Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Pulmonary Nodule Classification Aided by Clustering

S.L.A. Lee, A.Z. Kouzani, and G. Nasierding
School of Engineering
Deakin University
Waurn Ponds, VIC 3217, AUSTRALIA

E.J. Hu
School of Mechanical Engineering
Adelaide University
North Terrace, Adelaide, SA 5005, AUSTRALIA

***Abstract*—Lung nodules can be detected through examining CT scans. An automated lung nodule classification system is presented in this paper. The system employs random forests as its base classifier. A unique architecture for classification-aided-by-clustering is presented. Four experiments are conducted to study the performance of the developed system. 5721 CT lung image slices from the LIDC database are employed in the experiments. According to the experimental results, the highest sensitivity of 97.92%, and specificity of 96.28% are achieved by the system. The results demonstrate that the system has improved the performances of its tested counterparts.**

***Keywords*—classification aided by clustering, nodule, detection**

I. INTRODUCTION

Lung tissue abnormalities that are roughly spherical with round opacity and a diameter of up to 30 mm [1] are known as lung nodules. According to [2, 3], lung nodules are categorized into a number of groups including: juxta-vascular, well-circumscribed, pleural tail, and juxta-pleural. Currently, lung nodule can be diagnosed through lung imaging techniques. This includes chest radiography, computed tomography, and magnetic resonance imaging. In addition, bronchoscopy is used in medical imaging to confirm the abnormality findings in the three common imaging modalities.

Detection of lung nodules can be achieved through inspection of medical imaging scans. Recent studies, however, show that there may exist inter-reader variability in the detection of nodules by radiologists. Therefore, an automated detection system can provide preliminary nodule detection that may assist expert radiologists in their decision making. This automated system could improve the accuracy of lung nodule detection.

There are a number of existing systems that include a classification component. Li *et al.* [4] incorporated rule-based with composite feature to reduce the false positives. 117 scans from three different databases consisting of 153 nodules were included in the evaluation. An overall sensitivity of 86% with 6.6 FPs/scan was produced. The system used a dataset that had been obtained from the Japanese domain including only confirmed malignant and benign nodules which were usually larger than 30mm. It therefore created an artificial bias in predicting the sensitivity of the system. The false positive rate of this dataset was not included in the calculation. Severe diffuse disease images were excluded in the American datasets to avoid bias in estimation of false positives.

Combining fuzzy connectivity, originated by Udupa *et al.* [5], with adaptive contrast region growing was done by Dehmeshki *et al.* [6]. It identifies nodules of low and high contrast, juxta-pleural, and juxta-vascular. Fuzzy maps based rule was included to detect nodule attached to vessel which could not be dealt with the region growing. An initial result of 84% detection rate was accepted by radiologist and further 16% was obtained from alternative segmentation solutions. The issue with this method was that some manual intervention by expert radiologist was required if unsatisfactory results were produced initially.

Weighted multi-scale convergence index (WMCI) and fisher linear discriminant (FLD) were combined by Hardie *et al.* [7]. The system was evaluated on 154 chest radiographs from JRST database with 100 malignant and 54 benign nodules which result 78.1% sensitivity with 4 FPs/image. Also, a performance comparison between quadratic classifier and Gaussian Bayes linear classifier and FLD classifier was performed. The FLD's performance was superior to its studied counterparts.

Recently, ensemble classifiers [8] have emerged as a robust approach. It combines the decisions of multiple classifiers to form an integrated output. Pereira *et al.* [9] developed a multi-scale and multi-oriented filter bank used in a multi-layer perceptrons (MLP) to detect nodules. For 19 different classifiers in the MLP, the mean of 77.71% sensitivity and 87.18% specificity on the 154 nodules images were recorded. Recently, Ochs *et al.* [10] applied voxel-by-voxel classification method for airways, fissures, nodules, and vessels from CT images. AdaBoost algorithm was chosen to train and test 29 scans from the Lung Imaging Database Consortium (LIDC) database [11]. For nodule classification, the area under the receiver operating curve (ROC), known as A_z value, recorded as 0.945. Based on the literature, it seems that there is still room for improving the performance of lung nodule classification. One technique that could be used to improve image classification is through clustering [12] approach.

