Computer-Aided Detection of Acinar Shadows in Chest Radiographs

Tao Xu, Irene Cheng, Richard Long, and Mrinal Mandal

Abstract— Despite the technological advances in medical diagnosis, accurate detection of infectious tuberculosis (TB) still poses challenges due to complex image features and thus infectious TB continues to be a public health problem of global proportions. Currently, the detection of TB is mainly conducted visually by radiologists examining chest radiographs (CXRs). To reduce the backlog of CXR examination and provide more precise quantitative assessment, computer-aided detection (CAD) systems for potential lung lesions have been increasingly adopted and commercialized for clinical practice. CADs work as supporting tools to alert radiologists on suspected features that could have easily been neglected. In this paper, an effective CAD system aimed for acinar shadow regions detection in CXRs is proposed. This system exploits textural and photometric features analysis techniques which include local binary pattern (LBP), grey level co-occurrence matrix (GLCM) and histogram of oriented gradients (HOG) to analyze target regions in CXRs. Classification of acinar shadows using Adaboost is then deployed to verify the performance of a combination of these techniques. Comparative study in different image databases shows that the proposed CAD system delivers consistent high accuracy in detecting acinar shadows.

***Index Terms*— Textural and photometric classification, computer-aided detection (CAD), tuberculosis (TB)**

I. INTRODUCTION

Although effective therapies have reduced the mortality from infectious pulmonary tuberculosis (TB), TB continues to be a public health problem of global proportions especially in developing countries [1]. This is mainly due to the complex overlapping anatomical structures which often obscure the detection of TB features in the chest. In clinical practice, when signs or symptoms point to a lower respiratory tract illness, a chest radiograph (CXR) – an inexpensive and widely available tool – is typically used [2]. However, either because symptoms are non-specific or because a patient may not be considered at risk of TB, through inexperience or inadequate human resources, proper diagnosis of TB is often delayed or

Tao Xu and Mrinal Mandal are with the Department of Electrical and Computer Engineering, University of Alberta, T6G 2V4 CA (e-mail: tx1@ualberta.ca, mmandal@ualberta.ca). Irene Cheng is with the Dept. of Computing Science, University of Alberta, T6G 2E8 (e-mail: locheng@ualberta.ca), Richard Long is with the Department of Medicine, University of Alberta, T6G 2V2 (e-mail: richard.long@ualberta.ca).

missed by clinicians or radiologists. Digital radiography, which has replaced film-based chest units these days, has greatly facilitated advances such as computer-aided detection and diagnosis (CAD) systems to solve this problem. Since the first market launch of RapidScreen^{TM} RS-2000 system was approved in 2001 for clinical use by the Food and Drug Administration (FDA), lung CAD systems have been receiving increasing support in the radiology community. A wide variety of lung CAD systems have been reviewed in the literature [3-5]. Those CAD systems do not detect all potential lesions on CXRs but only aim at a single aspect, e.g. detection of lung cancer nodules, which would restrict the radiologist to focus only on the areas identified by the CAD system. The incorporation of CAD system as the second reader will help to screen significant cases and thus improve the analysis performance of radiologists. For instance, one recent evaluation of several commercial CAD software approved by FDA shows that CAD improves the sensitivity of inexperienced readers for the detection of small nodules [6].

CAD systems for nodule detection have so far been receiving most attention, but little work has been done beyond lung nodules detection [5], neglecting many aspects of infectious pulmonary TB. A recent study [7] reveals that typical infectious pulmonary TB (also know as post-primary TB or reactivation TB) are more likely to have transmission events or a public health consequence than atypical TB (those without typical CXR findings of post-primary TB). Therefore, the objective of our research is to develop a comprehensive CAD system for automatic recognition of typical radiographic patterns to identify highly infectious post-primary TB. Typical radiographic patterns as mentioned in many radiology handbooks include cavities, volume loss, acinar shadows (AS) and so on [8-10]. In the diagnosis of pulmonary TB on plain chest radiograph, AS reflect the presence of endobronchial spread of disease, the spread of tuberculous ‘caseous’ material within the bronchial tree. Such spread is a typical feature of pulmonary TB, and AS is found to be present on 68% of the plain CXRs [9]. The AS are either within the vicinity of the major focus of disease (for example a cavitated area), immediately dependent from it or occasionally in the contralateral lung – indicating position or posture-related drainage from the major

focus. The presence of acinar shadows in the vicinity of an upper lung zone infiltrate, especially if cavitary, further adds to the probability that the infiltrate reflects the presence of infectious pulmonary TB. Our previous work [11] focused on the extraction of TB cavities on CXRs. Thus, in current work, we focus on the automatic detection of AS on CXRs, which is defined as "round or ovoid poorly defined pulmonary opacities approximately 5-8 mm in diameter, presumed to represent an anatomic acinus rendered opaque by consolidation" [12]. An example of AS is shown in the 1st column in Fig 1.

We propose a hybrid computerized technique for AS detection in chest radiographs using the following steps. Before applying the detection technique, the CXR image is preprocessed as a downsampled and contrast-enhanced subimage which only contains the lung fields. The preprocessed image is densely scanned from the top left to the bottom right with non-overlapping rectangular sliding windows. Multiple features combining Local Binary Pattern (LBP), Grey Level Co-occurrence Matrix (GLCM) and Histogram of Oriented Gradients (HOG) are then extracted from each sliding window, and fed to the Adaboost classifier which is trained offline using labeled training data. The classifier denotes the sliding windows as positive windows containing AS or negatives. To further reduce the number of false positive windows, we introduce an efficient morphological operation to achieve high sensitivity, specificity and precision.

The rest of this paper is organized as follows. Section II introduces the materials used in this study. Section III describes our proposed method in detail. Section IV presents the experiments results and analysis. Conclusion and future work are discussed in Section V.

II. Materials and Methods

A. Materials and Preprocessing

Standard posterior-anterior (PA) CXRs from two image databases are used in this study. The first database obtained from the University of Alberta Hospital consists of CXRs on 58 cases of sputum smear positive pulmonary TB. 37 cases in this database are typical infectious TB with AS and 21 cases are

atypical TB without AS. Ground truth information was determined by a panel of three independent expert chest radiologists. One of the experts helped to draw the region of AS. The second database obtained from the Japanese Society of Radiological Technology (JSRT) [13] contains 93 CXRs of normal cases and 154 CXRs of abnormal cases with solitary nodule which is defined as a discrete, well-marginated, rounded opacity less than or equal to 3 cm in diameter [8]. Images in this database were also evaluated with the consensus of three chest radiologists. Finally, images in these two image databases are grouped into four datasets: D1 – typical infectious TB with AS (37 cases); D2 – atypical TB without AS (21 cases); D3 – Normal cases without AS (93 cases); D4 – Nodule cases without AS (154 cases).

Each image in the four datasets was preprocessed using the following procedure. First, to standardize the image resolution and for computational efficiency, each image is scaled to have pixel size equivalent to 0.8 mm with 8-bit intensity. For example, an original CXR image whose resolution is 2048×2048 with pixel size 0.2 mm in both horizontal and vertical direction will be resized as 512×512. Note that this scaling may cause some information lost but does not have significant effect on the AS detection outcome. Second, a rectangle region which only contains the lung field is cropped from the scaled image. Finally, this subimage is contrast-enhanced using adaptive histogram equalization technique [14]. Fig. 1 shows that the subimage qualities are improved after our contrast enhancement step.

