A Nodule Detection System for Postero-Anterior Chest Radiographs

Paola Campadelli*, Elena Casiraghi*

**Dipartimento di Scienze dell'Informazione*
Università degli Studi di Milano
Via Comelico, 39/41 20135 Milano, Italy.

campadelli,casiraghi@dsi.unimi.it

ABSTRACT *In this paper we describe a method processing Postero Anterior chest radiographs to extract a set of nodule candidate regions characterized by a low cardinality and a high sensitivity ratio. It is based on two consecutive multiscale procedures to first enhance the visibility of the nodules and then extract a first set of candidates. To reduce its cardinality two different methods, applied to a set of most representative features, are described and compared: a rule based system and a feed-forward neural network trained by back-propagation. Both the systems had similarly good results. The method has been developed and tested on a standard database containing* 154 *radiographs of patients with lung nodules and* 93 *with no nodules. The final set of candidates detected contains about* 10000 *regions, approximately* 40 *per image, and* 8 *true positives out of* 154 *are lost; the sensitivity ratio of the system is then equal to* 0.95%.

KEYWORDS CAD systems; Multiscale Image enhancement; Pattern classification

1 Introduction

In the field of medical diagnosis the chest radiography is the most common type of procedure for the initial detection and diagnosis of lung cancer, due to its noninvasivity characteristics, radiation dose and economic considerations. Studies such as [1] explain why chest radiograph is one of the most challenging radiograph to produce technically and to interpret diagnostically. Several studies in the last two decades, as [2] and [3], calculated an average miss rate of 30% for the radiographic detection of early lung nodules by humans, and demonstrated that 90% of peripheral lung cancers were visible in radiographs produced earlier than the date of the cancer discovery by the radiologist. These results showed the potentiality of early diagnosis improvement, suggesting the use of computer programs for radiographs analysis. These are the main reasons why in the last two decades a great deal of research work has been devoted to the study of systems aimed to lung nodules detection, and a wide variety of them have been already proposed [4], [5], [6], [7], [8], [9], [10], [11], [12], [13]. Most of the methods presented in the literature are based on a three stage

processing scheme. The purpose of the first two steps is to increase the visibility, also called spicularity, of the nodules and then to extract all the regions that may contain nodules; the last stage is focused on the selection of the real nodules within the set of the extracted candidates, whose cardinality is usually quite high; this is generally done by means of a two step classification method, to reduce the number of candidates and then extract the final ones.
The main problem of the presented schemes is the high number of false positives; such problem has been faced with two different strategies: either to reduce the number of candidates extracted by the first two steps [12], or to leave to proper classifiers the task of reduction [5], [9]; in both cases however many true positives are discarded, leaving the problem open.
In this paper we present an approach to enhance the visibility of the nodules, and then extract a set of candidate regions with a low cardinality and a high sensitivity ratio. A novelty with respect to the methods in the literature is the fact that both the steps described are based on a multiscale analysis of the image in order to handle all the possible sizes of the nodules. A comparison with results presented in the literature shows the good performance of the system developed.

2 Materials

The method has been developed and tested on a standard database acquired by the Japanese Society of Radiological Technology. It is a standard database containing a total of 247 radiographs: 154 containing lung nodules and 93 of patients with no disease. The images were digitized with a 0.165 mm pixel size, a matrix size of 2048 by 2048, and 4096 gray levels. The diameter of the nodules ranges from 5 to 35 mm. All the nodules in the images have been classified according to the difficulties encountered in their detection by the radiologists. They have been divided in 5 classes ranging from obvious to extremely subtle. The algorithms described in sections 4 and 5 work with images down-sampled to a dimension of 256 by 256 pixels (we will refer to them as the $Original\ Images$). This size has been chosen experimentally to reduce their computational costs without worsening the performances. The method described in section 6 works on images down-sampled to a dimension of 512 by 512 pixels.

