A comparative study of machine learning methods for lung diseases diagnosis by computerized digital imaging*

Suk Ho Kang** · Youngjoo Lee***

《目 次》

Abstract

In this study, we tested and compared several state-of-art machine-learning methods for automated classification of obstructive lung diseases based on the features from texture analysis using HRCT (High Resolution Computerized Tomography) images. HRCT can provide accurate images for the detection of various obstructive lung diseases, including centrilobular emphysema, panlobuclar emphysema and constrictive bronchiolitis. Features on the HRCT images, however, can be subtle, particularly in the early stages of disease, and image based diagnosis is subject to inter-observer variation. In order to support the clinical diagnosis and improve its accuracy, three different types of automated classification systems were developed and compared based on the classification performance and clinical applicability. Not only Bayesian classifier, a typical kind of statistic method, but also ANN (Artificial Neural Network) and SVM (Support

* This work was supported by the Second Stage of Brain Korea 21 Project in 2007.

** Department of Industrial Engineering, Seoul National University, Email: shkang@snu.ac.kr

*** Department of Industrial Engineering, Seoul National University, Email: spicio3@snu.ac.kr

Vector Machine) were employed. We tested these three classifiers for the differentiation of normal and three types of obstructive lung diseases. The ANN showed the best performance of 86.0% overall sensitivity and there is significant difference among other classifiers (one-way ANOVA, $p<0.01$). In discussion, we addressed what characteristic of each classifier made differences in the performance and which classifier was more suitable for clinical applications and proposed appropriate way to choose the best classifier and determine its optimal parameters to discriminate the diseases better. This result can be applied to the classifier for differentiation of other diseases.

I. Introduction

HRCT (*High Resolution Computerized Tomography*) can provide accurate images for the detection of various obstructive lung diseases, including centrilobular emphysema, panlobular emphysema and bronchiolitis obliterans. Features on the thin-section HRCT images can be subtle, however, particularly during early stages of disease, and diagnosis is subject to inter-observer variation. The main characteristics of images used to detect obstructive lung diseases are the presence of areas of abnormally low attenuation in the lung parenchyma, which, in the case of emphysematous destruction of the lung parenchyma, can be detected automatically by means of attenuation thresholding. However, areas of decreased attenuation are a feature of other obstructive lung diseases.[1] Thus, the accurate differentiation among these diseases may be difficult, even for expert thoracic radiologists.

Efforts have been made to develop computerized methods that could assist radiologists in improving diagnostic accuracy, by differentiating among obstructive lung diseases. Using a computer-aided diagnosis (CAD) scheme, radiologists could incorporate the output from the CAD into their decisions.[2] One automated computational scheme developed to classify obstructive lung diseases more accurately than radiologists made use of a naïve Bayesian classifier, which was trained to predict the likelihood of obstructive lung diseases based on quantitative texture features automatically extracted from the ROI (region of interest) of HRCT images.[2] Another classification system employs a Bayesian classifier and SVM (Support Vector Machine) to assess 3D texture features of lung parenchyma with abnormally low attenuation areas (LAA).[3]

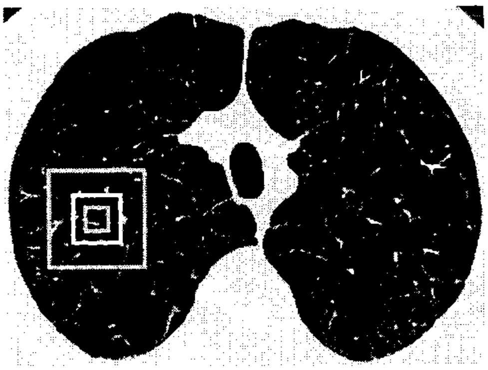
Several classification systems have been employed in medical CAD systems, and the selection of an appropriate classification scheme has been shown to be important for improving performance based on the characteristics of the data set.[4-5] In addition, modification or optimization of parameters for feature extraction, including the size of the ROI, is important. The motivation of this paper is to achieve a better understanding of the machine classification process for differentiating the obstructive lung diseases, to evaluate the classification in terms of sensitivity and specificity, and to analyze the strengths and weaknesses of the well-known classifiers for the clinical application.