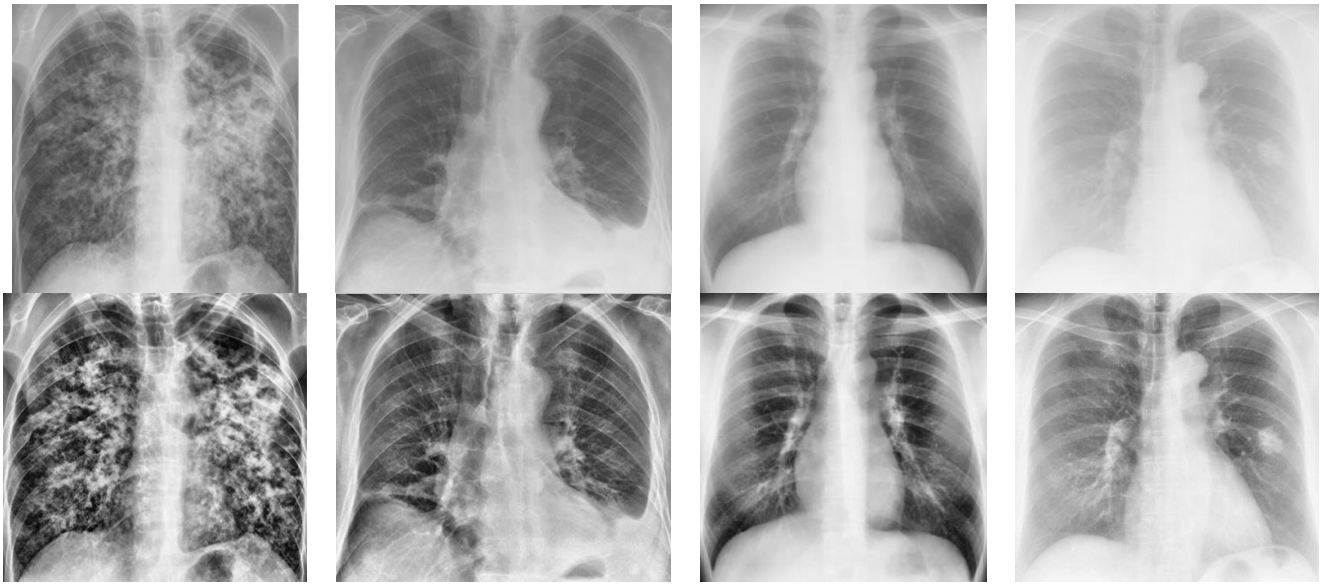

Fig. 1. Subimages from D1 to D4 (left to right) with (bottom row) and without (top row) contrast enhancement.

B. Methods

The proposed computer-aided AS detection system follows the state-of-the-art sliding window paradigm in object detection [15][16][22]. Each preprocessed CXR image is first divided into non-overlapping windows, whose size is fixed as 16×16 in our experiments. Multiple features are then calculated from each window to form a multi-dimensional feature vector. Based on these feature vectors, a classifier is trained offline to distinguish windows containing instances of AS from other windows. The test image scanned with sliding windows is then analyzed and classified using the model generated from the training set. Fig. 2 shows the schematic of the proposed AS detection technique.

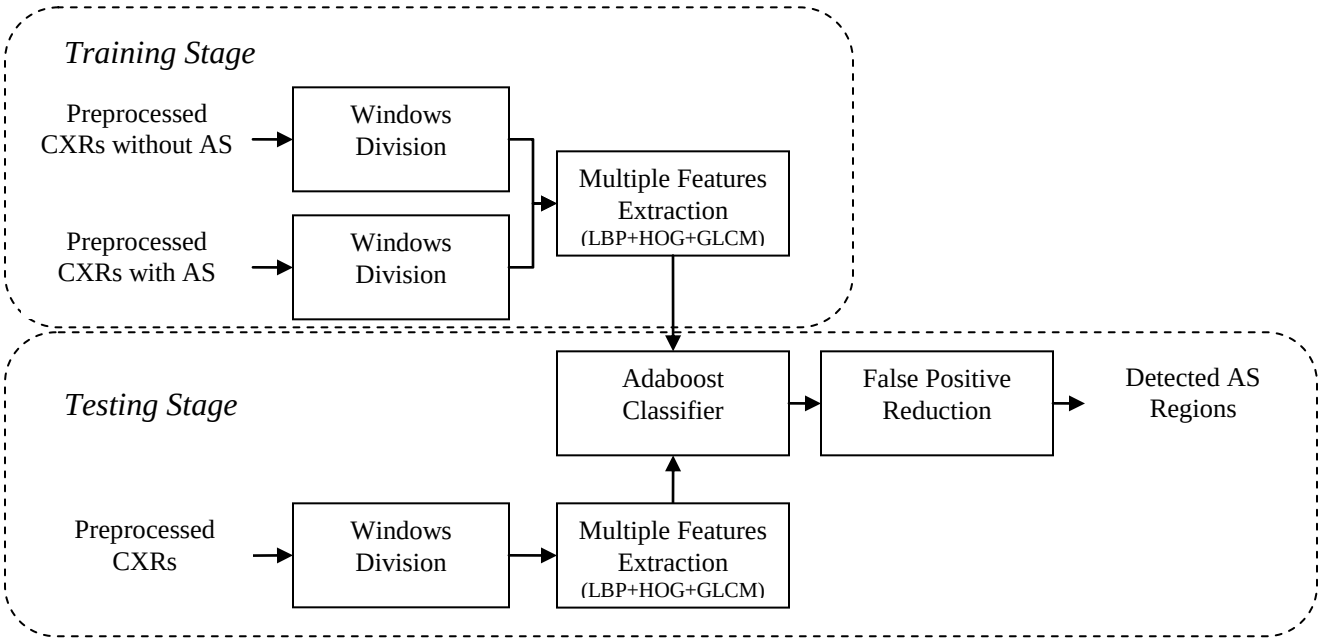

Fig. 2. Schematic of the proposed CAD system for AS detection.

1) Multiple Features Extraction

A variety of features are tested in our study, including features using histogram moments (HM) [17], Fourier spectrum (FS) [18], gray-level co-occurrence matrices (GLCM) [19], fractional dimension (FD) [20], local binary pattern (LBP) [21], histogram of oriented gradients (HOG) [22] and Tchebichef moments (TM) [23]. Finally, we combined GLCM, LBP and HOG to achieve the best performance.

a) Grey Level Co-occurrence Matrix (GLCM) based Features

Spatial GLCMs are one of the most well-known and widely used texture features. These second order statistics are accumulated into a set of 2D matrices. Given a displacement vector $(d,\theta)=(dx,dy)$, each co-occurrence matrix $P(i,j|d,\theta)$ measures the spatial dependency of two greylevels, *i* and *j*. It is calculated as:

$$P(i,j|d,\theta)=\left\|\{(x_1,y_1),(x_2,y_2):I(x_1,y_1)=i,I(x_2,y_2)=j\}\right\| \quad (1)$$

where (x_1,y_1) and (x_2,y_2) are the pixels' coordinates in the image *I*, $(x_2,y_2)=(x_1\pm dx,y_1\pm dy)$ and $\|\cdot\|$ is the cardinality of a set. Texture features, such as contrast, correlation, energy and homogeneity, are then derived from the co-occurrence matrix.

b) Local Binary Pattern (LBP) based Features

The LBP[21] is a hybrid texture feature widely used in image processing recently. It combines the traditionally divergent statistical and structural models of texture analysis. The LBP feature has some key advantages, such as its invariance to monotonic gray level changes and computational efficiency. The general LBP operator based on a circularly symmetric neighbor set of *P* members on a circle of radius *R*, denoted as $LBP_{P,R}$ is obtained by thresholding the neighborhood pixel values with the gray value of the center.

$$LBP_{P,R}=\sum_{p=0}^{P-1}s(g_p-g_c)2^p \qquad \left[\text{Note: } s(x)=\begin{cases}1 \text{ if } x\geq 0\\ 0 \text{ otherwise}\end{cases}\right] \quad (2)$$

where g_p, g_c are gray levels of the neighborhood pixels and center pixel, respectively, and $s(\cdot)$ is the unit-step function. See Fig. 3 for an example of 3×3 neighbourhood, i.e. $P = 8$, $R = 1$.

