Quantification of Nodule Detection in Chest CT: A Clinical Investigation Based on the ELCAP Study

Amal A. Farag, Shireen Y. Elhabian, Salwa A. Elshazly and Aly A. Farag
Computer Vision and Image Processing Laboratory
University of Louisville, Kentucky, USA

Abstract. This paper examines the detection step in automatic detection and classification of lung nodules from low-dose CT (LDCT) scans. Two issues are studied in detail: nodule modeling and simulation, and the effect of these models on the detection process. From an ensemble of nodules, specified by radiologists, we devise an approach to estimate the gray level intensity distribution (Hounsfield Units) and a figure of merit of the size of appropriate templates. Hence, a data-driven approach is used to design the templates. The paper presents an extensive study of the sensitivity and specificity of the nodule detection step, in which the quality of the nodule model is the driving factor. Finally, validation of the detection approach on labeled clinical dataset from the Early Lung Cancer Action Project (ELCAP) screening study is conducted. Overall, this paper shows a relationship between the spatial support of the nodule templates and the resolution of the LDCT, which can be used to automatically select the template size. The paper also shows that isotropic templates do not provide adequate detection rate (in terms of sensitivity and specificity) of vascularized nodules. The nodule models in this paper can be used in various machine learning approaches for automatic nodule detection and classification.

Keywords: Parametric Templates Matching, Lung Nodule Definitions, Sensitivity and Specificity of CAD systems.

1 Introduction

The goal of computer-based nodule analysis methods is to assist the radiologists in early detection of presumable nodules. Assistance means to be able to mimic what a physician does in detecting and judging doubtful nodules. In the United States, lung cancer accounts for over 30% of all cancer–related deaths, resulting in over 160,000 deaths per year [1]. That is more than the annual deaths from colon, breast, pancreatic, prostate, and ovarian cancers combined. Lung cancer survival is strongly dependent on the pathologic stage at the time of diagnosis [2][3]. The hope is that early detection of lung cancer can improve the survival rate of this disease, thus research studies to reach an optimal detection rate is important. Should the use of LDCT scans become a standard clinical practice (for example, as a component of annual physical exams), an automatic way to analyze the scans will lend great benefit for the entire healthcare system; e.g., [4]-[7] and extensive surveys in [8][9].

The generalized framework for a CAD system consists of four main steps (see, Figure 1): 1) Filtering and normalization of the LDCT scans. 2) Segmentation of the lung regions (parenchyma) from the surrounding tissue. 3) Detection of lung nodules and 4) Nodule classification as benign or malignant.

Fig. 1: A block diagram of the major steps involved in computer-based analysis of LDCT of the chest in order to detect and classify doubtful lung nodules.

This paper will focus on the third step; nodule detection. The ELCAP dataset [10] is used for nodule design and testing and the sensitivity and specificity of the template matching approach in terms of detection is studied.

Since the early 90's various approaches for automated pulmonary nodule detection have been introduced. These approaches can be categorized as follows [4]-[7]: model based and density-based approaches. Template matching is one technique for model-based approaches which utilize a priori information of the size, intensity and shape of the nodules. Density-based approaches uses the fact that the lung parenchyma has relatively lower CT attenuation (density) than those of the lung nodules, thus they utilize image processing algorithms that rely on multiple thresholding, region growing and clustering. Various other approaches from the computer vision literature have been employed as well (e.g. [14]), which will not be surveyed due to space limitations.

Since the components of the CAD system in Fig. 1 are serial, we will briefly mention the techniques we employed in the filtering of scan artifacts, and segmentation of the lung tissues. For extensive surveys of various approaches on LDCT CAD research, the reader is directed to [8][9].

This paper is organized as follows: section 2 briefly discusses filtering of the scanning artifacts and segmentation of the lung region from the surrounding tissues; section 3 discusses template modelling and simulation; section 4 discusses template matching; section 5 presents experimental results and validation; and section 6 concludes the paper.

