Journal of Medical Systems (2019) 43: 87
https://doi.org/10.1007/s10916-019-1222-8

SYSTEMS-LEVEL QUALITY IMPROVEMENT

Analysis of Tuberculosis in Chest Radiographs for Computerized Diagnosis using Bag of Keypoint Features

Satyavratan Govindarajan [1] · Ramakrishnan Swaminathan [1]

Received: 1 January 2019 / Accepted: 21 February 2019 / Published online: 28 February 2019

Abstract

Chest radiography is the most preferred non-invasive imaging technique for early diagnosis of Tuberculosis (TB). However, lack of radiological expertise in TB detection leads to indiscriminate chest radiograph (CXR) screening. A modest classification approach based on the local image description to detect subtle characteristics of TB using CXRs is highly recommended. In this work, an attempt has been made to classify normal and TB CXR images using Bag of Features (BoF) approach with Speeded-Up Robust Feature (SURF) descriptor. The images are obtained from a public database. Lung fields segmentation is performed using Distance Regularized Level Set (DRLS) formulation. The results of segmentation are validated against the ground truth images using similarity, overlap and area correlation measures. BoF approach with SURF keypoint descriptors is implemented to categorize the images using Multilayer Perceptron (MLP) classifier. The obtained results demonstrate that the DRLS method is able to delineate lung fields from CXR images. The BoF with SURF keypoint descriptor is able to characterize local attributes of normal and TB images. The segmentation results are found to be in high correlation with ground truth. MLP classifier is found to provide high Recall, Specificity (Spec), Accuracy, F-score and Area Under the Curve (AUC) values of 87.7%, 85.9%, 87.8%, 87.6% and 94% respectively between normal and abnormal images. The proposed computer aided diagnostic approach is found to perform better as compared to the existing methods. Thus, the study can be of significant assistance to physicians at the point of care in resource constrained regions.

Keywords Chest radiograph · Tuberculosis · Bag of features · Speeded-up robust feature descriptor · Level set

Introduction

Tuberculosis (TB) is a contagious lung disease caused by the bacteria *Mycobacterium tuberculosis*. This airborne communicable disease is the leading cause of deaths from a single infectious agent [1]. According to World Health Organization (WHO), 10.4 million people developed TB and an estimated amount of 1.6 million people died due to the disease in 2017 [2]. Early diagnosis with mass screening strategies are essential to prevent TB, especially in remote and resource poor regions [1, 3].

WHO highly recommends chest radiograph (CXR) imaging for the preliminary and immediate diagnosis of pulmonary TB. The common radiological indicators of pulmonary TB are opacity in lung lobes (air space consolidation), sand like patterns (military micronodular infiltrates), and indistinctness in medial and lateral pulmonary sulci (pleural effusion) [4, 5]. However, these image biomarkers appear subtle to be detected using CXRs and human interpretation leads to misdiagnosis [6].

Computer Aided Diagnosis (CAD) solutions use radiographic imaging data and machine learning techniques to assist radiologists in the diagnosis of TB [7]. It involves image segmentation, feature extraction and classification procedures. The task of delineation of lung borders from CXR images is highly challenging owing to inter and intra lung shape variations among subjects and, inconsistencies in the boundaries of significant landmarks [8, 9]. Different rule-based methods, pixel-classification, deformable models, pattern classification and hybrid methods have been applied for the segmentation of lung fields [4, 7–9, 11–13]. However, these methods tend to

This article is part of the Topical Collection on *Systems-Level Quality Improvement*

✉ Satyavratan Govindarajan
s.vratan@gmail.com

[1] Biomedical Engineering Group, Department of Applied Mechanics, Indian Institute of Technology Madras, Chennai, India

be stuck in local optima, relies on many internal parameters and generally require large training data to meet the real-time requirements of the CAD system in clinical setting. Image registration-based segmentation models have also been successfully implemented [9–11]. The drawback of this category is that the registration accuracy limits the accuracy of segmentation [14].

Level Set (LS) approaches are geometric curves which evolve as a dynamic contour to track complex and irregular topological variations in images. This method is commonly applied in segmentation due to its provision of an implicit boundary illustration which is non-parameterized, contain split and merge solutions, and the ability to naturally extend to any dimension [15, 16]. A level set evolution termed Distance Regularized Level Set (DRLS) eliminates the need for re-initialization which in turn mitigates the induced numerical errors caused by strong backward diffusion [15–17]. These models can achieve sub-pixel accuracy of object boundaries and can represent smooth closed contours. Due to the intrinsic distance regularization embedded in the level set function, the formulation uses relatively large time steps to significantly reduce number of iterations, whilst preserving the accuracy in both full and narrowband implementations. It has been applied in the segmentation of Corpus Callosum from brain MR images and kidney segmentation from ultrasound images [16, 17]. Combination of DRLS and a deep belief network has been applied in lung field segmentation in which semi-automated methods provided better results [18]. Recently, the DRLS method has been used as a post-processing step subsequent to feature learning to refine lung fields [7].

Several shape, edge, texture, fractal features and combination of them have been explored to characterize lung regions [1, 10–12, 19] and various classifiers such as Multilayer Perceptron (MLP), logistic regression, Support Vector Machines (SVM), decision trees, and Naïve Bayes, have been implemented [1, 4, 11, 12]. Speeded-Up Robust Feature (SURF) keypoint detector and descriptor method is analogous to scale-invariant feature transform, though it has a condensed descriptor length which is effectively faster in several image classification and object recognition tasks [20, 21]. However, the application of SURF algorithm in TB detection using CXR images has been limited [22] and not been explored comprehensively.

Bag of Features (BoF) approach provides an extremely compact image description as these methods are inspired by Bag of visual words concept used in visual categorization [23]. The method is computationally efficient, and intrinsically invariant to affine transformations and intra-class variations. This approach has been widely employed in image classification and retrieval [21, 23–25] and also in medical applications [26–28]. The technique has been applied for X-ray pathology image categorization and retrieval as a patch-based representation using principal component analysis and scale-invariant transform [29, 30]. However, a BoF with SURF descriptor approach has not been explored in the categorization of normal and TB CXRs.

