Chest X-ray Image View Classification

7 authors, including:

Zhiyun Xue
U.S. Department of Health and Human Services
59 PUBLICATIONS **906** CITATIONS

Daekeun You
University of Michigan
31 PUBLICATIONS **195** CITATIONS

Sema Candemir
The Ohio State University
39 PUBLICATIONS **510** CITATIONS

Stefan Jaeger
National Institutes of Health
80 PUBLICATIONS **2,030** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   R&D and API Development for Web-Deployable Spine X-Ray Classification, Boundary Estimation of Lumbar Vertebrae, and Medical Validation of Spine X-Ray Shapes View project

Project   Content-based Image retrieval View project

Chest X-ray Image View Classification

Zhiyun Xue, Daekeun You, Sema Candemir, Stefan Jaeger, Sameer Antani, L. Rodney Long, George R. Thoma
Lister Hill National Center for Biomedical Communications
National Library of Medicine
Bethesda, USA
{xuez, santani, rlong, gthoma}@mail.nih.gov

Abstract—**The view information of a chest X-ray (CXR), such as frontal or lateral, is valuable in computer aided diagnosis (CAD) of CXRs. For example, it helps for the selection of atlas models for automatic lung segmentation. However, very often, the image header does not provide such information. In this paper, we present a new method for classifying a CXR into two categories: frontal view vs. lateral view. The method consists of three major components: image pre-processing, feature extraction, and classification. The features we selected are image profile, body size ratio, pyramid of histograms of orientation gradients, and our newly developed contour-based shape descriptor. The method was tested on a large (more than 8,200 images) CXR dataset hosted by the National Library of Medicine. The very high classification accuracy (over 99% for 10-fold cross validation) demonstrates the effectiveness of the proposed method.**

Keywords—chest radiograph; view classification; contour-based shape feature

I. INTRODUCTION

Chest radiographs are frequently taken in the hospital as one crucial diagnostic imaging tool for identifying abnormalities in the chest. Very often, two views are taken: the *frontal view* and the *lateral view*. Figure 1 shows one patient's frontal and lateral chest radiographs, respectively. For computer-aided diagnosis (CAD) of lung diseases, segmenting the lung region out of the chest X-ray images is an essential component of the system. Atlas-based or shape-model-based image segmentation methods are one type of effective lung segmentation method [1]. For such methods, there is a need to know the view of the chest X-ray beforehand so that the correct model is applied. Not only is the external lung shape different according to the image view, but internal lung features may differ as well, as may be noted in Figure 1. Therefore, the classifier developed for categorizing the diseases may need to be trained differently, according to the image view [2, 3].
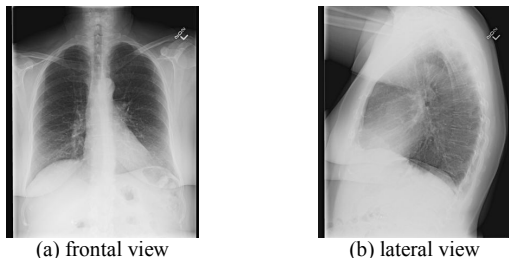
However, the view information may be unavailable in the text which accompanies the image, such as the large dataset of chest X-ray images we are working on (explained in detail in Section II). This motivated us to develop a high performance binary classification method to separate these two types of views for chest X-ray images based on image visual content only.

There are a few research efforts reported in the literature for identifying the chest radiograph image view [4-10]. Pietka et al. [4] developed a method to determine the image view based on minimum/maximum profile-length ratio. Boone et al. [5] applied a neural network to classify the image orientation. The features they used were projection profiles and four regions of interest. Arimura et al. [6] proposed a method to identify the frontal/lateral view using a template matching technique. The similarity measures were based on the cross-correlation coefficient. Lehmann et al. [7] also proposed a method to determine the image view based on the similarity of the image to reference images, but used four distance measures and K-nearest-neighbor classifier. In addition to the minimum/maximum profile-length ratio used in [4], Kao et al. [8] used another two features based on the analysis of the projection profile, namely body symmetry index and background percentage index, to identify frontal/lateral view. Kao et al. [9] also developed a method for distinguishing posteroanterior (PA) view from anteroposterior (AP) view. Luo et al. [10] classified the projection view of chest radiographs by using Bayes's decision theory. The features they used included the existence, shape, and spatial relation of medial axes of anatomic structures, as well as the average intensity of region of interest.