2 Artifacts Removal and Lung Segmentation in LDCT Scans

Filtering: During the acquisition of CT scans the generated slices can be affected by contrast, resolution, noise, artifacts and/or distortion. In particular, a slice can be corrupted by random fluctuations in image intensity which is considered a source of noise, thus, a CT slice is required to be filtered prior to further processing such as

segmentation. In this work we use the anisotropic diffusion filter (e.g., [11]), which is a well proven approach in various image analysis applications.

Image segmentation: as the name implies, connotes dividing or separating the information content in an image (or volume of images) into recognizable classes. It is a very fundamental and important step in image analysis. Various approaches have been used to segment the lungs from the rest of the organs that show the chest CT scans; again, the reader is referred to the surveys in [8][9]. The segmentation algorithm employed in this paper (similar to Armato et al [5]) exploits the intensity characteristics of lung CT scans, which appears as two distinct modes in the gray level intensity (Hounsfield Units) histogram. Multi-level thresholding is used to first isolate the thoracic region from the CT slice background; the lung parenchyma is then extracted from the segmented thoracic region, Fig. 2.

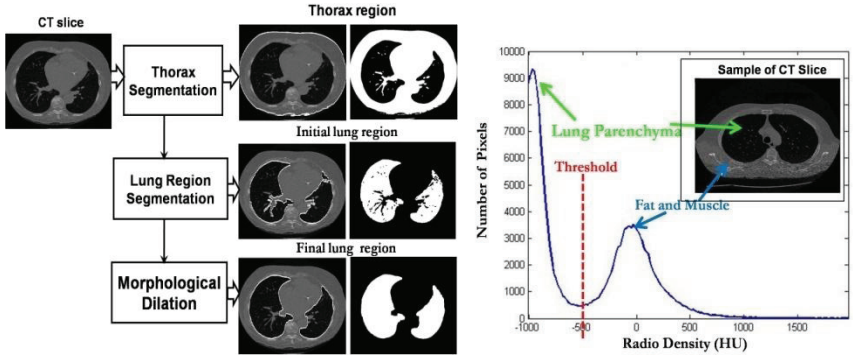

Fig. 2: *A Block diagram of the segmentation algorithm. Lung region (parenchyma) and fat/muscle region constitute two dominant peaks in the histogram, in order to separate the lung region a threshold is chosen to maximize the separation.*

Morphological dilation is then applied as a smoothing filter on the contour of the segmented lung region in order to avoid losing nodules that may be attached to the lung walls. To decrease the sensitivity of the segmentation result to the structuring element diameter, we apply it to the inner and outer lung region contour. After segmentation was completed small nodules attached to the pleural surface were found to no longer exist since these nodules were segmented as not belonging to the lung parenchyma. This operation resulted in 6.5% of the ground truth nodules to be excluded from further experimentations. Various improvements to the segmentation are possible; the algorithm employed here is 93.5% accurate with respect to retaining the nodules.

3 Nodule Modeling and Simulation

Proper nodule modeling is crucial for successful detection and classification. Modeling involves the shape, spatial support and the appearance (intensity) of the

template. The focus of this paper is on data-driven approaches for nodule modeling and simulation.

3.1 Nodule Types

A pulmonary nodule usually has a spherical shape; however, it can be distorted by surrounding anatomical structures such as vessels and the pleural surface. We shall use the nodule classification of Kostis et al [12][13], which groups nodules into four categories:

i. well-circumscribed where the nodule is located centrally in the lung without being connected to vasculature;
ii. vascularized where the nodule has significant connection(s) to the neighboring vessels while located centrally in the lung;
iii. pleural tail where the nodule is near the pleural surface, connected by a thin structure;
iv. and juxta-pleural where a significant portion of the nodule is connected to the pleural surface.