Deep learning techniques provided exceptional results in the classification of CXR images. However, they require very high-performance computational systems and large annotated training data. Availability of such systems are limited in medical field [7]. Therefore, a modest classification scheme involving robust feature extraction is necessary to decrease detection errors and to improve screening efficiency of CXRs. It could be valuable in the settings of limited data accessible for the early detection of TB [1, 4, 7].

In this framework, normal and TB CXR image classification is performed using BoF approach with SURF descriptors. The CXR lung fields are segmented using DRLS scheme and the results are validated. BoF scheme is employed and the resultant bag of keypoints as feature vectors are fed to MLP classifier and their performance is analyzed. The rest of the paper describes the methods, results and discussion, the conclusions followed by acknowledgement and references.

Materials and methods

The methods of this study are demonstrated in Fig. 1. The CXR images of normal and TB subjects are obtained from a public dataset. The images are subjected to DRLS segmentation to obtain lung field binary masks. The delineated masks are subjected to BoF approach with SURF keypoint extraction and vector quantization using K-means clustering. The results of the approach are the bag of keypoints from images represented as collection of frequency histograms. These feature bags are then fed to MLP classifier to classify the images as normal and TB. Performance of segmentation and classification is analyzed.

Image database

In this work, 138 digital CXRs in Posterior-Anterior view (Normal-80, Abnormal (TB)-58) are obtained from the publicly available TB database – Montgomery County Set (MC) with their corresponding ground truth images. Images are acquired in 12-bit grayscale and has been captured using Eureka stationary X-ray machine. The ground truth images marked by the expert radiologists have been provided. The abnormal CXR images consist of manifestations of TB, including cavitations, effusions and miliary patterns [31]. The size of the images given in the dataset is either 4020 × 4892 or 4892 × 4020 pixels. The images are reduced to a resolution of 512 ×

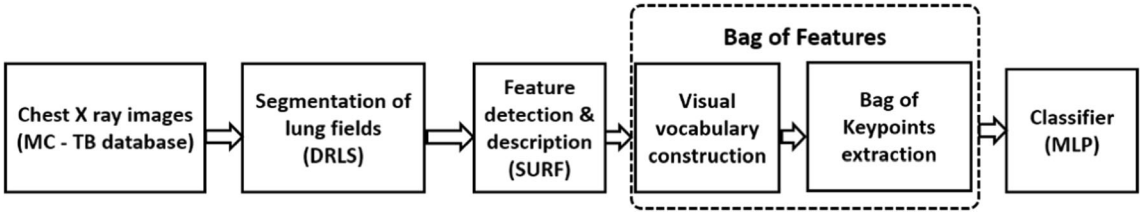

Fig. 1 Block diagram representation of proposed methodology

512 pixels using bicubic interpolation to yield smoother boundary, marginally better suppression of disruptive feature noise from ribs and bones and, for effective segmentation process [9]. This image size has been consistently used in the subsequent steps for classification.

Distance regularized level set evolution

Level set is an active contour model that utilizes a dynamic parametric contour model represented by $C(s,t):[0,1]\times[0,\infty)\rightarrow\Re^2$, s in [0, 1] and t in $[0,\infty)$. The standard curve evolution relation is given in Eq. 1.

$$\frac{\partial C(s,t)}{\partial t}=F\hat{N} \tag{1}$$

where F is the function that controls the rate of expansion/shrinkage and $\hat{N}$ is the normal pointing within the curvature C. It is expressed as $\hat{N}=-grad(\varphi)/|grad(\varphi)|$, where φ is level set function. The distance regularized level set formulation [15] is utilized to avoid re-initialization at periodic intervals and the level set contour is driven to the desired location using an additional external energy term added in the formulation.

The minimized energy equation of an edge based active contour model in DRLS formulation is provided in Eq. 2, given as initial level set function $\varphi(x,0)=\varphi_0(x)$.

$$\frac{\partial\varphi}{\partial t}=\mu div\left(d_p|\nabla\varphi|\nabla\varphi\right)+\lambda\delta_y(\varphi)div\left(g\frac{\nabla\varphi}{|\nabla\varphi|}\right)+\alpha g\delta_y(\varphi) \tag{2}$$

Here, μ and $\lambda>0$, constant, δ is a Dirac delta function, d_p is a function defined in [15] and α controls the velocity of the moving contour. g is the edge indicator having a Gaussian kernel with standard deviation G_σ [7, 15, 17] to smooth the edges of lung fields and reduce noise of image I is given in Eq. 3,

$$g=\frac{1}{1+|\nabla G_\sigma * I|^2} \tag{3}$$

The edge indicator approaches to zero when its gradient magnitude function is large. Thus, the curve propagation stops at lung fields to perform segmentation. Semi-automatic DRLS model is used in this study. The edge indicator takes smaller steps at object boundary locations.

The segmented lung fields of all CXR images are compared with the ground truth images by comparing them with similarity and overlap measures such as Accuracy, Sensitivity, Specificity, Jaccard index, Dice index and Volume similarity. The ratio of intersecting and non-intersecting areas obtained using these measures validate the reliability of DRLS segmentation against ground truth images. The segmented areas of normal and abnormal masks are correlated with ground truth area using Pearson correlation coefficient (R) measure [32].

Speeded-up robust feature (SURF) descriptor

To detect and characterize the most informative regions of the images, dense SURF local descriptors are implemented. SURF keypoints are extracted from image patches to describe the local intensity distributions. The patches are selected using grid method of patch sizes 8 × 8 and 16 × 16. SURF keypoint descriptor is a multi scale-space model which uses a Hessian blob detector to find keypoints in an image and the Hessian determinant for scale selection [20, 33]. The descriptor is formed within the neighborhood of keypoints which is computed using distribution of Haar wavelet responses.