In this paper, we propose a new method, which has three major components. First, we pre-process the image to enhance contrast and remove irrelevant regions. We then extract four features: (1) image profile, (2) body size ratio, (3) pyramid of histograms of orientation gradients (PHOG) [11], and (4) our newly developed contour-based shape feature (CBSF). We used a Support Vector Machines (SVM) supervised classifier to train and test, and achieved very high accuracy (above 99%) for the large dataset that contains approximately 8300 chest X-ray images. We also compared our method on another publicly available chest x-ray dataset reported by [7, 12].

II. DATA

The National Library of Medicine (NLM) has been maintaining a large dataset of chest X-ray DICOM images containing both frontal view and lateral view and related

(a) frontal view   (b) lateral view

Fig. 1. Chest radiographs

textual radiology reports. The data was collected by the medical school at the University of Indiana, and consists of about 4000 radiology reports and 8300 images. These images, as well as corresponding radiology report information, have been integrated into *OpenI*, a multi-modal biomedical literature retrieval system developed by NLM [13]. For these images, we examined both the textual radiology reports and the DICOM image headers for frontal/lateral view information. The radiology reports contain no information indicating which of the image pair is lateral and which is frontal. However, in the DICOM image header, there are fields named "ViewPosition", "PatientOrientation", and "Laterality" which might be thought to contain view information. To investigate this possibility, we extracted the values of these fields for all the images and examined them, obtaining the results in Tables 1-3. Table 1 lists the number of images whose header contains the specific field name (case insensitive) and the number of images having values (not empty) for the specified field name. As indicated by Table 1, although most of the images have the field "Laterality" in their headers, they all have empty values except one image. In addition, for the other two fields, over one third of the images have no values. Table 2 and Table 3 list all the values (case insensitive) extracted from the field "ViewPosition" and the field "PatientOrientation", and the corresponding number of images. The value of the single image with a non-empty "Laterality" field is "L". We then inspected each image with a non-empty field value to determine the actual image view. For each field value, the number of images that we determined are frontal and the number of images that are lateral are listed in Tables 2 and 3. They show that the values of "LL", "RL", "lateral", and "large lateral" for field "ViewPosition" and the values of "A\F" and "P\F" for field "PatientOrientation" indicate that the images are lateral (the one frontal case for the "LL" and "P\F" values may be due to human error when generating the DICOM image). Similarly, the value of "R\F" for field "PatientOrientation" indicates that the images are frontal. However, for the images that have values of "PA" or "AP" for field "ViewPosition" or value of "L\F" for field "PatientOrientation", there are not only frontal view images but also lateral view images. So, the information about whether an image is frontal or lateral is not always available in the text (i.e., the DICOM header or radiology report).

Table 1. Image header information on view

Field name	No. of images having the field	No. of images having values for the field
ViewPosition	8231	5313
PatientOrientation	8150	4850
Laterality	7079	1

Table 2. Values for field "ViewPosition"

Values	No. of images	No. of frontal	No. of lateral
PA	3871	2353	1518
AP	469	457	12
LL	804	1	803
RL	34	0	34
lateral	133	0	133
large lateral	2	0	2

Table 3. Values for field "PatientOrientation"

Values	No. of images	No. of frontal	No. of lateral
L\F	4064	2537	1527
A\F	556	0	556
P\F	216	1	215
R\F	14	14	0

III. METHOD

A. Pre-processing

The DICOM images from Indiana University comprise a heterogeneous set of x-ray images that were captured by several x-ray technicians with different x-ray machines. For some of these images, radiologists have manually optimized the intensity window to visually enhance the lung tissue region. This windowing information is available in the DICOM header. We used this information, when it was available, in the three pre-processing steps described below:

First, if windowing information is not available, we compute the minimum and maximum pixel values of the raw input x-ray, $min(I)$ and $max(I)$, as given in Hounsfield units. If windowing information is available, we compute $min(I)$ and $max(I)$ as follows:

$$\min(I) = wc - \frac{ww}{2}; \max(I) = wc + \frac{ww}{2} \qquad (1)$$

where *wc* is the window center and *ww* is the window width. All pixel values smaller than $min(I)$ will be displayed as black, and all pixel values larger than $max(I)$ will be displayed as white.