3.2 Nodule Simulation

In a CT scan the nodules can take various shapes and topologies, but the common characteristic amongst the nodules is the density distribution that tends to be concentrated around a region with an exponential decay (e.g., [7][15][16]). To illustrate this behavior, Fig. 3 shows the radial distance for each of the nodule types. Figure 4 shows the average distribution of HU densities for each nodule type. It can be observed that nodules' density distribution exhibit a bi-modal distribution (two dominant peaks) where the lower density mode represents regions surrounding the nodules and the higher density mode represents nodules densities. A linear model can be used to fit the joint density.

Using Gaussian kernels in the linear model, we employ the EM algorithm to obtain the joint and marginal density estimates. Figure 5 shows the marginal density estimates of the four nodule types. The range of intensity (HU) values is within two quantities q_{min} and q_{max} which will be used to estimate the density of a nodule template. From the knowledge of the nodule density (HU) vs. radial distance, and the nodules' marginal probabilities, we can compute the probability density function of the nodule's intensity (HU) with respect to radial distances from nodule centroid. These densities have been shown to be concentrated in radial distances around 5 pixels.

The behavior of the nodule intensity (HU) vs. radial distance provides three clues: 1) the intensity (HU) decays exponentially from the centroid; 2) the probability density of the intensity (HU) is concentrated within a range q_{min} and q_{max}; and 3) the probability density of the HU vs. radial density is concentrated below 10 pixels (hence, the spatial support of the templates is within 21x21 pixels). The first two clues

enable us to devise a formula for estimating the density of a nodule model (template). The third clue provides a figure of merit of how big the nodule template should be.

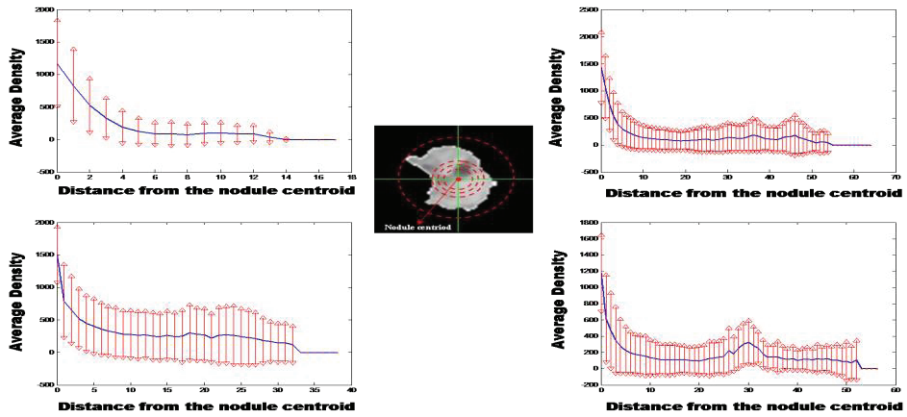

Fig. 3: *The distribution of the radial distance from the centroid of the nodules. The bars are one standard deviation off the mean values of each nodule type. Note the exponential behavior of the radial distance, and that it diminishes after a distance of 10.*

Given the range q_{min} and q_{max} of nodule density distribution (Fig. 5), obtained from the ground truth nodules (an ensemble of nodules outlined by the radiologists), the intensity or HU, at a distance *r* from the centroid, can be estimated by the following equations.

$$q(r) = q_{max} e^{-(r/\rho)^2}, \; 0 \leq r \leq \mathrm{R} \qquad (1)$$

$$\rho = R(ln(q_{max}) - ln(q_{min}))^{-1/2} \qquad (2)$$

These equations can be used to simulate parametric and non-parametric templates. For example, to simulate a circular template, we simply need to specify the radius R, and the intensity (HU) of the inside of the template at a distance *r* from the center can be calculated from Eq. 1 and 2. In this case, any of the four nodule intensity distributions in Fig. 5 may be used to obtain q_{min} and q_{max}. We should point out that this approach is valid for 2D and 3D non-parametric template designs as well.

