

# Foreign Object Detection in Chest X-rays

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

**Abstract**—Automatic analysis of chest X-ray images is one important approach for screening/identifying pulmonary diseases. The existence of foreign objects in the images hinders the performance of such processing. In this paper, we focus on one type of foreign objects that is often shown in the images of a large dataset of chest X-rays we are working on—the buttons on the gown that the patient is wearing. The method we propose involves four major steps: intensity normalization, low contrast image identification and enhancement, segmentation of lung regions, and button object extraction. Based on the characteristics of the button objects, we applied two methods for the step of button object extraction. One was based on the circular Hough transform; the other was based on the Viola-Jones algorithm. We tested and compared both methods using a ground truth dataset containing 505 button objects. The results demonstrate the effectiveness of the proposed method.

**Keywords**—chest X-ray; object detection; computer aided diagnosis; lung segmentation

## I. INTRODUCTION

The National Library of Medicine (NLM) has been maintaining a large dataset of chest radiographs as well as related radiological reports. This dataset, collated by the medical school at the University of Indiana, contains about 4000 pairs (frontal and lateral) of chest X-ray DICOM images and 4000 corresponding textual reports [1]. It has been integrated into *Open-i*, a multi-modal biomedical literature retrieval system developed by NLM [2]. This dataset is also valuable for the development/evaluation of computer-aided diagnosis (CAD) algorithms for pulmonary diseases. In addition, we have been developing a tuberculosis (TB) screening system [3]. This TB screening system aims to screen patients who live in resource limited areas by automatically identifying chest X-rays showing signs of TB. The algorithm of TB detection consists of three major steps: 1) identify the lung regions; 2) extract features (texture and shape descriptors) from the area within the lungs; and 3) classify the image into one of two classes (TB vs normal). We have tested the performance of the system on three other TB datasets. Although there are few TB cases in the NLM Indiana chest X-ray dataset, the dataset is useful for chest X-ray CAD work due to its large quantity of data. For example, we evaluated our new method for classifying the views (frontal view vs the lateral view) of the chest X-ray images using this dataset and reported the results in our previous study [4].

Some chest X-rays in the NLM Indiana dataset have foreign objects. The most frequently appeared foreign objects

shown in the lung area are buttons on the gown that the patients were wearing. Figure 1 shows a couple of example images containing several buttons. Figure 1 also shows the close-up of the button areas in the images. The existence of these button objects (especially those of them that are located within lung regions) may impede the CAD system performance as they are foreign objects and therefore should not be considered. In this paper, we applied and compared two methods for extracting this type of foreign objects. Based on our best knowledge, we are the first one to address this type of foreign objects on chest X-ray image analysis although it may be common to have buttons (from patient gowns) shown on the chest X-rays taken in the hospitals.

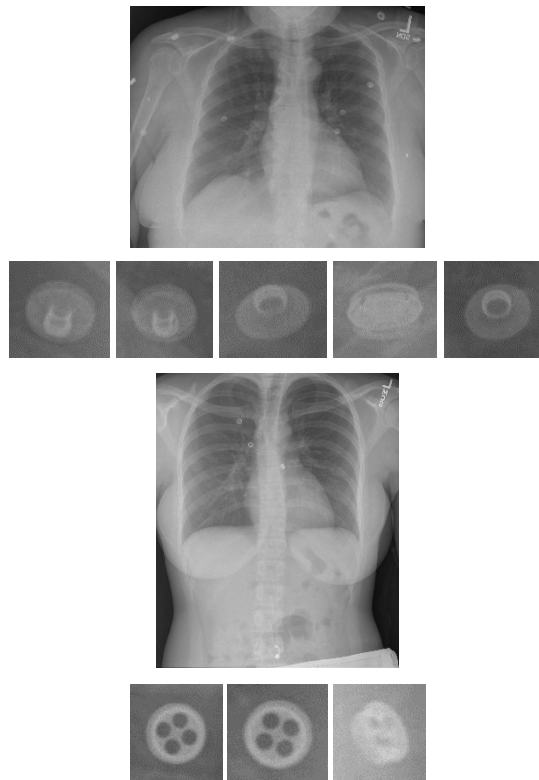


Fig. 1. Button objects in the chest X-rays.