II. CLASSIFICATION-AIDED-BY-CLUSTERING

Clustering [13] is an unsupervised learning algorithm in which no labeled data is provided in the learning phase. The goal of clustering is to group objects of similar set of finite unlabelled data belonging to the same cluster when a fixed number of clusters are given. Previously, classification aided by clustering has been applied across various research areas

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

including protein [14], medical [15], and remote sensing [16]. It has proven to improve the classification accuracies in those research areas.

For urban land use/cover mappings, Lo and Choi [17] described a system which incorporated Iterative Self-Organizing Data Analysis Techniques (ISODATA) clustering algorithm with supervised fuzzy classification. The datasets was clustered into 60 classes by ISODATA and a homogenous cluster that consists of specific land use/cover are labelled and retained. The inhomogeneous clusters with mixed pixels characteristics are fed to the fuzzy classifier for further classification. The combination of ISODATA and fuzzy classification achieved 91.5% accuracy where single fuzzy classification achieved only 77.75%. Fig. 1 shows the block diagram description of the fuzzy classification enhanced by ISODATA clustering.

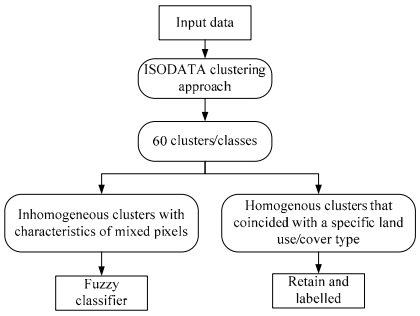

Figure 1. Fuzzy classification enhanced by ISODATA clustering [17].

For lung nodule classification-aided-by-clustering (CAC), Kawata *et al.* [18] proposed a linear discriminant classification boosted by k-means clustering using malignant and benign pulmonary nodules datasets based on topological histogram features. The A_z value under the ROC curve of the hybrid approach was recorded to be higher than that of the single linear discriminant classifier. Fig. 2 displays the block diagram description of this system.

Gurcan *et al.* [19] used curvature analysis and candidate detection scheme using k-means clustering, rule-based followed by LDA which reported to reduce the false positive rate. 1454 image slices from 34 patients consisting of 63 nodules were utilised in the study, and rule based classifier produced 84% sensitivity with 5.48 FPs/slice. When LDA was applied after rule-based, the false positive rate falls to 1.74 per slice for the same sensitivity.

Based on the reviewed literature, the CAC method has demonstrated an ability to improve the classification accuracy. In the next section, the random forest (RF) aided by expectation maximization (EM) clustering combination is employed to the lung nodule dataset.

III. Proposed Random Forest Classification Aided by EM Clustering

Clustering-based classification approach can be implemented in a number of ways. Depending on the architecture used, varying performance outcomes could be achieved. As reported above, very few attempts have been made to apply clustering-based classification to lung nodule classification. In this section, the author proposes a unique architecture for random forests (RF) based CAC for lung nodule classification. Fig. 3 illustrates the proposed architecture including training and test stages. In the training stage, the nodule and non-nodule parts of the training set are merged and the entire data is clustered into *M* clusters. This is done to fully take advantage of the similarity among features of nodule and non-nodule instances. Each cluster is then divided into two groups: nodule (N) and non-nodule (NN) instances. Using the original labels of the training set instances, *M* clusters are formed named cluster 1, cluster 2 … cluster M that includes *N1, N2, ...NM,* and *NN1, NN2, ..., NNM* instances, respectively. *M* classifiers are trained consisting of *N1, N2, ...NM,* and *NN1, NN2, ..., NNM* classes, accordingly. In the test stage, on the other hand, the test set instances are presented to the developed cluster model in order to determine which classifier is used to test the instance. *M* is selected as 2 due to the fact that our lung nodule dataset is not a very large and balanced (nodule/non-nodule ratio) dataset.