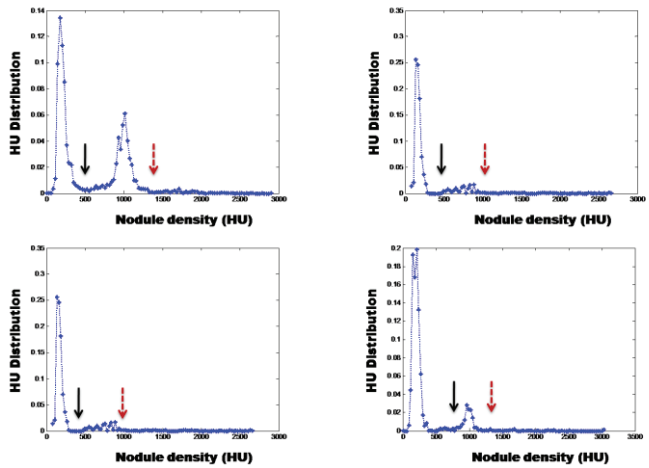

Fig. 4: *Average distribution of HU densities for each nodule type. The middle region corresponds to the nodule. The solid arrow is q_{min} and the dashed arrow (red) is q_{max} in Eq. 1 and 2.*

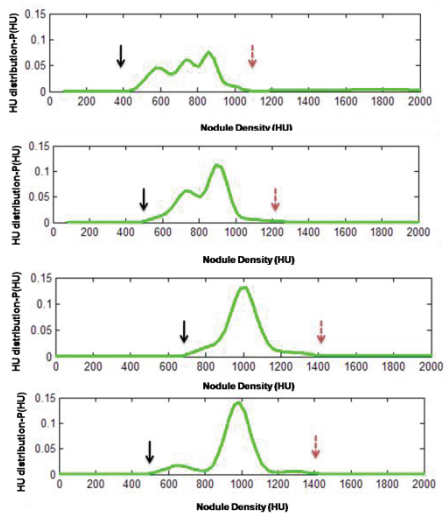

Fig. 5: *Probability density of the Radial distance of the nodules: Top left: Well-Circumscribed nodule; Top right: Vascular nodule; Bottom left: Juxta-Pleural; and Bottom right: Pleural-Tail nodule. Arrows show q_{min} and q_{max}. This behavior may be described by a single Gaussian (to resemble the circular-symmetric Gaussian in [7][15][16]).*

To test the overall approach of nodule detection, we used a set of 2D scans from the ELCAP clinical data set with known ground truth in terms of nodule type and location. Therefore, only 2D templates are used in this paper. The most common parametric nodule models in 2D are circular and semi-circular. Fig. 6 shows a few examples of such templates.

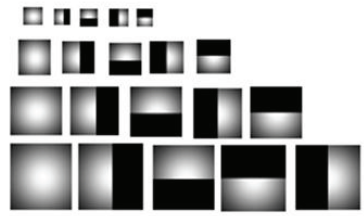

Fig. 6: *An ensemble of generated circular and semi-circular templates with radii ranging from 5 to 20 pixels with orientation from 0^o to 360^o with step size 90^o, in the semi-circular cases.*

The isotropic templates are defined in terms of the radius (size), *R*, and the gray level distribution as a circular symmetric Gaussian function [7][15][16] while non-isotropic templates are defined by radius, gray level distribution and orientation. Automatic generation of the gray level distribution of the nodules with known radius and histogram of nodule prototypes can be generated, for a given shape, using Eq. 1 and 2 above. This is particularly simple to perform in the case of parametric templates (e.g., Fig. 6, where given a radius *R* we only need to estimate the intensity of HU from these equations.

