PROBABILISTIC BRANCHING NODE DETECTION USING ADABOOST AND HYBRID LOCAL FEATURES

Tatyana Nuzhnaya[2], Michael Barnathan[2], Haibin Ling[1], Vasileios Megalooikonomou[2], Predrag R. Bakic[3], Andrew D.A. Maidment[3]

[1]Center for Information Science and Technology, [2]Data Engineering Laboratory (DEnLab), Department of Computer and Information Sciences, Temple University, 1805 N. Broad St. Philadelphia, PA 19122, USA

[3]Department of Radiology, University of Pennsylvania, 3400 Spruce St., Philadelphia, PA 19104, USA

ABSTRACT

Probabilistic branching node inference is an important step for analyzing branching patterns involved in many anatomic structures. Based on an approach we have developed previously, we investigate combining machine learning techniques and hybrid image statistics for probabilistic branching node inference, using adaptive boosting as a probabilistic inference framework. Then, we use local image statistics at different image scales for feature representation, including the Harris cornerness, Laplacian, eigenvalues of the Hessian, and Harralick texture features. The proposed approach is applied to a breast imaging dataset consisting of 30 images, 7 of which were previously reported. The use of boosting and the Harralick texture feature further improves upon our previous results, highlighting the role of texture in the analysis of the breast ducts and other branching structures.

Index Terms— Branching Structure, Breast Imaging, AdaBoost.

1. INTRODUCTION

Studies have associated morphological variability of the breast ductal network with subsequent development of breast cancer, suggesting that analysis of branching structures within the human breast can assist in diagnosing malignancy or estimating cancer risk [1]. Node detection is a very important first step towards the ductul tree segmentation. The motivation for this paper is to improve the approach for automatic probabilistic branching node detection proposed in [2].

However, this task is very challenging because anatomic tree-like structures are usually very complex in both topology and pattern of appearance. The imaging process often introduces more obstacles, such as blurring, noise (in some modalities), and the vessel occlusion and intersection caused by 3D to 2D projection. Furthermore, the modalities for acquiring images of natural branching structures differ in their degree of sensitivity in visualizing the tree. In unenhanced mammography, the branching topology of a tree structure may be barely visible or even absent from an image, but still contributes to the image texture of its surroundings [3,4]. The maximum depth of tree-branching that is captured in the image may also vary, depending on a modality's ability to extract a branching structure from its complex surroundings. Modalities which offer visualization of higher levels of branching are usually more prohibitive in terms of cost, health hazard, or comfort. Alternatively, modalities that can capture only the indirect effects of the presence of a branching structure on its complex surroundings are more easily available.

Motivated by these challenges, we have previously investigated novel approaches for studying branching structures [5], the use of branching patterns in classification of disease [3] and their influence on corresponding texture in medical images [4]. We have also previously proposed combining machine learning techniques and hybrid local features for probabilistic branching node inference and detection [2].

In this paper, we extend our previous results using Harralick's texture features [6] in conjunction with adaptive boosting (AdaBoost) on a larger dataset of images, strengthening our previous findings and further illustrating the importance of texture in branching structure analysis.

2. BACKGROUND

Many effective methods have been developed to detect and analyze the branching anatomic structures. Machine learning techniques play important roles in some recent systems for vessel anatomy study [7]. In [7], Adaboost [8] is applied on features for classification of lung bronchovascular anatomy. A thorough survey on vessel detection is given in [9]. In [2] we used a support vector

machine as a probabilistic inference framework. Then, we used local image statistics at different image scales for feature representation, including the Harris cornerness, the Laplacian, and the eigenvalues of the Hessian. The proposed approach was applied to a seven breast imaging dataset. We achieved very encouraging results, which were helpful for further analysis of the breast ducts.

Breast images are often highly textured and voxel based analysis is an interesting path to explore. Moreover, as color and intensity are not as important in medical images as in photographs, texture analysis becomes crucial in medical image retrieval. According to [6], texture refers to visual patterns which have properties of homogeneity and cannot result from the presence of only a single color or intensity.

3. METHODOLOGY

3.1. Problem formulation

The focus of our research was to investigate the use of Haralick's texture features and adaptive boosting algorithm to improve the vessel detection results previously reported in [20]. Instead of directly detecting branching nodes, we are interested in the probability of any given location being a branching node. Specifically, we start with an normalized input image $I : D \rightarrow [0,1]$, where $D=[1..m]\text{x}[1..n]$ is the lattice on which I is defined. For any $(x,y)\in D$, the intensity $I(x,y)$ is normalized from the original image by

$$I(x,y) = (I(x,y) - I_{\min}) / (I_{\max} - I_{\min}) ,$$

where $I_{\max}$ and $I_{\min}$ are the maximum and minimum intensities over all original un-normalized images. Our task is to find a node probability estimation $P(x,y;I) : D \rightarrow [0,1]$, such that for any (x,y) in D, $P(x,y)$ measures the probability that a tree node exists at pixel (x,y).