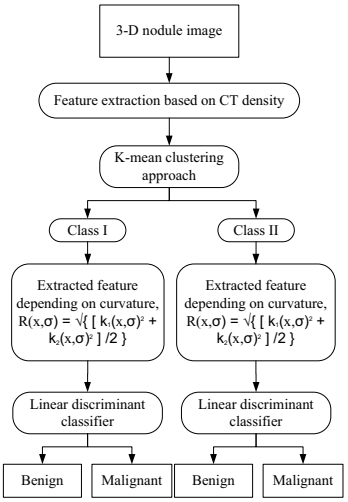

Figure 2. Block diagram of the malignant and benign classification aided by k-means clustering [18].

5721 lung images were obtained from the LIDC database. All images were of the size 512×512. A converter that uses the information in the XML file and extracts out the nodule regions from the lung images was developed. For nodule patterns that could fit within a 30×30 region, we extracted from the image such a region surrounding the nodule pattern. On the other hand, for nodule patterns that could not fit within a 30×30 region, we extracted the entire nodule pattern first, and then resized the extracted pattern into a 30×30 region. In total, we created 1203 30×30 nodule files. In addition, we developed a program that searched through all 5721 image files and randomly captured 1203 30×30 regions that did not contain any nodule patterns. Thus, we formed a two-class dataset consisting of 1203 30×30 nodule and 1203 30×30 non-nodule patterns. The details of the experimental procedure of our work can be found in [20].

The performance of the RF classifier was compared against that of the SVM and DT classifiers. With regard to the random

forest classifier, (i) different number of trees to grow, and (ii) different number of variables that are randomly sampled as candidates at each split were explored. Concerning the SVM classifier, we used the SVM with the radial basis function (RBF) kernel. For the SVM, we varied (i) cost, and (ii) gamma, γ, parameters. For DT, (i) pruning confidence parameter, and (ii) setting minimum number of instances per leaf were investigated.

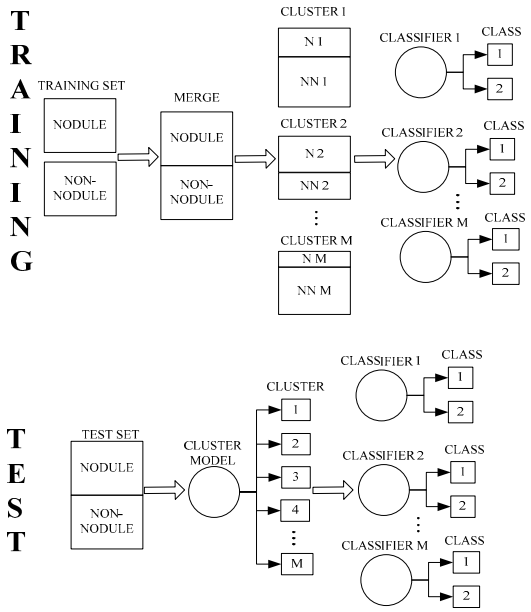

Figure 3. Architecture of the proposed classification-aided-by-clustering method.

Four experiments were conducted: 20-80/fixed, 50-50/fixed, 80-20/fixed, and 50-50/varied. For the first three experiments, the parameters for each classifier were fixed based on the consistency of the classifiers' performance (see Table I). On the other hand, for the forth experiment with 50% training and 50% test data, the parameters of the three classifiers were varied (see Table II). Note that the DT function will execute if the pruning confidence parameter lies between 0 (zeros) to 1 (one). In each experiment, the following parameters are calculated and presented: true positive rate or sensitivity (TPR), false positive rate (FPR), specificity (SPC), and accuracy (ACC).

The CAC approach merged the training nodule and non-nodule patterns into entity. The parameters used in expectation maximization (EM) are as follows.

- Maximum iterations: 100
- Minimum standard deviation : 1×10^{-6}
- Number of clusters: 2
- Seed: 100

The non-cluster based classification was compared against the CAC.