4. Template Matching

Among the widely used approaches for computerized nodule detection is the template matching approach, in which the nodule model (template) is swept across the scan (2D slices or the 3D volume) in a raster fashion and a similarity measure is calculated between the intensity (or HU) of the template and the region of the CT data underneath. If the result of the matching process between an unknown object in the CT data and the template is sufficiently high, the unknown object is labeled as resembling the template (i.e., a nodule candidate), however due to image noise, spatial, amplitude quantization effects, and a priori uncertainty of the exact shape and structure of the object to be detected, an exact match rarely occurs. Hence, a subsequent step of authenticating the detected nodule candidates is performed to reduce the false positives. Among the widely used similarity measures is the normalized cross correlation (NCC), which has a maximum value of unity that occurs if and only if the image function under the template exactly matches the template. The normalized cross-correlation of a template, *t(x,y)* with a sub-image *f(x,y)* is:

$$NCC = \frac{1}{n-1}\sum_{x,y}\frac{(f(x,y)-\bar{f})(t(x,y)-\bar{t})}{\sigma_f \sigma_t}, \qquad (3)$$

where *n* is the number of pixels in template *t(x,y)* and sub-image *f(x,y)* which are normalized by subtracting their means and dividing by their standard deviations.

The probability density functions (pdf) of nodule and non-nodule pixels are computed using the normalized cross correlation coefficients resulted from templates with varying parameters (shape, size and orientation if applicable). Based on the Bayesian

classification theory, the intersection between the pdf's of the two classes is chosen as the threshold separating the correlation coefficients resulted from nodule and non nodule pixels.

A decision fusion approach (Fig. 7) was also conducted in this paper using the circular and semi-circular templates in the detection process for radius 10 and 20 pixels with different orientations (semi-circular case) in a serial fashion and the final decision is Exclusive-OR (XOR) of the binary outputs. Template matching is performed as before using the same NCC threshold and the output of the template matching from each nodule model is a binary image (NCC values rank from zero to 1; after thresholding the zeros are NCC values below the specified threshold and the ones are otherwise), representing where the candidate nodules for the corresponding parametric template are located. Binarization is performed onto these images to give any pixels that are black a value of 0 and any white pixels a value of 1. The images are then Exclusively-ORed (XOR) together to receive a final black and white image that represents the candidate nodule locations for all of the nodule types. Fig. 8 summarizes this process.

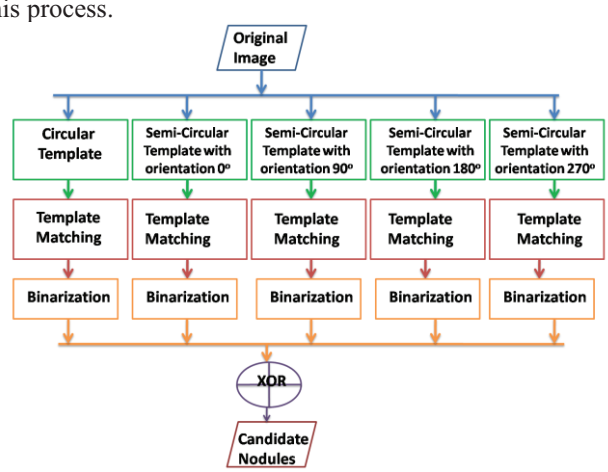

Fig. 7: *Block diagram of decision fusion template matching approach.*

XOR-ing the images after template matching and binarization provides us with locations where one of the template has a nodule detected for that pixel value will be taken as a true candidate nodule. We experimented with several other logic operators and studied the sensitivity and specificity results for each case before deciding on the XOR operator.

5 Experimental Results

This study is based on the ELCAP public database [10], which contains 50 sets of low-dose CT lung scans taken at a single breath-hold with slice thickness 1.25 mm. The locations of the 397 nodules are provided by the radiologists, where 39.12% are

juxta-pleural nodules, 13.95% are vascularized nodules, 31.29% are well-circumscribed nodules and 15.65% are pleural-tail nodules. The official reports indicate the mean nodule diameter to be 8.5 mm with standard deviation 3.6. The ELCAP database is of resolution 0.5x0.5mm [10].