By introducing an uniformity measure *U* which corresponds to the number of spatial transitions (bitwise 0/1 changes) in the pattern, the rotation symmetric and multiscale LBP, denoted as $LBP_{P,R}^{riu2}$, is calculated as follows:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \le 2 \\ P+1 & \text{otherwise} \end{cases} \tag{3}$$

where $U(LBP_{P,R}) = \left|s(g_{P-1} - g_c) - s(g_0 - g_c)\right| + \sum_{p=1}^{P-1} \left|s(g_p - g_c) - s(g_{p-1} - g_c)\right|$. Eq. (3) assigns a unique label to each of the "uniform" patterns corresponding to the number "1" bits in the pattern, while the "nonuniform" patterns are grouped under the label P+1. For example in Fig. 3, the clockwise 8-neighbor pixels' intensities are thresholded as 10010111 which is 151 for the $LBP_{P,R}$. Since the $U(LBP_{P,R}) = 4$, $LBP_{P,R}^{riu2} = 9$ belongs to the "nonuniform" patterns.

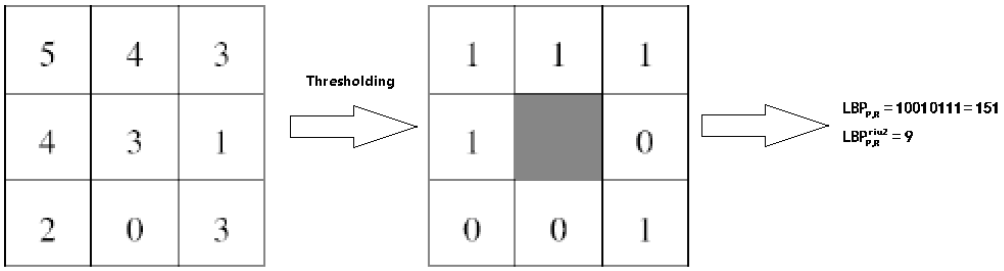

Fig. 3. An example of calculating LBP values in a 3×3 neighborhood.

Based on the $LBP_{P,R}^{riu2}$ values of pixels in each sliding window, a uniform LBP histogram is then generated. The final textural features are obtained as six statistical measurements (mean, standard deviation, skew, kurtosis, entropy and energy) of the histogram.

c) Histogram of Oriented Gradients (HOG) based Features

The HOG feature [22], similar to Lowe's scale-invariant feature transform (SIFT) feature, is regarded as an excellent descriptor to capture the edge or local shape information. It has a great advantage of being robust to changes in illumination or shadowing. The HOG feature for each 16×16 window is calculated as follows:

Step1. Gradient Computation: The gradient of each pixel in the window is calculated using two filter kernels: [-1, 0, 1] and $[-1, 0, 1]^T$. Let the magnitude and orientation of the gradient of the i^{th} pixel

$(1 \le i \le 256)$ be denoted by m_i and ϕ_i, respectively.

Step2. Orientation Histogram: Each window is first divided into non-overlapping cells of equal dimension, e.g., a rectangular cell of 8×8. The orientation histogram is then generated by quantizing ϕ_i into one of the 9 major orientations: $\frac{(2k-1)\pi}{9} \pm \frac{\pi}{9}$, $1 \le k \le 9$. The vote of the pixel is weighted by its gradient magnitude m_i. Thus, a cell orientation histogram H_c is a vector with dimension of 1×9.

Step3. Block Normalization: In order to account for changes in illumination and contrast, the cell histogram must be locally normalized, which requires grouping the cells together into larger, spatially-connected blocks. We treat the 16×16 window as one block (i.e. 2×2 cells). Therefore, the feature vector of one block H_b is concatenated by four cell histograms: $H_b = [H_{c1}\ H_{c2}\ H_{c3}\ H_{c4}]$. The normalized HOG feature vector is then calculated as:

$$\hat{H}_b = \frac{H_b}{\|H_b\|} \tag{4}$$

where $\|\cdot\|$ represents the L^2 norm, $\hat{H}_b$ is a vector with dimension of 1×36.

Combining the GLCM, LBP and HOG features, a feature vector of size 1×46 is obtained for each window. These features vectors are fed to the classifier, explained below, for AS detection.

2) Adaboost Classifier

Classifier plays an important role in a CAD system design. Currently, the Support Vector Machine (SVM) [24] and variants of boosting [25] are two leading classifiers for their good performance and efficiency. Boosting is a general technique for improving performance of any given classifier. It can effectively combine a number of weak classifiers into a strong classifier which can achieve an arbitrarily low error rate given sufficient training data [26]. Compared to SVM, boosting techniques such as Adaboost [25], the most popular boosting method, have less parameters to tune, are more resistant to overfitting problem and do not require prior knowledge of the features. In our study, comparison tests

between SVM and Adaboost with same features were performed. Based on the performance, Adaboost is selected as the final classifier for AS detection. See more details in Section IV for the comparison experiments.

The Adaboost algorithm forms a strong classifier by combining a set of weak learners linearly in an iterative manner. A single level decision tree called decision stump is used as the weak classifier. The pseudo-code of the Adaboost combing *M* decision stumps is shown in Fig. 4. The final AS detection in a test CXR is done using the strong classifier achieved from Adaboost.

Given N training examples $(x_1, x_2, ..., x_N)$ and corresponding labels $(y_1, y_2, ..., y_N)$ with $y_i \in \{-1, 1\}$:

1. Initialize the observation weights, $w_i^+ = \frac{1}{2N^+}, w_i^- = \frac{1}{2N^-}$,

where $i = 1, 2, ..., N$, $N^+ + N^- = N$, '+' and '-' represent positive and negative samples, respectively.

2. For *m* = 1 to *M*

(a) Fit a decision stump $h_m(x)$ to the training data using weights $w_i^{(m)}$, where $h_m(x_i) = sign(x_i - t_m)$, t_m is some feature value chosen as a threshold for the decision stump.

(b) Compute $err_m = \frac{\sum_{i=1}^{N} w_i^{(m)} I(y_i \neq h_m(x))}{\sum_{i=1}^{N} w_i^{(m)}}$, where I is an indicator function.

(c) Compute $\alpha_m = \frac{1}{2} \ln(\frac{1 - err_m}{err_m})$

(d) Update $w_i^{(m+1)} = w_i^{(m)} \exp(\alpha_m I(y_i \neq h_m(x))$

3. Combine weak learners into a strong classifier $f(x) = sign(\sum_{m=1}^{M} \alpha_m h_m(x))$

Fig. 4. Adaboost algorithm used in the proposed CAD system

III. Results and Discussion

Given a randomly selected subset images from datasets D1-D4, comparison experiments of different features and classifiers are performed to verify the effectiveness of the proposed detection system. Although individual feature analysis techniques have their merits in specific applications, we propose to use a combination of GLCM, LBP and HOG features with Adaboost for AS detection based on our performance analysis outcome. To further improve the accuracy, we propose an efficient technique using

morphological operations. The entire CAD system was implemented in MATLAB 2007b on an Intel Pentium 4 CPU 2.8G Hz with 2G RAM computer.