## II. METHOD

### A. Pre-processing

- Image intensity normalization

We first converted the intensity of DICOM images from Hounsfield units to the range of [0, 1]. 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 manually optimized the intensity window to visually enhance the lung tissue region. This windowing information may be available in the DICOM header. If the windowing information was not available, the minimum and maximum intensity values of the input X-ray,  $min_I$  and  $max_I$ , were computed from all the pixels (in Hounsfield units). If the windowing information was available, the values of  $min_I$  and  $max_I$  became

$$min_I = I_{wc} - \frac{I_{ww}}{2}; max_I = I_{wc} + \frac{I_{ww}}{2} \quad (1)$$

where  $I_{wc}$  is the value of the window center and  $I_{ww}$  is the value of the window width. All pixel values smaller than  $min_I$  became  $min_I$ , and all pixel values larger than  $max_I$  became  $max_I$ . After deciding the  $min_I$  and  $max_I$ , all pixels values between  $min_I$  and  $max_I$  were linearly scaled to the interval between 0 and 1. Lastly, the photometric interpretation field in the DICOM header was checked to determine whether the image intensities needed to be inverted.

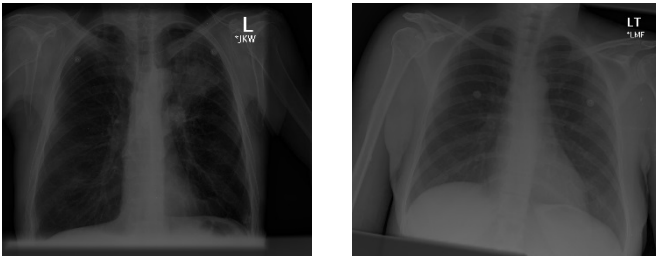


Fig. 2. Low contrast chest X-rays.

- Low contrast image identification and enhancement

Some images among those having no windowing information in the DICOM header had low contrast after the step of image intensity normalization. Figure 2 shows two such low contrast images. To identify these images, we calculated the mean value and standard deviation value of the image intensities. If they were both less than a certain threshold, as in (2), then the image was considered to have low contrast. This criterion was based on the observation that low contrast images usually had low mean intensity value and the intensity values were clustered closely around the mean as indicated by the histograms. Figure 3 shows the histograms of the corresponding images in Figure 2. The threshold values ( $T_1$  and  $T_2$ ) were selected heuristically based on experimental testing.

$$mean < T_1 \text{ and } std < T_2 \quad (2)$$

After extracting the low contrast images, we enhanced them. This was done by mapping the intensities to new values such that 1% of the pixels' intensities were saturated at the lowest and highest intensities of the image. Figure 4 shows the results after contrast enhancement for the images in Figure 2. The step of contrast enhancement improved the performance of

both subsequent steps: lung segmentation and button extraction.

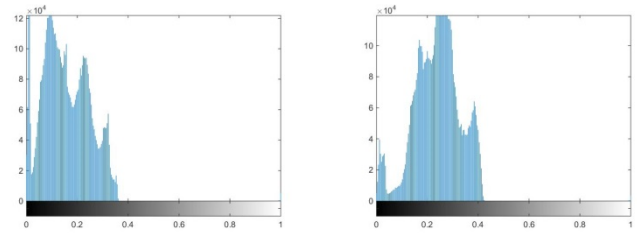


Fig. 3. Histograms of the images in Fig. 2.

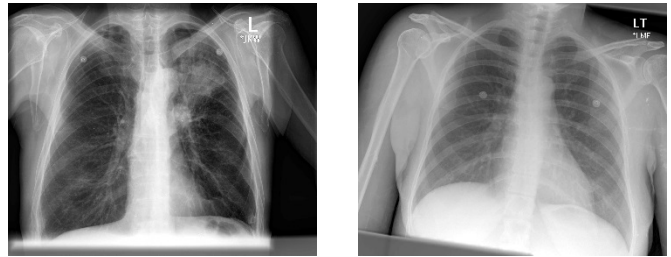


Fig. 4. Images after contrast enhancement.

