

EURRAD 00243

Image analysis methods for solitary pulmonary nodule characterization by computed tomography

D. Cavouras¹, P. Prassopoulos² and N. Pantelidis²

¹Department of Medical Instruments, School of Technological Applications, Technological Educational Institute of Athens and ²Department of Diagnostic Radiology, Airforce General and V.A. Hospital, Athens, Greece

(Received 29 July 1991; accepted after revision 18 October 1991)

Key words: Image, analysis; Image, processing; Thorax, solitary pulmonary nodule; Thorax, image analysis

Abstract

Computer software was designed for classifying solitary pulmonary nodules (SPNs) into benign and malignant from their CT images, using image analysis methods. The system made use of three features, computed from the CT density matrix of the SPN, and a class-discriminating algorithm. System evaluation was performed on 51 histologically confirmed SPNs of indeterminate CT diagnosis. Overall classification accuracy in distinguishing benign and malignant SPNs was 90.2%, while 83.3% of the benign and 93.9% of the malignant SPNs were correctly classified. The proposed system may be of value to the radiologist in assessing the probability of malignancy in patients with a solitary pulmonary nodule.

Introduction

The solitary pulmonary nodule (SPN) is a common diagnostic problem especially in asymptomatic patients [1–3]. Patient management depends mainly on the probability that the SPN is malignant. The diagnostic procedures for estimating the SPN, such as radiography, CT, bronchoscopy and aspiration biopsy, do not always provide reliable diagnosis and, thus, patients undergo a thoracotomy [1,4].

Computed tomography is the best non-interventional method for assessing the SPN, because of its high contrast resolution and capacity for measuring attenuation values [5]. Several reports have attempted to estimate the benignity of the SPN from the fat and calcium distribution in the nodule [1,6–8].

In this study we applied pattern recognition methods to analyze by computer the internal structure of the SPN, from its CT density matrix, in order to discriminate benignity from malignancy.

Material and Methods

The study comprised 95 patients, 73 males and 22 females (age: 39–72 years) who had been referred for CT examination because of a solitary pulmonary nodule over the last 8 years. The nodules were considered indeterminate on the basis of well-established CT criteria [1,2,4–6]. In our study we did not include patients with an SPN larger than 2 cm, cavitation, pleural effusion, hilar adenopathy, atelectasis, and the presence of extensive calcification or fat.

Thirty-seven nodules proved benign (28 granulomas, 5 fibroses, 2 hamartomas, 2 focal inflammations) and 58 malignant. Nodule size, as determined by the largest diameter of the SPN, was between 0.4–1 cm in 18, 1–1.5 cm in 32, and 1.5–2 cm in 45 pulmonary nodules. The presence of calcification(s) was recognized in the CT images of 11 nodules.

The patients underwent a routine CT chest examination (120 kVp, 50 mA, 5 s scan time, 9 mm slice thickness, 1.5 mm pixel size) in deep expiration. Additional scans of the nodule (120 kVp, 50 mA, 0.5 mm pixel size, 3–5 mm slice thickness) were obtained, and the slice through the center of the SPN was selected for further processing.