A. Ground truth setting and other Configurations

Although radiologist has helped to extract the ground truth region containing AS, some sliding windows are vague to be defined as positive windows with AS due to the incomplete coverage of the target. Fig. 5 shows an example where the red block is one of the sliding windows and white closed contour is the groudtruth drawn by radiologist. To resolve such ambiguity, we define an area ratio R = (area of AS in the window) / (area of the window) to divide the positive and negative windows. Given a threshold value t, a positive window should satisfy $R \geq t$, and vice versa. For example in Fig. 5, if $t = 3/4$, the red block will not be defined as the positive windows with AS.

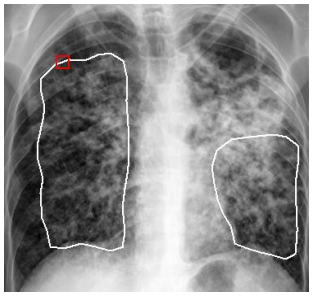

Fig. 5. An example of a window (red) containing part of AS (white contour).

The selected multiple features used for comparison include 6 features based on HM [17], 5 features based on FS [18], 4 features based on GLCM [19], one feature of FD value [20], 6 features using uniform LBP histogram [21], 36 features based on HOG [22] and 6 features based on TM [23]. Details of these features are listed in Table 1. Due to the huge number of existing texture features, we only select the above seven types which are state-of-the-art and are adopted in various studies. For example, HM based features belong to statistical features, FS based features belong to signal processing based features, GLCM based features belong to structural features, FD value belongs to model based feature, LBP based features are both statistical and structural features. LBP, HOG and TM are the state-of-the-art features in the object detection area. All these features have at least one of the rotation, scale and translation invariance characteristics.

Table 1 Multiple Features Used for Comparison

Multiple Features from	No. of Features	Used Features
HM [17]	6	Mean, standard deviation, relative smoothness, third moment, uniformity, entropy
FS [18]	5	Energy, variance, entropy, low frequency energy ratio, low/high frequency energy ratio
GLCM [19]	4	Contrast, correlation, energy, homogeneity
FD [20]	1	Fractional dimension
LBP [21]	6	Mean, standard deviation, skewness, kurtosis, entropy, energy
HOG [22]	36	4 cells of 9 major orientation of the gradient
TM [23]	6	Moments of T_{00}, T_{01}, T_{10}, T_{11}, T_{12}, T_{21}

Classifiers selected for comparison are LIBSVM [27] and Adaboost [25]. LIBSVM is an integrated software for support vector classification, regression, and other work. In LIBSVM, linear, polynomial, radial basis function (RBF) and sigmoid are widely used kernels. In our experiments, we found that the RBF kernel function outperform other kernels. Parameters are all set to the default values in LIBSVM. As for Adaboost, only one parameter, *M*, is set to be 100 because no significant performance improvement is achieved while increasing *M*.

We use sensitivity, specificity and precision to evaluate the performances of classification, which are widely used in medical domain. These parameters are defined as follows:

$$\text{Sensitivity} = \frac{\text{No. of true positives}}{\text{No. of true positives} + \text{No. of false negatives}} \tag{5}$$

$$\text{Specificity} = \frac{\text{No. of true negatives}}{\text{No. of true negatives} + \text{No. of false positives}} \tag{6}$$

$$\text{Precision} = \frac{\text{No. of true positives}}{\text{No. of true positives} + \text{No. of false positives}} \tag{7}$$

B. Comparison Experiments

Since the region of AS only occupies a small part of a CXR image, to make the sample size between normal and abnormal cases relatively balanced in training, we randomly select 25 preprocessed CXR images in D1 to obtain positive windows and 6 preprocessed CXR images from D2 to D4 to get negative windows. For testing, we select the rest from D1 and randomly select 10 images from D2 to D4. The classification results using only one type of features with SVM are listed in Table 2. The corresponding

receiver operating characteristic (ROC) curves by tuning threshold *t* are shown in Fig. 6. The Area Under the Curve (AUC) is also calculated and listed in Table 2. In comparison, the classification results of using individual features with Adaboost and the corresponding ROC curves are illustrated in Table 3 and Fig. 7.

Table 2. SVM classification results using one type of features.

Results	*t*	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1	Avg	AUC
Only HM	Sen	94.5%	94.1%	93.2%	92.4%	90.4%	89.6%	89.2%	90.1%	91.7%	0.729
	Spe	34.7%	37.1%	41.7%	42.4%	48.3%	50.6%	52.7%	55.6%	45.4%	
	Pre	37.5%	36.6%	36.7%	35.5%	36.1%	35.0%	33.9%	33.4%	35.6%	
Only FS	Sen	97.9%	98.3%	98.2%	98.1%	98.2%	52.8%	6.3%	0.6%	68.8%	0.796
	Spe	26.1%	27.7%	29.0%	30.3%	31.8%	93.2%	99.4%	100.0%	54.7%	
	Pre	35.8%	34.8%	33.7%	32.9%	32.0%	70.0%	75.4%	80.0%	49.3%	
Only GLCM	Sen	93.6%	93.7%	93.1%	92.5%	92.7%	91.9%	92.0%	91.7%	92.6%	0.913
	Spe	80.9%	82.3%	84.0%	85.2%	85.8%	87.5%	88.3%	90.0%	85.5%	
	Pre	67.4%	67.5%	68.1%	68.5%	68.1%	68.9%	68.5%	69.7%	68.3%	
Only FD	Sen	74.6%	71.1%	69.2%	66.7%	64.1%	62.5%	60.0%	56.5%	65.6%	0.804
	Spe	79.1%	82.7%	84.6%	86.7%	88.8%	91.1%	91.8%	93.1%	87.2%	
	Pre	56.0%	57.7%	58.4%	59.7%	61.2%	64.1%	63.2%	63.6%	60.5%	
Only LBP	Sen	97.4%	97.8%	97.9%	98.3%	98.6%	98.8%	97.9%	98.7%	98.2%	**0.946**
	Spe	89.3%	89.1%	89.1%	90.3%	90.6%	90.7%	91.4%	91.7%	90.3%	
	Pre	77.9%	76.4%	76.6%	77.0%	76.7%	75.6%	75.8%	75.0%	76.4%	
Only HOG	Sen	95.3%	95.7%	96.3%	96.2%	96.5%	96.5%	95.5%	95.9%	96.0%	0.885
	Spe	72.5%	73.7%	75.7%	77.3%	78.0%	79.2%	80.8%	82.0%	77.4%	
	Pre	55.3%	54.7%	55.3%	55.5%	54.8%	54.1%	53.8%	53.1%	54.6%	
Only TM	Sen	97.7%	90.0%	67.7%	51.0%	42.2%	15.9%	1.0%	0.3%	45.7%	0.745
	Spe	34.3%	47.7%	63.6%	76.8%	84.5%	98.1%	99.9%	100.0%	75.6%	
	Pre	34.7%	36.4%	36.7%	39.3%	42.9%	68.3%	77.8%	100.0%	54.5%	