TABLE I. FIXED CLASSIFIER PARAMETERS

Random Forests	Support Vector Machine	Decision Tree
No-of-trees-grown: 25 No-of-variables-at-each split: 100	Cost: $2^{4.5}$ γ: 2^{-29}	Pruning confidence: 0.25 Instances/leaf: 2

TABLE II. VARIED CLASSIFIER PARAMETERS

Random Forests	Support Vector Machine	Decision Tree
No-of-trees-grown: 5 to 100 No-of-variables-at-each split: 5 to 100	Cost: 2^{1} to $2^{20.5}$ γ: 2^{-35} to $2^{-15.5}$	Pruning confidence: 0.1 to 0.95 Instances/leaf: 1 to 18

A. *Experiment I: 20-80/Fixed*

In this experiment, 20% (240 nodules and 242 non-nodule patterns) of the images of the available dataset were used to form the training set, and the other 80% (963 nodules and 961 non-nodule patterns), of the images were used to form the test set. Table III and Table IV presented the result for non-cluster based classification and CAC.

TABLE III. NON-CLUSTER BASED CLASSIFIERS PERFORMANCE

Classifier	TPR	FPR	SPC	ACC	Execution Time (sec)
Decision Tree	90.55	13.01	86.99	88.77	10.43
Support Vector Machine	90.55	10.82	89.18	89.86	7.25
Random Forests	94.50	11.55	88.45	91.48	6.97

TABLE IV. CLASSIFIERS AIDED BY EM CLUSTERING PERFORMANCE

Classifier	TPR	FPR	SPC	ACC	Execution Time (sec)
Decision Tree	92.42	14.36	85.64	89.03	120.31
Support Vector Machine	91.17	11.24	88.76	89.97	115.22
Random Forests	95.85	10.93	89.07	92.46	119.87

B. *Experiment II: 50-50/Fixed*

In this experiment, 50% (601 nodules and 602 non-nodule patterns) of the images of the available dataset were used to form the training set, and the other 50% (602 nodules and 601 non-nodule patterns) of the images were used to form the test set. Table V and Table VI presented the result for non-cluster based classification and CAC.

TABLE V. NON-CLUSTER BASED CLASSIFIERS PERFORMANCE

Classifier	TPR	FPR	SPC	ACC	Execution Time (sec)
Decision Tree	90.70	8.99	91.01	90.86	15.29
Support Vector Machine	89.70	5.66	94.34	92.01	10.86
Random Forests	94.68	5.99	94.01	94.34	9.05

TABLE VI. CLASSIFIERS AIDED BY EM CLUSTERING PERFORMANCE

Classifier	TPR	FPR	SPC	ACC	Execution Time (sec)
Decision Tree	92.86	7.99	92.01	92.44	187.96
Support Vector Machine	90.53	5.16	94.84	92.68	178.64
Random Forests	95.51	5.49	94.51	95.01	180.48

C. Experiment III: 80-20/Fixed

In this experiment, 80% (963 nodules and 961 non-nodule patterns) of the images of the available dataset were used to form the training set, and the other 20% (240 nodules and 242 non-nodule patterns) of the images were used to form the test set. Table VII and Table VIII presented the result for non-cluster based classification and CAC.

TABLE VII. NON-CLUSTER BASED CLASSIFIERS PERFORMANCE

Classifier	TPR	FPR	SPC	ACC	Execution Time (sec)
Decision Tree	91.67	9.92	90.08	90.87	65.68
Support Vector Machine	90.00	4.96	95.04	92.53	41.27
Random Forests	95.00	3.72	96.28	95.64	45.42

TABLE VIII. CLASSIFIERS AIDED BY EM CLUSTERING PERFORMANCE

Classifier	TPR	FPR	SPC	ACC	Execution Time (sec)
Decision Tree	94.17	7.02	92.98	93.57	240.95
Support Vector Machine	92.50	2.89	97.11	94.81	217.79
Random Forests	97.92	3.72	96.28	97.10	221.38

D. Experiment IV:50-50/Varied

To further improve the classification accuracy, sensitivity, and specificity, the classifier parameters were varied (refer to Table II) within key allowable ranges to obtain the best parameters that improve the results. This experiment used the same data set as described in Experiment II: 50-50/Fixed. Table IX and X present the best results achieved for the non-cluster based classification and CAC.