II. Material and Methods

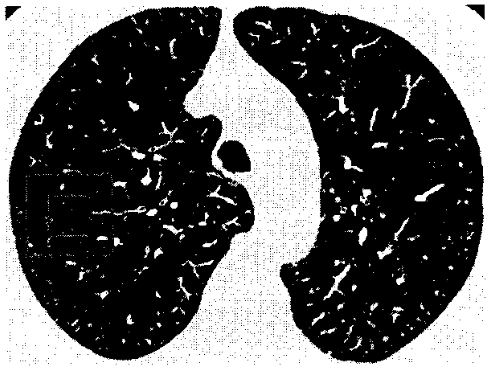
The images were selected from HRCT obtained in 17 healthy subjects (n=66), 26 patients with bronchiolitis obliterans (n=69), 28 patients with mild centrilobular emphysema (n=64), and 21 patients with panlobular emphysema or severe centrilobular emphysema (n=62).

Every 265 ROI was selected just one at each half lung for avoidance of the redundancy of images. All patients were recruited at the department of radiology, Asan Medical Center. The scanned field of view cover whole chest.

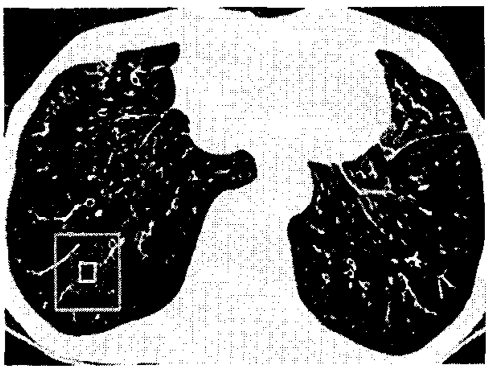
Automated segmentation of the lung was performed. The major pulmonary vessels of the lung parenchyma were removed by simple thresholding below-400 Hounsfield Unit. This step is important in that macroscopic structures, such as the major pulmonary vessels and chest wall, are of a size approaching that of the ROI used, and their statistically significant features cannot be obtained by textural extractors. Performance of automated anatomic segmentation allowed texture analysis of the finest structures of the lung parenchyma, and each feature value of texture analysis was normalized relative to the clipped pixels) for categorization into parenchymal ROI area.



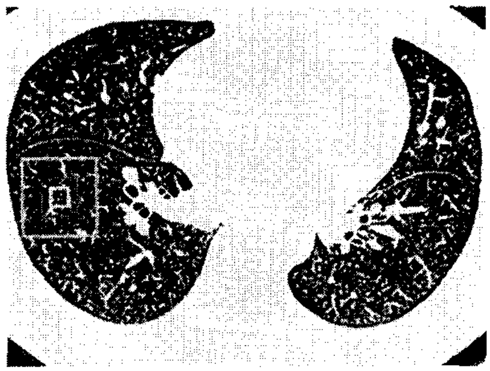
(a) Panlobular emphysema or severe centrilobular emphysema



(b) Mild centrilobular emphysema



(c) Bronchiolitis obliterans



(d) Normal lung parenchyma

〈Figure 1〉 Cross-sectional thin-section CT scans of the chest (window level, -850 HU; width, 400 HU). On each image, the three different sizes of rectangular (16x16, 32x32, and 64x64) highlights region of interest (ROI) that is typical of a particular condition.

For each image, two experienced thoracic radiologists each selected three sizes of rectangular ROI (16x16, 32x32, and 64x64 one of three diseases or normal lung tissue. Areas with HU between-400 and-1024 were segmented for clipping the major pulmonary vessels or chest wall. Since the bin size could influence the performance of the classifier[6], we tested various bin sizes (Q-bin size 16, 32, 64, 128, 144, 196, and 256) of run length encoding and the co-occurrence matrix. Overall sensitivities of the system, using each combination of variable ROI and bin sizes, were calculated and compared.