3 Enhancing the spicularity of the nodules

At first, the lung area is identified using the algorithm described in [14] and extended in order to include the parts behind the heart, near the spinal column and behind the diaphragm, where lung nodules may be present. The method adopted to extend this area is described in [15]. Note that errors due to the segmentation algorithm may influence the performances of the overall system:

we loose one nodule which is located outside the detected lung area.
The algorithm used to enhance the spicularity of the nodules is based on a difference approach to search for circular areas surrounded by a darker ring, but it also considers several background images created at consecutive scales; the enhanced image is then created by combining all the difference images obtained at the several scales. We think that this multiscale approach is the missing part of the methods presented in the literature: they are able to highlights nodules with a size belonging to a limited range related to the shape of the smoothing filter used to get the only background image used.
In our scheme we produce several smoothed version of the $Original\ Image$ by convolving it with gaussian filters whose standard deviation, s, takes values in the range $2-12$, according to the minimum and maximum possible pixel size of the nodule radius. For each scale, s, we then subtract from the $Original\ Image$ its smoothed version, to get a resulting $Difference\ Image$ where the details visible at that scale are enhanced. Since the distribution of gray levels in a nodule sub-image can be approximated by a gaussian, the result of subtracting to a nodule sub-image its smoothed version is usually an image with a positive peak in the central part of the nodule, and negative values in the neighborhood. Moreover the histogram of the $Difference\ Image$ shows that most of the pixels take negative values while, on the set of positive values, a peak can always be identified. We create a binary image by selecting all the pixels with a value bigger than the one corresponding to the peak; these pixels are the ones corresponding to the highest frequencies, that is the details, that can be identified at the scale s. Summing up all the binary images obtained at different scales we get a final $Sum\ Image$ (figure 1).

4 Extracting the nodule candidates

In the $Sum\ Image$ the nodules often appear as regions with circular shape of different sizes, characterized by the highest gray levels at the center and surrounded by a much darker ring. Based on this observation we process this image looking for a measure which helps in selecting the pixels corresponding to the centers of nodules. To handle all the possible sizes of the nodules the procedure described below is repeated for all the possible radius values $r = 1, .., 12$ and all the results are then combined.
Having fixed the radius r, we calculate for each pixel $P(x, y)$ a coefficient $P_r(x, y)$ which measures the contrast between a circular region with center $P(x, y)$ (and radius r) and its surrounding; $P_r(x, y)$ is thus defined:

$$P_r(x, y) = MEAN(Circle_r(P(x, y))) - MEAN(Ring_r(P(x, y))) \quad (1)$$

where $MEAN(X)$ is the mean of the gray values of the pixels inside a generic region X; $Circle_r(P(x, y))$ is the region composed by the pixels contained in the circle of radius r and centered in $P(x, y)$; $Ring_r(P(x, y))$ is the region

composed by the pixels in the 2-pixel-thick ring around the $Circle_r(P(x,y))$. Note that the thickness of the ring is fixed to 2 for every radius. This is because what allows to identify a circular region is a darker ring surrounding it, no matter which is the thickness of the ring itself.
To select the pixels which are potential nodule centers, we automatically define a threshold on the set of the coefficients $\{P_r(x,y)\}$, by means of the algorithm described by Kapur in [16], thus obtaining a *binary image*.
For each connected region in it, we calculate the circularity as defined in [4] and the biggest diagonal, D, of the minimum ellipse containing the region itself. We then discard a candidate either if its circularity is lower than 0.5 or D is bigger than $2r$. The regions left correspond to the candidate nodules with radius r. Repeating the procedure for every possible radius we obtain a set of 11 *binary images* $B(r)$, each containing a set of candidate nodules. All these images are then combined to determine the final set of candidates. First, all the regions appearing in only one of the *binary images* are taken as candidates. For the others the following procedure is employed: when two regions, X and Y, belonging to $B(r_1)$ and $B(r_2)$ (r_1 and r_2 being two consecutive radius values) intersect, their union U is at first considered. If the biggest diagonal D of the minimum ellipse containing it is less than $2r_2$, then U is taken as representative; otherwise we calculate the means