Consider, $X=(p,q)$is the given representative location in an image M, the Hessian matrix $H(X,\sigma)$ in X at scale σ can be determined as in Eq. 4,

$$H(X,\sigma)=\begin{pmatrix}\mathrm{Lxx}(X,\sigma) & \mathrm{Lxy}(X,\sigma)\\ \mathrm{Lyx}(X,\sigma) & \mathrm{Lyy}(X,\sigma)\end{pmatrix} \tag{4}$$

where Lxx(X,σ), Lxy(X,σ) and Lyy(X,σ) are the results of convolution of the second order derivative of Gaussian filter with the image M in point X. It finds keypoints along both the horizontal and vertical directions that correspond to sudden changes in the intensity variations. These points are investigated over a subspace of {x, y, σ}, where σ represents the Gaussian scale space at which the keypoint occurs. The vector length of SURF descriptor is 64 and is formed by histogram representations of gradient orientations in the local neighborhood around each keypoint.

Bag of features approach and classification

Bag of features (BoF) stands analogous to the concept of Bag of Visual Words used in text mining. It is used to characterize an image as unordered compilation of local features [23, 24]. This procedure involves two major steps, (i) extraction of local features from images, (ii) construction of visual vocabulary by

reducing the number of features and representation of bag of keypoint vectors through histograms. Extraction of local features is performed using SURF descriptors and the method is already explained.

Vocabulary construction and BoF representation The descriptors constructed from the images are grouped into clusters of visual words using K-means clustering where K is vocabulary size [21] and cluster centroids are found using Euclidean distance metric. For an image, each descriptor is mapped into its nearest cluster centroid. An encode method is used to count the visual word occurrences in an image based on the cluster centroids. It produces a reduced representation of an image through histograms. This feature vector is the Bag of keypoints and the approach is referred to as BoF illustration of the image.

Classification from BoF approach of normal and abnormal images is performed using Multilayer Perceptron (MLP) classifier. MLP is a supervised learning network belongs to the family of feed-forward neural nets trained with back propagation tuning [34, 35]. MLP is a powerful function approximator of prediction and classification, arguably the most commonly used and well-studied artificial neural network architecture due to their fast operation, ease of implementation, and smaller training set requirements [36]. MLPs contain non-linear activation function with multiple layers. The sigmoid function is used as an activation function in the study. The procedure is computationally inexpensive and can be implemented with local information. In this work, MLP with three hidden layers, learning rate of 0.3 and momentum of 0.2 is considered. Performance of classifiers are analyzed using Recall, Specificity (Spec), Accuracy and F-score measures [37]. Receiver operating characteristic (ROC) analysis is employed and Area Under the Curve (AUC) measure is computed to discriminate normal and TB images [1]. Ten-fold cross validation for ten number of iterations is performed on the classifier and average values of the folds are reported [1, 2]. The segmentation, feature extraction and Bag of keypoints representation and classification are performed using MATLAB 9.5.0 (R2018b).

Results and discussion

Figure 2 (a-d) shows a representative set of normal and TB CXR images containing original images in Row 1, segmented binary masks at Row 2 and ground truth masks in Row 3. From Fig. 2, Row 1, it is observed that the lung field boundaries appear to be indefinite at significant landmarks such as aortic arch, cardiophrenic and costrophrenic angles, hila, and, in surrounding rib cage contours. The images suffer from high contrast variations and intensity disparities along the radiolucent ribs in lung field regions due to overlay of sub- structures. The lung fields are also found to have structural dissimilarities owing to inter-subject variability posing diagnostic challenges in lung field segmentation and abnormality detection.

DRLS method is initiated to segment lung field boundaries from CXR images. The initial contour is dynamically evolved to shrink so as to find the lung field boundaries using edge indicator as the stopping condition. Figure 2, Row 2 shows segmented lung binary masks through DRLS method. It can be observed that the method is able to delineate lung fields along all the significant landmarks of CXRs. The ground truth images provided in the dataset are shown in Fig. 2, Row 3.

The segmented lung binary masks of all the images are validated by comparing them with the ground truth images. The similarity and overlap measures such as Accuracy, Sensitivity, Specificity, Jaccard index, Dice index and Volume similarity are calculated for overall validation of segmentation and are found to be 96.08%, 85.02%, 99.68%, 84.29%, 91.3% and 92.2% respectively as shown in Table 1. The segmentation validation measures for normal and abnormal images are also computed and are found to be 96.59%, 86.49%, 99.74%, 85.71%, 92.29%, 93.32% and, 95.55%, 83.48%, 99.61%, 82.85%, 90.44%, 91.25% respectively as shown in Table 1. It can be seen that high validation performance is achieved for both normal and abnormal CXRs. However, the performance of segmentation in abnormal images is observed to be comparatively low than normal images. This could be due to the irregular boundaries and pixel intensity variations of lung fields in the case of abnormalities which produce more opacity such as effusions and major cavitations present in some of the CXR images.

The scatterplot representation of segmented normal and abnormal lung areas against ground truth areas are depicted in Fig. 3 (a & b). The Pearson correlation coefficient values are found to be $R = 0.98$ and $R = 0.96$ for segmented normal and abnormal lung areas against the ground truth area. It can be observed that the segmented lung areas are in maximum correlation with ground truth areas. The high validation results suggest that DRLS method is able to accurately delineate the lung fields of CXR images.

Furthermore, the segmented binary masks are multiplied with the original images and are subjected to dense SURF keypoint feature extraction and BoF approach for image category classification. Firstly, dense SURF descriptors are extracted as patches from the mask multiplied images to detect and localize keypoints which provides information on the position and the corresponding local intensity distribution of lung field regions as shown in Fig. 4. The images in Fig. 4 (a & b) are segmented mask multiplied normal and abnormal images of Fig. 2 (a & c) respectively. It can be observed that the SURF feature descriptors are able to localize keypoints and its corresponding descriptors which represent significant normal and subtle abnormal characteristics.