### B. Lung segmentation

We limited the search space within the region of the lungs, as our goal for identifying foreign objects in a chest X-ray was to improve the performance of automatic analysis of pulmonary diseases. Limiting the search space to the lung or chest area also improved the performance of button extraction as, generally speaking, the shape of buttons in this area appears more regular than that of buttons near shoulders; the front surface (having a circular-like shape) of the buttons in the chest region faces the X-ray direction. The next step was lung segmentation. To locate the lung regions, we used our atlas-based lung segmentation algorithm [5]. The system employed lung models which were expert delineated boundaries of lung on chest X-rays of other patients. We used the JSRT set [6], which is a publicly available set with reference boundaries [7]. When a patient's X-ray was presented, the system first found the most similar X-rays in the model set to the patient's X-ray. The rough shape of the lungs was modeled with intensity histograms in the horizontal and vertical directions. The similarity of X-rays was measured by comparing the histograms with a similarity measure. After the model selection, the system computed the correspondence map between the model X-ray and patient X-ray. Correspondence map computation was conducted by modeling the patient X-ray with local image features, and matching the most similar locations. For the correspondence map computation, we employed the SIFT-flow algorithm [8]. The computed correspondence map was considered as a transformation mapping from model X-rays to the patient X-ray, and applied to the model masks to transform them to the approximate lung model for the patient X-ray. The system then combined CXR intensity values and the lung model with an objective function to decide the final boundary. The system solved the objective

function with a graph-cut energy minimization approach [9]. Figure 5 shows two examples of the lung segmentation results. The boundary of lung regions is marked in red.

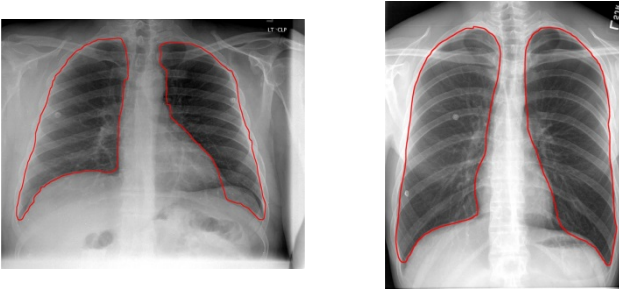


Fig. 5. Lung segmentation results.

### C. Button region extraction

In general, the buttons shown on chest X-rays have a circle-like/eclipse-like/annulus-like shape. Compared to nodules, they have relatively sharper/clearer boundaries and their shapes are more regular. Based on these observations, we applied two methods for extracting button objects within the lung areas. One is based on the Circle Hough Transform. The other is based on the Viola-Jones algorithm.

#### 1) Method based on the Circle Hough Transform

The Circle Hough Transform (CHT) is a technique used for identifying circular-like objects in images. It has been applied to many image processing applications. The conventional CHT algorithm can be summarized by the following steps. First extract edges from the image. This is usually done by applying Sobel or Canny detector. Then for each edge point, draw a circle of radius  $r$ , with the center being the edge point, onto an accumulator space. A peak value in the accumulator space will be at or close to the location of the center of a circular object of radius  $r$  in the image. There are many variations of the conventional algorithm of CHT aiming to improve performance and reduce computation and storage cost. For a recent survey on the theory, techniques, and applications of Hough transforms, please refer to [10]. For our application, we applied the size invariant circle detection method presented in [11]. This method uses a combination of several modifications to the CHT which includes the use of edge orientation, simultaneous consideration of a range of circle radii, and the use of a complex accumulator array with the phase proportional to the log of radius. It shows that a specific combination of these modifications to the CHT is equivalent to a scale invariant kernel operator and that the method is robust at low signal-to-noise ratio [11].

#### 2) Method based on the Viola-Jones algorithm

The Viola-Jones algorithm [12] is a well-known method for detecting faces. It can also be trained to detect various other types of objects. The Viola-Jones algorithm consists of four major components:

- A set of rectangle features which are reminiscent of *Haar* basis functions and overcomplete;
- The integral image which is an image representation that can make the computation of rectangle features very fast;

- A classifier built on *AdaBoost* learning algorithm which can also be used as a feature selector;
- The combination of classifiers in a cascade architecture which allows background regions to be discarded quickly [12].

After training the Viola-Jones detector with positive and negative samples of an object, given a test image, the detector scans across the image at multiple scales and locations to find the sub-windows that contain the object. The detector also combines the overlapping multiple detections of one object into a single detection of the object in the post-processing stage. Although the training of Viola-Jones detector may be slow, the detection is very fast and can be used for real-time analysis.

