Automatic Rapid Segmentation of Human Lung from 2D Chest X-Ray Images

Zhennan Yan, Jing Zhang, Shaoting Zhang, and Dimitris N. Metaxas

Department of Computer Science, Rutgers University, USA
{zhennany,zhangj81,shaoting,dnm}@cs.rutgers.edu

Abstract. In this paper, we propose a complete framework that segments lungs from 2D Chest X-Ray (CXR) images automatically and rapidly. The framework includes two main steps: First, given a set of manually segmented training data, some landmark detectors are obtained using learning techniques. Second, using these detected landmarks as boundary indicators, a statistical shape inference is applied to fit the contours and get a rough segmentation of lung. This framework could also be applied to different applications for other organs or medical imaging modalities. Our major goal is to segment the lung rapidly with a reasonable accuracy. In the experiments, we compared the performance of Boosted Cascade with Haar features and Support Vector Machine (SVM) with Gabor features.

Keywords: Segmentation, Detection, Lung, CXR, Haar, Boosted Cascade, Gabor, SVM

1 Introduction

Accurately segmenting organs from medical images is a fundamental yet challenging task. Good segmentation results can be used for clinical or research analysis. For example, computer aided diagnostic (CAD) programs usually need the segmentation result as input to analyze potential diseases. Manual segmentation has been shown to be sufficiently robust and accurate, but is labor intensive. Thus, it is practically necessary to develop an automatic and rapid segmentation framework.

In this paper, we study a framework for 2D segmentation of lung from CXR images, and speed-up the whole process while archiving good performance. We use a coarse-to-fine strategy. First, we train some boundary landmark detectors and learn a statistical shape model to provide prior information. Then, based on the detected landmarks and prior shape constraints, use a learning-based statistical model to fit the boundary contours. This framework can be used for creating an atlas of lung, and could also be extended to do whole-body organs segmentation.

This work focus on the following aspects: (1) Implemented a complete framework for auto segmentation of lung from CXR images. (2) Evaluated two methods in the detection: Boosted Cascade [12] with Haar feature [9]; SVM [3] with Gabor feature [14]. (3) Coarse-to-fine strategy for rapid detection. (4) Comparison of different methods regarding to the detection results. (5) Applied learning-based statistical model to infer the contour of lung. (6) Extensively evaluated the segmentation results using sensitivity, specificity and dice similarity coefficient.

2 Related work

Our segmentation framework is relevant to two types of methods: the appearance-based models and the shape-based models. Appearance-based models use low-level intensity information, like brightness gradients and texture gradients. For example, Snakes [7], Gradient Vector Flow [13], Metamorphs [6] and Level-set [10] methods. However, images are usually corrupted by several artifacts, such as the image noise, missing or occluded parts, intensity inhomogeneity, which cause misleading cues. The shape-based models incorporate higher-order shape information about the objects. Prior statistical shape knowledge may disambiguate this misleading information in segmentation process. But the shape-based models usually need a good initialization. Such models (e.g., Active Shape Models (ASM) [1], Active Appearance Models (AAM) [2]) have been used in different applications for shape prior. However, due to the less reliable appearance information, the detected landmarks could be misleading for the shape models. Zhang et al. proposed a sparse shape representation [16, 17] in their segmentation framework and achieved promising results since the sparse shape representation can avoid misleading info caused by sparse landmark outliers. A complete segmentation framework has been proposed in [15]. They used 3D data and 3D models for segmenting liver and distal femur from CT and MRI images respectively. Some frameworks similar to ours have been applied to organ segmentation, but for different organs or different imaging modalities. In our study, we aim to speed up the whole framework to segment the lung from 2D CXR images automatically and rapidly while still achieved a reasonable accuracy.

For the object detection, there have been some successful approaches using different features, such as Haar [9], Gabor [14], SIFT [8], etc. Viola,P. and Jones, M. have proposed a boosted cascade algorithm for rapid object detection in [12]. These algorithms have been employed in our framework.