Note that function P is more general than the commonly used detection function that provides a binary output. The probabilistic output of P is very flexible. It provides a local confidence that can be fused in the future steps involving semantic (usually global) information. Second, as shown in the following sections, it can be used for candidate node detection.

We use a learning-based approach to automatically build function P. This involves two issues: probabilistic branching node inference framework and feature representation.

3.2. Probabilistic branching node inference

An adaptive boosting (AdaBoost)[11] machine learning algorithm was chosen for this task for its strong theoretical basis and ability to concurrently select and combine relevant features from the feature set during the training of each independent classifier.

In this framework, let $z= f (x,y;I)$ be the feature vector extracted from image I at location (x,y). The algorithm concurrently selects and combines relevant features from the feature set during the training of each independent classifier, thus avoiding a separate feature selection process common with other classification methods. The basic premise of the algorithm is that any number of weak classifiers with an error rate less than 50% can be combined to form an ensemble classifier whose error rate approaches that of an optimal classifier. The algorithm determines the optimal combination of T weak classifiers $\{h_1(x), \ldots, h_T(x)\}$, chosen from any number of possible weak classifiers, when training the ensemble classifier for each image instance. The classification boundary is then defined by the following equation,

$$\sum_{i=1..T_s} \alpha_i h_i(z) = \Delta ,$$

where T is the number of weak classifiers $h_i(z)$, α_i are parameters estimated by the learning procedure, and Δ is the threshold that will be adjusted for trading off the false positive and false negative rates..

Since our goal is a probability function that measures the likelihood of a given pixel being a node, we use the probabilistic output of AdaBoost. In particular, a confidence output (or margin) from the learnt AdaBoost model is converted to a probability using a sigmoid function.

3.3. Hybrid local features

To find local feature representation $z= f (x,y; I)$, we use four kinds of image statistics: Harris cornerness [14] $h(x,y)$, Laplacian $l(x,y)$, eigenvalues $(\lambda_1(x,y), \lambda_2(x,y))$ of the Hessian matrix $H(x,y)$, and Harralick's texture features[6] of $G(x,y)$.

The Harris cornerness is derived to measure the divergence of local principal directions, which is therefore useful to distinguish branching nodes. It is defined as

$$h(x,y) = \frac{\Phi(I_x^2)g(I_y^2) - \Phi(I_x I_y)^2}{\Phi(I_x^2) + \Phi(I_y^2) + \varepsilon} ,$$

where Φ is a local smoothing function using a Gaussian kernel, ε is used to avoid underflow, and I_x and I_y denote image gradients.

The image Laplacian is defined as $l(x,y)=I_{xx}+I_{yy}$, where I_{xx} and I_{yy} denote second derivatives of image I. The Laplacian is known to relate to local "blob-like" structures such as nodules [17].

The Hessian matrix is often used for vessel analysis [8]. It is defined as

$$H(x,y) = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix},$$

where I_{xx}, I_{yy}, and I_{xy} denote second derivatives of image I.

The basis for Haralick's texture features is the gray-level co-occurrence matrix (GLCM) [6]. Those matrices model spatial dependencies between gray levels of an image. Given a distance d and an orientation θ, the (i,j) coefficient of the corresponding matrix is the probability of going from

a gray level i to a gray level j with an intersample spacing of d along the axis making an angle with the x axis.
After the computation and normalization of the GLCM, a set of six $g_k(x,y)$ most used in practice features is computed: Energy, Entropy, Inverse Differential Moment, Shade, Prominence and Inertia.

To combine these hybrid features together, each feature is multiplied by a coefficient for roughly normalization, specifically, we have $l'(x,y)=3l(x,y)$, $\lambda_i'(x,y)=2\lambda_i(x,y)$, and $h'(x,y)=5h(x,y)$, $g'_k(x,y)= g_k(x,y)$.

Furthermore, a hierarchical scheme is used to capture over scale information. In our implementation three scales are used, resulting in a 30-dimensional features space, i.e.,

$$z=(\ldots, l^{(s)},\lambda_1^{(s)},\lambda_2^{(s)},h^{(s)},g_1^{(s)},g_2^{(s)}, g_3^{(s)}, g_4^{(s)}, g_5^{(s)}, g_6^{(s)},\ldots)^T,$$
$$s = 1,2,3,$$

where (s) indicates that the feature is extracted at the image of scale s (image I smoothed by a Gaussian with standard deviation 2^{s-1}), and (x,y) in the feature vectors are omitted for notation simplicity. By this hierarchical scheme, the feature vector implicitly captures neighborhood image statistics at different scales.