Second, we linearly scale all pixels values between $min(I)$ and $max(I)$ so that they fit the interval between zero and one. We determine the slope and intercept of this linear mapping by solving the following linear equation:

$$\begin{pmatrix} \min(I) & 1 \\ \max(I) & 1 \end{pmatrix} \begin{pmatrix} slope \\ intercept \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad (2)$$

We then compute the new pixel values $I_{new}(x, y)$ from the old values $I_{old}(x, y)$ by applying the following transformation:

$$I_{new}(x, y) = slope * I_{old}(x, y) + intercept \qquad (3)$$

Third, we check the photometric interpretation field in the DICOM header to determine whether the image intensities need to be inverted, with $I_{new} = 1 - I_{old}$.

The images are then contrast-enhanced because the contrast may be low in some images as shown in Figure 2(a). This is done by mapping the intensities to new values such that 1% of the pixels' intensities are saturated at the lowest and highest intensities of the image. Sometimes, there is a large background region surrounding the chest region in the image, as in the examples shown in Figure 2(a). Since this background region provides no meaningful information and may worsen the classification accuracy as the features are extracted from the entire image, it is removed next. The image is binarized by thresholding and then is cropped using the bounding box of the white (foreground) region. The threshold is set as the median value of the intensities of all the pixels in the image. Figure 2(b) is the result after contrast enhancement and background border removal for the image in Figure 2(a). The average size

of the original image is over 2000 pixels in each dimension. In order to increase the computing efficiency, the images are resized to a quarter of their original size while keeping the width/height ratio unchanged.
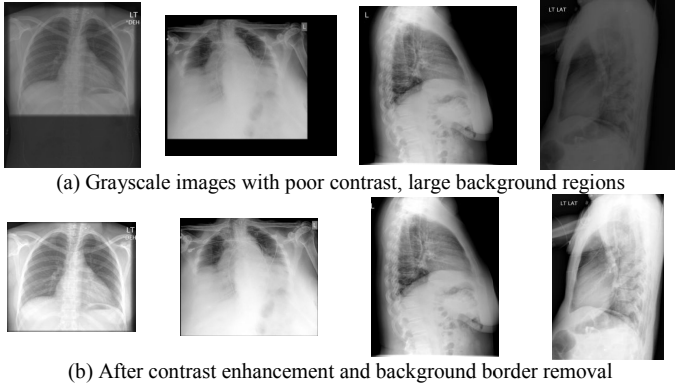

(a) Grayscale images with poor contrast, large background regions

(b) After contrast enhancement and background border removal

Fig. 2. Images before and after pre-processing

B. Features

After pre-processing, the following features are extracted.

1) Image profile

The image profile feature is obtained by separately projecting the intensity of the grayscale image in the vertical and horizontal directions. In order to obtain a feature vector of uniform length, the images are resized to N by N since the size of the images in the dataset varies. Therefore, the length of the horizontal intensity projection f_{hor} and the length of the vertical intensity project f_{ver} are both N. Figure 3 shows examples of frontal and lateral Chest X-rays and their corresponding vertical and horizontal projection profiles for N = 200 (this size is determined empirically). The image profile feature aims to represent the different distributions of dark and bright pixels along two directions in frontal and lateral views. For example, as shown in Figure 3, the vertical profile of the frontal image (image 1) often has a high peak (corresponding to the bright spine) between two large valleys (corresponding to the two dark lungs) while that of the lateral image (image 2) often does not exhibit this characteristic.

2) Body size ratio

Based on the observation that the lateral view of the human chest is usually narrower than its frontal view, we propose the *body size ratio* feature, which we define as the ratio between the median length of the horizontal body cross sections and the maximal length of the vertical body cross sections. We believe this is a more discriminative feature than the *image size ratio* (the ratio of image width to height), since there are some lateral images in which the body does not stand straight up or the arms are parallel to the ground and are partially included in the image. Figure 4 shows two such examples (Images 2 and 3). The red lines indicate locations of the cross sections that are used to calculate the body size ratio (there might be multiple such cross sections, the ones shown here are the middle ones among all the matching cross