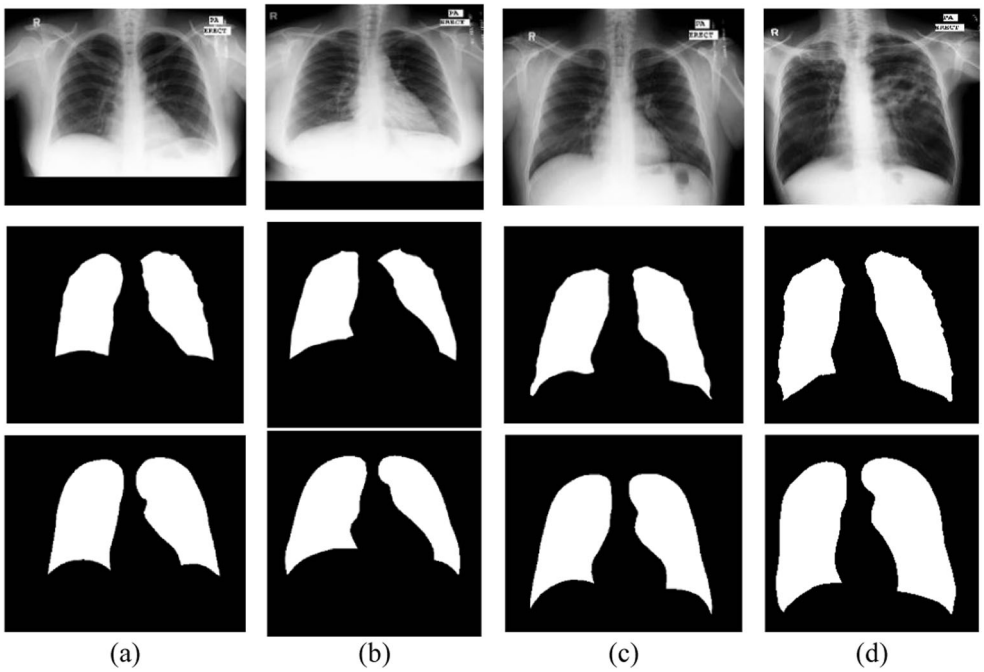

Fig. 2 Representative (a & b) Normal and (c & d) TB CXR images; Row 1: Original images, Row 2: DRLS Segmented masks, Row 3: Ground truth images

In this study, the local feature detection is performed with patch sizes 8 × 8 and 16 × 16. The number of detected SURF features are reduced using vector quantization by K-means clustering. This enables to construct a visual vocabulary with different vocabulary sizes 100, 300, 500, 700 and 900 to be examined. The reduced representation of normal and abnormal images is encoded to obtain frequency histograms and forms a bag of keypoint feature vectors.

The average visual word histogram representation for normal and abnormal (TB) images with vocabulary size 500 is shown in Fig. 5 (a & b) respectively. It can be observed that average histogram of abnormal images is found to be different from the average histograms of normal images. It is also noticed that there are visual words which are more common than others. This could be due to intrinsic nature of intensity distributions across TB CXR images. The most frequent visual word is found to be present in all the abnormal (10th bin, highlighted in pink) and normal images (298th bin, highlighted in blue). Hence, these detected local image patches correspond to the normal and abnormal characteristics of images. Therefore, this information of frequency distribution of features is vital for classification process.

Normal and abnormal (TB) image classification is performed using MLP classifier for considered patch sizes and vocabulary sizes. Figure 6 (a) shows the trend of classifier accuracy as the values increase from vocabulary sizes of 100 to 500 and thereafter decreases for both the patch sizes. It is observed that MLP classifier provided maximum average accuracy of 87.8% for patch size 8 × 8 and vocabulary size of 500. The performance measures such as Recall, Spec and F-score for patch sizes 8 × 8 and 16 × 16 with vocabulary size of 500 are analyzed as shown in Table 2. For patch size 8 × 8, it can be observed that MLP provided high Recall, Spec and F-score values of 87.7%, 85.9% and 87.6% when compared to the performance values for patch size 16 × 16. This could be due to the ability of MLP in better differentiating normal and TB images based on the detection of subtle local image characteristics with small grid sizes. Using 8 × 8 with vocabulary size 500, it is found that out of 80 normal images, 74 images have been correctly classified and for 58 abnormal images, 47 images have been correctly classified.

The high classification performance for 8 × 8 patch size and vocabulary size of 500 could be due to the detection of robust local features by SURF descriptors from images. This allows characterization of subtle abnormal features from

Table 1 Performance validation of DRLS segmentation method

Measures	Accuracy	Sensitivity	Specificity	Jaccard index	Dice index	Volume similarity
Overall (%)	96.08	85.02	99.68	84.29	91.38	92.27
Normal (%)	96.59	86.49	99.74	85.71	92.29	93.32
Abnormal (%)	95.55	83.48	99.61	82.85	90.44	91.25