The ELCAP expert identified a point on the nodule not the entire spatial support; this point may not necessarily coincide with the center or the centroid of the nodule. Indeed, the NCC measure decays as we move away from the spatial support of the nodules. The sensitivity and specificity of template matching was studied for templates of radii ranging from 0.5mm (1 pixel) to 10mm (20 pixels). The orientations of the templates in the semi-circular case are from 0 to 360 with step size of 45 degrees. For each detected nodule, let the coordinates of its centroid be $\boldsymbol{c}$ = (x,y) , the nodule is considered correctly detected and counted as true positive (TP) when the distance between the detected point and the closest ground truth point is smaller than the template radius. All other detected points are considered false positives (FP). True negative (TN) is the number of points which are not detected as candidate nodules and when compared to the ground truth they are not nodules. False negative (FN) is when no point is detected in the neighborhood of the ground truth nodule. The sensitivity and specificity are defined in terms of the false positive (FP) and the true positive (TP) nodules.

$$\text{Sensitivity } SN = TP/(TP+FN), \qquad \text{specificity } SP = TN/(FP+TN) \qquad (4)$$

These values are defined with respect to the NCC matching criterion. Fig. 8 shows the NCC distribution for parametric templates of radius 10 pixels (i.e., template size 21x21). As expected, higher values of NCC leads to reduced detection rates, while smaller values increases the detection rates, and consequently the number of FPs. A threshold of NCC value was set at 0.5 which have been shown, empirically, to be a good compromise between TP and FPs. This will be considered a hard classifier. The receiver operating characteristic (ROC) curve is used to plot the fraction of true positives (TPR = true positive rate) vs. the fraction of false positives (FPR = false positive rate), where TPR = TP = (TP+FN) and FPR = FP = (FP+TN). We experimented with different template sizes ranging from 1 to 20 pixels (0.5mm-5mm). The ROC curves showed us that variations in the template shape and/or orientation have minimal affect on well-circumscribed nodule detection. The semi-circular template with orientation 180 degrees best simulates juxta-pleural as well as pleural-tail, while the circular template least represents these nodule types.

Experiments revealed that larger template radii yield higher sensitivity; while the performance tends to decrease for smaller template sizes, yet, the specificity in the larger template radii overall decrease. The rate of performance decay depends on template shape and nodule type, the pleural-tail nodules are the most sensitive nodule type to the template size, while the well-Circumscribed nodule showed the least sensitivity to template size. At radius of 10, we found the overall sensitivity using the decision fusion approach to be higher than the results obtained without using the decision fusion approach. In the individual nodule cases the templates that yielded relatively smaller marginal differences from the decision fusion approach in terms of

sensitivity rates was the semi-circular template with orientation 315° representing the Well-Circumscribed nodule type and the circular template representing the Juxta-Pleural nodule type. The overall specificity was similar in the circular template cases. Semi-circular templates with orientations 0° to 135° (without decision fusion) provided higher specificity but this also means there are more false positives.

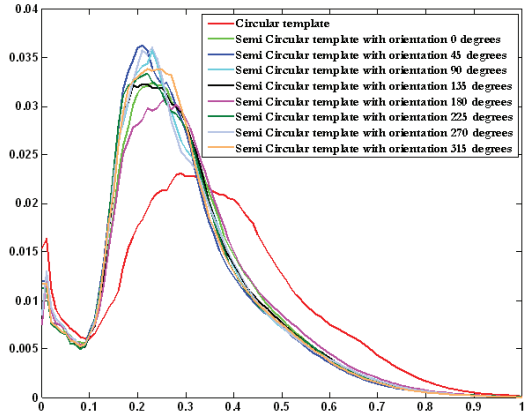

Fig. 8: *Distribution of the Normalized Cross-Correlation (NCC) for parametric nodules (circular and semi-circular) with radius 10 pixels. Higher NCC values results in less FPs while smaller values provide more FPs. At a threshold of 0.5 the range of NCC is {0.007, 0.012}.*

Tables below show the results for the sensitivity and specificity for radii 10 and 20 pixels. The sensitivity and specificity were analyzed for both the hard classifier and the decision fusion classifier. ROC curves are not included due to space limitations.