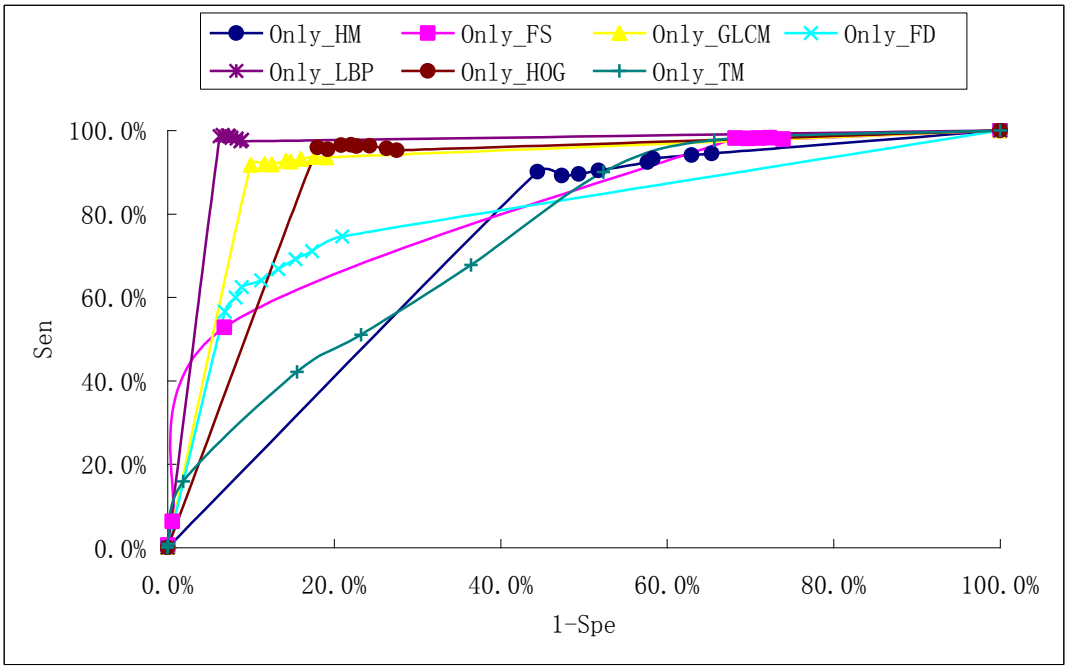

Fig. 6. ROC curves of different features using SVM.

Table 3. Adaboost classification results using one type of features.

Results	*t*	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1	Avg	AUC
Only HM	Sen	86.9%	84.0%	85.4%	79.9%	79.7%	73.4%	70.4%	68.8%	78.6%	0.737
	Spe	49.2%	53.5%	53.4%	60.1%	60.7%	66.9%	70.3%	72.4%	60.8%	
	Pre	56.7%	54.1%	52.9%	49.8%	48.2%	44.1%	41.0%	38.1%	48.1%	
Only FS	Sen	91.6%	86.7%	86.0%	83.6%	81.9%	82.2%	77.7%	76.7%	83.3%	0.858
	Spe	71.8%	77.3%	79.1%	80.4%	80.3%	81.6%	83.5%	85.0%	79.9%	
	Pre	72.0%	69.4%	67.8%	66.0%	64.1%	62.3%	58.9%	56.2%	64.6%	
Only GLCM	Sen	93.4%	93.6%	92.7%	92.4%	92.3%	92.0%	92.4%	92.6%	92.7%	0.917
	Spe	83.6%	85.3%	84.7%	87.3%	88.0%	89.6%	90.1%	90.8%	87.4%	
	Pre	81.0%	79.9%	78.7%	77.7%	76.6%	75.0%	73.5%	71.5%	76.7%	
Only FD	Sen	81.6%	82.3%	78.7%	73.6%	74.2%	74.5%	72.9%	66.6%	75.5%	0.822
	Spe	73.8%	73.8%	77.4%	81.2%	81.3%	81.7%	83.4%	89.6%	80.3%	
	Pre	73.7%	72.4%	70.2%	67.6%	66.4%	64.6%	62.1%	57.6%	66.8%	
Only LBP	Sen	96.2%	96.3%	96.3%	96.4%	96.7%	96.6%	96.6%	96.4%	96.4%	**0.959**
	Spe	94.1%	94.4%	94.4%	95.4%	95.7%	96.2%	96.4%	95.8%	95.3%	
	Pre	91.4%	90.9%	90.1%	89.8%	89.3%	88.4%	87.5%	85.5%	89.1%	
Only HOG	Sen	89.7%	91.1%	90.4%	90.8%	92.1%	90.9%	87.7%	90.5%	90.4%	0.885
	Spe	81.0%	81.4%	82.9%	83.7%	84.9%	86.2%	87.7%	87.7%	84.4%	
	Pre	72.3%	71.2%	69.7%	68.6%	67.5%	65.4%	62.6%	61.1%	67.3%	
Only TM	Sen	80.7%	78.1%	76.1%	75.7%	72.9%	71.2%	67.4%	64.4%	73.3%	0.785
	Spe	67.9%	70.2%	72.8%	73.6%	76.2%	77.7%	80.2%	85.0%	75.5%	
	Pre	65.7%	63.4%	61.3%	59.8%	57.4%	54.8%	51.3%	47.7%	57.7%	

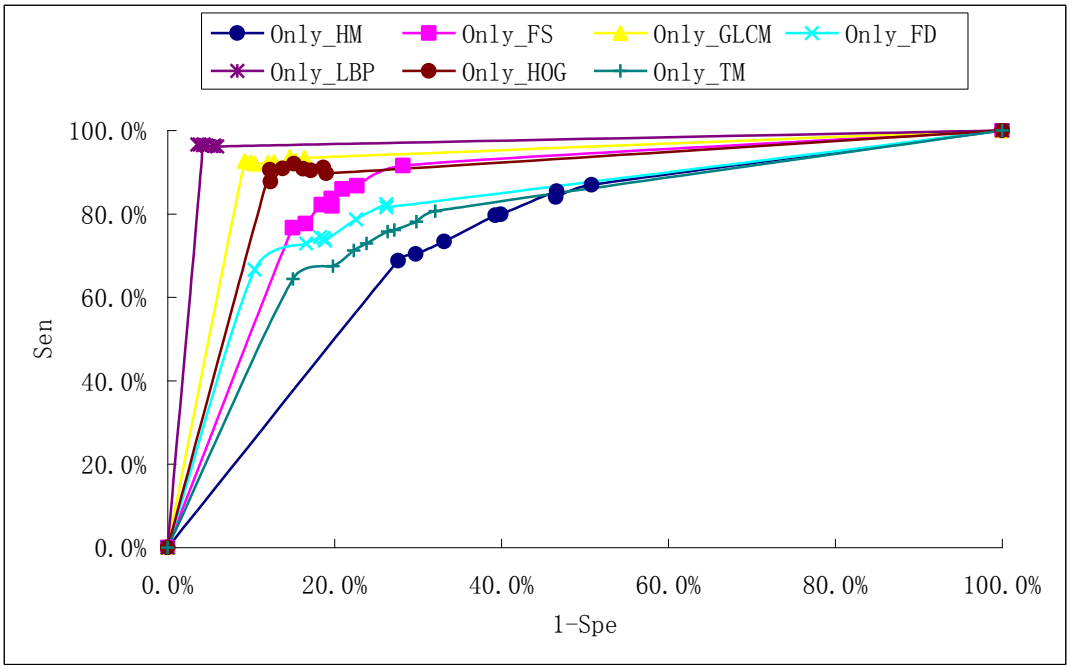

Fig. 7. ROC curves of different features using Adaboost.