4. EXPERIMENTAL RESULTS

4.1. Experimental setup

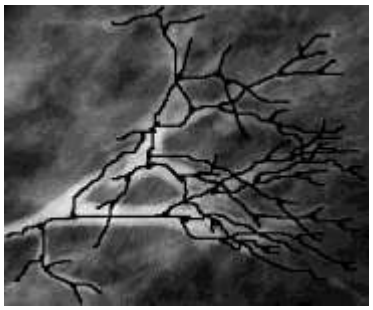
To test the proposed approach, we use a dataset containing 30 breast images. All the images have been manually annotated by experts; these annotations are used in both training and evaluation. An example annotation is shown in Figure 2 (b).

For evaluation, we conduct a leave-one-out experiment on the dataset. In the training stage, the annotated nodes are used as positive samples, and negative samples are randomly selected pixels that are at least eight pixels far from any positive samples. In the testing phase, we applied the learnt AdaBoost model to all image pixels and output their node probabilities.

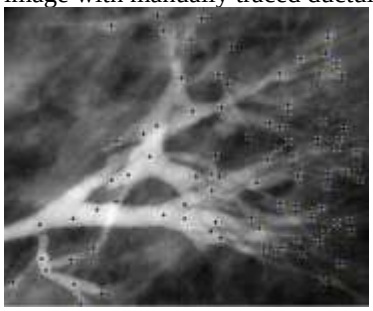
4.2. Results

The result on one image is shown in Figure 1. In addition to the probability map (Figure 1(d)), we also output the detected top candidates (Figure 1(c)). This is achieved by first finding all local maximums from the probabilistic map, and then picking from these maximums the top 80 with largest probabilities.

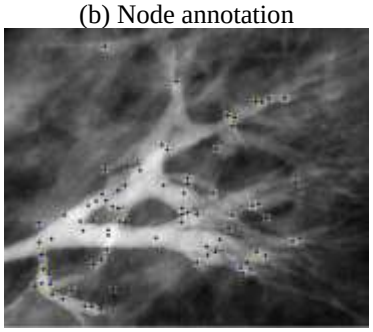
For a quantitative study, we output the average (over all images) number of correct nodes among top N candidates picked according to the learned probabilities, for N=20, 40, 60, 80. We compare the hybrid features with other features. The results are summarized in Table 1, which shows the superiority of the proposed approach. From Table 1 we see that about one third of the selected candidates are correct, which can be used for further tree-structure detection steps.



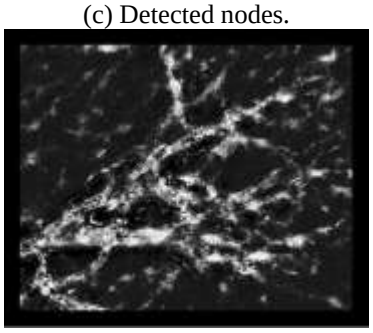
(a) Input image with manually traced ductul network



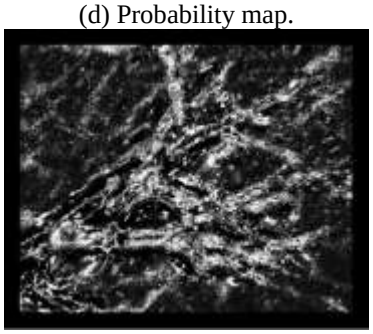
(b) Node annotation



(c) Detected nodes.



(d) Probability map.



(e) Probability map from preliminary experiments.

Fig. 1: (a) An example image, (b) its annotation, (c) node detection, (d,e) probability maps for branching nodes.

From the probability maps figures, we can see the large variation in topology and appearance among breast ductal systems. In addition, some annotated nodes have very similar local appearances to non-node pixels.

The experimental result is very promising considering that only thirty images are used for training and the large appearance variation among them. We expect that more training samples and introduction of new characteristic features will boost the performance.

	7 images, SVM		30 images, ADABOOST	
N	40	80	40	80
Hybrid	**14.57**	**26.86**	**18.52**	**38.97**
Laplacian	11.57	24.71	11.57	24.71
Hessian	12.86	25.00	12.86	27.33
Harris cornerness	9.43	12.57	10.23	13.63
GLCM Hybrid	-	-	8.57	12.38

Table 1: Average number of correct nodes among top *N* detected candidates.