Nodule Type	Template Radius = 10 Hard Classifier		Template Radius = 10 Decision Fusion Classifier	
	Sensitivity	Specificity	Sensitivity	Specificity
Well-Circumscribed	49.44 %	81.15 %	58.43%	75.55%
Vascularized	70.73 %	78.54 %	73.17%	72.36%
Juxta-Pleural	83.48 %	66.73 %	93.04%	71.80%
Pleural-Tail	91.30 %	77.80 %	97.83%	72.81%

Nodule Type	Semicircular Template Radius 10; Orientation = 0° Hard Classifier		Semicircular Template Radius 10; Orientation = 90° Hard Classifier	
	Sensitivity	Specificity	Sensitivity	Specificity
Well-Circumscribed	37.08%	93.44%	31.46 %	93.70%
Vascularized	48.78 %	94.00 %	29.27 %	93.23%
Juxta-Pleural	75.65 %	88.41 %	73.04 %	86.70 %
Pleural-Tail	60.87%	92.50%	67.39%	92.50%

Nodule Type	Semicircular Template Radius 10; Orientation = 180° Hard Classifier		Semicircular Template Radius 10; Orientation =270° Hard Classifier	
	Sensitivity	Specificity	Sensitivity	Specificity
Well-Circumscribed	31.46 %	92.88 %	30.34%	93.64%

Vascularized	48.78 %	93.94 %	46.34%	93.36%
Juxta-Pleural	77.39%	86.30 %	75.65%	86.27%
Pleural-Tail	78.26 %	92.74 %	63.04%	92.80%

Nodule Type	Template Radius = 20 Hard Classifier		Template Radius = 20 Decision Fusion Classifier	
	Sensitivity	Specificity	Sensitivity	Specificity
Well-Circumscribed	76.40 %	61.17%	87.80%	36.58%
Vascularized	87.81%	53.71%	87.80%	36.58%
Juxta-Pleural	89.57 %	54.99%	92.17%	33.16%
Pleural-Tail	95.65 %	50.63%	97.83%	35.18%

Nodule Type	Semicircular Template Radius 20; Orientation = 0° Hard Classifier		Semicircular Template Radius 20; Orientation = 90° Hard Classifier	
	Sensitivity	Specificity	Sensitivity	Specificity
Well-Circumscribed	53.93 %	76.99%	57.30%	72.51%
Vascularized	75.61%	73.75%	73.17%	71.58%
Juxta-Pleural	82.61%	64.72%	80.87%	65.93%
Pleural-Tail	75.61%	73.75%	86.96%	66.32%

Nodule Type	Semicircular Template Radius 20; Orientation =180° Hard Classifier		Semicircular Template Radius 20; Orientation =270° Hard Classifier	
	Sensitivity	Specificity	Sensitivity	Specificity
Well-Circumscribed	59.55%	71.28%	53.93%	74.35%
Vascularized	70.73%	71.85%	73.17%	73.40 %
Juxta-Pleural	75.65%	64.24%	83.48%	66.30%
Pleural-Tail	86.96%	68.65%	91.30%	69.19%

6 Conclusion

In this paper, a data-driven approach was devised to model and simulate typical lung nodules. We studied the effect of template shape on detection of different nodules types. Variations in the template shape and/or orientation has minimal affect on well-circumscribed nodule detection. The hard classifier (based on thresholding the NCC) showed that the semi-circular template best simulates juxta-pleural and pleural-tail nodules, while the circular template least represents these nodule types. Similar conclusion also was achieved with the vascular nodules. Experiments revealed that larger template radii yield higher sensitivity, while the performance tends to decrease for smaller template sizes, yet, the specificity decreased with larger template. The overall performance depends on template shape and nodule type. The pleural-tail nodules are most sensitive to the template size, while the well-circumscribed nodule was the least sensitive. Overall, the decision fusion classifier provided best performance for templates of radius 10 pixels (i.e., nodule sizes of 1.5 mm), in terms of sensitivity and specificity. Current efforts will extend the nodule models into 3D using large clinical data set. Significant efforts are also directed towards designing non-parametric templates in 2D and 3D with attributes of the real data.