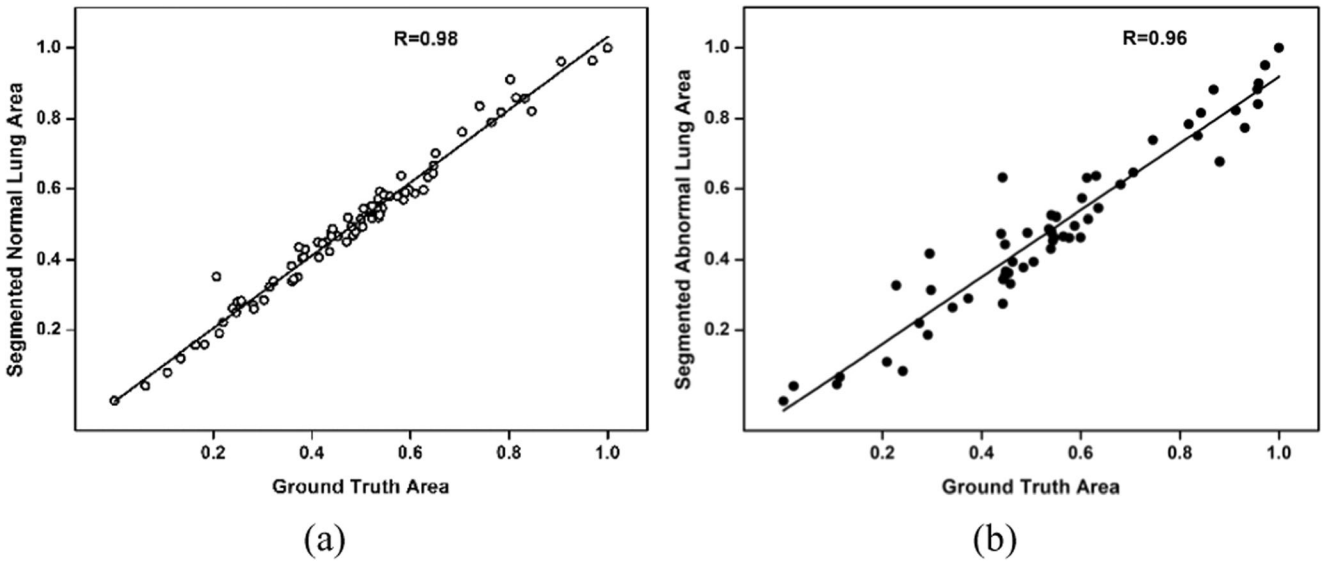

Fig. 3 Scatter plot illustration of segmented lung area versus ground truth for (**a**) normal and (**b**) abnormal (TB) subjects

images for better differentiation of normal and TB CXR images. Better performance of MLP could be due to the ability of neural networks to handle complex feature data and better estimate the decision surface.

Figure 6 (b) depicts ROC analysis of classifiers. MLP provided high AUC value of 94% for patch size 8 × 8 with that of 88.6% for 16 × 16. This shows the ability of MLP classifier to better discriminate normal and abnormal (TB) subjects for patch size 8 × 8. The obtained performance of classification from the proposed methodology is compared with the other methods that are reported to have been implemented particularly on the MC dataset and is tabulated in Table 3. It is observed that the proposed method achieved high Recall, Spec, Accuracy and AUC values than the other methods for the considered dataset.

Jaeger et al. [31] performed an automatic tuberculosis screening in chest radiographs using two feature sets. Graph cuts are used to segment the lung fields and the features are extracted. The feature sets consisted of combination of intensity, texture, shape and edge features such as intensity histograms, shape descriptor histograms, histogram of oriented gradients, Tamura texture descriptor, Hu moments and others. Further, a feature selection strategy is applied to identify optimal features. These optimal features are fed to classifiers namely, MLP, Support Vector Machine (SVM), decision trees and logistic regression. The best results were obtained for SVM with AUC of 86.9% in the MC dataset with an overall classification accuracy of 78.3%. However, use of global features in this work restrict higher performance in the classification of normal and TB CXR images.

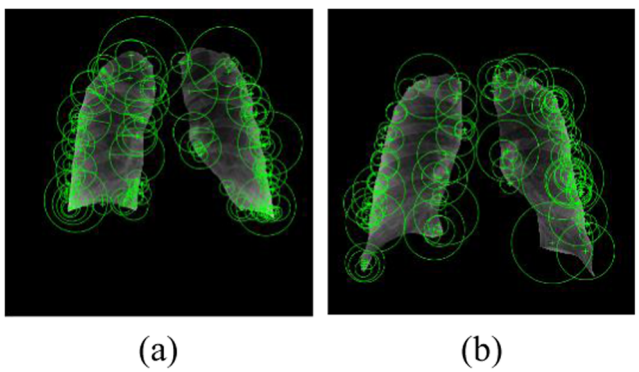

Fig. 4 Extraction of SURF keypoints from representative (**a**) normal and (**b**) TB lung field images

Santosh et al. [36] performed edge map analysis for automatic pulmonary TB screening in CXRs. It is based on the observation that abnormal images tend to exhibit deformed thoracic edge maps. Pyramid Histograms of oriented gradient (PHOG) features of thoracic edges is studied using five different regions of interest selection based on automatic, manual, relaxed and other techniques. A fully connected MLP was used. Accuracy of 79.2% and AUC value 88% is obtained. The drawback of the method is the spatial position of the chest cavity in the CXR image affect PHOG features as it is not translational and rotational invariant. This occurs mostly in abnormal images, and allows for disparity in normal and abnormal CXRs.

Lopes & Valiati [5] investigated three proposals using pretrained convolutional neural networks (CNN) as feature detectors in the analysis of TB screening. The first proposal evaluated simple CNN features to train the model. The second proposal implemented bag of CNN features as multiple instance learning for disease classification. The third proposal involved classifier ensembles trained using CNNs. The authors used Scale invariant feature transform based atlas model and graph cuts to segment lung regions. The best AUC of 90.8% and accuracy of 82.6% in MC set were obtained when vocabulary size K = 500 using bag of CNN features. The