$$M_X = \frac{1}{\mid X \mid} \sum_{p \in X} P_{r_1} \quad \textit{and} \quad M_Y = \frac{1}{\mid Y \mid} \sum_{p \in Y} P_{r_2} \tag{2}$$

and take as representative region the one with the higher value. We then build a gray level image, called *Regions Image* (figure 1), by assigning to each pixel in each candidate region the value

$$G(x,y) = \max_{r \in [2,12]} (P_r(x,y)) \tag{3}$$

and then scaling it in the range $[0, 255]$.
With this extraction scheme we get a set of about 24000 regions on the 247 images of the database, with an average of about 100 regions per image and only 7 true positives lost out of 153. These results have been compared with those described [9] and [6]. The first method has been tested on the same database and applied to the lung area defined by [14], bringing to a loss of 12 true positives, out of 154, before the candidate extraction. The result of the extraction scheme is a set of 33000 candidates and a loss of other 8 true positives. We implemented the second method and applied it to the same lung area used in [14] obtaining really poor results. A more extensive comparison with other methods would need an implementation of all them, since none has been tested on this database. In terms of number of candidates obtained at this first stage it can be pointed out that we use a lung area that is about 1.5 times bigger than the one commonly considered, and this has some influence on the cardinality of the obtained set.

5 Pruning the candidate set

To reduce the number of the extracted candidates we searched for a set of rules which could describe the main characteristics of the real nodules, hence allowing us to discard some false positives. To this end we calculated for each region a set of 40 features and studied their distribution. The statistical analysis allowed us to select a set of 16 most representative features, whose combination by means of simple rules, has proved to be effective for a first candidates selection. The created rule based scheme is indeed able to detect and discard more than half false positives. In the following we will describe just the selected features. They are based on the shape and position of the region, the gray level distribution in the original radiograph down-sampled to the dimension of 512×512 pixels, the values of the gray levels in the *Regions Image*, and the set of coefficients $P_r(x, y)$ associated to each pixel for each radius value.
Six features are based on the shape; they are:
circularity, as defined in [4]; effective radius, that is the radius of the circle with an area equivalent to the one of the region; the ratio between the perimeter of the region and the perimeter of the circle with radius of 12 pixels (we are considering $r = 12$ as maximum radius value); the ratio between the area of the region and the area of the circle with radius of 12 pixels; the ratio between the two dimensions of the minimum bounding box including the region; the ratio between the two dimensions of the maximum bounding box including the region.

Two features are based on the gray level distribution of the pixels in the original radiograph down-sampled to the dimension of 512×512 pixels: they are simply the mean and the standard deviation of the gray levels in each region.
Three features are calculated on the gray level of the pixels in the *Regions Image*: they are the mean, the contrast and the maximum value of the gray levels. The contrast is calculated as the difference between the maximum and minimum value of the gray levels in the region.
The position feature has been introduced to eliminate false positives detected on the rib cage boundaries, which are characterized by the fact that they are attached to the lung borders and have an elongated shape. It is calculated considering the external contour of the region and it is the fraction of the number of pixels of the contour which lay outside the lung area with respect to the total number of the pixels in the contour itself.
Two features are calculated as an estimate of the most characteristic radius value to be associated to a generic region X. We use two different methods to get it and hence obtain two values that can be compared. One method first calculates for each pixel $P(x, y)$ a most eligible radius $rad(x, y)$. This is done by considering all the $P_r(x, y)$ obtained for that pixel and then calculating:

$$rad(x, y) = Argmax_{i \in [2,12]}(P_i(x, y)) \tag{4}$$

The first radius, R_X^1, associated to X is then $R_X^1 = \max_{(x,y)\in X} rad(x,y)$.
The second method calculates the radius R_X^2 by considering for each different value of r the sum