The size invariant circle detector, which is based on the CHT, outputs boundaries of button objects, while the Viola-Jones object detector outputs image sub-windows that contain button objects. In addition, the Viola-Jones object detector is a classification method which requires training data, while the size invariant circle detector does not.

## III. EXPERIMENTAL RESULTS AND DISCUSSION

We extracted 278 chest X-ray DICOM images from the NLM Indiana dataset which contain button objects in lung areas using visual inspection. All of the selected DICOM images were first converted to PNG images using image intensity normalization described in the pre-processing step. We then generated ground truth data by using the image labeling tool provided in Matlab (Training Image Labeler). The regions of button objects within lung regions or on the boundary of lung regions were marked by drawing a rectangle enclosing each button object. The rules for marking were: 1) the box should contain the entire button object; 2) the object should be located approximately at the center of the box; 3) there should be a gap between the boundary of the button object and the boundary of the box. Figure 6 gives one example of the marked images (the button objects are marked by blue rectangles). The total number of marked button objects was 505.

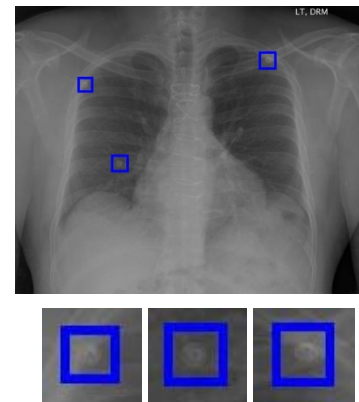


Fig. 6. Examples of ground truth markings.

Among the 278 images, we identified 79 low contrast images using the method described in Section II.A. The threshold values in (2) were set as  $T_1 = 0.4$  and  $T_2 = 0.2$  and



were tested on the entire NLM Indiana chest X-ray dataset. These 80 low contrast images were then enhanced. After pre-processing, the lung regions were then extracted. Because the area of the lungs was the search space for the objects, the performance of lung segmentation affected the final performance of the button detector. This is discussed in later paragraphs.

For circular object detection, we applied two methods. One was the size invariant circle detector [11]. The other was the Viola-Jones object detector [12]. Both methods were implemented and provided by the toolboxes of Matlab [13, 14].

We applied the circle detector to all 278 images and obtained the boundaries of all the identified button objects within the automatically segmented lung regions. To evaluate the performance of the circle detector, we used the following criterion: if the bounding box of the extracted boundary of the button object was included inside the corresponding ground truth rectangle, then it was a true positive; otherwise, it was considered as a false positive. There are several important parameters for the method [13]. One is the edge gradient threshold which is used to determine the edge pixels from the gradient image obtained by applying Sobel operator to the image. Generally speaking, more circular objects with weak edges would be detected with lower value of this threshold. Another important parameter is the range of the radii of the circles the algorithm is going to search for. It was decided based on the size of the buttons with respect to the size of images. The images were downscaled to  $\frac{1}{4}$  of their original size (the average original size was around 2500 by 2500 pixels). The third important parameter is the sensitivity for the accumulator array. High sensitivity value extract more circular objects but may have had more false detections. We set the value of edge gradient threshold, radii range, and the sensitivity as 0.02, [5-20] and 0.8, respectively. Table I lists the evaluation results of the circle detector. The number of button objects extracted by the circle detector was 518. Among them, 430 were true positive and 88 were false positive. The number of false negative was 75. Therefore, the precision was 0.83 and the recall was 0.85. Figure 7 shows several examples of the results obtained from the size invariant circle detector. The extracted lung region, the button boundaries output by the circular detector, and the ground truth marking are denoted by red, green, and blue color, respectively. Besides those that are located outside of the extracted lung area, most of the false negatives are located on the lung boundary as pointed by arrows in Figure 8. The reason they were missed may be because their partial contour is not significant enough with respect to either the degree of being circular or the extent of edge gradient magnitude. Examples of false positives are given in Figure 9. There are basically three types of false positives. One is due to the unsatisfactory lung boundary extraction which may have included button objects that are not labeled in the ground truth (Figure 9(a)). The second type is the end components of catheters or tubes (Figure 9(c) and 9(d)). They were extracted because of their roundish shape. Although these end components are considered as being unwanted for button object extraction, it is beneficial to extract them as they are also foreign objects which are often shown in the Indiana chest X-

ray dataset. The third type may be vessels or nodules that have relatively stronger edges (Figure 9(b)).

