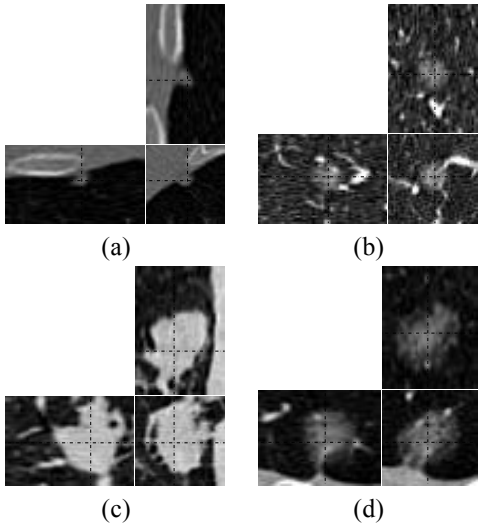

Fig. 1. Illustration of different types of nodules: (a) a small nodule attached to lung wall or pleural; (b) a vascularized nodule; (c) a solid large nodule with irregular non-spherical shape; (d) a non-solid nodule.

The lung nodules can be mainly categorized into 3 types: solid, part-solid and non-solid ones. Most of the previous research have focused on detection of solid nodules [2, 3, 4, 5, 6, 7, 8], while fewer attention is paid to part- and non-solid ones [9]. However, recent medical study indicated that non-solid or ground glass opacity (GGO) nodules, which have more irregular shapes and vague boundaries, are more likely to develop to malignant cancer than solid ones [10].

Besides the aforementioned difficulties, challenges for lung nodule detection also include vast data to be processed and

lack of positive samples. Typical thin slice CT screening has $512 \times 512 \times 400$ 12-byte data in one volume, therefore, in terms of efficiency, nodule detection task is usually split to two steps: 1) lung nodule candidate generation (CG) which aims to achieve high sensitivity, and 2) false positive reduction which aims to achieve high specificity. The CG algorithm is fast but may generate several hundred candidates per volume. The purpose of this paper is to study the false positive reduction problem in lung GGO nodule detection. Two examples of false positives are shown in Fig. 2.

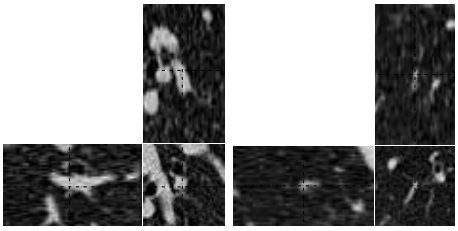

Fig. 2. Illustration of sample false positives.

Inspired by [11], we propose to employ a learning-based method to differentiate GGO with general 3D shape descriptors which do not explicitly assume the ellipsoidality of the GGO nodules. As the GGO nodules have much irregular shapes and vague boundaries, the features should be rough and not rigid. The new descriptor is based on weighted spatial histograms of gradient orientations, which is robust to different intensity levels, shapes and appearances. For each volume of interest (VOI) given by CG, the orientations of 3D gradients within a rough segmentation mask are represented by two angles, i.e. the angle when the gradient is projected to X-Y plane and the angle between the Z-axis and the gradient. The weighted 2D histograms of these two angles for 8 quadrants with the segment's geometrical center as the origin are concatenated to a vector as the 3D shape descriptor of this VOI. The weights of gradients are determined by the product of gradient magnitudes and a spatial kernel. An SVM classifier trained on these descriptors is used to classify the candidates to nodules and non-nodules so as to reduce the false positives. The proposed method achieves promising results on a mixture dataset including 324 nodules (81 GGOs) for 216 patients [1].

2. RELATED WORK

Quite a lot of different approaches have been proposed towards lung nodule detection since early of 90s. There are two primary factors in most of algorithms: features and classification models. Features extracted from raw radiography data summarize the 3D data in concise form and provide the evidences for classification models. Classification models grasp the essential discriminative power inside the features and generalize to unseen data.

[1]We would like to acknowledge Dr. M.Ujita (Jikei Medical University, Japan), Dr. J.Ko and Dr. D.Naidich (NYU Med Ctr, USA) for their clinical contributions to this work.

The features have been applied to lung nodule detection are extensive, ranging from simple image statistics, e.g. edges, gradient magnitudes, hough transform, circularity of region, contrastness and area, to sophisticated ones e.g. wavelets, radial volume distribution [6], and surface normal overlap [5] et al. [2] presented a comprehensive survey about lung nodule detection methods prior to 2001.