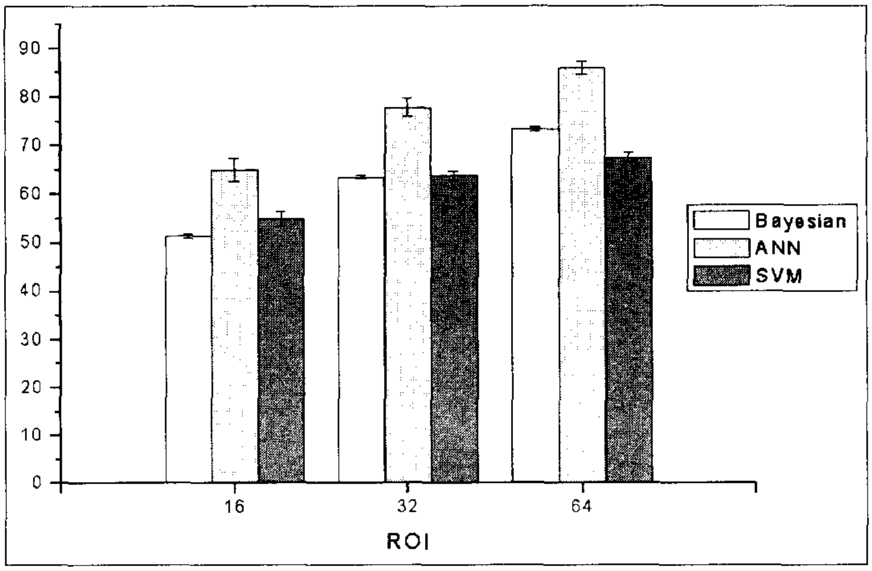
〈Table I〉 Summary of 13 textural features that represent each ROI

Descriptor	Dimension
Histogram	Mean
	S.D.
	Skewness
	Kurtosis
Gradient	Mean
	S.D.
Run-length matrix	Short primitive emphasis (SPE)
	Long primitive emphasis (LPE)
Co-occurrence matrix	Angular second moment (ASM)
	Contrast
	Correlation
	Inverse difference moment (IDM)
	Entropy

The machine learning methods, we consider, were: Bayesian classifier, artificial neural network[7-8] (ANN), and support vector machine[9] (SVM). The features employed in this study are listed in Table I.

Ⅲ. Results

The overall sensitivity is presented in Figure 2. The overall sensitivity of ANN with 64x64 ROI discriminates 86.0% obstructive lung diseases obviously better than any other case in this experiment. The performance of ANN, however, is not stable so that has a large variance, while the variance of Bayesian is very small. The ANN shows significantly better performance than other two classifiers in every ROI (t-test, $p<0.01$). ROI size has a significant effect on overall sensitivity. As ROI increase, the overall sensitivity becomes higher in every classifier (t-test, $p<0.01$).

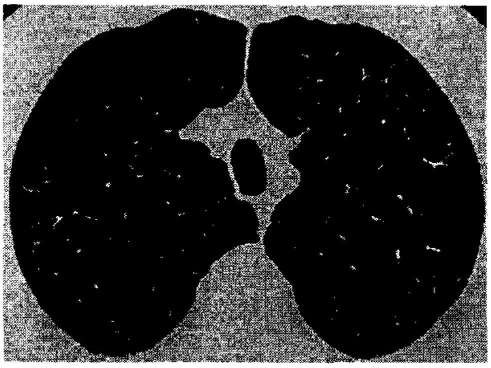


⟨Figure 2⟩ Overall sensitivity basis on the texture feature set

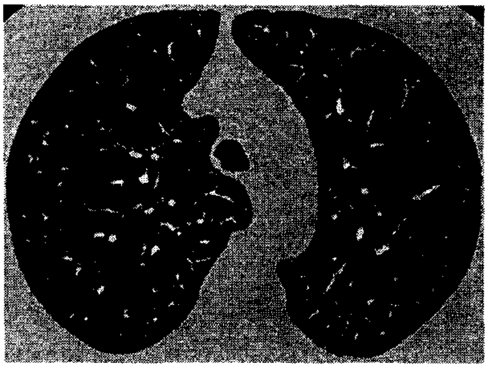
For specificity, it is varied in proportion to sensitivity; ANN shows the best specificity in given ROI size and larger ROI leads higher specificity (t-test, $p<0.01$). Among four classes, mild emphysema is most hard to detect in almost classifiers and ROI, while severe emphysema is easy to detect comparatively. In the term of testing time, SVM is the fastest - just spent less than 0.4 second to test all samples 20 times. ANN needs about 1.2 second to finish testing. The Bayesian, however, consumes more than 7.0 second and it is very poor performance in a comparison to other two classifiers. All tests were performed on an Intel Pentium IV PC containing 3.0 GHz and 1.0 GB of main memory.