$$Sum_X(r) = \sum_{(x,y)\in X} P_r(x,y) \tag{5}$$

R_X^2 is then calculated as

$$R_X^2 = Argmax_{i\in[2,12]}(Sum_X(r)) \tag{6}$$

The fact that R_X^1 and R_X^2 are very similar in case of true positive elements and significantly different for many false positives allows us to recognize some regions as false positives.
The features described above are calculated by considering each region separately. The remaining *two* features instead have been calculated in order to make a comparison between the regions belonging to the same image. The underlying assumption is that a nodule should be somehow different from the other false positives in the same lung area. Let us suppose that in the *Regions Image* there are N elements. For each radius r, we build an array of length N whose components are the values $S_X(r)$. Sorting the 11 arrays we can associate to each region X the 11 indexes that describe its position in each array. We then characterize a region X with the maximum among its 11 indexes and the corresponding value $Sum_X(r)$.
Applying 8 simple rules which basically describe the relationships observed between pairs of features, we can easily discard 14011 false positives without loosing any true positives, hence reaching a sensitivity ratio equal to 0.95 and a total number of candidates equal to 10352, approximately 40 candidates per image.
In figure 1 we show an examples of the results of the pruning method on a patient with an extremely subtle nodule.
Our results can be compared with the ones of the method described in [9] in which the authors apply a classifier that selects 5028 candidates from the first set composed by 33000 regions, but loosing other 15 true positives in addition to the 20 already lost, for a total of 35 false negatives. Our pruning method is able to discard less false positives but does not loose any nodules. Although the rules chosen are very intuitive and simple they need some thresholds to be experimentally set; therefore they may depend on the images in the database.

For this reason we also tested a feed-forward neural network trained by standard back-propagation. The advantage of this system with respect to the rule based one is that, once the network is trained, it does not need any threshold to be set.
The input of the network is represented by a vector $X = [x_1, ..., x_{16}, x_{17}, x_{18}]$ where $x_1, .., x_{16}$ represent the features previously described, and x_{17}, x_{18} are the spatial coordinates of the center of mass of the nodule; they are expressed in a local coordinate system which has its origin in the center of mass of the lung fields and it is scaled with respect to the width and length of the lung

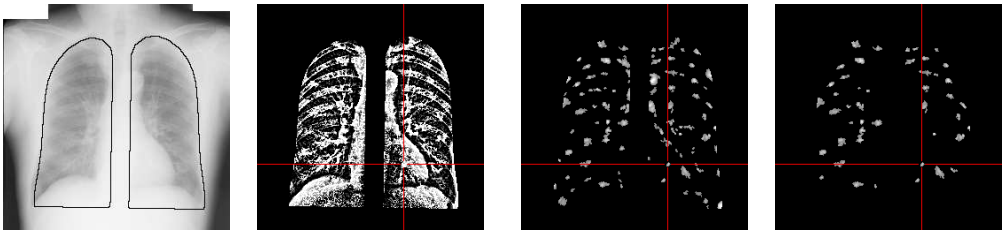

Figure 1: *Original Image*, *Sum Image*, *Regions Image* and Pruned set of Regions

area. We apply to the input vector a preprocessing aimed to data normalization, followed by a scaling that brings all the values to the range $[0.0, 1.0]$. Since the two classes to be recognized, "*nodules*" and "*NOT* "*nodules*", are highly unbalanced (the "*nodules*" are 146, the "*NOT nodules*" are 24000) our choice was to train the neural network in order to recognize and discard the regions whose set of features is totally different from the one of the nodules. To this end we classified the data obtained as "*possible nodules*" and "*NOT nodules*". This classification was realized considering as "*NOT nodules*" the ones discarded by the rule based system and as "*possible nodules*" all the others, hence obtaining two classes with respectively 14011 and 10352 elements. We made experiments using both a training and a validation set. They were formed by randomly choosing from both the classes, the 50% of the elements for the training, and the 10% for the validation. The remaining elements (40% of the total) were used for testing. We used a neural network with 1 hidden layer composed of 8 neurons and an output layer with 2 neurons, and made several experiments changing the parameters of the learning algorithm, the maximum number of epochs for the training, the minimum error allowed on the training set, and the elements in the input set. During these experiments an input data was classified as belonging to a class when the corresponding output neuron had a value bigger than 0.7. The obtained results are comparable with those of the rule based system: the network does not loose any true positives and in the worst case the number of false positives detected is never less than the 99% of the number recognized by the rule based system.

