EFFECTS OF BONE SHADOW REMOVAL ON LESION DETECTION 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 only effective in the early and symptomless stage of the disease. Screening can help early diagnosis, but a sensitive – where sensitivity is the fraction of correctly diagnosed positive cases and all positive cases –, cheap and side effect-free method has to be used to enable mass usage. Standard chest radiography meets these requirements, except that current methods have moderate sensitivity. Efficiency can be improved by using a Computer Aided Detection (CADe) system. The most important problem of existing CADe systems is the low positive predictive value. In other words high sensitivity can only be reached at the cost of many false detections. Recently published systems can detect 60-70% of cancerous tumours, while they also mark approximately four false positive regions on each image [1], which allows them to be used only as a second reader.

Usability of CADe systems can be improved either by reducing the number of false detections – to give the examiner less extra work –, or by finding more nodules – to increase sensitivity. The detections of CADe should be also complementary to the findings of radiologists, to better improve sensitivity when radiologists and CADe work in cooperation.

Common causes of detection errors are shadows of the ribcage and the clavicles. On one hand, these structures can conceal the shadows of abnormalities by darkening the image thus reducing contrast. On the other hand, rib crossings on the radiographs sometimes mimic convex structures appearing to be real lesions. Both of these effects seriously affect human examiners and CADe systems too; however, the hiding effect is problematic mostly to radiologists and physicians while the CADe system suffers rather from false positives due to rib crossings. Previous studies showed that the elimination of bones can improve the accuracy of physicians [2]. Our hypothesis was that properly compensating bone shadows can be also used to reduce the number of the CADe system's false findings.

We expected that "cleaning" the image from bone shadows helps our existing CADe scheme to produce less false positives. As the area originally hidden by bone shadows cannot be perfectly restored based on one radiograph, we also created features which help a supervised classifier to eliminate falsely detected structures caused by bones. In the sequel, we describe the existing CADe scheme, the bone shadow removal algorithm and the modifications to improve detection. Afterwards, we demonstrate the results and compare the system to the one without shadow removal.

II. Materials and Methods

For our existing lesion finder solution we used a three-step scheme. The first step segments the viewable area of the lung. The viewable area reduces the possible set of locations for lesions, as we only target the ones partly or completely inside this area.

The second step, called lesion enhancement, highlights suspicious structures like the target lung nodules and infiltrated areas by using image processing algorithms. After normalization and resizing, various 2D filters are used to enhance different types of lesions. Afterwards, the candidate collection converts the enhanced images into separate and segmented nodule candidates.

The last step reduces the number of false positive findings on the enhanced image with the help of a

classifier. It begins with the calculation of features of candidates serving as an input for the classifier. These features mostly describe texture and geometry, and are used by a supervised learning algorithm, namely a support vector machine.

In the updated system, before the second step ribs and clavicles are first segmented based on a previously calculated rough model, afterwards the segmentation is refined to fit edges while satisfying various constraints. Finally the objects are cleared from the image to restore the background. The segmentation steps are not exposed here, a comprehensive description can be found in [3].

The segmentation data is used to remove the bone shadows from the images in order to enhance the structure of the lung. The elimination is based on creating intensity profiles on vertically differentiated images, which are subtracted from the differentiated original image. An integration step returns to the original domain and produces the bone shadow free image.

To better exploit the results of bone segmentation, three new features were created based on the following observations. First, density variations in bone structure can appear as intensity extrema and may produce false positive findings. Second, areas where bones overlap on the 2 dimensional summation images appear as dark, approximately rhomboid shaped structures frequently recognized by the system as lesions. Third, in the cases where bone shadows cause false positive candidates, the border of the candidate follows the edge of the bone in a relatively large portion. These three observations motivated the following three features respectively.