Acknowledgements: This research has been supported by the Kentucky Lung Cancer Program and The University of Louisville.

References

1. Kloecker, G., et al.: Lung Cancer in the US and in Kentucky. KMA **105** pp.159--164(2007)
2. Mountain, C. F.: Revisions in the international system for staging lung cancer. Chest **111** pp.1710--1717 (1997)
3. Gajra, A., et al.: Impact of tumor size on survival in stage IA non-small cell lung cancer: a case for subdividing stage IA disease. Lung Cancer **42** pp.51--57 (2003)
4. Zaho, B., Gamsu, G., Ginsberg M. S., Jiang, L., Schwartz, L. H.: Automatic Detection of small lung nodules on CT utilizing a local density maximum algorithm. Journal of Applied Clinical Medical Physics **4** (2003)
5. Armato, S. G. 3rd, Giger, M. L., Moran C. J., Blackburn, J. T., Doi, K., MacMahon H.: Computerized detection of pulmonary nodules on CT scans. Radio Graphics **19** pp.1303--1311 **(**1999)
6. S. Hu, E. A. Hoffman and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," IEEE Transactions on Medical Imaging, Vol. 20, pp. 490–498, 2001.
7. Y. Lee, T. Hara, H. Fujita, S. Itoh and T. Ishigaki, " Automated Detection of Pulmonary Nodules in Helical CT Images Based on an Improved Template-Matching Technique," IEEE Transactions on Medical Imaging, Vol. 20, 2001.
8. B. Ginneken, B. Romeny and M. Viergever, "Computer-Aided Diagnosis in Chest Radiography: A Survey," IEEE Transactions on Medical Imaging, Vol. 20, 2001.
9. I. Sluimer, A. Schilham, M. Prokop, and B. van Ginneken, "Computer Analysis of Computed Tomography Scans of the Lung: A Survey," IEEE Transactions on Medical Imaging, vol. 25, No. 4, pp. 385–405, April, 2006.
10. ELCAP public lung image database, http://www.via.cornell.edu/databases/lungdb.html
11. G. Grieg, O. Kubler, R. Kikinis, and F. A. Jolesz, "Nonlinear Anisotropic Filtering of MRI Data," IEEE Transactions on Medical Imaging, Vol. 11, No. 2, pp. 221-232, June 1992.
12. W. J. Kostis, et al., "Small pulmonary nodules: reproducibility of three-dimensional volumetric measurement and estimation of time to follow-up," Radiology, Vol. 231, pp. 446-52, 2004.
13. W. J. Kostis, A.P. Reeves, D. F. Yankelevitz and C. I. Henschke, "Three dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images," Medical Imaging IEEE Transactions Vol. 22, pp. 1259—1274, 2003.
14. Asem Ali and Aly A. Farag, "Automatic Lung Segmentation of Volumetric Low-Dose CT Scans Using Graph Cuts," 4th International Symposium on Visual Computing (ISVC-08), Las Vegas, December 1-3, 2008, pp. 258-267.
15. A. A. Farag, A. El-Baz, G.Gimelfarb, R. Falk and S. G. Hushek, "Automatic detection and recognition of lung abnormalities in helical CT images using deformable templates," Proc. of International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI-2004, September 26-29, 2004, pp. 856-864.
16. Shireen Y. Elhabian, Amal A. Farag, Salwa Elshazly and Aly A Farag, "Sensitivity of Template Matching for Pulmonary Nodule Detection: A Case Study," (CIBEC 2008), Cairo, Egypt December 18-20, 2008, pp. 110 -114.