PRUNING THE NODULE CANDIDATE SET IN POSTERO ANTERIOR CHEST RADIOGRAPHS

Paola Campadelli and Elena Casiraghi
Universitą degli Studi di Milano,
Dipartimento di Scienze dell'Informazione,
via Comelico 39/41,
20135, Milano
{Campadelli, Casiraghi}@dsi.unimi.it

`http://homes.dsi.unimi.it/∼campadel/LAIV/index.htm`

Abstract In this paper we describe and compare two different methods to reduce the cardinality of the set of candidates nodules, characterized by an high sensitivity ratio, and extracted from PA chest radiographs by a fully automatized method. The methods are a rule based system and a feed-forward neural network trained by back-propagation. Both the systems allow to recognize almost the 75% of false positives without losing any true positives.

Keywords: CAD Systems, Neural Networks, Support Vector Machines

Introduction

In the field of medical diagnosis the chest radiography is by far 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. Several studies in the last two decades (see [Austin et al., 1992] and [Forrest and Friedman, 1982]) calculated an average miss rate of 30% for the radiographic detection of early lung nodules by humans. In a large lung cancer screening program 90% of peripheral lung cancers have been found to be visible in radiographs produced earlier than the date of the cancer discovery by the radiologist. This explains why in the last two decades a great deal of research work has been devoted to the development of Computer Aided Diagnostic (CAD) systems aimed to lung nodules detection, and a wide variety of them have been already proposed and reviewed in [van Ginneken et al., 2001]. Most of the methods presented in the literature are based on a two stage processing scheme. At first, the image is processed in order to increase the visibility(conspicuity) of the nodules; the regions that may contain nodules are then extracted by means of rule based schemes which exploit the main nodule characteristics, such as the gray level and the circular shape. The second step aims to the selection of the real nodules within the set of the extracted candidates; this is generally done by means of classification methods, which first reduce the number of candidates, usually quite high, and then extract the final ones. The main issue regarding this step is the choice of a proper set of features to describe and represent the data so that the samples belonging to the two classes may easily distinguished. Moreover 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 ([Wu et al., 1994], [Xu et al., 1998], [Yoshida et al., 1995]), or to leave to proper classifiers the task of reduction ([Penedo et al., 1998], [Lin et al., 1996], [Lo et al., 1993], [A. Schilham and Loog, 2003]); in both cases however many true positives are discarded, leaving the problem open.

We developed a fully automatized method using multiscale approaches to segment the lungs (see [Columbano, 2004]) and then enhance the visibility of the nodules in the area detected (see [Campadelli and Casiraghi, 2004]). The multiscale analysis of the image is a fundamental key for both the task of segmentation and nodule detection; in the first case it allows to capture the details of the lung borders which belong to different scales, and in the second case it is used to handle all the possible sizes of the nodules.

This multiscale approach is an additional element with respect to the method presented in the literature and it has proven to be necessary and effective. First of all, it allows to increase the visibility of also the most subtle nodules, facilitating the extraction task which indeed loses very few true positives and creates a set with an high sensitivity ratio. Second, it is employed to extract a set of features which is different from the ones commonly presented in the literature since it captures the properties of the candidates at various scales; in this paper we present our experiments using these features as input to trim the set of candidates initially extracted; to this aim we experimented two different methods and compared their performances: they are a rule based system and a feed-forward neural network trained by back-propagation. Both the systems had similar results.

1. Materials and methods

The method has been tested on a standard database acquired by the Japanese Society of Radiological Technology. It contains 247 radiographs: 154 containing lung nodules and 93 of patients with no disease. The images were digitized with a 0.175 mm pixel size, a matrix size of 2048×2048, 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 algorithm for the candidate extraction works on images down-sampled to an experimentally set dimension of 256×256 pixels to reduce its computational cost without worsening the performances. The features used as input for nodule classification are instead calculated on the images with 512×512 pixels to have a more detailed representation.

2. Extraction and pruning of the nodule candidate set

The method starts with the segmentation of the images to extract the lung fields (see [Columbano, 2004]) and the enhancement of the conspicuity of the nodules as described in [Campadelli and Casiraghi, 2004]. In [Paola Campadelli and Columbano, 2004] we describe a multiscale method analyzing the *enhanced images* to extract a unique set of candidate regions; it also allows to create an image I_r for different nodule radiuses, $r = 2, .., 12$. The value $I_r(x, y)$, associated to each pixel (x, y), is a measure which represents the pixel as a potential center of a nodule of radius r. The combination of the 11 images I_r is a gray level image $G(x, y)$, obtained by assigning to each pixel in each candidate region the value

$$G(x, y) = \max_{r \in [2,12]} (I_r(x, y)) \tag{1}$$

and then scaling it in the range $[0, 255]$. All the details of the method are described in [Campadelli and Casiraghi, 2004].
With this extraction scheme we get a set of about 32000 regions on the 247 images of the database, with an average of about 130 regions per image and only 5 true positives lost out of 154. The comparison of these results with those presented in the literature and described in [A. Schilham and Loog, 2003] and [Keserci and Hiroyuki, 2002] showed the better performance of our method (see [Campadelli and Casiraghi, 2004]).