TABLE I. RESULTS OF CIRCLE DETECTOR

Ground Truth	Extracted	TP	FP	FN	Precision	Recall
505	518	430	88	75	0.83	0.85

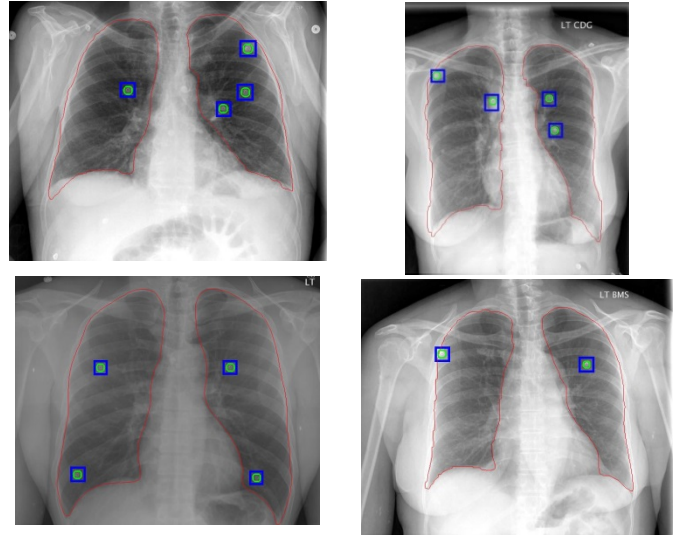


Fig. 7. Results of button extraction based on CHT.

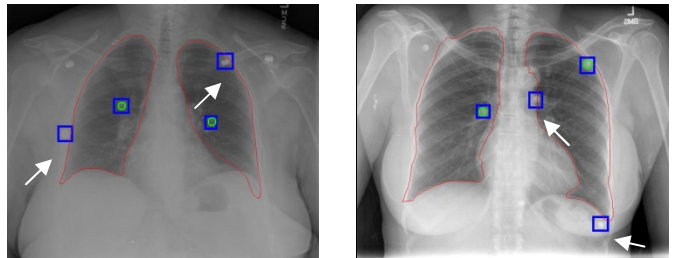


Fig. 8. False negatives of circular detector.

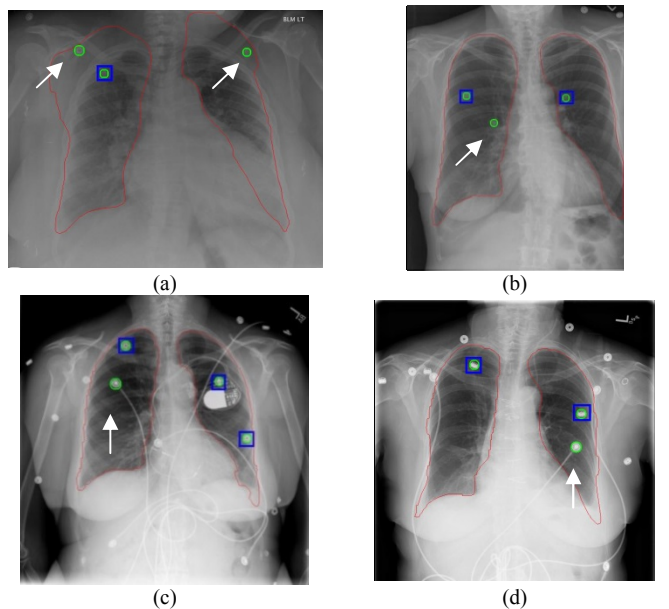


Fig. 9. False positives of circular detector.