5. CONCLUSION

We improved the results previously reported using combining machine learning tools with hybrid local features for branching node inference. The proposed improvements demonstrate promising results on a dataset containing thirty breast images.

6. ACKNOWLEDGMENTS

This work was supported in part by NSF Research Grant IIS-0237921. The funding agency specifically disclaims responsibility for any analyses, interpretations and conclusions.

7. REFERENCES

[1] H. P. Dinkel, A. Trusen, A. M. Gassel, M. Rominger, S. Lourens, T. Muller, and A. Tschammler, "Predictive value of galactographic patterns for benign and malignant neoplasms of the breast in patients with nipple discharge," Brit. J. Radiol., vol. 73, pp. 706–714, 2000.

[2] H. Ling, M. Barnathan, V. Megalooikonomou, P. Bakic, A.Maidment. Probabilistic branching node detection using local hybrid features. Proceedings of the 6th IEEE International Symposium on Biomedical Imaging (ISBI), 2009.

[3] V. Megalooikonomou, M. Barnathan, D. Kontos, P. R. Bakic, A. D.A. Maidment, " A Representation and Classification Scheme for Tree-like Structures in Medical Images: Analyzing the Branching Pattern of Ductal Trees in X-ray Galactograms," IEEE Transactions on Medical Imaging, Vol. 28, No, 4, pp. 487-493, April 2009.

[4] M. Barnathan, J. Zhang, D. Kontos, P. Bakic, A. Maidment, V. Megalooikonomou, "Analyzing Tree-Like Structures in Biomedical Images Based on Texture and Branching: An Application to Breast Imaging", Proceedings of the International Workshop on Digital Mammography (IWDM), Tucson, AZ, Digital Mammography, Lecture Notes in Computer Science, Vol. 5116, pp. 25-32, 2008.

[5] D. Kontos, Bakic, P.R., Maidment, A.D.A., "Analysis of Parenchymal Texture Properties in Breast Tomosynthesis Images," SPIE Medical Imaging: Computer Aided Diagnosis, San Diego, USA, 2007.

[6] R. M. Haralick, K. Shanmugam, and I. Dinstein, Textural Features of Image Classification, IEEE Transactions on Systems, Man and Cybernetics, vol. SMC-3, no. 6, Nov. 1973

[7] Robert A. Ochs, Jonathan G. Goldin, Fereidoun Abtin, Hyun J. Kim, Kathleen Brown, Poonam Batra, Donald Roback, Michael F. McNitt-Gray, Matthew S. Brown, Automated classification of lung bronchovascular anatomy in CT using AdaBoost, Medical Image Analysis, Volume 11, Issue 3, June 2007, Pages 315-324.

[8] Yoav Freund, Robert E. Schapire: A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. J. Comput. Syst. Sci. 55(1): 119-139 (1997)

[9] Kirbas, C. and Quek, F. 2004. A review of vessel extraction techniques and algorithms. ACM Comput. Surv. 36, 2 (Jun. 2004), 81-121.

[10] Agam, G, Armato, SG, Wu C. Vessel Tree Reconstruction in Thoracic CT Scans With Application to Nodule Detection, IEEE Transactions on Medical Imaging, 24(4): 486-499 (2005), 2005

[11] Charles Florin, Nikos Paragios, James Williams: Globally Optimal Active Contours, Sequential Monte Carlo and On-Line Learning for Vessel Segmentation. ECCV (3) 2006: 476-489

[12] Michal Sofka, Charles V. Stewart: Retinal Vessel Centerline Extraction Using Multiscale Matched Filters, Confidence and Edge Measures. IEEE Trans. Med. Imaging 25(12): 141-126 (2006)

[13] Carrillo J.F., Hernández Hoyos M., Dávila E.E., Orkisz M., Recursive tracking of vascular tree axes in 3D medical images, International Journal of Computer Assisted Radiology and Surgery, 2007, 2:6, 331-339

[14] Harris, C. and Stephens, M., A combined corner and edge detector, Alvey Vision Conference, 8-13, 1988.

[15] Prasad, M., Sowmya, A. Detection of Bronchovascular pairs on HRCT Lung Images, IEEE International Symposium on Biomedical Imaging, pp. 1135 – 1138, IEEE, USA, 2004.

[16] Prasad, M., Sowmya, A. Automatic detection of bronchial dilatation in HRCT lung images, Journal of Digital Imaging, 2007.

[17] Okada, K., Comaniciu, D., and Krishnan, A. Robust anisotropic Gaussian fitting for volumetric characterization of Pulmonary nodules in multislice CT. IEEE Trans. Med. Imaging 24(3): 409-423, 2005.