TABLE IX. NON-CLUSTER BASED CLASSIFIERS BEST PERFORMANCE

Classifier	TPR	FPR	SPC	ACC	Execution Time (sec)
Decision Tree	92.52	8.99	91.01	91.77	24.35
Support Vector Machine	94.02	5.16	94.84	94.43	12.78
	94.35	5.49	94.51		
Random Forests	96.01	5.16	94.84	95.43	10.13
	95.51	4.66	95.34		

TABLE X. CLASSIFIERS AIDED BY EM CLUSTERING BEST PERFORMANCE

Classifier	TPR	FPR	SPC	ACC	Execution Time (sec)
Decision Tree	94.52	7.32	92.68	93.60	189.11
Support Vector Machine	94.68	4.99	95.01	94.84	173.54
Random Forests	96.84	4.66	95.34	96.09	179.89

IV. DISCUSSIONS

The motivation of this study was to explore the techniques that can improve lung nodule classification performance. The author proposed to employ the classification-aided-by-clustering approach incorporating ensemble base learner. The experiments were trained and tested on an Intel Xeon CPU 5130 @ 2.00GHz on-board of a Dell Precision Workstation 490. In addition, the total average execution time of both the training and the test operations were recorded.

Three experiments (20-80/Fixed, 50-50/Fixed, and 80-20/Fixed) were carried out on the non-cluster based classifier, each using a different set of training and test datasets. The RF, the SVM, and the DT were trained and tested in the three experiments using different data sizes to examine the influence of the training and test dataset size on the performance of the systems.

Comparing Table III, V and VII, the results demonstrate that the non-CAC based RF performs better than the non-CAC based SVM as well as the DT in all experiments. The best result was obtained where 80% of the images were used in the training, and the remaining 20% employed in the testing. The highest classification accuracy of 95.64%, sensitivity of 95.00%, specificity of 96.28%, and false positive rate of only 3.72% were produced by the non-CAC based RF where the training set contained 80% of images and the test set consisted of the remaining 20%. The highest classification accuracy achieved by the tested non-CAC based SVM and the DT classifiers were 92.53% and 90.87% for the same training and test sets, respectively.

By comparing the execution time in each experiment, the non-CAC based DT recorded the highest execution time in all experiments. The non-CAC based RF classifier recorded the lowest execution time for 20-80 and 50-50 dataset sizes. For the overall execution time, non-CAC based SVM classifier recorded 59.38 seconds, which was slightly less than RF of 61.44 seconds for the three experiments.

For the forth experiment (50-50/varied), the parameters of the non-cluster based classifiers were varied to improve the classifier's accuracy, sensitivity, and specificity. As shown in Table IX, the RF classifier recorded 95.43% accuracy where the best trees and variables were 65 and 65, and 65 and 90 respectively. An improvement of 1.17% was achieved from the previous non-varying parameter of 100 trees and 25 variables. For the SVM classifier, the accuracy of 94.43% was recorded with an increment of 1.91% compared to the previous parameter where cost, and γ, is set to$2^{4.5}$, and 2^{-29},

respectively. The best parameters are cost, 2^{2}, γ, 2^{-26}, and cost, $2^{2.5}$, γ, 2^{-26}. The DT base classifier only attained an improvement of 0.91% comparing to 91.77 % accuracy previously acquired with confidence, 0.25, and instance per leaf, 2. The best parameters for decision trees were 0.10 for cost, and 4 for instance per leaf, respectively. The RF has shown to be an accurate classifier as it performs well for the lung nodule classification.