The classification methods can be mainly categorized to two types: template-based [3, 4, 12] and learning-based [6, 7, 8]. Template-based methods have to be able to tolerate the deviation from the template using robust matching criteria. In addition, there may be some unusual nodules or non-nodules which do not satisfy the template. For learning-based models, the training data have to cover the possible variances for the nodules and non-nodules as much as possible, which is also very hard to be satisfied. The state of the art of lung nodule detection for solid nodules was achieved by asymmetric cascade of sparse hyperplane classifiers [8] with 87.5% sensitivity *vs.* 1 false positive per volume in a particular dataset.

GGO nodules are more difficult than the solid ones for both template-based and learning-based methods mainly due to the large variations. It is not only difficult to obtain a template to tolerate the variations but also hard to gather sufficient training data. The intensities and appearances of GGO nodules vary greatly, therefore most of the conventional simple image statistics cannot capture the essentials of GGO and it is hard to design templates. [9] tried Gaussian reference models to detect the GGO candidates and applied artificial neural networks to reduce the false positives. 25 GGO were detected with approximately 543 false positives in 715 CT slices which amounts to more than 30 false positives per volume. In this paper, we employ SVM classifier trained on a new general 3D shape descriptors based on limited assumptions that the shapes of the GGO nodules, i.e. the spatial distributions of gradient orientations, are similar.

3. 3D VOLUME SHAPE DESCRIPTOR

Since clear prior knowledge about GGO nodules is not available, strong assumptions such as ellipsoidal shapes with clear surfaces are not proper. From our observation, Hounsfield values of GGO nodules and even the magnitudes of the gradients vary a lot. Therefore, with the assumption that the shape patterns of the GGO nodules are similar, we propose to construct the shape descriptor based on the gradient orientations as features for false positive reduction. With a rough estimation of the nodule segment, the histograms of gradient orientations for 8 quadrants are concatenated to represent the shape. The contribution of each gradient is weighted by the gradient magnitude and a spatial kernel to emphasize the voxels on the edges and those close to the segment boundaries. This shape descriptor has the advantages of generality, sparseness, and roughness. It can describe the rough shapes of the nodules without explicit assumptions. Even the nodule is attached to lung wall or other structures, majority of the histograms of

gradient orientations will agree with that of other GGOs.

Given the marker points of nodule candidates generated by CG, we first need to estimate the segments of these candidates. In our approach we perform the nodule segmentation with an efficient divergent gradient field response (DGFR) [13] method. For each marker of nodule candidate, a 43-voxel cubic VOI is extracted. 43 is selected empirically since it is sufficiently large for most GGO nodules. DGFR segmentation followed by a watershed algorithm labels all voxels in the VOI. With the marker as the seed we employ breadth-first search to obtain the segmentation by region growing. In case, the segmentation cannot provide a valid segments, we regard this candidate as non-nodule.

With the binary mask of nodule candidate given by rough segmentation, we align the candidate before calculating the descriptor. By performing PCA on the 3D coordinates of the points inside the binary mask, we rotate the VOI and the mask according to their primary orientations. The flow chart of the algorithm is summarized in Fig. 3. The details about the entire procedure are presented in next section as well as the classifier training and cross-validation scheme.

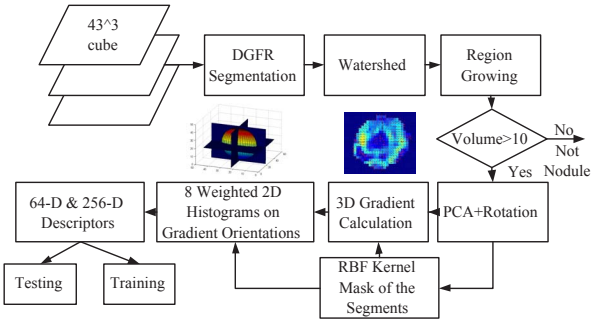

Fig. 3. Flowchart of the algorithm procedure.

4. FALSE POSITIVE REDUCTION PROCEDURE

4.1. Segmentation

Given the $43 \times 43 \times 43$ 3D VOI $\mathbf{V}$ and the marker point $s = \{s_x, s_y, s_z\}$, we need to know the scope of the candidate which will be used to generate the shape descriptor. The watershed algorithm is applied to estimate the segmentation based on DGFR [13] and followed by a region growing procedure with breadth first search. The segmentation module gives a binary mask $\mathbf{M}$ of the nodule candidate, where

$$\mathbf{M}(x,y,z) = \begin{array}{ll} 1 & \text{voxel } (x,y,z) \text{ is on the segment} \\ 0 & \text{voxel } (x,y,z) \text{ is not on the segment.} \end{array} \quad (1)$$

Note, if the DGFR cannot provide a reasonable mask, i.e the volume is less than 10 voxels, this candidate will be regarded as non-nodule and excluded from training and testing.