From the above experimental results, it is observed that LBP outperforms the other features. GLCM and HOG also achieve good performance. While keeping the similar sensitivity, classifier Adaboost greatly increase the precision and specificity outperforming SVM. Our hypothesis is to use multiple strong features to deliver better outcome. Thus, we perform other comparison experiments using

different combination strategies of LBP, GLCM and HOG features. The SVM and Adaboost classification results using LBP combined with other features are listed in Table 4 and 5, respectively. The corresponding ROC curves with different *t* are shown in Fig. 8 and 9. Since it is difficult to discriminate different curves, the rectangular regions in Fig. 8 and 9 are enlarged and shown in the middle of the Figs. It could be observed that the combination of LBP, GLCM and HOG achieves the overall best performance in both SVM and Adaboost classification. For the performance comparison between SVM and Adaboost classifiers, Adaboost greatly outperforms SVM in specificity and precision while keeping similar high sensitivity. The Adaboost classifier using LBP, GLCM and HOG based features provides around 5% and 12% improvement in average specificity and precision, which means it not only reduces the false positives but also increases the accuracy on the total detected positives significantly.

Table 4. SVM classification results using combined features of LBP, GLCM and HOG.

Results	*t*	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1	Avg	AUC
LBP +HOG	Sen	97.0%	97.7%	98.0%	97.6%	98.4%	98.9%	98.6%	98.8%	98.1%	0.947
	Spe	87.8%	88.1%	89.3%	90.1%	90.8%	90.9%	91.7%	92.1%	90.1%	
	Pre	74.0%	73.1%	74.0%	74.3%	74.7%	73.6%	73.7%	72.8%	73.8%	
LBP +GLCM	Sen	97.5%	97.9%	98.2%	98.1%	98.3%	98.2%	98.2%	98.1%	98.1%	0.950
	Spe	88.5%	88.7%	89.7%	90.4%	90.8%	91.6%	92.1%	92.5%	90.5%	
	Pre	78.3%	77.5%	77.9%	78.2%	78.0%	78.1%	77.7%	77.7%	77.9%	
LBP +GLCM +HOG	Sen	98.5%	98.6%	98.7%	98.6%	99.2%	99.4%	99.5%	99.4%	99.0%	**0.956**
	Spe	88.5%	88.9%	89.4%	90.4%	90.7%	91.5%	92.0%	92.6%	90.5%	
	Pre	78.4%	77.7%	77.4%	78.1%	77.7%	77.8%	77.5%	77.0%	77.7%	

Table 5. Adaboost classification results using combined features of LBP, GLCM and HOG.

Results	*t*	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1	Avg	AUC
LBP +HOG	Sen	95.9%	96.2%	96.2%	97.0%	96.7%	96.5%	96.7%	96.2%	96.4%	0.962
	Spe	94.4%	95.1%	95.3%	95.7%	96.0%	96.7%	96.4%	96.7%	95.8%	
	Pre	91.2%	90.7%	90.1%	89.7%	89.1%	88.2%	87.3%	86.1%	89.1%	
LBP +GLCM	Sen	96.1%	96.3%	96.8%	97.0%	97.6%	97.3%	97.4%	96.6%	96.9%	0.963
	Spe	93.4%	93.6%	94.0%	94.7%	95.0%	95.9%	96.3%	96.5%	94.9%	
	Pre	92.4%	91.9%	91.5%	91.0%	90.5%	89.7%	88.9%	87.9%	90.5%	
LBP +GLCM +HOG	Sen	97.1%	97.5%	97.1%	97.4%	97.8%	98.4%	97.4%	97.7%	97.5%	**0.968**
	Spe	94.2%	94.7%	94.8%	95.0%	95.4%	95.6%	96.5%	96.5%	95.3%	
	Pre	92.0%	91.4%	91.0%	90.5%	90.0%	89.3%	88.4%	87.4%	90.0%	

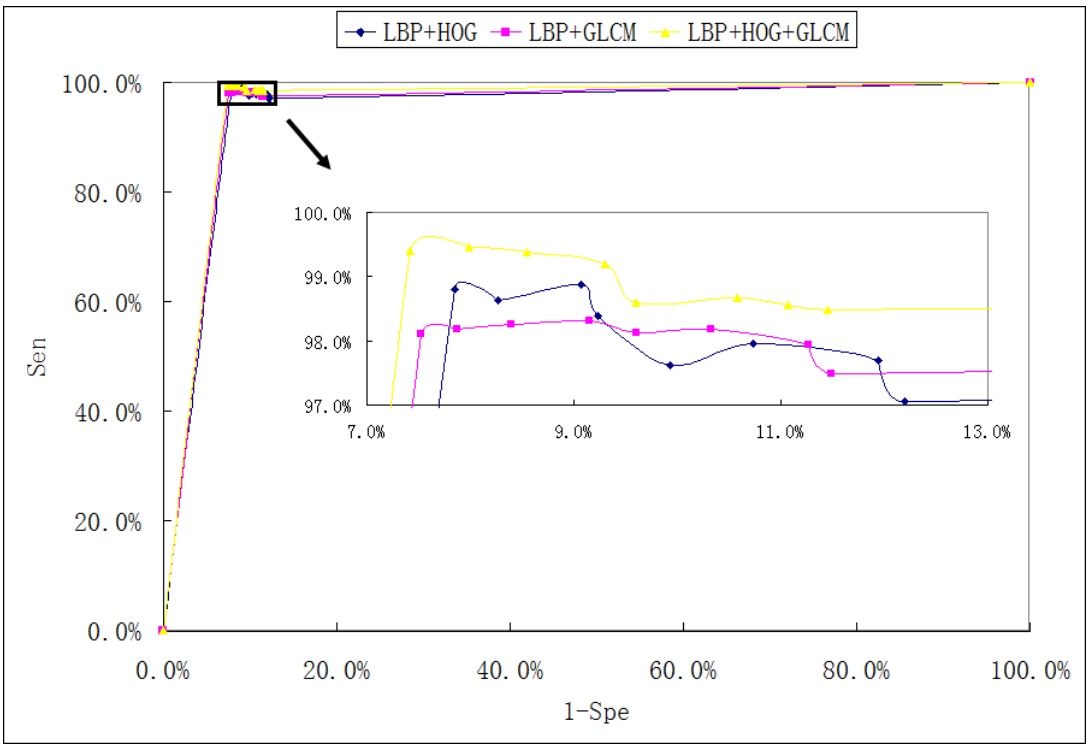

Fig. 8. ROC curves using different combination of LBP, GLCM and HOG features with SVM. To better discriminate different curves, the rectangle area has been enlarged and shown in the middle.