Further improvement for the performance of lung nodule classification is done by incorporating CAC. The proposed RF CAC was compared to SVM and DT CAC EMs using the author devised architecture (see Table IV, VI, VIII, and X). In this architecture, the randomly selected nodule and non-nodule training patterns were merged before performing the EM clustering. After clustering, the number of instances in each cluster is examined. The results reveal that the two-cluster approach is preferable because the number of instances in some clusters for the three-cluster approach were insufficient. Therefore, the two-cluster approach was used in the experiments. After identifying the nodule and non-nodule patterns in each cluster, the classifiers were trained for each cluster. The test nodule and non-nodule patterns which were not included in the training phase were merged, and then tested. The resultant data in each cluster was presented to the relevant classifier's model.

For the three CAC experiments, the highest classification accuracy was achieved by the RF CAC EM of 97.10% with 80-20 dataset size. The SVM and the DT CAC EMs only achieved 94.81% and 93.57% accuracy on the same dataset size. Comparing the execution times, SVM CAC EM results the lowest execution time for all the three dataset sizes with an average of 170.55 seconds. The execution time of RF CAC for the three different dataset sizes was slightly more than the SVM-based classifier with an average of 173.71 seconds

For the forth CAC experiment, variation of parameters was employed (see Table II). Fig. 4 and Fig. 5Figure 5 illustrate that RF CAC gives the best classification accuracy of 96.09% for 21 trees and 85 variables , followed by SVM based CAC of 94.84% for cost = 2^{5} and $\gamma = 2^{-27}$, and DT based CAC of 93.6% for pruning confidence = 0.25 and instances/leaf = 9. It is believed that the accuracy in Experiment III could be further improved by varying the classifier parameters.

The randomly selected 50-50 train and test data size were utilized in this experiment to calculate the A_z value. Comparing the non-CAC RF, SVM, and DT against RF, SVM, and DT CACs, the highest A_z value of 0.9735 was achieved by RF CAC (see Fig. 6). The DT CAC has improved the A_z value from 0.9343 compared with $A_z = 0.9194$. Overall, the CAC performs better than non-CAC. Among the CAC based methods, the RF CAC performs better.

V. CONCLUSION

A classification approach was implemented for lung nodule detection. For the non-CAC experiments, random forests based classifier performs better than support vector machine and decision tree. Random forests based CAC recorded the highest accuracy of 97.10% with a total average execution time of 173.71 seconds. The two other counterparts only achieved an accuracy of 93.57% and 94.81% for the same training and test sets. A maximum accuracy improvement of 2.70% was achieved by the CAC over non-CAC. It can be concluded that the random forests based CAC proved to perform well for the lung nodule classification problem.

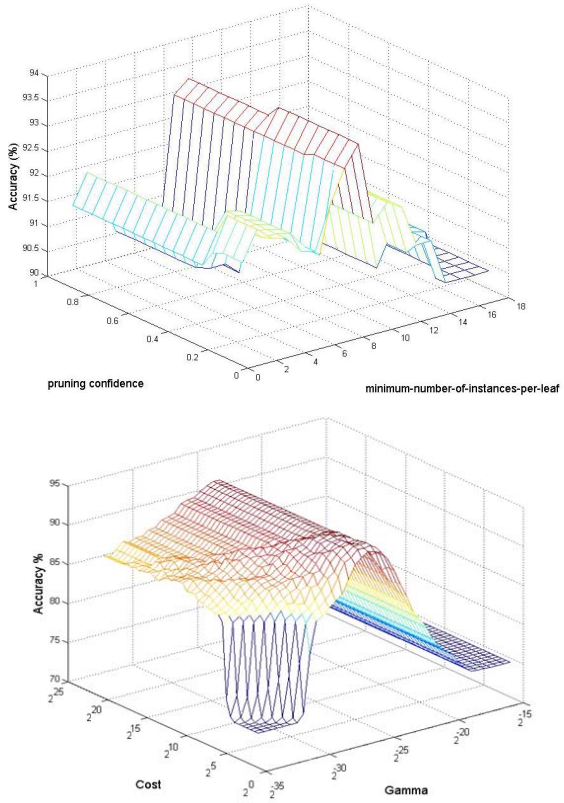

Figure 4. Classification accuracies for the DT and SVM CACs for varying classifier parameters.