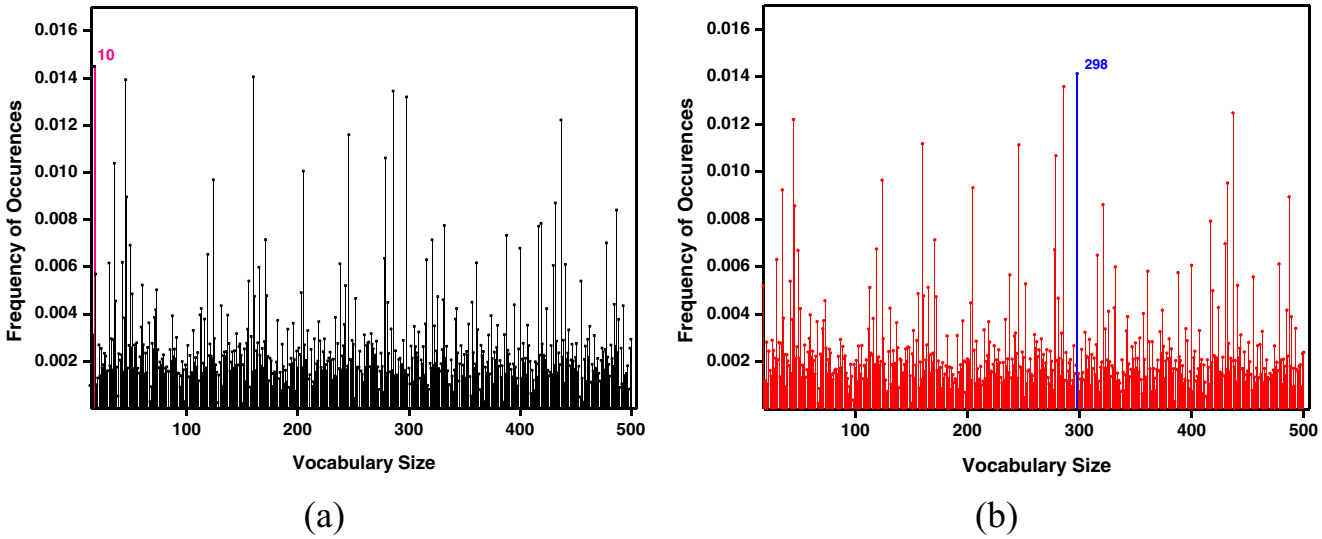

Fig. 5 Average visual word histogram representation of encoded bag of keypoints for (**a**) abnormal and (**b**) normal images. Most frequent visual word is highlighted (pink for abnormal and blue for normal)

authors discuss about the complexity of bag model with pre-trained CNNs which are still deep architectures and concludes that deep networks are often not required to improve the performance of disease classification.

Santosh et al. [1] analyzed the symmetry of lung regions using low and high-level shape, edge and texture features (R-signature and Gist) for the automated TB screening considering resource-constrained regions. Atlas based graph cuts are used to segment the lung fields. Classification of normal and TB have performed using voting method based on the combination of three different Bayesian network, MLP, and random forest classifiers. For MC set, the AUC value was found to be 90%.

Vajda et al. [2] performed feature selection strategies for the multitude of image features. Lung segmentation was carried out using atlas-based graph cuts method. The feature sets contained combination of edge, texture, shape, moment features similar to the work of Jaeger et al. [31]. A wrapper type feature selection model was used. Fully connected neural network classifier was used and found to yield better results than SVM. Feature selection method with neural nets provided accuracy of 84.75% and AUC value of 91%.

Conclusion

CAD systems are essential to improve the diagnostic efficacy of CXRs for TB screening in remote and resource poor settings. The subtle manifestations of TB in lung regions are hard to detect even by an experienced radiologist. Due to interruptions in radiologists' workflow while case read, memory decay cause errors in the diagnostic findings during screening

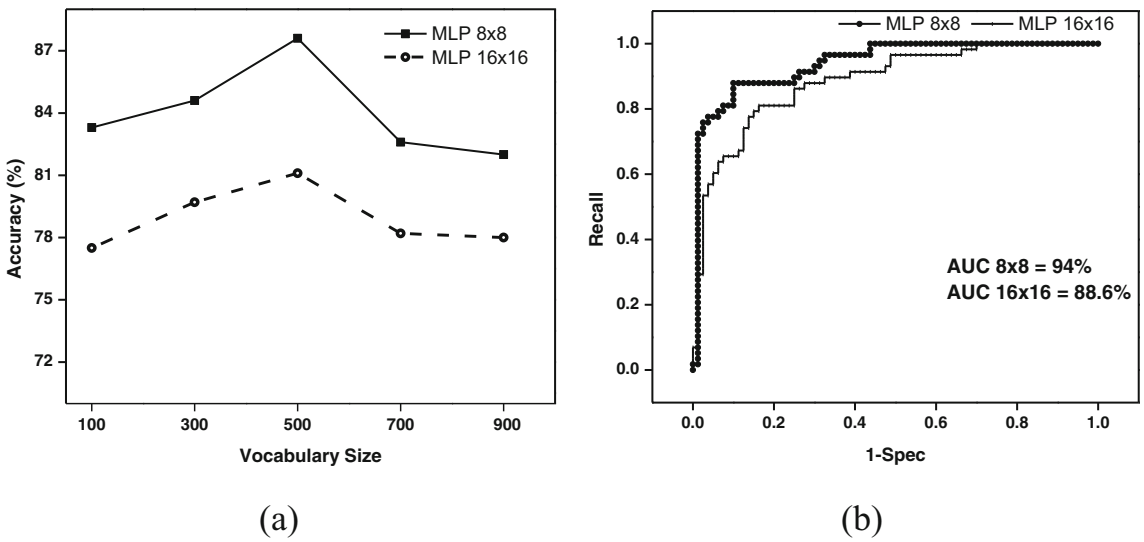

Fig. 6 (**a**) Accuracy of classifiers for different vocabulary sizes and (**b**) ROC analysis of MLP classifier

Table 2 Performance analysis of MLP classifier

Classifier	Patch size	Recall %	Spec %	F-score %
MLP	8 × 8	**87.7**	**85.9**	**87.6**
	16 × 16	81.2	79.2	81.1

*Bold numbers refer to best values for that category

process. Feature descriptor methods could provide localized image information in CXR images to detect subtle abnormal findings.

In this work, an attempt to classify normal and TB CXRs based on the local description of images by bag of keypoints using BoF approach with SURF descriptors is proposed. The results of DRLS segmentation are in high similarity and correlation with that of ground truth images. The dense SURF feature extraction technique is able to detect and localize keypoints along both the horizontal and vertical directions of image patches. These keypoints correspond to sudden changes in the intensity variations which characterize normal and TB CXR images. Since these features are scale invariant, rotation invariant and robust to noise, the subtle local characteristics of abnormal images can be detected and distinguished from other regions of the image and across various images.