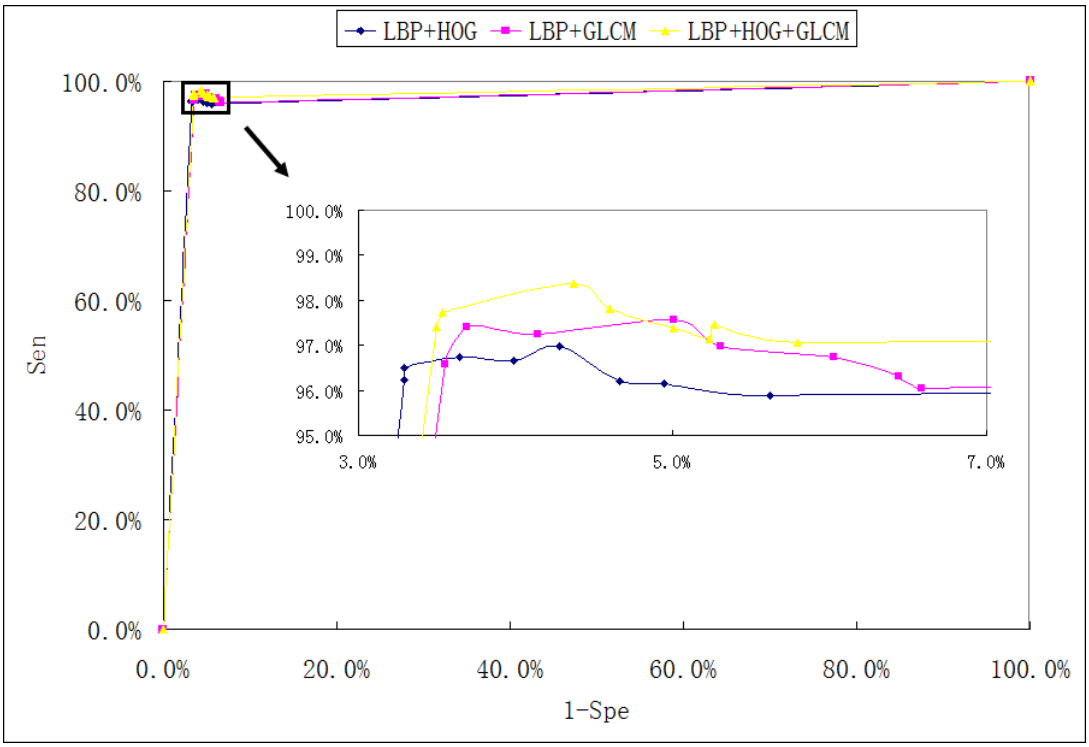

Fig. 9. ROC curves using different combination of LBP, GLCM and HOG features with Adaboost. To better discriminate different curves, the rectangle area has been enlarged and shown in the middle.

Based on the outcome of our comparison analysis, we apply LBP+GLCM+HOG features and Adaboost classifier in the final AS detection system. Considering different characteristics among D2, D3 and D4

(atypical TB without AS, normal cases without AS and nodule cases without AS, respectively), we conduct tests for the datasets D1 with D2 (D1D2), D1 with D3 (D1D3), and D1 with D4 (D1D4), respectively. Notice that in Table 5, the best performance of using LBP+GLCM+HOG with Adaboost is achieved when t = 3/4, thus the threshold for the whole datasets tests is chosen as t = 3/4. The final AS detection performance is shown in Table 6. Quantitative analysis shows that the proposed CAD system achieves both high sensitivity and specificity. Examples of AS detection results of CXRs from D1 are shown in Fig. 10. It could be observed that the detected positive windows are quite consistent with the ground truth drawn by radiologist. However, the specificity and precision in the test of D1D2 is relatively lower. It is because lots of false positive windows are detected in the images of D2 comparing to D3 and D4. This higher false positive rate (lower specificity) in D1D2 test could be explained as more lesions similar to AS caused by atypical TB are observed in CXRs in D2. Although the specificity in tests of D1D3 and D1D4 are close to 100%, there are still several images in D3 and D4 containing false positive windows. Examples of those false positive windows are shown in Fig. 11.

Table 6. Final AS detection results

Results	D1D2	D1D3	D1D4
Sen	98.4%	98.1%	98.3%
Spe	93.5%	99.9%	99.9%
Pre	63.3%	99.5%	98.1%

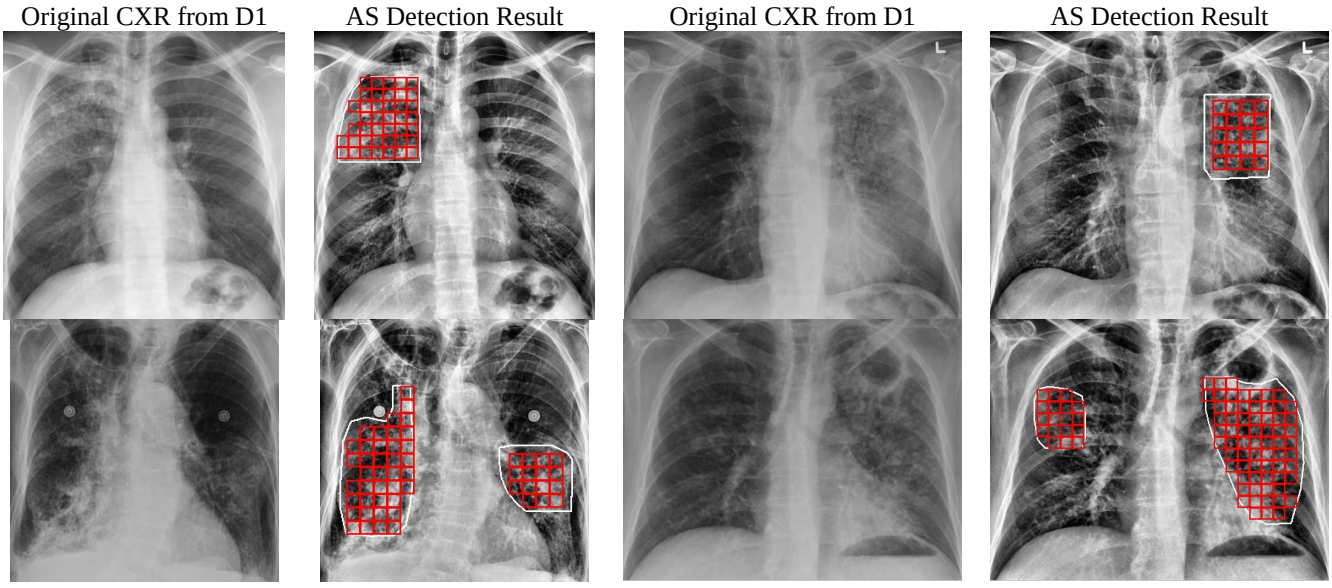

Fig. 10. True positive windows detected in D1 comparing with the ground truth (white contour).

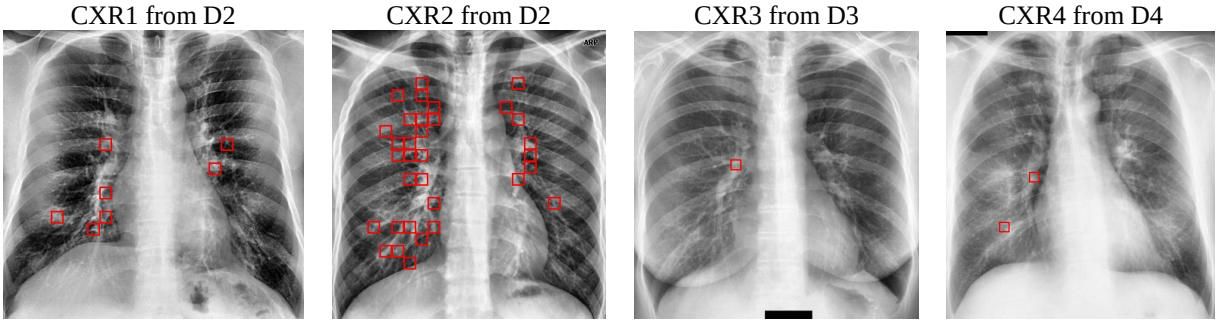

Fig. 11. False positive windows detected in D2, D3 and D4.