sections). As given in Figure 4, the image size ratios (the second row) for the frontal image and the two lateral images are 1, 0.95, and 1, respectively, while the body size ratios (the third row) are 0.91, 0.57 and 0.61, respectively. This indicates that, as previously noted, body size ratio is a more discriminative feature compared to image size ratio. (We extract the body region as the largest connected component in the binarized image obtained in the step of background border removal.)
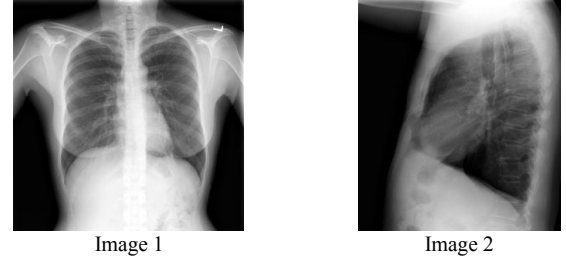

| Image 1 | Image 2 |

Fig. 3. Image profiles for two different image views, showing different "signatures" both for horizontal and for vertical

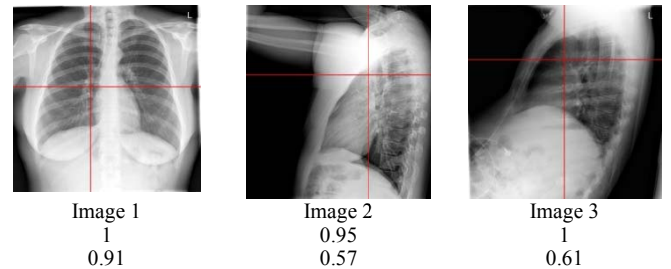| Image 1 | Image 2 | Image 3 |
| 1 | 0.95 | 1 |
| 0.91 | 0.57 | 0.61 |

Fig. 4. Body size ratio. Our definition aims to discriminate frontal/lateral view even when arms are included (Image 2) or the body is not vertical (Image 3).

3) PHOG feature

The PHOG (Pyramid of Histograms of Orientation Gradients) feature was proposed by Bosch et. al. [12]. It is mainly inspired by the image pyramid representation of Lazebnik et al. [14] and the HOG (Histograms of Oriented Gradients) descriptor [15]. The PHOG feature represents both local shape and its spatial layout. The extraction of the PHOG feature (as shown in Figure 5) consists of two main steps:

1) Edge contours are extracted using a Canny edge detector [16], and orientation gradients are computed using Sobel masks.

2) Edge gradients of each contour point are then counted to extract a HOG descriptor from each cell at each level of the pyramid. All HOG descriptors are then concatenated to one descriptor.

We set the PHOG parameters as follows: number of bins (K) is 8, orientation range is [0 360], and pyramid level (L) is 3. So the length of the PHOG feature is $K \sum_{i=0}^{L} 4^i = 680$.
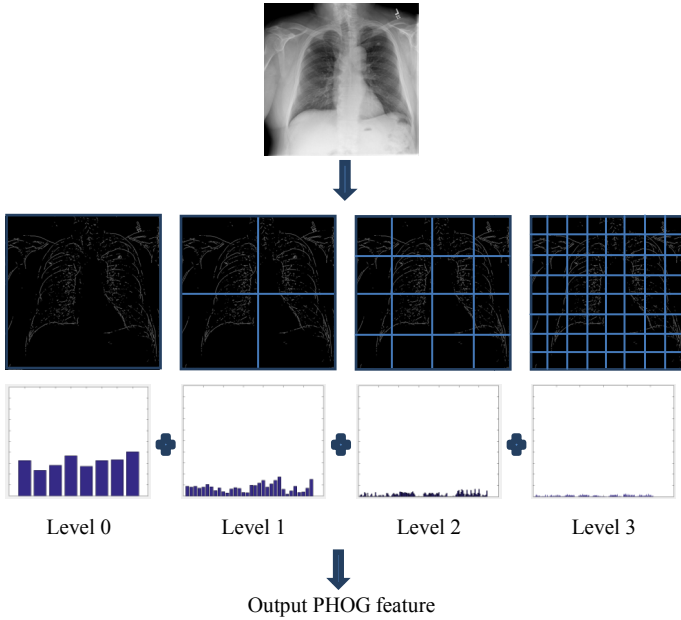

Fig. 5. PHOG feature

4) Contour-based shape feature (CBSF)