The segmentation is carried out for both full and half resolutions. In terms of scale selection, we have tried several schemes and find empirically the scheme that selects the one which gives larger difference of Gaussian (DOG) [11] response on s works well, which is the approximation of Laplacian operator. If the segment is small, i.e. the volume is less than 512 voxels, to facilitate calculation of the histograms, we up-sample the segment and the mask to $16 \times 16 \times 16$.

4.2. Descriptor Generation

For the segment $\mathbf{M}(x,y,z) = 1$, we align to its 3 primary directions with PCA by rotating around the geometrical center $g = \{g_x, g_y, g_z\}$. The rotated VOI and mask are denoted as $\mathbf{V}_r$ and $\mathbf{M}_r$. The 3D gradients of voxels inside the segment are calculated with 3D Sobel operators. Each gradient can be represented as a 3D vector $\{dx, dy, dz\}$ with the magnitude $\mathbf{D_m}(x,y,z)$. The orientation can be represented by two angles θ_{XY} and θ_Z:

$$\begin{aligned} \theta_{XY}(x,y,z) &= \arctan(\frac{dy}{dx}) \\ \theta_Z(x,y,z) &= \arctan(\frac{\sqrt{dx^2+dy^2}}{dz}), \end{aligned} \quad (2)$$

where θ_{XY} is the angle of the 3D gradient when projected to X-Y plane and θ_Z is the angle between the gradient and Z-axis.

Using the geometrical center g as the origin, the 3D volume is divided to 8 quadrants, as illustrated in Fig. 3. A weighted 2D histogram $H_i, i = 1, \cdots, 8$ on $\{\theta_{XY}, \theta_Z\}$ is calculated for each quadrant. The weight is the product of the gradient magnitude $\mathbf{D}_m(x,y,z)$ and a spatial kernel

$$\mathbf{K}_m(x,y,z) = 1 - G(g_x,\sigma_x)G(g_y,\sigma_y)G(g_z,\sigma_z), \quad (3)$$

which gives more weights to the gradients near the segment boundaries. The bin of the 2D histogram is calculated as

$$\begin{aligned} H_i(u) &= \sum \mathbf{D}_m(x,y,z)\mathbf{K}_m(x,y,z) \\ &\quad \delta(\theta_{XY}(x,y,z) - u_{XY}, \theta_Z(x,y,z) - u_Z), \end{aligned} \quad (4)$$

where δ is the Kronecker delta function and θ_{XY} and θ_Z are quantized to bin (u_{XY}, u_Z). In the spatial Gaussian kernel $G(g_x, \sigma_x)$, the variance σ_x is set to the radius of the segment in X direction, so on so forth.

The weights are mainly determined by the gradient magnitudes. So to avoid orientations of gradients with extremely large magnitudes dominate the histograms, we set a threshold to alleviate their influences. If one bin exceeds 0.2 we truncate it to 0.2 and iteratively re-normalize the histogram.

The 8 2D histograms are concatenated to construct the final shape descriptor. Two kinds of descriptors with different lengths are implemented. In the long descriptors, θ_{XY} is quantized to 8 bins where each bin stands for 45^o and 4 bins for θ_Z, so the length $8 \times 32 = 256$. In the short descriptors, 4 bins for θ_{XY} and 2 bins for θ_Z, that is $8 \times 8 = 64$ dimensional.

4.3. Classification

SVM classifier, specifically LibSVM [14], is employed to classify the candidates based on the proposed 3D shape descriptor. The GGO nodules are rare compared with the false positives. To deal with the unbalanced training data, besides setting higher costs for positive data in training, we also rotate the GGO samples to generate more positive training data. In testing, the shape descriptors are calculated for the remaining GGO nodules and false positives and classified with the trained SVM classifiers.

5. EXPERIMENTAL RESULTS

5.1. Settings

We test the performance of 64- and 256-dimensional descriptors on a mixture dataset including 216 volumes with 324 nodules (81 GGOs) and 9590 false positives given by the CG. Each volume has 400 slices on average and the sizes of nodules range from 2mm to 18mm. The dataset is divided to training and validation sets as shown in Table. 1. P(T) and N(T) indicate the numbers of positive and negative samples used in training, while P(V) and N(V) represent the numbers of positive and negative samples in validation set. Note that each positive training sample is rotated 5 times to generate more positive samples for training. For each setting, we evaluate the performance with the average sensitivity (Sens.) and false positives per volume (FPs.) for 100 tests as well as the standard deviation of sensitivity (Std.) as in Table. 1.

Table 1. Cross validation performance.