The first feature calculates the area fraction of the nodule overlapping with a bone structure. The feature can be described formally as

$$\mathrm{BoneOverlap}_i = \frac{|\{(x,y)|(x,y) \in C_i \bigcap (\bigcup_j B_j)\}|}{|\{(x,y)|(x,y) \in C_i\}|}, \tag{1}$$

where C_i contains the points of the i^{th} candidate while B_j means the set of points inside the j^{th} bone structure. $|.|$ means the size of the sets. The second feature calculates the overlap of candidates with bone crossings – the areas where two or more bone segments overlap each other. Using the previous notation, the formula of the feature is

$$\mathrm{BoneXOverlap}_i = \frac{|\{(x,y)|(x,y) \in C_i \bigcap (\bigcup_{j,k} (B_j \bigcap B_k))\}|}{|\{(x,y)|(x,y) \in C_i\}|}. \tag{2}$$

The third feature calculates the fraction of the nodule perimeter that runs near to a bone shadow border. It can be calculated as

$$\mathrm{FollowBoneEdge}_i = \frac{|\{(x,y)|(x,y) \in \delta C_i, \min_j d((x,y), \delta B_j) < 1.5\mathrm{mm}\}|}{|\{(x,y)|(x,y) \in \delta C_i\}|}, \tag{3}$$

where δC_i and δB_j are the endpoint set of C_i and B_j, while d is the Euclidean distance.

III. Results

We tested the system on a private chest X-ray database containing images of 285 patients where 134 of the cases contained at least one malignant lung nodule. Nodule diameter ranged from 2 mm to 98 mm, the average was 23 mm. Most of the malignant cases were validated by CT. The images came from a digital X-ray machine working in daily practice at a Hungarian clinic.

10-fold cross validation with 30 repetitions was used for evaluation. Results are demonstrated on free-response receiver operating characteristic (FROC) curves, showing the sensitivity as a function of average number of false positives produced for each image. On the final output, a CADe drawing was considered to be correct if its centroid – center of gravity – was located inside a physician's marker; otherwise it was labeled as a false positive.

We compared the original CADe system – working on unprocessed images – with two new versions utilizing bone information. The first version of the new system consumed bone-shadow-free images on its input, but involved no other modifications compared to the original system. The second version used –besides the cleared radiographs – the three new features mentioned above. The results can be seen in Figure 1.

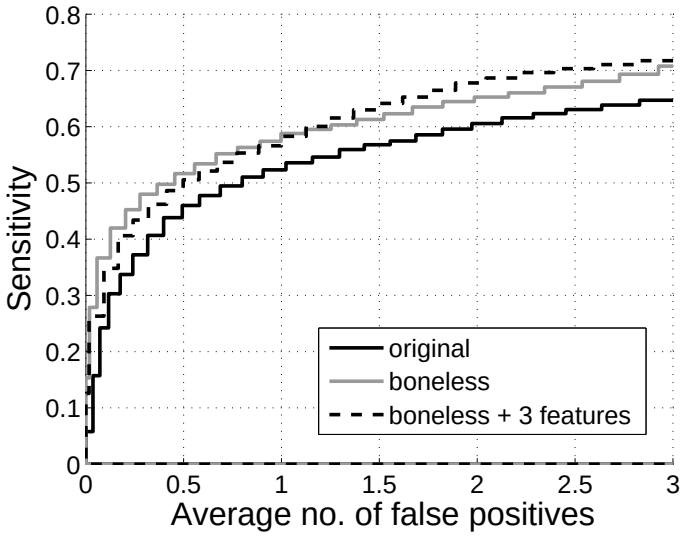

Figure 1: FROC curves of the original nodule detector system (black line), the same system using preprocessed images without bone shadows (gray line) and the system utilizing both the new features and preprocessed images (dashed line). The usage of cleared images greatly improves detection accuracy, while the new features show a slight benefit at important working points.

The improvement when using rib-shadow-free images is clear. At constant 65% sensitivity the number of false positives could be reduced from 2.83 to 1.9. Utilizing the new features, the number of false positives falls further to 1.6. This means that 43% of false detections could be eliminated which is a great benefit for radiologists and physicians having to examine much less false CADe results at screening. Examining the working points at lower sensitivities, the benefit of bone-shadow-free images are still clear; however, the new features cannot improve the accuracy further. Moreover using the new features has a negative effect under 57%, where the output depends stronger on the classifier performance. This can be an effect of more features being used that can reduce the generalization capability of the classifier. However, in everyday practice the working points with higher sensitivities considered to be more useful.