7 Future work

The results shown, although fairly good from the point of view of the reached sensitivity ratio, are still not satisfactory with respect to the number of remaining false positives. Their number is still too high to be helpful for radiologists during their decision making process. We are therefore looking for other learning strategies and trying to combine different features and image data acquired from the original (i.e. not down-sampled) radiographs.

REFERENCES

[1] Cj Vyborny, 1997.,*The AAPM/RSNA physics tutorial for residents: Image quality and the clinical radiographic examination.,* Radiographics, **17**, 479-498

[2] J. Forrest and P. Friedman, 1982.,*Radiologic errors in patient with lung cancer.,* West Journal on Med., **134**, 485-490

[3] J.H.M. Austin and B.M. Romeny and L.S. Goldsmith, 1992.,*Missed bronchogenic carcinoma: radiographic findings in 27 patients with a potentially resectable lesion evident in retrospect.,* Radiology, **182**, 115-122

[4] M. Giger and K. Doi and H. Mac Mahon, 1988.,*Image feature analisys and computer-aided diagnosis in digital radiography:Automated detection of nodules in peripheral lung fields.,* Medical Physics, **15**, 158-166

[5] J.-S. Lin and S.-C. Lo and M. Freedman and S.Mun, 1996.,*Reduction of false positives in lung nodule detection using a two-level neural classification.,* IEEE Trans. Med. Imag., **15**, 206-217 **[6]** Keserci,Bilgin and Yoshida,Hiroyuki, 2002.,*Computerized detection of pulmonary nodules in chest radiographs based on morphological features and wavelet snake model.,* Medical Image Analysis, **6**, 431-447

[7] T. Matsumoto and H. Yoshimura and K. Doi and M. Giger and A. Kano and H. MacMahon and M. Abe and S. Montner , 1992.,*Image feature analisys of false-positives diagnosis produced by automated detection of lung nodules.,* J. of applied statistics, **27**,587-579

[8] M. Penedo and M. Carreira and A. Mosquera and D. Cabello, 1998.,*Computer aided diagnosis: A Neural Network based approach to lung nodule detection.,* IEEE Trans. Med. Imag.,**27**,872-880

[9] A. Schilham,B.Van Ginneken and M. Loog, 1998.,*Multi-scale nodule detection in chest radiographs.,* Proc. MICCAI,

[10] H. Yoshimura and M. Giger and K.Doi and H. MacMahon and S. Montner, 1992.,*Computerized scheme for the detection of pulmonary nodules: A non linear filtering technique.,* Investigative radiol.,**27**,124-127

[11] H. Yoshida and X. Xu and K. Doi and M. Giger, 1995.,*Computer-aided Diagnosis (CAD) scheme for detecting pulmonary nodules using wavelet transforms.,* Investigative radiol.,**27**,124-127

[12] X.-W. Xu and S. Katsuragawa and A. Ashizawa and H. MacMahon and K. Doi, 1998.,*Analisys of image features of histograms of edge gradient for false positive reduction in lung nodule detection in chest radiographs.,* Proc SPIE ,**3338**,318-326

[13] J.Wei and Y. Hagihara and H. Kobatake, 1999.,*Detection of cancerous tumors on chest X-ray images- Candidates detection filtering and its evaluation.,* Proc ICIP'99,**3338**,318-326

[14] B. van Ginneken and B.M. ter H. Romeny, 2000.,*Automatic segmentation of lung fields in Chest Radiographs.,* Medical Physics, **27**, 2445-2455

[15] Paola Campadelli and Elena Casiraghi, 2003.,*Lung Edge Detection in Postero Anterior Chest Radiographs.,* Proc. of Image2003 **15**, 27-37

[16] J. N. Kapur and P. K. Sahoo and A.K. C.Woong, 1985.,*A new method for gray level picture thresholding using the entropy of the histogram.,* Computer Vision Graphics and Image Processing, **29**, 273-285

NOTE - *work financed by University of Milan (grant FIRST* 2003*)*