Bag of features provides effective multiple instance learning framework by operating on the collective response of obtained keypoints. An appreciable difference in the average visual word histograms generated from bag of feature model of abnormal and normal images is observed. The most frequent visual word is present in all the normal and abnormal images even though the local regions are described by different histograms. These detected local image patches correspond to the normal and abnormal characteristics of images.

MLP with non-linear transformation function is able to assess the complex bag of keypoint feature data to classify normal and TB CXRs. MLP provides high performance values in differentiating normal and TB images based on the detection of subtle local image characteristics with small grid sizes. The powerful discriminating ability of MLP classifier is also observed with high AUC value obtained using ROC analysis. This shows that MLP provide an important support for the medical diagnostic decision. The classification results obtained from this study exceed the performance of existing methodologies for the considered dataset. Hence, the proposed method based on local description of an image, could be applied as a means to integrate medical expertise while developing an automated TB classification system.

Table 3 Performance evaluation of proposed methodology with other methods

Method	Recall %	Spec %	Accuracy %	AUC%
Jaeger et al. [31]	–	–	78.3	86.9
Santosh et al. [38]	–	–	79.2	88
Lopes & Valiati [5]	–	–	82.6	90.8
Santosh et al. [1]	86	81	83	90
Vajda et al. [2]	–	–	84.75	91
Proposed work	**87.7**	**85.9**	**87.8**	**94**

*Bold numerals refer to best values for that category

Acknowledgements The authors extend sincere thanks to the Science and Engineering Research Board, Department of Science and Technology, (Government of India) for supporting this study.

Funding This study is funded by Science and Engineering Research Board, Department of Science & Technology, Government of India (SERB-SB/S3/EECE/291/201).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Santosh, K. C., and Antani, S., Automated chest x-ray screening: Can lung region symmetry help detect pulmonary abnormalities? IEEE Trans. Med. Imaging 37:1168–1177, 2018. https://doi.org/10.1109/TMI.2017.2775636.
2. Vajda, S., Karargyris, A., Jaeger, S., Santosh, K. C., Candemir, S., Xue, Z., Antani, S., and Thoma, G., Feature selection for automatic tuberculosis screening in frontal chest radiographs. J. Med. Syst. 42:146, 2018. https://doi.org/10.1007/s10916-018-0991-9.
3. Skoura, E., Zumla, A., and Bomanji, J., Imaging in tuberculosis. Int. J. Infect. Dis. 32:87–93, 2015. https://doi.org/10.1016/j.ijid.2014.12.007.
4. Jaeger, S., Karargyris, A., Candemir, S., Siegelman, J., Folio, L., Antani, S., and Thoma, G., Automatic screening for tuberculosis in chest radiographs: A survey. Quant. Imaging. Med. Surg. 3:89–99, 2013. https://doi.org/10.3978/j.issn.2223-4292.2013.04.03.
5. Lopes, U. K., and Valiati, J. F., Pre-trained convolutional neural networks as feature extractors for tuberculosis detection. Comput. Biol. Med. 89:135–143, 2017. https://doi.org/10.1016/j.compbiomed.2017.08.001.
6. Melendez, J., Hogeweg, L., Sánchez, C. I., Philipsen, R. H., Aldridge, R. W., Hayward, A. C., Abubakar, I., van Ginneken, B., and Story, A., Accuracy of an automated system for tuberculosis detection on chest radiographs in high-risk screening. Int. J. Tuberc. Lung. Dis. 22:567–571, 2018. https://doi.org/10.5588/ijtld.17.0492.
7. Hooda, R., Mittal, A., and Sofat, S., Segmentation of lung fields from chest radiographs-a radiomic feature-based approach. Biomed. Eng. Lett.:1–9, 2018. https://doi.org/10.1007/s13534-018-0086-z.
8. Shao, Y., Gao, Y., Guo, Y., Shi, Y., Yang, X., and Shen, D., Hierarchical lung field segmentation with joint shape and