2.1 Chest X-ray imaging

The chest radiograph, or CXR, is a projection radiograph of the chest. CXR imaging is an important way to diagnose pathological changes of the chest, including lung, trachea, bronchia, and other nearby structures. Segmentation of lung in CXR images is the first step in much disease diagnosis. However, a typical problem with projective radiographs is the lack of contrast between overlapping objects, such as bones and soft tissues. This can cause ambiguity about the true location of the object boundaries in the image.

2.2 Gabor feature

Features based on Gabor filter have been widely used in image processing since Gabor kernels are characterized as localized, orientation, and frequency selective. The Gabor wavelet representation of an image not only describes spatial frequency structure in the image but also preserves information about spatial relations. Convolving the lung image with complex Gabor filters with 5 spatial frequency and 8 orientation captures the whole frequency spectrum, both amplitude and phase. Feature extraction algorithm mainly has two steps: feature point localization and feature vector computation.

Feature point localization: Feature points can be located by searching the peaks on the responses of the lung image to Gabor filters with a window size of $W \times W$. Here, we use $W = 9$.

Feature vector generation: Feature vectors are generated from feature points as a composition of Gabor wavelet transform coefficient.

2.3 Haar feature

Stated in [9], there are many extended Haar-like features can be extracted from images. Fig. 1 shows 14 feature prototypes could be used. For a window size of 25×25, there could be more than 157,000 features. These features can be computed rapidly by using some intermediate (auxiliary) images [9, 12]. Such large feature vectors can cause computational complexity during learning process. But this complexity can be avoid by feature selection based methods, such as Adaboost [4].

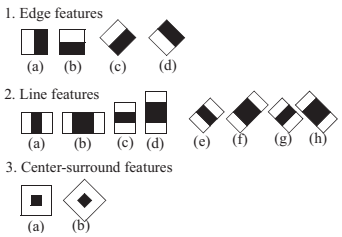


Fig. 1. 14 Haar feature prototypes [9]

2.4 SVM

SVMs are a set of related supervised learning methods widely used for classification and regression. It exhibit good accuracy of classification if the data set is small. By applying kernel function, SVM can efficiently map input data which may not be linearly separable to a high dimensional feature space where linear methods can be used.

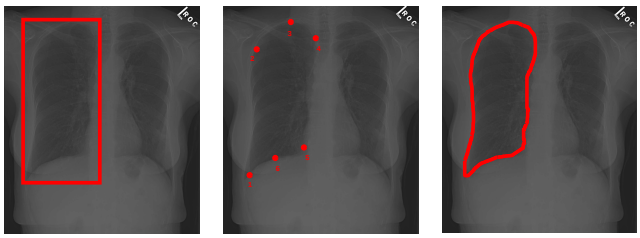
SVM has been widely used in many applications. One of the main reasons is that its decision function only depends on the dot product of the input feature vector with the Support Vectors (SVs). Theoretically, features with any dimension can be fed into SVM for training since it has no requirements on the dimension of the feature vector. But in practice, the huge dimension of Gabor features demands large computation complexity and memory costs, optimized Gabor features [11] are proposed with a general SVM classifier can achieve good performance.

2.5 Boosted Cascade

Given an input image, detectors usually have a large amount of sub-windows to check. This large number of sub-windows could cause unacceptable false positive rate and cost a lot of time. In [12], a cascade of classifiers was proposed to achieve low false positive rates while greatly reducing computation time. The series of classifiers can eliminate a large number of negative sub-windows. After several stages of processing the number of sub-windows have been reduced radically. This can significantly reduce the computation time and improve the performance by lower the false positive rate.