In our previous research [17], we developed a contour-based shape descriptor (feature) (CBSF) and applied it to biomedical image modality classification. We apply adaptive binarization followed by morphology operations (dilation and erosion) to an input image and then extract contours of white connected components (blobs). The contour image is then divided into cells (for example, 7x7) and 4-directional chain code features are extracted from the contours (in each cell) to represent the overall shape of the image content. Chest X-rays with different views can be characterized by their different shapes of body (outline) and lung, and hence this descriptor may be a significant encoder of the shape difference that is essential for classification. Figure 6 shows contours in two different views that are extracted from the images and used to extract the shape descriptor.
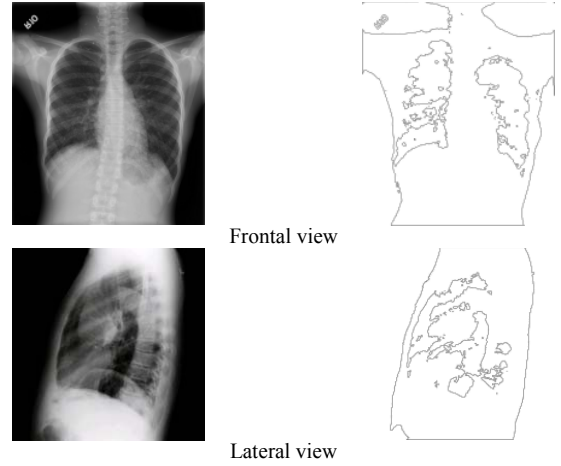

Frontal view

Lateral view

Fig. 6. The contour based shape feature extracted from two different views

We experimentally assessed the effectiveness of each feature, as described below.

C. Classification

As previously noted, we used the support vector machine (SVM) supervised classifier. Specifically, we used the SVM trained using the sequential minimal optimization (SMO) [18] algorithm and implemented in the WEKA software [19]. We used the linear kernel and default values for all parameters.

IV. EXPERIMENTAL TEST

A. Data

We tested our method on the NLM Indiana chest radiographs dataset. We created the ground truth dataset by visually inspecting and classifying each image as frontal or lateral, resulting in 4143 frontal and 4090 lateral image labels. The set has images in various sizes that range from a minimum dimension of 1024 pixels and a maximum dimension of 4248 pixels. We also tested our method on the IRMA chest x-ray dataset [12]. The IRMA dataset that we obtained contains 1266 frontal radiographs and 601 lateral radiographs. The size of frontal radiographs in this dataset ranges between 345 to 512 in width and 357 to 512 in height, while that of lateral radiographs ranges between 296 to 512 in width and 390 to 512 in height.

B. Classification Result

1) NLM Indiana dataset

Table 4 lists the classification accuracies of the NLM Indiana chest radiographs dataset for each feature and overall combined features. It contains the accuracy for 10-fold cross-validation (CV) using the whole dataset as well as the accuracy when the dataset is split into a training set (2/3 of the number of total images) and testing set (the remaining 1/3 of the images). We found that each of the features is effective, especially the PHOG and CBSF. The individual classification accuracy for each feature is about 98.4% (IP), 90.2% (BSR), 99.7% (PHOG), and 99.7% (CBSF) respectively. PHOG and CBSF both obtained the highest accuracy, but the feature dimension of CBSF is much lower than that of PHOG (196 vs 680). It is also noteworthy that, the single-length BSR feature

achieves about 90% classification accuracy by itself alone. The classification accuracy of the combined feature (IP+BSR+PHOG+CBSF) is 99.9%. The confusion matrix (10 fold CV) for the combined feature is shown in Table 5. We also examined the classification performance of the combined feature when adding a step of attribute selection to reduce the feature vector length. We used the meta-classifier "AttributeSelectedClassifier" in WEKA to encapsulate the attribute selection process with the classifier itself. Therefore both the attribute selection method and the classifier only get access to the data in the training set (or folds if cross-validation is performed). For the attribute selection method used in the "AttributeSelectedClassifier", we selected the feature evaluator as "CfsSubsetEval" (which evaluates the value of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them) and the search method as "BestFirst" (which searches the space of feature subsets by greedy hill-climbing augmented with a backtracking facility). For the classifier used in the "AttributeSelectedClassifier", we selected SMO. We performed 10-fold CV. The attribute selection process reduced the feature length 1277 to 149, yet the classification accuracy was still 99.9%.

Table 4. Classification results on NLM data

Feature	Dimension	Accuracy	
		10-fold CV	Testing set
Image profile (IP)	400	98.4%	98.4%
Body size ratio (BSR)	1	90.5%	90.2%
CBSF	196	99.7%	99.7%
PHOG	680	99.7%	99.6%
IP +BSR +CBSF+PHOG	1277	99.9%	99.9%
IP +CBSF+PHOG	1276	99.9%	99.2%

Table 5. Confusion matrix for the combined feature IP+BSR+CBSF+PHOG. Reducing the feature size to 149 yielded the same classification accuracy.