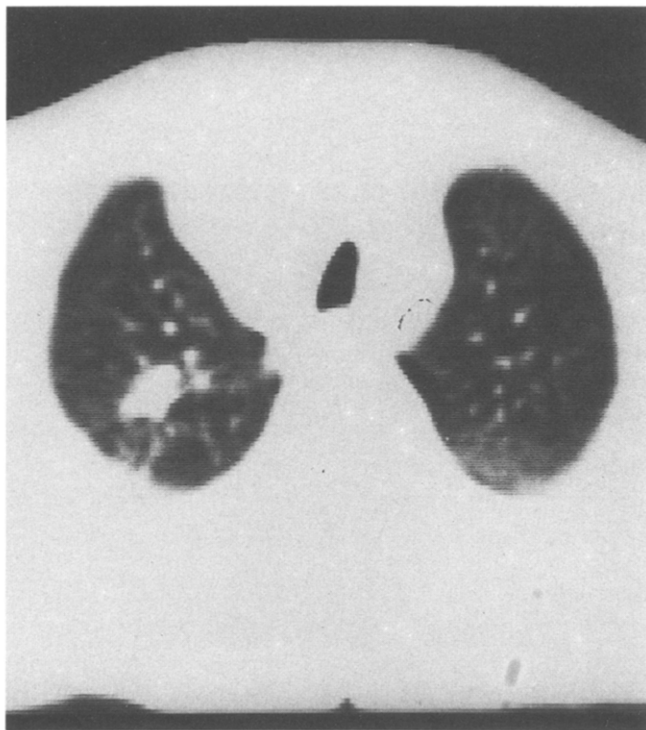
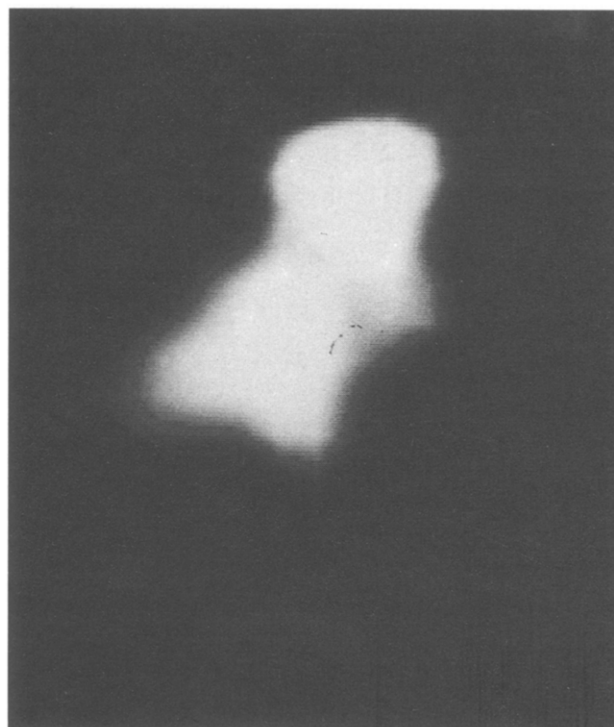
**a****b**

Fig. 1. (a) Right upper lobe CT scan revealing a solitary pulmonary nodule in an asymptomatic patient. (b) Nodule image localization.

The CT images of the nodules (Fig. 1a and b) were transferred to a PC/AT computer, where the software for discriminating benign and malignant SPNs, the Solitary Pulmonary Nodule Discriminating (SPND) system, was developed. The SPNs were separated into two groups: the nodules of the first 44 patients formed group A (19 benign and 25 malignant) and were used for the design of the SPND system, while 51 SPNs formed group B (18 benign and 33 malignant) and were employed in the evaluation of the system.

The design of the SPND system was performed in two stages. In the first, 20 features were computed from the CT density matrix of the SPN. Four were determined on the basis of the criteria commonly used to indicate SPN benignity in CT [1,2,4-6] as follows: (1) the mean value of the 32 highest density values; (2) the least distance of the 4 highest adjacent density values from the SPN margin; (3) the ratio of the number of highest density values (>150 H.U.) over the total number of SPN density values; and (4) the ratio of the highest (>150 H.U.) and lowest (<0 H.U.) density values over the total number of SPN density values. The next four features were computed from the SPN density histogram, which gives the frequency of density values in the SPN image matrix: mean value, variance, skewness and kurtosis [9,10]. The remaining 12 textural

features were calculated from the co-occurrence matrix [9-12] of the SPN, which is a two-dimensional histogram describing the frequency that two adjacent density values occur in the SPN image matrix.