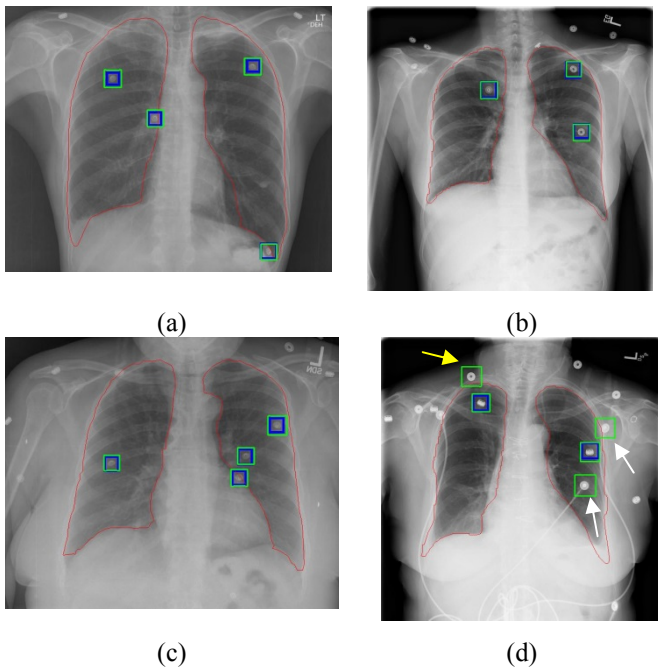


Fig. 10. Results of button extraction based on Viola-Jones algorithm.

For the Viola-Jones object detector, we used the cropped rectangular regions of 360 button objects from 200 images of the 278 ground truth images as the positive training samples. We used 2096 images that do not contain any button object as the negative training images. The Viola-Jones object detector provided by Matlab automatically generates negative training samples from the negative images by using sliding windows. There are several important parameters for the Viola-Jones object detector: the number of cascade stages; the object size for training; the false alarm rate; the true positive rate; and the feature type [14]. Besides the originally used *Haar*-like features, the LBP (Local Binary Patterns) features [15] and the HOG (Histograms of Oriented Gradients) features are two other commonly used features for the Viola-Jones object detector. For the definition and trade-off of the rest of the parameters, please refer to the Mathworks website [14]. We used images of original size for both training and testing. We set the object training size as [100, 100] which is the average size of the cropped ground truth rectangular regions, the false alarm rate as 0.2, the number of cascade stages as 8, and used the default value for the true positive rate which is 0.995. After training the object detector, we used the remaining 78 ground truth images (unused in training), which contained 145 button objects, as the testing images. The Viola-Jones object detector scans a whole testing image and outputs boundary boxes of identified objects in it. If the entire boundary box is located outside the extracted lung region, this boundary box is removed. To evaluate the performance of the Viola-Jones object detector, we applied the following rule: if the center of the extracted object boundary box is inside the boundary of the ground truth box, it is considered as a correct detection. Table II lists the corresponding testing results. The number of extracted objects was 159. 137 of them were true positives and 22 were false positives. The number of false negatives was 8. Therefore, the precision and recall for the Viola-Jones detector was 0.86 and 0.94, respectively. Figure 10 shows several

examples of detection results by the Viola-Jones detector. The object boundary extracted by the detector is marked with green color (the ground truth marking is shown in blue color). All the button objects in Figure 10(a)-(c) are correctly detected. Figure 10(d), shows that there are three false positives. Two of them are the end components of medical tubes (pointed to by white arrows). The third one is indeed a button object (pointed to by a yellow arrow). It is located outside the boundary of the lungs so it is not included in the ground truth labeling, but the algorithm keeps it because its extracted box intersects with the extracted lung boundary. To compare the performance of the Viola-Jones detector with that of the circle detector, we measured the performance of the circle detector on the images in the testing set only. The results are given in Table III which indicates that the Viola-Jones detector performs better for this dataset. During our parameter tuning, we found that the circle detector also seems to be more sensitive to the value of parameters.

TABLE II. RESULTS OF VIOLA-JONES DETECTOR

Ground Truth	Extracted	TP	FP	FN	Precision	Recall
145	159	137	22	8	0.86	0.94

TABLE III. RESULTS OF CIRCLE DETECTOR (FOR THE TESTING SET USED BY VIOLA-JONES DETECTOR)

Ground Truth	Extracted	TP	FP	FN	Precision	Recall
145	152	127	25	18	0.84	0.88

#### IV. CONCLUSIONS AND FUTURE WORK

In this paper, we focused on identifying one common type of foreign object shown in chest X-rays—buttons on gowns. The method consists of four major steps: image intensity normalization, low contrast image identification and enhancement, lung region segmentation, and button objects extraction. We applied two methods for the step of button objects extraction. One is based on the circular Hough transform, the other is the Viola-Jones object detector. The method was tested using a dataset with 278 images and labeled with 505 button objects. For the circular Hough transform-based detector, we obtained 0.83 for precision and 0.85 for recall when applying the method to the whole dataset. For the Viola-Jones object detector, we used 360 objects from 200 images as the training set and 145 objects from 78 images as the testing set. We obtained 0.86 for precision and 0.94 for recall.