Classified as →	Frontal	Lateral
Frontal	4141	5
Lateral	4	4083

2) IRMA dataset

Table 6 lists the classification accuracies of the IRMA chest radiographs dataset for each feature and overall combined features. For CBSF, the number of cells is set as 3 x 3 for the IRMA dataset (7 x 7 for the NLM Indiana dataset), due to the small image size of the IRMA dataset compared to the NLM Indiana dataset. The classification accuracy of the combined features is 99.9% (the best correctness reported in [7] is 99.7%). The performance of each feature is similar for both datasets except the BSR. The unsatisfactory performance of BSR for the IRMA dataset may be attributed to the bright peripheral bands commonly seen in the IRMA lateral images (as the examples shown in Figure 7) which causes the failure of body segmentation with thresholding. Table 7 shows the confusion matrix (10 fold CV) for the combined feature (IP+CBSF+PHOG).

Table 6. Classification results on IRMA data

Feature	Accuracy
	10-fold CV
Image profile (IP)	98.0%
Body size ratio (BSR)	66.9%
CBSF	99.8%
PHOG	99.6%
IP + BSR + CBSF+PHOG	99.9%
IP + CBSF+PHOG	99.9%

Table 7. Confusion matrix for the combined feature IP+CBSF+PHOG for IRMA data.

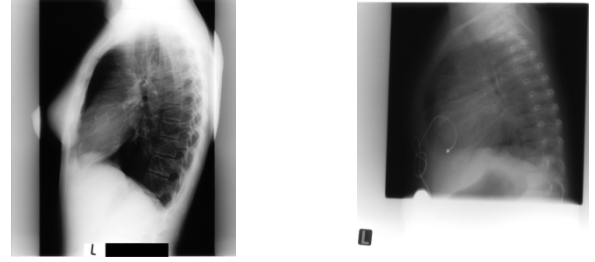| Classified as →	Frontal	Lateral
Frontal	1266	0
Lateral	1	600

Fig. 7. Lateral chest radiographs in IRMA dataset.

V. CONCLUSION

We proposed a new method for classification of frontal and lateral chest X-rays. We believe that this is one important initial step in a chest radiograph CAD system. We extracted several effective features consisting of image profile, body size ratio, and our newly developed contour-based shape descriptor and used an SVM classifier to train and test our algorithm.. The proposed method was evaluated on a dataset of over 8000 chest radiographs and achieved very high classification accuracy (above 99%).

This frontal/lateral CXR classifier will be included in our CAD system for screening tuberculosis disease in resource-poor areas. We have been developing algorithms for segmenting the lung region [1] and extracting color, texture and shape features within the lung region for classifying the images into normal and abnormal (for any lung disease) classes [20]. These algorithms have been integrated into a system running on a portable computer (Mac Mini) connected to the x-ray machine workstation; our program listens to the workstation and automatically receives and processes (segments and classifies) x-ray DICOM images as they are acquired from patients. The frontal/lateral view classification module will be added in front of the lung segmentation module.

Our frontal/lateral classification algorithm can also be applied to regular non-DICOM images. We plan to add it to modality filters used in journal article figure searching in our OpenI system. As Figure 8 shows, sometimes the journal text information relating to the figure (such as caption and mention) does not indicate if the figure is frontal or lateral, a situation where the proposed method in this paper can help.
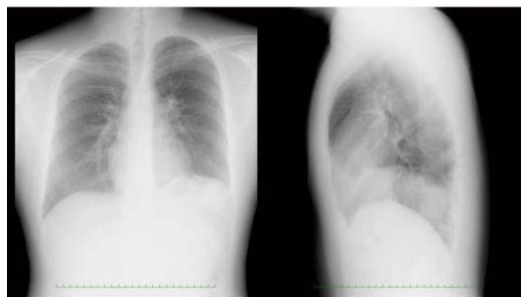
Figure Caption: Preoperative chest radiograph. The chest radiograph shows elevation of the left diaphragm. In this case, the lateral chest radiography was important for the detection of an abnormality in the thoracic cavity.

Fig. 8. Figure of CXR from journal article in OpenI system. Frontal/lateral identifiers may not be given in the article text.