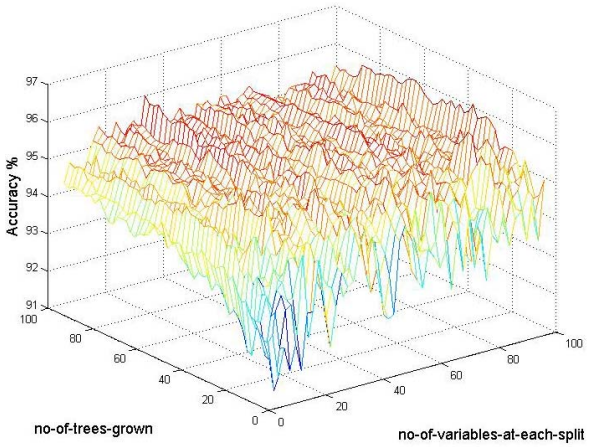

Figure 5. Classification accuracies for the RF CAC for varying classifier parameters.

REFERENCES

[1] J. H. Austin, N. L. Mueller, and P. J. Friedman, "Glossary of terms for CT of the lungs: recommendations of the nomenclature," *Committee of the Fleischner Society. Radiology*, vol. 331, pp. 200:327, 1996.

[2] 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," *IEEE Trans. On Medical Imaging*, vol. 22, no. 10, pp. 1259-1274, 2003.

[3] S. Diciotti, G. Picozzi, M. Falchini, M. Mascalchi, N. Villari, and G. Valli, "3D Segmentation Algorithm of Small Lung Nodules in Spiral CT Images," 2008.

[4] Q. Li, F. Li, and K. Doi, "Computerized Detection of Lung Nodules in Thin-Section CT Images by use of Selective Enhancement Filters and an Automated Rule-Based Classifier," *Acad Radiol*, vol. 15, no. 2, pp. 165–175, 2008.

[5] J. K. Udupa, P. K. Saha, and R. A. Lotufo, "Relative fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 11, pp. 1485–1500, 2002.

[6] J. Dehmeshki, H. Amin, M. Valdivieso, and X. Ye, "Segmentation of pulmonary nodules in Thoracic CT scans: A region growing approach," *IEEE Transaction on Medical Imaging*, vol. 27, no. 4, pp. 467-480, 2008.

[7] R. C. Hardie, S. K. Rogers, T. Wilson, and A. Rogers, "Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs " *Medical Image Analysis* vol. 12, pp. 240-258, 2008.

[8] J. Lu, K. N. Plataniotis, A. N. Venetsanopoulos, and S. Z. Li, "Ensemble-based discriminant learning with boosting for face recognition," *IEEE Trans. on Neural Networks*, vol. 17, no. 1, pp. 166--178, 2006.

[9] C. S. Pereira, L. A. Alexandre, A. M. Mendonça, and A. Campilho, "A Multiclassifier Approach for Lung Nodule Classification," in *Image Analysis and Recognition*, vol. 4142/2006, *Lecture Notes in Computer Science*, A. Campilho and M. Kamel, Eds. Berlin: Springer Berlin / Heidelberg, 2006, pp. 612-623.

[10] R. A. Ochs, J. G. Goldin, A. Fereidoun, H. J. Kim, K. Brown, P. Batra, D. Roback, M. F. McNitt-Graya, and M. S. Brown, "Automated classification of lung bronchovascular anatomy in CT using AdaBoost," *Medical Image Analysis* vol. 11, no. 3, pp. 315-324, 2007.

[11] "Lung Imaging Database Consortium (LIDC)." [Online]. Available: http://imaging.cancer.gov/programsandresources/InformationSystems/LIDC.

[12] D. Lu and Q. Weng, "A Survey of Image Classification Methods and Techniques for Improving Classification Performance," *International Journal of Remote Sensing*, vol. 28, no. 5, pp. 823-870, 2007.