3 Proposed Approach

In this experiment we aim to roughly locate the right lung from CXR, using landmark detections and sparse dictionary learning. CXR imaging is widely used because of the fast speed and low cost. Computer Aided Diagnostic (CAD) programs can detect various pathologies including abnormal cardiac sizes, pneumonia shadow and mass lesions by analyzing X-ray images. To automatically detect such pathology, one needs to locate lungs robustly and efficiently. In our application, the size of input images is about 2500×2500 pixels. The ground truths are binary masks of manual segmentation results, labeled by a clinical expert. From the binary masks, 2D contours can be easily extracted. We then manually select six specific points (e.g., corner points, top and bottom points) on the contour as the anatomical landmarks for training purpose. The six landmarks are defined as fig. 2(b). After that, a fixed amount of points between two neighboring landmarks are evenly and automatically interpolated along the contour. Thus, one-to-one correspondence is obtained for all landmarks and shapes. When a new image comes, we first detect the anatomical landmarks (which may contain errors), and then use shape priors to infer a shape from these detected landmarks.



(a) Lung (b) Landmarks (c) Contour

Fig. 2. Coarse to fine strategy

In the first step, we train one lung detector and six landmark detectors. Because of the large resolution and low contrast of CXR images, direct landmark detection could be time-consuming and have too many false alarms. Here we proposed a coarse-to-fine framework for lung segmentation. Fig. 2 shows the strategy: (a) detect the bounding box of right lung; (b) detect landmarks in the bounding box with some location constraints; (c) infer contour based on landmarks (second step).

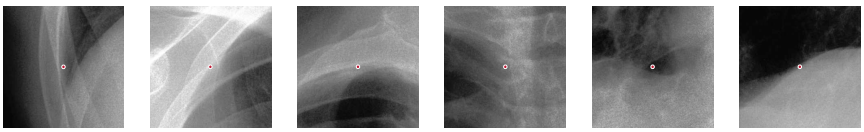
After necessary image processing, we scale the original image to lower resolution (about 40×40) and extract patches of size 27×18 for whole lung as positive samples and negative samples which don't contain the lung or contain a small portion of lung. This set of samples is used for training lung detector. Then scale the original image to about 200×200, and extract patches of size 25×25 for training the 6 landmarks. Positive landmark samples are illustrated in fig. 3. These sets of samples are used for training six landmark detectors.

We apply two different methods to create active models to detect the lung and landmarks: Gabor+SVM and Haar+Boosted Cascade. For Gabor+SVM, gabor features are

extracted as input of SVM classifier according to section 2.2. In experiment, we use polynomial degree-2 as SVM kernel function. For Haar+Boosted Cascade, it is introduced in section 2.3 that the haar feature vector of a sample image usually has very high dimension, so we use Adaboost method in each cascade classifier to do feature selection. This can speed up the detection procedure while maintaining a good performance. Similar methods have been proved very successful in face detection.

The trained detectors can only detect sub-windows that may contain the expected objects. For the landmarks, however, we need six points rather than rectangles. A very simple hypothesis is that picking the central point of the sub-window as the landmark. A more reasonable hypothesis is selecting a corner as the landmark in that sub-window. So we use the Harris operator [5] to compute the responses of the detected sub-window and locate the corner as the local maxima in the response image.

In the second step, we first learn statistical models (ASM or sparse shape composite) for the right lung boundary shape from training set. Given a new image, we interpolate a prototype boundary based on the detected 6 landmarks. Then, we use the learned shape prior information to refine the prototype boundary and get an approximate contour for the right lung. Fig. 4 shows this procedure, in which the blue points and lines are the ground truth, the red points and lines are the expected results.



(a) LM 1 (b) LM 2 (c) LM 3 (d) LM 4 (e) LM 5 (f) LM 6

Fig. 3. Example of landmark patches. (LM stands for landmark)

4 Experiments

4.1 Set-up

The data set is a total of 372 2D CXR images with labeled lung. 188 of them have manually labeled landmarks, which are used to train the landmark detectors and to construct the prior-shape repository. The rest 184 of them are used for testing. Due to the similarity of left lung and right lung, we only work on right lung segmentation.