	P(T)	N(T)	P(V)	N(V)	Sens.	Std.	FPs.
64D	224	567	100	9590	79.9%	4.86%	5.33
	254	567	70	9590	80.8%	4.78%	5.63
256D	224	567	100	9590	80.2%	3.97%	5.18
	254	567	70	9590	81.0%	4.94%	4.31

5.2. Performance

For the 9590 false positives, 4905 are excluded by the DGFR segmentation since the volumes of segments are not valid. We test the performance on 2 validation sets with randomly selected 70 and 100 nodules respectively. The ROC curves are obtained by adjusting the variance of the kernel in SVM training. The costs for positive samples are set to 100 versus 1 for negative samples and the regularization term is 1000. As shown in Fig. 4, the lines link the mean sensitivities and the error bars show their standard deviations, which demonstrate that the performance of 64-dimensional and 256-dimensional descriptors are comparable.

6. CONCLUSION

In this paper, we propose to use new 3D volume shape descriptors based on spatial histograms of gradient orientations as features to reduce the false positives in lung GGO nodule detection. The advantages of the proposed shape descriptors are their generality and the strong generalization capability to remove false positives. The descriptors can be combined with other features to further improve the performance and have the potential to be applied in other medical image analysis applications, e.g. colon polyp and tumor detection.

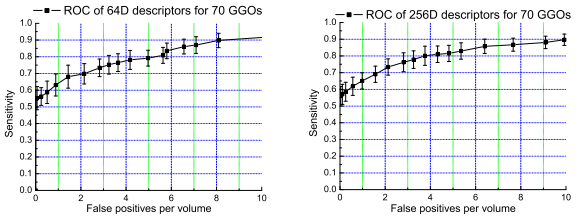

Fig. 4. ROC curves of 64D and 256D descriptors for testing 70 GGO nodules, respectively.

7. REFERENCES

[1] M. Wolf and et al., "CAD performance analysis for pulmonary nodule detection on thin-slice mdct scans," in *CARS*, June 2005.

[2] B. van Ginneken, B.M.ter H. Romeny, and M.A. Viergever, "Computer-aided diagnosis in chest radiography: A survey," *IEEE Trans. Med. Imag.*, vol. 20, no. 12, pp. 1228 – 1241, dec. 2001.

[3] Y.Lee, T.Hara, H. Fujita, S. Itoh, and T. Ishigaki, "Automated detection of pulmonary nodules inhelical CT images based on an improved template matching techniques," *IEEE Trans. Med. Imag.*, vol. 20, no. 7, pp. 595 – 604, jul. 2001.

[4] M.S. Brown, M.F. McNitt-Gray, and et al., "Patient-specific models for lung nodule detection and surveillance in CT images," *IEEE Trans. Med. Imag.*, vol. 20, no. 12, pp. 1242 – 1250, dec. 2001.

[5] D.S. Paik, C.F. Beaulieu, and et al., "Surface normal overlap: A computer-aided detection algorithm with application to colonic polyps and lung nodules in helical ct," *IEEE Trans. Med. Imag.*, vol. 23, no. 6, pp. 661 – 675, jun. 2004.

[6] X. Lu, G. Wei, J. Qian, and A.K. Jain, "Learning-based pulmonary nodule detection from multislice CT data," in *CARS*, 2004.

[7] P. Campadelli, E. Casiraghi, and G. Valentini, "Lung nodules detection and classification," in *ICIP*, sept. 2005, vol. 1, pp. 1117 – 1120.

[8] J. Bi, S. Periaswarmy, K. Okada, T. Kubota, G. Fung, M. Salganicoff, and B. Rao, "Computer aided detection via asymmetric cascade of sparse hyperplane classifiers," in *SIGKDD*, Philadelphia, PA, Aug. 2006, pp. 837–844.

[9] H.A. Bastawrous, T. Fukumoto, N. Nitta, and M. Tsudagawa, "Detection of ground glass opacities in lung ct images using gabor filters and neural networks," in *IMTC*, May 2005, pp. 251 – 256.

[10] C.I. Henschke and et al., "CT screening for lung cancer: Frequency and significance of part-solid and nonsold nodules," *American Journal of Roentgenology*, vol. 178, no. 5, pp. 1053–1057, 2002.

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

[12] A. Farag and et al., "Detection and recognition of lung nodules in spiral CT images using deformable templates and bayesian post-classification," in *ICIP*, oct. 2002, vol. 5, pp. 2921–2924.

[13] L. Bogoni, J. Liang, and S. Periaswamy, "System and method for toboggan based object segmentation using divergent gradient field response in images," 2005, US patent: 20050185838.

[14] C.-C. Chang and C.-J. Lin, *LIBSVM: a library for suppoert vector machines*, 2001, Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.