appearance sparse learning. IEEE Trans. Med. Imaging 33:1761–1780, 2014. https://doi.org/10.1109/TMI.2014.2305691.
9. Candemir, S., Jaeger, S., Palaniappan, K., Musco, J. P., Singh, R. K., Xue, Z., Karargyris, A., Antani, S., Thoma, G., and McDonald, C. J., Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. IEEE Trans. Med. Imaging 33: 577–590, 2014. https://doi.org/10.1109/TMI.2013.2290491.
10. Candemir, S., Jaeger, S., Palaniappan, K., Antani, S., Thoma, G., Graph-cut based automatic lung boundary detection in chest radiographs. *IEEE Healthcare Technology Conference: Translational engineering in health & medicine*: 31–34, 2012.
11. Qin, C., Yao, D., Shi, Y., and Song, Z., Computer-aided detection in chest radiography based on artificial intelligence: A survey. Biomed. Eng. Online 17:113, 2018. https://doi.org/10.1186/s12938-018-0544-y.
12. Hooda, R., Mittal, A., Sofat, S., A Survey of CAD Methods for Tuberculosis Detection in Chest Radiographs. Soft Computing: Theories and Applications. *Adv. Intell. Syst. Comput*. Springer: 273–282, 2019. https://doi.org/10.1007/978-981-13-0589-4_25.
13. Lee, W. L., Chang, K., and Hsieh, K. S., Unsupervised segmentation of lung fields in chest radiographs using multiresolution fractal feature vector and deformable models. Med. Biol. Eng. Comput. 54:1409–1422, 2016. https://doi.org/10.1007/s11517-015-1412-6.
14. Maintz, J. A., and Viergever, M. A., A survey of medical image registration. Med. Image Anal. 2:1–36, 1998. https://doi.org/10.1016/S1361-8415(01)80026-8.
15. Li, C., Xu, C., Gui, C., and Fox, M. D., Distance regularized level set evolution and its application to image segmentation. IEEE Trans. Image Process. 19:3243–3254, 2010. https://doi.org/10.1109/TIP.2010.2069690.
16. Anandh, K.R., Sujatha, C.M., Ramakrishnan, S., Atrophy analysis of corpus callosum in Alzheimer brain MR images using anisotropic diffusion filtering and level sets. 36th IEEE Annual International Conference on Engineering in Medicine and Biology Society: 1945–1948. 2014.
17. Yang, F., Qin, W., Xie, Y., Wen, T., and Gu, J., A shape-optimized framework for kidney segmentation in ultrasound images using NLTV denoising and DRLSE. Biomed. Eng. Online 11:82, 2012. https://doi.org/10.1186/1475-925X-11-82.
18. Ngo, T.A., Carneiro, G., Fully Automated Segmentation Using Distance Regularised Level Set and Deep-Structured Learning and Inference. *Deep Learning and Convolutional Neural Networks for Medical Image Computing*. Springer:197–224, 2017. https://doi.org/10.1007/978-3-319-42999-1_12.
19. Karargyris, A., Siegelman, J., Tzortzis, D., Jaeger, S., Candemir, S., Xue, Z., Santosh, K. C., Vajda, S., Antani, S., Folio, L., and Thoma, G. R., Combination of texture and shape features to detect pulmonary abnormalities in digital chest X-rays. Int. J. Comput. Assist. Radiol. Surg. 11:99–106, 2016. https://doi.org/10.1007/s11548-015-1242-x.
20. Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L., Speeded-up robust features (SURF). Comput. Vis. Image Underst. 110:346–359, 2008. https://doi.org/10.1016/j.cviu.2007.09.014.
21. Schmitt, D., and McCoy, N., Object classification and localization using SURF descriptors. CS 229:1–5, 2011.
22. Alfadhli, F.H., Mand, A.A., Sayeed, M.S., Sim, K.S., Al-Shabi, M., Classification of tuberculosis with SURF spatial pyramid features. IEEE International Conference on Robotics, Automation and Sciences: 1–5. 2017.
23. Csurka, G., Dance, C., Fan, L., Willamowski, J., and Bray, C., Visual categorization with bags of keypoints. ECCV Workshop Stat. Learn. Comput. Vis. 1:1–22, 2004.
24. O'Hara, S., Draper, B.A., Introduction to the bag of features paradigm for image classification and retrieval. arXiv preprint arXiv: 1101.3354. 2011.
25. Dardas, N. H., and Georganas, N. D., Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques. IEEE Trans. Instrum. Meas. 60:3592–3607, 2011. https://doi.org/10.1109/TIM.2011.2161140.
26. Rueda, A., Arevalo, J., Cruz, A., Romero, E., González, F.A., Bag of features for automatic classification of Alzheimer's disease in magnetic resonance images. *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. Springer: 559–566, 2012. https://doi.org/10.1007/978-3-642-33275-3_69.
27. Cruz-Roa, A., Caicedo, J. C., and González, F. A., Visual pattern mining in histology image collections using bag of features. Artif. Intell. Med. 52:91–106, 2011. https://doi.org/10.1016/j.artmed.2011.04.010.
28. Islam, M., Dinh, A. V., and Wahid, K. A., Automated diabetic retinopathy detection sing bag of words approach. J. Biomed. Sci. Eng. 10:86–96, 2017. https://doi.org/10.4236/jbise.2017.105B010.
29. Avni, U., Greenspan, H., Sharon, M., Konen, E., Goldberger, J., X-ray image categorization and retrieval using patch-based visual words representation. IEEE International Symposium on Biomedical Imaging: From Nano to Macro: 350–353. 2009.
30. Avni, U., Goldberger, J., Sharon, M., Konen, E., and Greenspan, H., Chest x-ray characterization: From organ identification to pathology categorization. ACM Proc. Int. Conf. Multimed. Inform. Retriev.:155–164, 2010.
31. Jaeger, S., Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F. M., Xue, Z., Palaniappan, K., Singh, R. K., Antani, S. K., and Thoma, G. R., Automatic tuberculosis screening using chest radiographs. IEEE Trans. Med. Imaging 33:233–245, 2014. https://doi.org/10.1109/TMI.2013.2284099.
32. Anandh, K. R., Sujatha, C. M., and Ramakrishnan, S., A method to differentiate mild cognitive impairment and Alzheimer in MR images using eigen value descriptors. J. Med. Syst. 40:25, 2016. https://doi.org/10.1007/s10916-015-0396-y.
33. Kashif, M., Deserno, T. M., Haak, D., and Jonas, S., Feature description with SIFT, SURF, BRIEF, BRISK, or FREAK? A general question answered for bone age assessment. Comput. Biol. Med. 68:67–75, 2016. https://doi.org/10.1016/j.compbiomed.2015.11.006.
34. Bishop, C.M., Neural networks for pattern recognition. Oxford university press. 1995.
35. Yan, H., Jiang, Y., Zheng, J., Peng, C., and Li, Q., A multilayer perceptron-based medical decision support system for heart disease diagnosis. Expert Syst. Appl. 30:272–281, 2006. https://doi.org/10.1016/j.eswa.2005.07.022.
36. Lin, C. C., Ou, Y. K., Chen, S. H., Liu, Y. C., and Lin, J., Comparison of artificial neural network and logistic regression models for predicting mortality in elderly patients with hip fracture. Injury 41:869–873, 2010. https://doi.org/10.1016/j.injury.2010.04.023.
37. Kavitha, J. C., and Suruliandi, A., Feature extraction using dominant local texture-color patterns (DLTCP) and classification of color images. J. Med. Syst. 42:220, 2018. https://doi.org/10.1007/s10916-018-1067-6.
38. Santosh, K. C., Vajda, S., Antani, S., and Thoma, G. R., Edge map analysis in chest X-rays for automatic pulmonary abnormality screening. Int. J. Comput. Assist. Radiol. Surg. 11:1637–1646, 2016. https://doi.org/10.1007/s11548-016-1359-6.