Ⅳ. New Work to be Presented

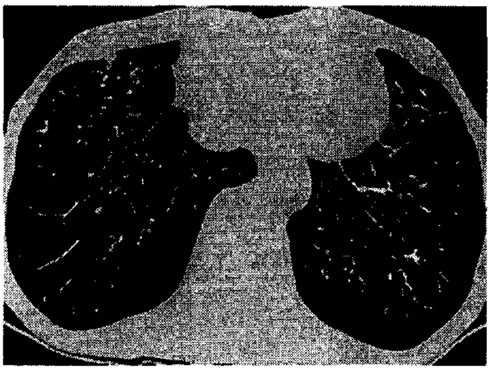
Understanding classifier characteristic and classification result is important in clinical applications. In this paper, we have implemented these three promising machine-learning tools and have investigated whether theses state-of-the-art learning methods could further lead to more accurate classification of obstructive lung diseases. The detailed analysis allows us to compare the results in the terms of not only their accuracy, but also other properties, such as testing time, robustness of different ROI sizes, which are important to the CAD application of.



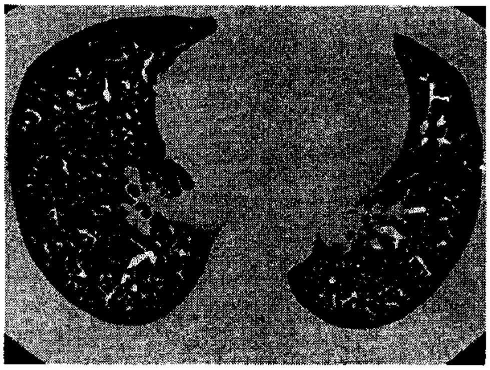
(a) Panlobular emphysema or severe centrilobular emphysema (PLE)



(b) Mild centrilobular emphysema (CLE)



(c) Bronchiolitis obliterans (BO)



(d) Normal lung parenchyma (NL)

Panlobular emphysema or severe centrilobular emphysema (PLE)

Mild centrilobular emphysema (CLE)

Bronchiolitis obliterans (BO)

Normal lung parenchyma (NL)

⟨Figure 3⟩ An application of the machine classifier to differentiation of lung parenchyma. At every pixel, the semi-transparent color was coded by the classification result.

machine classifiers and to clinicians and researchers who would like to get an understanding of the classification process and analysis. An automatic classification system is displayed in Figure 3 (Bayesian classifier, 64x64 pixels of ROI)

V. CONCLUSION

In this study, we investigated the use of the Bayesian classifier, ANN, and SVM for differentiation of obstructive lung diseases. These different machine learning methods were trained and tested to classify a given ROI into the normal or three types of diseases, based on quantitative image features extracted from the HRCT images. Results obtained from three different ROI sizes of experiments demonstrated that the ANN yielded the best performance, outperforming that of the Bayesian and SVM; however, the performance of ANN has a largest variance among three classifiers. In the term of computational cost, the SVM is advantageous and the Bayesian is extremely poor.

References

1. S. Copley, A. Wells, N. Muller, M. Rubens, N. Hollings, J. Cleverley, D. Milne, D. Hansell., "Thin-section CT in obstructive pulmonary disease: discriminatory value," *Radiology* 223(3), 812-819 (2002).
2. F. Chabat, G. Yang, D. Hansell, "Obstructive lung diseases: texture classification for differentiation at CT," *Radiology* 228(3), 871-877 (2003).
3. Y. Xu, E. van Beek, Y. Hwanjo, J. Guo, G. Mclennan, E. Hoffman, "Computer-aided Classification of Interstitial Lung Diseases Via MDCT: 3D Adaptive Multiple Feature Method (3D AMFM)," *Academic Radiology* 13(8), 969-978 (2006).
4. K. Chan, T. Lee, P. Sample, M. Goldbaum, R. Weinreb, T. Sejnowski, "Comparison of machine learning and traditional classifiers in glaucoma diagnosis," *IEEE Transactions on Biomedical Engineering* 49(9), 963-974

(2002).

5. L. Wei, Y. Yang, R. Nishikawa, Y. Jiang, "A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications," *IEEE transactions on medical imaging* 24(3), 371-380 (2005).
6. S. Handrick, B. Naimipour, D. Raicu, J. Furst, "Evaluation of binning strategies for tissue classification in computed tomography images," *Proceedings of SPIE Medical Imaging: Image Processing* 6144, 61444L (2006).
7. C. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, U.K., 1995.
8. B. Ripley, *Pattern recognition and neural networks*, Cambridge Univ. Press, Cambridge, U.K., 1996.
9. V. Vapnik, *The nature of statistical learning theory*, Springer-Verlag, New York, 2006.