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 12 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 image $G(x,y)$, and the set of coefficients $I_r(x,y)$ associated to each pixel for each radius value.
Six features are based on the shape; they are:

- circularity, as defined in [Giger et al., 1988];
- 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;

One feature is based on the gray level distribution of the pixels in the original radiograph down-sampled to the dimension of 512×512 pixels: it is simply the mean of the gray levels of the pixels in each region.
Two features are calculated on the gray level of the pixels in the image $G(x,y)$: they are the mean and the maximum value of the gray level of the pixels.
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 (x,y) a most eligible radius $rad(x,y)$. This is done by considering all the $I_r(x,y)$ obtained for that pixel and then calculating:

$$rad(x,y) = Argmax_{r \in [2,12]}(I_r(x,y)) \tag{2}$$

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

$$Sum_X(r) = \sum_{(x,y)\in X} I_r(x,y) \tag{3}$$

Again R_X^2 is calculated according to:

$$R_X^2 = Argmax_{r \in [2,12]}(Sum_X(r)) \tag{4}$$

We note that R_X^1 and R_X^2 are very similar in case of true positive elements and significantly different for many false positives; based on this fact we are allowed to recognize some regions as false positives.
The remaining feature is the maximum value of the various $Sum_X(r)$, hence the one associated with R_X^2.

This set of features is the input of a rule based system composed by 12 simple rules: 8 of them basically describe the relationships observed between pairs of features, meanwhile 4 are based on simple tresholding. The system can easily discard 19500 false positives without loosing any true positives, hence reaching a sensitivity ratio equal to 0.97 and a total number of candidates that is about 12000, approximately 50 candidates per image.
Our results can be compared with the ones of the method described in [A. Schilham and Loog, 2003] 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 36 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.

3. Employing Neural Network Classifiers for candidates selection

In this section we describe experiments aimed to the use of a feed-forward neural network trained by standard back-propagation to trim the first set of extracted candidates. 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. Moreover we expect that learning algorithms working with the global set of features as input, could have better performances since they could learn more complicated relationships than the simple ones of our rule based system. In the following we describe our experiments aimed to this purpose.
The input of the network is represented by a vector $X = [x_1, ..., x_{12}, x_{13}, x_{14}]$ where $x_1, .., x_{12}$ represent the features previously described, and x_{13}, x_{14} 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 area. We apply to the input vector a preprocessing aimed to data normalization, fol-

lowed by a scaling that brings all the values to the range $[0.0, 1.0]$.
One observation to be done is that the two classes to be recognized, *"nodules"* and *"NOT nodules"*, are highly unbalanced (the *"nodules"* are 149, the *"NOT nodules"* are about 32000). Our choice was then 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 about 20000 and 12000 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 momentum and the learning rate) the maximum number of epochs for the training, the minimum error allowed on the training set, and especially 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.
The final number of candidates obtained by both systems described above is still quite high, and needs a further reduction in order to be useful for clinical purposes. To perform this reduction we experimented neural networks whose input is the sub-image of the candidate itself created by extracting from the original radiograph an area of dimension 400 by 400 centered on the centroid of the candidate region considered. Several experiments have been done using as input down-sampled versions of the sub-images created. We tried with images down-sampled to 40x40, 20x20, 10x10, 8x8 and trained architectures of different complexities. The best results obtained on the test set composed of 11000 false positives samples and 30 true positives allowed to maintain only 1300 false positives but caused the loss of 10 true positives. Since this results are not acceptable we are now looking for other features and different or more complex classifiers, such as Support Vector Machines.

References

A. Schilham, B.Van Ginneken and Loog, M. (2003). Multi-scale nodule detection in chest radiographs. *Proc. MICCAI*.

Austin, J.H.M., Romeny, B.M., and Goldsmith, L.S. (1992). Missed bronchogenic carcinoma: radiographic findings in 27 patients with apotentially resectable lesion evident in retrospect. *Radiology*, 182:115–122.

Campadelli, Paola and Casiraghi, Elena (2004). Nodule detection in postero anterior chest radiographs. *Submitted to MCO2004*.

Columbano, Simon (2004). Analisi di immagini toraciche e relativa segmentazione. *Master Thesis, Universitàdegli Studi di Milano*.

Forrest, J. and Friedman, P. (1982). Radiologic errors in patient with lung cancer. *West Journal on Med.*, 134:485–490.

Giger, M., Doi, K., and Mahon, H. Mac (1988). Image feature analisys and computer-aided diagnosis in digital radiography:automated detection of nodules in peripheral lung fields. *Med. Phisycs*, 15:158–166.

Keserci, B. and Hiroyuki, Y. (2002). Computerized detection of pulmonary nodules in chest radiographs based on morphological features and wavelet snake model. *Medical Image Analysis*, pages 431–447.

Lin, J.-S., Lo, S.-C., Freedman, M., 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.

Lo, S.-C., Freedman, M., Lin, J.-S., and S.Mun (1993). Automatic lung nodule detection using profile matching and backpropagation neural network techniques. *Journal Digital Imaging*, 1:48–54.

Paola Campadelli, Elena Casiraghi and Columbano, Stefano (2004). Lung segmentation and nodule detection in postero anterior chest radiographs. *Accepted to GIRPR2004*.

Penedo, M., Carreira, M., Mosquera, A., and Cabello, D. (1998). Computer aided diagnosis: A neural network based approach to lung nodule detection. *IEEE Trans. Med. Imag.*, 17:872–880.

van Ginneken, B., ter H. Romeny, B.M., and Viergever, M.A. (2001). Computer-aided diagnosis in chest radiography: A survey. *IEEE Trans. On Med. Imag.*, 20:1228–1241.

Wu, Y., Doi, K., Giger, M., Metz, C., and Zhang, W. (1994). Reduction of false positives in computerized detection of lung nodules in chest radiographs using convolution neural network techniques and application for lung nodule detection. *J. Digital Imaging*, 7:196–207.

Xu, X.-W., Katsuragawa, S., Ashizawa, A., MacMahon, H., and Doi, K. (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.

Yoshida, H., Xu, X., Doi, K., and Giger, M. (1995). Computer-aided diagnosis (cad) scheme for detecting pulmonary nodules using wavelet transforms. *Proc. SPIE*, 2434:621–626.