We will use the confusion matrix to analyze the specificity, sensitivity, precision and accuracy of the lung detectors: $Sensitivity = \frac{TP}{TP+FN}$, $Specificity = \frac{TN}{TN+FP}$, $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$. True Positive (TP): the object detected by the detector is truly labeled object. False Positive (FP): the object detected by the detector is not truly labeled object. True Negative (TN): the object not detected by the detector is not truly labeled object. False Negative (FN): the object not detected by the detector is truly labeled object.

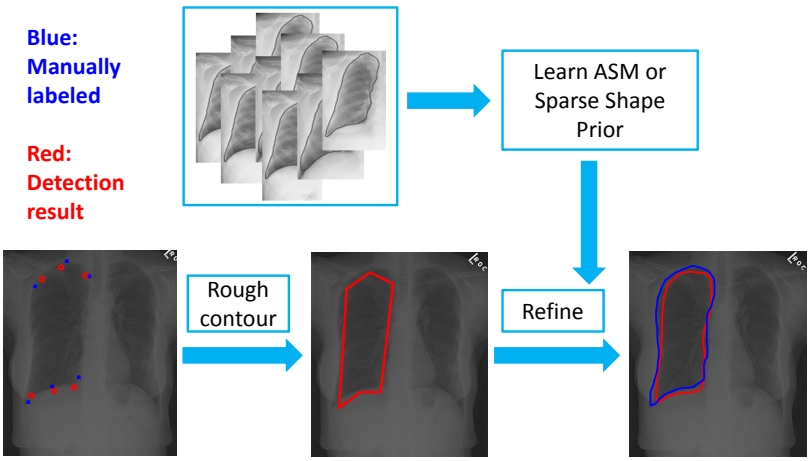
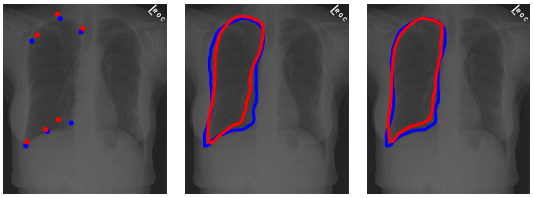


Fig. 4. Boundary contour refinement with shape prior

For the final segmentation results (right lung contours), the mean values and the standard deviations of sensitivity (P), specificity (Q), and Dice Similarity Coefficient (DSC) are reported. As a base line, we obtain the landmarks using the central points of the detected rectangles. And then compare the refined contour results with that using Harris corner detector.



(a) Landmark (b) PCA refine (c) Sparse refine

Fig. 5. Boundary contour refinement results

4.2 Results

In table 1, the first two rows are results for training set, while the last two rows contain results for testing set. Here, the last column represents the percentage of FP cases that detect the other lung from input images. For Haar+Boost, despite the low specificity score, high value of left lung hit means out of the FP cases this method usually detects the left lung rather than some random regions. So this low Specificity should be acceptable.

Table 1. Right lung detection results

Method	Sensitivity	Specificity	Accuracy	Time	Left lung hit
Haar+Boost	1.0	0.601064	0.800532	2.90 sec	0.92
Gabor+SVM	0.728723	0.590426	0.659574	264.72 sec	0.36
Haar+Boost	0.983696	0.61413	0.798913	2.44 sec	0.83
Gabor+SVM	0.581522	0.532609	0.557065	185.89 sec	0.13

Table 2. Boundary contour refinement results

Method	mean(P)	std(P)	mean(Q)	std(Q)	mean(DSC)	std(DSC)
PCA	0.8439	0.1000	0.9108	0.0989	0.8595	0.0619
Sparse	0.8740	0.1016	0.9137	0.1032	0.8914	0.0661

Table 2 shows the final boundary contour results refined by PCA and sparse shape composite. From the results, it can be concluded that the sparse shape prior is better than PCA prior. The comparison is shown in fig. 5.

Regarding to the detection time, Haar+Boost method is far superior to Gabor+SVM. Overall, the Haar+Boost performs better in our experiment. So we applied Haar+Boost in the landmark detection, the detection time for each CXR image is about 0.3 seconds. And the shape refinement by PCA costs about 0.02 seconds per image, while sparse shape refinement costs about 1.8 seconds.