[13] R. Xu and D. Wunsch II, "Survey of Clustering Algorithms," *IEEE Transaction on Neural Network*, vol. 16, no. 3, pp. 645-678, 2005.

[14] J. Weston, C. Leslie, D. Y. Zhou, A. Elisseeff, and W. S. Noble, "Semi-supervised protein classification using cluster kernels," *Bioinformatics*, vol. 21, no. 15, pp. 3241-3247, 2005.

[15] S. Wang, M. Zhou, and G. Gen, "Application of Fuzzy Cluster Analysis for Medical Image Data Mining," presented at *Proceedings of the IEEE International Conference on Mechatronics & Automation*, Niagara Falls, Canada, pp. 631-636, 2005.

[16] G. Droj, "The Applicability of Fuzzy Theory in Remote Sensing Image Classification," *Informatica*, vol. L II, no. 1, pp. 89-96, 2007.

[17] C. P. Lo and J. M. Choi, "A Hybrid Approach to Urban Land Use/Cover Mapping Using Landsat 7 Enhanced Thematic Mapper Plus (ETMz) Images," *International Journal of Remote Sensing*, vol. 25, no. 14, pp. 2687-2700, 2004.

[18] Y. Kawata, N. Niki, H. Ohmatsu, M. Kusumoto, R. Kakinuma, K. Mori, H. Nishiyama, Y. Eguchi, M. Kaneko, and M. Moriyama, "Hybrid Classification Approach of Malignant and Benign Pulmonary Nodules Based on Topological and Histogram Features," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*, vol. 1935/2000, *Lecture Notes in Computer Science*. Berlin: Springer Berlin / Heidelberg, 2000, pp. 297-306.

[19] M. N. Gurcan, B. Sahiner, N. Petrick, H. P. Chan, E. A. Kazerooni, P. N. Cascade, and L. Hadjiiski, "Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system," *Medical Physics*, vol. 29, no. 11, pp. 2552-2558, 2002.

[20] A. Kouzani, S. Lee, and E. J. Hu, "Lung Nodules Detection by Ensemble Classification," presented at *Proceedings of 2008 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2008)*, Singapore, pp. 324-329, 2008.

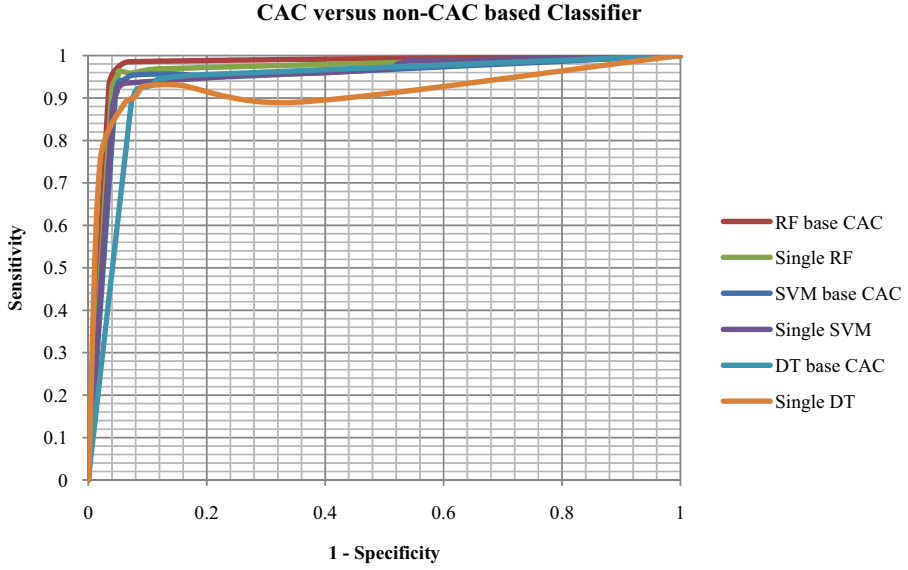

Figure 6. ROC results.