In the second stage of the SPND system design, a commonly used class discriminating algorithm, the Least Squares Minimum Distance (LSMD) classifier [13,14], was employed to automatically characterize the SPNs on the basis of the 20 selected features. The latter were combined in every possible way (i.e., 2, 3, 4 etc. feature combinations) by the LSMD classifier in order to determine the highest classification accuracy with the minimum number of features. Thus, the final SPND system included a section for the computation of the best determined features from the CT image density matrix and a section for the classification of the nodule with the LSMD classifier, employing that best feature combination.

SPND system evaluation was performed by testing its classification accuracy on the 51 SPNs of group B, which were not involved in the design of the system. The CT density matrix of each nodule was transferred to the computer, the best features were computed, the LSMD algorithm classified the SPN into benign or malignant, and the result was compared with the histological diagnosis of the SPN.

Results

In the design stage of the SPND system, every possible combination of the 20 features was employed to classify the 44 SPNs of group A. Best results obtained were 93.2% with the 'mean value-covariance-sumvariance' feature combination, 90.2% with 'mean value-covariance-sumvariance-autocorrelation', 88.6% with 'mean of 32 highest density values-covariance-sumvariance', and 86.4% with 'covariance-sumvariance'. Thus, the first feature combination was selected for use by the SPND system. The mean value was computed from the density histogram of the nodule's CT matrix, and the covariance and sumvariance from the co-occurrence matrix [11,14] of the SPN image.

Table 1 shows the results of the application of the SPND system to the 51 nodules of group B. The SPND system classified correctly 15 of 18 benign (83.3%) and 31 of 33 malignant SPNs (93.9%) and the overall system classification accuracy in discriminating benign and malignant SPNs was 90.2%.

Table 2 shows the mean value and standard deviation of the best three features. Each one showed a highly significant difference ($P < 0.001$) in the means of the benign and malignant SPN groups, using the statistical *t*-test.

TABLE 1

Two-way truth table showing the results of the SPND system classification of 51 SPNs into benign and malignant

| | | SPND | | | % Accuracy |
|----|-----------|--------|-----------|-------------------|------------|
| | | Benign | Malignant | Total Sums | |
| HC | Benign | 15 | 3 | 18 | 83.3 |
| | Malignant | 2 | 31 | 33 | 93.9 |
| | | | | Overall accuracy: | 90.2 |

SPND: Solitary Pulmonary Nodule Discriminating system.

HC: Histological Confirmation.

TABLE 2

Mean value, standard deviation and statistical significance of the best three features computed from the 51 SPNs of group B

| | Benign (18 SPNs) | Malignant (33 SPNs) | Statistical significance (<i>P</i>) |
|-----------------|---------------------|------------------------|--|
| Mean value (HU) | 76.9 ± 11.8 | 37.5 ± 8.9 | <0.001 |
| Covariance | 4.4 ± 1.1 | 1.8 ± 0.8 | <0.001 |
| Sumvariance | 72.6 ± 25.6 | 40.3 ± 18.4 | <0.001 |

The SPND system classified each nodule in less than 3 s on an AT/286 personal computer.

Discussion

Computed Tomography is the best imaging procedure in evaluating the benignity of the solitary pulmonary nodule [1,2,5]. Several studies have attempted to assess SPN nature from the calcium and fat content [1,6-8]. Furthermore, several reports have tried to get better insight by examining the 'print-out' of the nodule's CT density matrix [4,7]. Our work was based on the analysis of that density matrix by computer, so that additional features of the SPN, which are difficult to perceive visually, could be used as well. Thus, textural properties, such as the rate of density transitions, homogeneity, or contrast, were precisely calculated using statistical parameters [11,14]. Pictorial information, which can only be described mathematically by such parameters as the sumvariance and covariance, were also evaluated. To our knowledge, no reports have been published that employ similar methods on CT to discriminate SPNs.

Our study indicates that the application of image analysis methods to SPN characterization may be of value. The results obtained were comparable to those achieved by fine needle aspiration biopsy. SPND system classification accuracy was 83.3% for the benign nodules, which is close to 88% of the aspiration biopsy [15], while for the malignant SPNs the results were similar, 93.9% for the SPND system and 95% for the aspiration biopsy.