REFERENCES

[1] S. Candemir, S. Jaeger, K. Palaniappan, J.P. Musco, R.K. Singh, Z. Xue, A. Karargyris, S. Antani, G. Thoma, C. J. McDonald, "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," IEEE Transactions on Medical Imaging, vol.33, no.2, pp.577-590, February 2014

[2] G. Coppini, M. Miniati, M. Paterni, S. Monti, E. M. Ferdeghini, Computer-aided diagnosis of emphysema in COPD patients: Neural-network-based analysis of lung shape in digital chest radiographs, Medical Engineering & Physics, Volume 29, Issue 1, January 2007, Pages 76-86.

[3] F.M. Carrascal, J.M. Carreira, M. Souto, P.G. Tahoces, L. Gómez, J.J. Vidal, "Automatic calculation of total lung capacity from automatically traced lung boundaries in postero-anterior and lateral digital chest radiographs" Med Phys. 1998 Jul;25(7 Pt 1):1118-31.

[4] E. Pieka and H. K. Huang, "Orientation correction for chest images," J. Dig. Imag., vol. 5, no. 3, pp. 185–189, 1992.

[5] J. M. Boone, S. Seshagiri, and R. M. Steiner, "Recognition of chest radiograph orientation for picture archiving and communication systems display using neural networks," J. Dig. Imag., vol. 5, no. 3, pp. 190–193, 1992.

[6] H. Arimura, S. Katsuragawa, Q. Li, T. Isguda, and K. Doi, "Development of a computerized method for identifying the postero-anterior and lateral views of chest radiographs by use of a template matching technique," Med. Phys., vol. 29, no. 7, pp. 308–315, 2002.

[7] T. M. Lehmann, O. Guild, D. Keysers, H. Schubert, M. Kohnen, and B. B. Wein, "Determining the view of chest radiographs," J. Dig. Imag., vol. 16, no. 3, pp. 281–291, 2003.

[8] E. F. Kao, C. Lee, T. S. Jaw, J. S. Hsu and G. C. Liu, "Projection profile analysis for identifying different views of chest radiographs," Acad. Radiol., vol. 13, pp. 518–525, 2006.

[9] E. F. Kao, W.C. Lin, J. S. Hsu, M.C. Chou, T. S. Jaw, G. C. Liu, "A computerized method for automated identification of erect posteroanterior and supine anteroposterior chest radiographs," Phys Med Biol., vol. 56, no. 24, pp.7737-7753, 2011.

[10] H. Luo, W. Hao, D. H. Foos and C. W. Cornelius, "Automatic image hanging protocol for chest radiographs in PACS," IEEE Trans. Inf. Technol. Biomed., vol. 10, pp. 302–311, 2006.

[11] A. Bosch, A. Zisserman, X. Muñoz, "Representing shape with a spatial pyramid kernel", International Conference on Image and Video Retrieval. Amsterdam, The Netherlands, July 2007.

[12] http://ganymed.imib.rwth-aachen.de/irma/datasets.php?SELECTED=00003#00003.dataset

[13] D. Demner-Fushman, S.K. Antani, M. Simpson M, G.R. Thoma, "Design and development of a multimodal biomedical information retrieval system", JCSE, vol. 6, no.2, pp.168-177, June 2012.

[14] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: spatial pyramid matching for recognizing natural scene categories," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 2169-2178, 2006.

[15] N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), vol.1, pp.886-893, 2005.

[16] J. Canny, "A computational approach to edge detection", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 8, no. 6, pp. 679–698, 1986.

[17] D. You, S.K. Antani, D. Demner-Fushman, G.R. Thoma, "A contour-based shape descriptor for biomedical image classification and retrieval", Proceedings of SPIE 2014, vol. 9021, Document Recognition and Retrieval XXI, February 2014.

[18] J. C. Platt, "Sequential minimal optimization: a fast algorithm for training support vector machines", Microsoft Research, 1998.

[19] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I. H. Witten, "The WEKA data mining software: an update," SIGKDD Explorations, vol. 11, no. 1, pp. 10-18, November 2009.

[20] S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z. Xue, K. Palaniappan, R. Singh, S. Antani, "Automatic Tuberculosis Screening Using Chest Radiographs", IEEE Transactions on Medical Imaging, vol. 33, no. 2, pp. 233-245, 2014.