In addition to buttons, another type of foreign object that often appears in chest X-rays is tubes. These tubes include nasogastric tubes, endotracheal tubes, central venous catheters, and intercostal tubes. Their appearance in an image also impairs the performance of the computer-aided diagnosis algorithm. Two examples of images having tubes are shown in Figure 11. As mentioned in Section III, these tubes usually have roundish end components which can be detected by using the same detector for buttons. In the future, we will work on extracting the line segment of tubes.

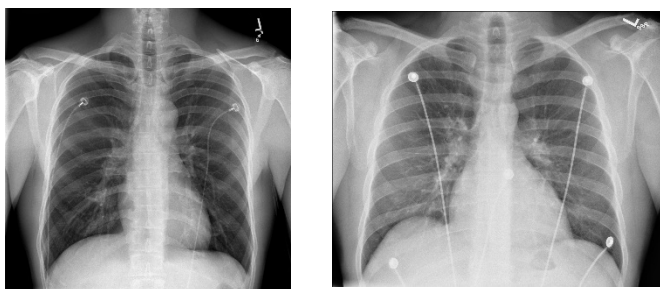


Fig. 11. Images having tubes.

#### ACKNOWLEDGMENT

This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM), and the Lister Hill National Center for Biomedical Communications (LHNCBC).

#### REFERENCES

- [1] D. Demner-Fushman, M. D. Kohli, M. B. Rosenman, S. E. Shooshan, L. Rodriguez, S. Antani, G. R. Thoma, C. J. McDonald, "Preparing a collection of radiology examinations for distribution and retrieval," *Journal of the American Medical Informatics Association*, July 2015.
- [2] D. Demner-Fushman, S. Antani, M. Simpson, G.R. Thoma, "Design and development of a multimodal biomedical information retrieval system," *Journal of Computing Science and Engineering*, vol. 6, no. 2, pp. 168-177, 2012.
- [3] 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.
- [4] Z. Xue, D. You, S. Candemir, S. Jaeger, S. Antani, L. R. Long, G. R. Thoma, "Chest X-ray image view classification," *Proceedings of the 28<sup>th</sup> IEEE International Symposium on Computer-Based Medical Systems*, June 2015.
- [5] S. Candemir, S. Jaeger, K. Palaniappan, J.P. Musco, R.K. Singh, Z. Xue, A. Karargyris, S. Antani, G. R. 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.
- [6] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K. Komatsu, M. Matsui, H. Fujita, Y. Kodera, K. Doi, "Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules," *American Journal of Roentgenology*, vol. 174, no. 1, pp 71-74, 2000.
- [7] B. Van Ginneken, M. B. Stegmann, M. Loog, "Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database," *Medical Image Analysis*, vol. 10, no.1, pp. 19-40, 2006.
- [8] C. Liu, J. Yuen, and A. Torralba, "Sift flow: Dense correspondence across scenes and its applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no.5, pp. 978-994, 2011.
- [9] Y. Boykov, O. Veksler, R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no.11, pp.1222-1239, 2001.
- [10] A. S. Hassanein, S. Mohammad, M. Sameer, M. E. Ragab, "A survey on Hough transform, theory, techniques and applications," *International Journal of Computer Science Issues*, IJCSI vol. 12, no. 1, January 2015.
- [11] T. J. Atherton and D. J. Kerbyson, "Size invariant circle detection," *Image and Vision Computing*, vol. 17, no. 11, pp. 795-803, 1999.
- [12] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Conference on Computer Vision and Pattern Recognition*, 2001.
- [13] <http://www.mathworks.com/help/images/ref/imfindcircles.html>
- [14] <http://www.mathworks.com/help/vision/ref/traincascadeobjectdetector.html>
- [15] N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol., pp. 886-893, 2005.
- [16] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, July 2002.