In a study by Cummings et al. [16], a successful attempt has been made to estimate the likelihood of SPN malignancy by combining parameters from the patient history and the SPN size in a scheme based on Bayes theorem. The authors suggested that new parameters, such as CT density measurements of the SPN, could also be incorporated into their scheme. We believe that the three best features determined in our study could be used for that purpose.

Our work was based only on features computed from the density values of the SPN CT image. If additional features were employed, such as the nodule's contour and size [5], then the SPND system could possibly achieve higher classification accuracy.

Acknowledgement

We gratefully acknowledge Mr. S. Goulinopoulos for his assistance in the collection of the material of this work.

References

- 1 Caskey CI, Templeton PA, Zerhouni EA. Current evaluation of the solitary pulmonary nodule. *Radiol Clin North Am* 1990; 28:511-520.
- 2 Sagel SS, Glazer HS. Lung, pleural, chest wall. In: Lee JK, Sagel SS, Stanley RJ, ed. *Computed body tomography*. New York: Raven Press: 1989; 295-387.
- 3 Cummings SR, Lillington GA, Richard RJ. Managing solitary pulmonary nodules: the choice of strategy is a 'close call'. *Am Rev Respir Dis* 1986; 134: 453-460.
- 4 Siegelman SS, Zerhouni EA, Leo FP, Khouri NF, Stitik FP. CT of the solitary pulmonary nodule. *AJR* 1980; 135: 1-13.
- 5 Siegelman SS, Khouri NF, Leo FP, Fishman EK, Braverman RM, Zerhouni EA. Solitary pulmonary nodules: CT assessment. *Radiology* 1986; 160: 307-312.
- 6 Siegelman SS, Khouri NF, Scott WW, Leo FP, Hamper UM, Fishman EK, Zerhouni EA. Pulmonary hamartoma: CT findings. *Radiology* 1986; 160: 313-317.
- 7 Proto AV, Thomas SR. Pulmonary nodules studied by computed tomography. *Radiology* 1985; 156: 149-153.
- 8 Fraser RG, Sanders C, Niklason LT, Sabbagh EA, Luna RF, Alexandra CB, Robinson CA, LA Katzenstein, Barnes GT. Calcification in pulmonary nodules: detection with dual energy digital radiography. *Radiology* 1986; 160: 595-601.
- 9 Chen PC and Pavlidis T. Segmentation by texture using a co-occurrence matrix and a split-and-merge algorithm. *Comput Graph Image Process* 1979; 10: 172-182.
- 10 Raeth U, Schlaps D, Limberg B, Zuna I, Lorenz A, Van Kaik G, Lorenz WJ, Kommerell B. Diagnostic accuracy of computerized B-scan texture analysis and conventional ultrasonography in diffuse parenchymal and malignant liver disease. *J Clin Ultrasound* 1985; 13: 87-89.
- 11 Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973; SMC-3(6): 610-621.
- 12 Haralick RM. Statistical and structural approaches to texture. *Proc IEEE* 1979; 67: 786-804.
- 13 Ahmed N, Rao KR. Feature selection in pattern recognition. In: *Orthogonal transforms for digital signal processing*. New York: Springer-Verlag, 1975; 225-253.
- 14 Lerski RA, Smith MJ, Morley P, Barnett E, Mills PR, Watkinson G, MacSween RN. Discriminant analysis of ultrasonic texture data in diffuse alcoholic liver disease. *Ultrasonic Imaging* 1981; 3: 164-172.
- 15 Khouri NF, Stitik FP, Erosa YS. Transthoracic needle aspiration of benign and malignant lung lesions. *AJR* 1985; 144: 281-288.
- 16 Cummings SR, Lillington GA, Richard RJ. Estimating the probability of malignancy in solitary pulmonary nodules. A Bayesian approach. *Am Rev Respir Dis* 1986; 134: 449-452.