DETECTION OF LUNG NODULES ON CHEST RADIOGRAPHS

Gergely ORBÁN
Advisor: Gábor HORVÁTH

I. Introduction

Lung cancer is one of the most common causes of cancer death. Many cures are known, but most of them are effective only in the early and symptomless stage of the disease. Screening can help early diagnosis, but an accurate, cheap and side effect free method has to be used to enable mass usage. Standard chest radiography mostly meets these requirements, except that current methods have a moderate accuracy. If someone suffering from cancer undergo the screening procedure, only has an approximately 30% probability of being diagnosed as positive. Efficiency can be improved by analysing the radiographs using a CAD (Computer Aided Detection) system. The most important problem of existing CAD systems is the high number of false detections. Although they can detect 60-70% of cancerous tumours, they also mark approximately four healthy regions on each image [1].

The goal of my work is to create a lung CAD being more effective than existing ones. For my current solution I used a common two step scheme utilizing my own improvements in each step. The first step is the enhancement of the target lung nodules by using various image processing algorithms. The second step selects suspicious areas on the enhanced image with the help of a classifier.

II. Constrained Sliding Band Filter

The aim of the first step in the scheme is the enhancement of nodules on chest radiographs. These nodules are darker than the surroundings and round shaped. A commonly used filter family called CI (Convergence Index) enhance areas based on their shape. A common property of round shaped objects is the radial direction of gradient vectors along their border, so the filters consider the surrounding of the centre. The output depends on the angle of the gradient vectors and the vectors connecting the centre and the given points. One of the most successful realization is the SBF (Sliding Band Filter) using the following idea [2] and illustrated on figure1. For a given centre it slides a band in each direction within given bounds, while the band has a fixed width. For each band the algorithm sums the cosine of the angles of radial and gradient vectors. The final position of the band for a direction will be the one with the highest sum. Finally it sums the maximal band values in each direction. A high final sum indicates a nodule. A weakness of the algorithm is the independence of the bands in each direction, enhancing very spiculated and distorted objects. My proposed algorithm the CSBF (Constrained Sliding Band Filter) links the position of the bands allowing smaller distortion. It ensures that the final band positions satisfy a circularity constraint controlled by a coefficient. The enhanced pixel values can be calculated with the following formula.

$$CSBF(x,y) = \max_{R_{min} \leq r \leq \frac{R_{max}-d}{c}} \frac{1}{N} \sum_{i=1}^{N} Cmax_{ir}, \tag{1}$$

$$Cmax_{ir} = \max_{r \leq n \leq r*c} \frac{1}{d} \sum_{m=n}^{n+d} \cos\theta_{im}, \tag{2}$$

where R_{min}, R_{max} are the bounds of the target object radius, c is the shape constraint coefficient, N is the number of directions concerned, d is the width of the band and θ_{im} is the angle of the radial unit

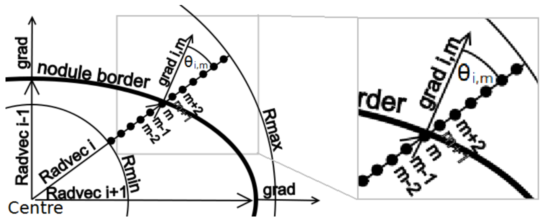

Figure 1: The CSBF filter.

vector and the mth gradient vector along the ith radial direction. The same result can be achieved by running several SBF filterings with different bounds and taking the minimum for each centre, however the execution times would be much greater, while CSBF does not take longer than one SBF run. The CSBF is a more general algorithm, thus for certain c values it works as a standard SBF. Furthermore for $c = 1$ the CSBF is identical to the Iris filter, another realization of the CI family.

III. False Positive Reduction

The second step of the CAD scheme concerns the areas with high CSBF value. A good practice is to collect many areas and select the suspicious ones with the help of a classifier, here an SVM (Support Vector Machine)[3]. The training sample set is extracted from a radiograph database with validated nodules. The raw input of the classifier is a 140 dimension vector containing various features that describe the shape, texture and symmetry of the area. Before handing the vector to the SVM, it is reduced to approximately 10 uncorrelated and relevant dimensions. For the SVM itself, I use a radial kernel function and my own parameter selection techniques.

IV. Results

For testing I used a database of 247 images widely used for benchmarking created by the JSRT (Japanease Society of Radiological Technology) and a private database of 150 images originating from a Hungarian clinic. As a test method I used 4-fold cross-validation. On the JSRT database 59% of the real nodules were found while producing on average 2.5 false detections per image which equals to the performance of the most efficient method in my scope and using the JSRT database [1]. The CSBF contributed to the performance by an approximately 5% increase in sensitivity and same specificity compared to my previous solution using SBF. On the other hand the lack of accurate nodule segmentation in my system is the probable reason why it cannot outperform the existing best solution of others. The performance on the private database are somewhat worse. The sensitivity is 60% at a false positive rate of 4. The main reason behind the results is the greater variety of the private database.

V. Conclusion

In conclusion the proposed system is capable of helping lung cancer diagnosis. To prove the usability, the system is built in the software of an X-ray machine and used experimentally at a clinic. The performance has to be improved further to finally provide the radiologists a really efficient tool.

References

[1] 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*, 12(3):240–258, June 2008.

[2] C. S. Pereira et al., "Evaluation of contrast enhancement filters for lung nodule detection," in *ICIAR 2007 Proceedings*, pp. 878–888, Montreal, Canada, Aug. 22–24 2007.

[3] M. Altrichter, G. Horváth, B. Pataki, G. Strausz, G. Takács, and J. Valyon, Eds., *Neurális Hálózatok*, Panem, 2006.