After seeing the benefits in detection accuracy we wanted to see how bone shadows and their removal affect the output distribution of the CADe system. We assumed that bone shadows and especially bone crossings cause many false positives in the original system and less or not at all in the new version. We conducted two tests to support our hypothesis.

In the first test we calculated the fraction of the total area of false findings that overlaps with bones. If we assume no connection between bone shadows and false positive location the fraction should be approximately the same as the fraction of bone shadow area inside the lung area as findings are also restricted to this area. As it can be seen in Table 1 false positives tend to overlap more with bone shadows using the original system while overlap even less than expected on rib shadow free images. This means that the new system is less likely to produce false positives on bone shadows than on other areas. We also checked how often physician markers overlap with bone shadows. The corresponding row of the table shows that physicians tend to avoid bone shadows when marking nodules, just like the new CADe system. We conducted the same measurements for rib crossings and came to similar

conclusions. While the original system puts more false findings on rib crossings, the new system and physicians tend to put less.

Table 1: Area fraction of false positive candidates overlapping with bone shadows.

	Original image	Rib-shadow-free image
Lung area overl. bones	0.489	0.489
FP area overl. bones	0.527	0.446
Phy. marker area overl. bones	0.471	0.471
Lung area overl. bone Xs	0.114	0.114
FP area overl. bone Xs	0.166	0.103
Phy. marker area overl. bone Xs	0.079	0.079

In the second test we analyzed the independence of a candidate being false positive and its overlap with bone shadows as two random variables. The first one is already binary while the overlap feature was quantized to three bins to discretize it. We used Pearson's chi-squared test to check independence and demonstrated the resulting p-values in Table 2. We ran the test for the original system, the system using rib-shadow-free images and the system also using the new features. Using a predefined significance level of 1% the first two cases are clearly significant meaning that false positiveness depends on – or at least correlates with – the overlap with bones while in the third case we can no longer be sure that the relation exists. However, we should note that a p-value of 0.049 suggests further investigation. We ran the same test using the bone crossing overlap as the second variable. Here, using the original system shows a clear correlation, while the results for the versions exploiting bone information are insignificant. These results do not contradict with our original hypothesis; however, the proof of independence needs other studies.

Table 2: Resulting p-values of independence tests between bone overlapping and false positiveness.

	Original image	Rib-shadow-free image	Rib-free + 3 features
Overlap w. bones	0.0001	0.0002	0.049
Overlap w. bone crossings	0.0015	0.61	0.2

IV. Conclusion

Our measurements proved that the removal of bone shadows can improve lesion detection accuracy on chest radiographs. The absence of these dark shadows reduces the number of misleading structures on the image while some newly created features help the classifier to distinguish between remainders of bones and real lesions. We also showed that the output of the newly created CADe system seems to have less dependency on bone location compared to our previous solution.

References

[1] S. Chen, K. Suzuki, and H. MacMahon, "Development and evaluation of a computer-aided diagnostic scheme for lung nodule detection in chest radiographs by means of two-stage nodule enhancement with support vector classification," *Medical Physics*, 38:1844, 2011.

[2] F. Li, T. Hara, J. Shiraishi, R. Engelmann, H. MacMahon, and K. Doi, "Improved detection of subtle lung nodules by use of chest radiographs with bone suppression imaging: Receiver operating characteristic analysis with and without localization," *American Journal of Roentgenology*, 196(5):W535–W541, 2011.

[3] S. Juhász, Á. Horváth, L. Nikházy, G. Horváth, and Á. Horváth, "Segmentation of Anatomical Structures on Chest Radiographs," in *XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010*, R. Magjarevic, P. D. Bamidis, and N. Pallikarakis, Eds., vol. 29 of *IFMBE Proceedings*, pp. 359–362. Springer Berlin Heidelberg, 2010, 10.1007/978-3-642-13039-7_90.