5 Conclusions and Future Work

In this work, an automatic rapid framework for lung segmentation from CXR images was implemented and validated. The problem of landmark detection was solved by Haar+Boost method efficiently and refined by Harris corner detection. Then the learned shape prior was used to infer the shape of the right lung. This framework could be extended to other applications, such as lesion detection, organ segmentation from CT or MRI medical images. In the future, we would like to extend this framework to 3D and also to solve other problems, such as tracking the cardiac [18].

References

1. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active Shape Models-Their Training and Application. Comput. Vis. Image Underst. 61(1), 38–59 (Jan 1995)
2. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. 23(6), 681–685 (Jun 2001)
3. Cristianini, N., Shawe-Taylor, J.: An introduction to support Vector Machines: and other kernel-based learning methods. Cambridge University Press, New York, NY, USA (2000)
4. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55(1), 119–139 (Aug 1997)
5. Harris, C., Stephens, M.: A combined corner and edge detector. In: Proceedings of the 4th Alvey Vision Conference. pp. 147–151 (1988)

6. Huang, X., Metaxas, D., Chen, T.: Metamorphs: Deformable shape and texture models. In: IEEE Conf. on Computer Vision and Pattern Recognition. pp. 496–503 (2004)
7. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision 1(4), 321–331 (1988)
8. Kozakaya, T., Shibata, T., Yuasa, M., Yamaguchi, O.: Facial feature localization using weighted vector concentration approach. Image Vision Comput. 28(5), 772–780 (May 2010)
9. Lienhart, R., Maydt, J.: An extended set of haar-like features for rapid object detection. In: IEEE ICIP 2002. pp. 900–903 (2002)
10. Osher, S., Paragios, N.: Geometric Level Set Methods in Imaging,Vision,and Graphics. Springer-Verlag New York, Inc., Secaucus, NJ, USA (2003)
11. Shen, L., Bai, L., Ji, Z.: A svm face recognition method based on optimized gabor features. In: Proceedings of the 9th international conference on Advances in visual information systems. pp. 165–174. VISUAL'07, Springer-Verlag, Berlin, Heidelberg (2007)
12. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: CVPR. pp. 511–518 (2001)
13. Xu, C., Prince, J.L.: Snakes, shapes, and gradient vector flow. IEEE Transactions on Image Processing 7(3), 359–369 (1998)
14. Yang, P., Shan, S., Gao, W., Li, S.Z., Zhang, D.: Face recognition using ada-boosted gabor features. In: Proceedings of the 16th International Conference on Face and Gesture Recognition. pp. 356–361 (2004)
15. Zhan, Y., Dewan, M., Zhou, X.S.: Cross modality deformable segmentation using hierarchical clustering and learning. In: Proceedings of the 12th International Conference on Medical Image Computing and Computer-Assisted Intervention: Part II. pp. 1033–1041. MICCAI '09, Springer-Verlag, Berlin, Heidelberg (2009)
16. Zhang, S., Zhan, Y., Dewan, M., Huang, J., Metaxas, D.N., Zhou, X.S.: Deformable segmentation via sparse shape representation. In: Proceedings of the 14th international conference on Medical image computing and computer-assisted intervention - Volume Part II. pp. 451–458. MICCAI'11, Springer-Verlag, Berlin, Heidelberg (2011)
17. Zhang, S., Zhan, Y., Dewan, M., Huang, J., Metaxas, D.N., Zhou, X.S.: Towards robust and effective shape modeling: Sparse shape composition. Medical Image Analysis 16(1), 265–277 (2012)
18. Zhou, Y., Yeniaras, E., Tsiamyrtzis, P., Tsekos, N., Pavlidis, I.: Collaborative tracking for mri-guided robotic intervention on the beating heart. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010 pp. 351–358 (2010)