C. False Positive Reduction

To further reduce the false positive windows so that radiologist diagnosis can be more effective, we introduce a morphological smoothing operation. Notice that most of those false positive windows are discrete, thus, we use a simple smoothing technique based on morphological operation to reduce the false positives. Considering a block of 3×3 windows, for each center window, if five or more of its 8-neighbor windows have the different classification label as the center window, the center window is smoothed as the same label as the majority of its neighbourhood. By applying this smoothing technique, the final AS detection results are greatly refined for the specificity and precision, while keeping the high sensitivity. See Table 7 for the refined results. Examples of false positive windows in CXR1, CXR3 and CXR4 in Fig. 11 are all removed except a few left in CXR2 (See Fig. 12 for the rest of false positive windows).

Table 7. Final AS detection results

Results	D1D2	D1D3	D1D4
Sen	92.4%	91.7%	92.5%
Spe	97.4%	100%	100%
Pre	80.4%	100%	100%

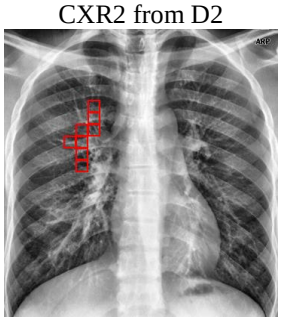

Fig. 12. False positive windows detected after the smoothing technique.

IV. Conclusions

A hybrid intelligent system is proposed to detect acinar shadow regions in chest radiographs. This novel CAD system takes advantages of integrating multiple features of LBP, GLCM and HOG into the Adaboost classifier. False positives are further reduced by introducing a morphological smoothing technique. The proposed CAD system shows an outstanding performance with more than 92% sensitivity, 97% specificity and 80% precision, which make it an effective tool to improve diagnostic performance. Our CAD system is efficient to eliminate a large number of irrelevant cases so that the radiologists can focus on a smaller set of significant cases. Note that accurate and automatic lung field segmentation technique such as ERF-ASM proposed in [28] used in conjunction of the CAD system proposed in this paper will help not only fully automate the detection but also remove false positives out of the lung region. Future work includes exploring other features for validating the classification accuracy between typical and atypical TB. For example, the co-occurrence of other abnormalities in the vicinity of AS might be studied.

Acknowledgements

The authors would like to thank Dr. Dirk Jan Kroon, University of Twente, for sharing the source code of Adaboost. They would also like to acknowledge Mr. Halim Elamy for coordinating the CXR image database from the University of Alberta Hospital.

Conflict of Interests

The authors declare that no conflict of interest exits in the submission of this manuscript, and they have no financial and personal relationships with other people or organizations that can inappropriately influence their work.

References

[1] World Health Organization, “WHO Report 2011 Global Tuberculosis Control,” Geneva, Switzerland, WHO Press, 2011, pp. 9-27.

[2] P. C. Hopewell, M. Pai, D. Maher, M. Uplekar, and M. C. Raviglione, "International standards for tuberculosis care," Lancet Infect. Dis., 2006, 6(11), pp. 710-725.

[3] K. Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential," Comput. Med. Imaging Graph., 2007, 31(4-5): 198-211.

[4] D.W. De Boo, M. Prokop, M. Uffmann, B.V. Ginneken, C.M. Schaefer-Prokop, "Computer-aided detection (CAD) of lung nodules and small tumours on chest radiographs," Eur. J. Radiol., 2009, 72 (2): 218-225.

[5] B.V. Ginneken, L. Hogeweg, and M. Prokop, "Computer-aided diagnosis in chest radiography: beyond nodules," European Journal of Radiology, 2009, vol 72, no. 2, pp. 26-30.

[6] D.W. De Boo, M. Uffmann, M. Weber, S. Bipat, E.F. Boorsma, M.J. Scheerder, et al., "Computer-aided detection of small pulmonary nodules in chest radiographs: an observer study," Acad. Radiol., 2011, 18(12):1507-14.

[7] A. Lau, R. Long, et al, "The public health consequences of smear positive pulmonary tuberculosis in patients with typical and atypical chest radiographs," 15th Annual International Union Against Tuberculosis and Lung Disease – North American Region Conference, February 24-26, 2011. Vancouver, British Columbia.

[8] S. Lange and G. Walsh, *Radiology of Chest Diseases (2nd ed.)*, Stuttgart: Thieme, 1998.

[9] R. Long, B. Maycher, A. Dhar, et al, "Pulmonary tuberculosis treated with directly observed therapy: serial changes in lung structure and function," *Chest*, 1998, 113(4):933-43.

[10] C. L. Daley, M. B. Gotway, R. M. Jasmer, *Radiographic Manifestations of Tuberculosis: A Primer for Clinicians*, Francis J. Curry National Tuberculosis Center, 2009.

[11] T. Xu, I. Cheng, R. Long and M. Mandal, "Novel coarse-to-fine dual scale technique for tuberculosis cavity detection in chest radiographs," EURASIP Journal on Image and Video Processing 2013, 2013:3.

[12] D.M. Hansell, A.A. Bankier, H. MacMahon, T.C. McLoud, N.L. Müller, and J. Remy, "Fleischner Society: glossary of terms for thoracic imaging," Radiology, 2008, 246(3):697-722.

[13] Japanese Society of Radiological Technology. http://www.jsrt.or.jp/english.html.

[14] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization", Chapter VIII.5, *Graphics Gems IV*. P.S. Heckbert (Eds.), Cambridge, MA, Academic Press, 1994, pp. 474-485.

[15] H. Harzallah, F. Jurie, and C. Schmid, "Combining efficient object localization and image classification," in Proc. of ICCV, 2009, pp.237-244.

[16] P.F. Felzenszwalb, R.B. Girshick, D.A. McAllester, and D. Ramanan, "Object Detection with Discriminatively Trained Part-Based Models," IEEE Trans. Pattern Anal. Mach. Intell., 2010, pp.1627-1645.

[17] R.C. Gonzalez and R.E. Woods, "Representation and Description", Chapter 11, *Digital Image Processing (2nd ed.)*, Prentice Hall, Upper Saddle River, NJ, 1992, pp.666-669.

[18] K.D. Donohue, L. Huang, T. Burks, F. Forsberg, and C.W. Piccoli, "Tissue Classification with Generalized Spectrum Parameters," Ultrasound in Medicine and Biology, 2001, vol. 27, no. 11, pp. 1505-1514.

[19] R. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification", IEEE Transactions on Systems, Man, and Cybernetics, 1973, vol. 3, no. 6, pp. 610–621.

[20] A.P. Pentland, “Fractal-Based Description of Natural Scenes”, IEEE Trans. Pattern Anal. Mach. Intell., 1984, pp.661-674.

[21] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary pattern,” IEEE Trans. Pattern Anal. Mach. Intell., 2002, vol. 24, no. 7, pp. 971–987.

[22] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection”, in Proc. of IEEE CVPR, 2005, vol. 1, pp. 886–893.

[23] R. Mukundan, S.H. Ong, and P.A. Lee, “Image analysis by Tchebichef moments”, IEEE Trans. Image Processing, 2001, pp.1357-1364.

[24] C. J. C. Burges, “A tutorial on support vector machines for pattern recognition,” Data mining and Knowledge Discovery, 1998, pp.121-167.

[25] R.E. Schapire, “A Brief Introduction to Boosting,” in Proc. IJCAI, 1999, pp.1401-1406.

[26] H. Zhang and C. Gu, “Support Vector Machines versus Boosting,” SVM report, 2008, Electrical Engineering, UC Berkeley.

[27] C. C. Chang and C. J. Lin, “LIBSVM: A Library for Support Vector Machines,” 2011, software available at: http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[28] T. Xu, M. Mandal, R. Long, I. Cheng, and A. Basu, “An edge-region force guided active shape approach for automatic lung field detection in chest radiographs,” CMIG, 2012, in press.