# Left Atrium Segmentation in CT Volumes with Fully Convolutional Networks

Honghui Liu[1], Jianjiang Feng[1], Zishun Feng[1], Jiwen Lu[1], and Jie Zhou[1]

[1] Department of Automation, Tsinghua University, Beijing, China
liuhh15@mails.tsinghua.edu.cn

**Abstract.** Automatic segmentation of the left atrium (LA) is a fundamental task for atrial fibrillation diagnosis and computer-aided ablation operation support systems. This paper presents an approach to automatically segmenting left atrium in 3D CT volumes using fully convolutional neural networks (FCNs). We train FCN for automatic segmentation of the left atrium, and then refine the segmentation results of the FCN using the knowledge of the left ventricle segmented using ASM based method. The proposed FCN models were trained on the STACOM'13 CT dataset. The results show that FCN-based left atrium segmentation achieves Dice coefficient scores over 93% with computation time below 35s per volume, despite of the high variation of LA.

## 1 Introduction

Among the most common and hazardous cardiovascular diseases, atrial fibrillation (AF) is usually characterized by abnormally rapid and irregular heart rhythm. In recent years, surgical treatment for AF, typically ablation procedure (AP), has gradually become the mainstream [9]. AP is minimally invasive, which is the main consideration for some patients. Computer aided atrial detection and precise segmentation can help doctors gain valuable preoperative information. Computed tomography (CT) has been widely used for diagnosis and treatment for cardiovascular disease. However, automated LA segmentation from CT data is still a non trivial task. Due to larger shape variations of LA than other organs, especially four pulmonary veins (PV) and left atrial appendage (LAA). Some existing methods have achieved good results, but the training cost is high [13]. In this paper, we propose a fully automatic LA segmentation system on C-arm CT. Based on state-of-the-art fully convolution network (FCN), and statistical shape models, the proposed method is efficient, robust, and is able to obtain good results with small training dataset.

## 2 Related Work

Image segmentation is a fundamental task in medical image analysis. Producing accurate segmentation is difficult due to many influencing factors: noise,

pathology, occlusion, and object shape complexity. Some semi-automated or automated algorithms have been applied to LA segmentation problem. Daoudi et al. [3] proposed an algorithm based on active contour, with region growing and snakes. This type of classical methods are simple and fast, but sensitive to image quality. Sandoval et al. [6] proposed an algorithm based on multi-atlas. Multi-atlas has remarkable advantages of robustness and making good use of a priori anatomical information. The main shortcoming of multi-atlas is computing cost. Image registration for 3D volume is quite time consuming even though GPU parallel acceleration strategy has been applied. Zuluaga et al. [15] proposed another multi-atlas propagation based segmentation. Nowadays, statistical shape models are widely used in image segmentation. Zheng et al. [14] proposed an algorithm based on a multi-part shape model and marginal space learning, which divides LA into six-parts: LA body, LAA and four PVs. This algorithm is efficient, and robust to image quality. However, a large manually labelled data set is required. For medical image application, this limitation cannot be ignored till now. Moreover, compared to LA body, the LAA has larger anatomical variations, thus method using strong shape priors is not suitable [5].

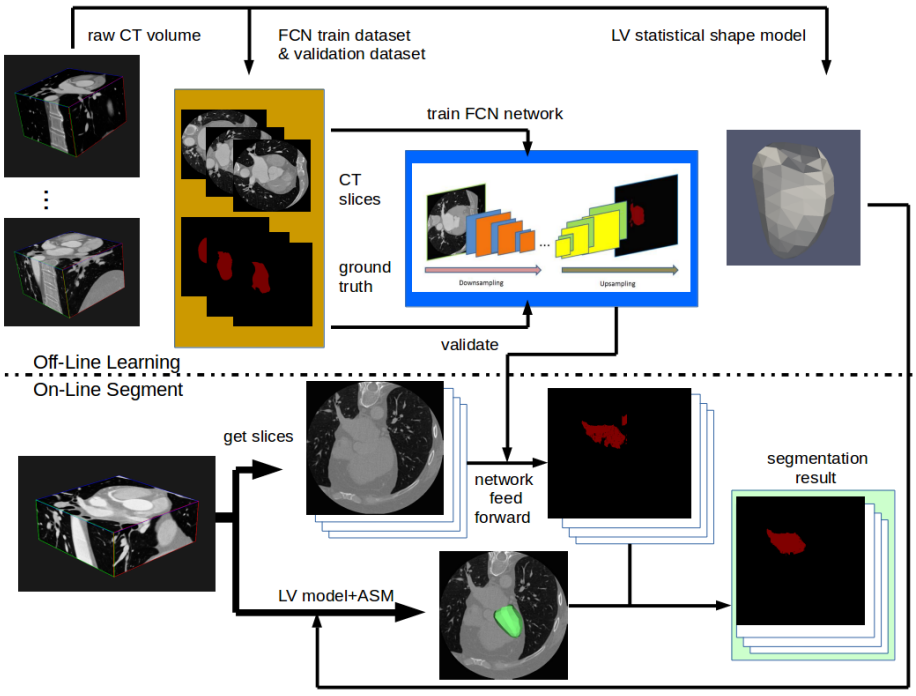



**Fig. 1.** Pipeline of the proposed approach: off-line training and on-line testing

Our work draws on recent progress of deep neural nets [12]. Image segmentation can be viewed as a pixel-wise classification task. In recent years, convolutional neural networks (CNN) achieve success on image classification problem. Therefore, taking advantage of coarse but highly abstract hidden layer output, pixel-wise classification can also be solved with CNN [10]. For segmentation task, the target is a dense label map, which is of the same size as the input. Specifically, to achieve this coarse-to-fine processing, we can add upsampling layer and deconvolution layer upon CNN, and then obtain the so-called fully convolutional network (FCN). Owing to its powerful ability to learn both local and global information, FCN has made great progress in image segmentation.

In this work, we present an automatic LA segmentation method for 3D CT. We transferred deep neural network architectures for natural scenes to medical task, thus good generalization ability can be expected. Furthermore, to achieve both strong deformable ability and shape constraints, we combined FCN and statistical shape models, and thus improve the accuracy while preserving the efficiency.

In the following sections, we will demonstrate our proposed pipeline, including preprocessing (Section 3.1), FCN (Section 3.2) and shape model based postprocessing (Section 3.3), report experiments (Section 4), and summarize the paper (Section 5).

## 3 Method

We trained a hourglass-shaped architecture with a per-pixel logistic loss, and validated with the standard pixel intersection over union metric. Our network includes 12 stacked convolution layers, each followed by max pooling layer and ReLU activation layer. Two deconvolution layers were added on the top of the CNN architecture.

Training FCN with small training dataset has been an awkward problem. To resolve this problem, we adopted two measures. First, we pre-train our network on a big dataset [4] with supervision. Next, the network was fine-tuned on our medical dataset. We selected slices from raw volume data along different axes. Thereby an abundant training dataset with more than 1000 slices can meet the demand for network training. This treatment may inevitably lose some 3D structure informatinon. However, with postprocessing from the 3D perspective, this 3D information lost will be minimized. Fig. 1 shows the steps of our proposed approach.

### 3.1 Preprocessing

Our preprocessing consists of two aspects. First, the image contrast is increased with histogram equalization. Next, we convert the gray scale image to pseudo color image. Taking advantage of three color channels, network training can be improved.

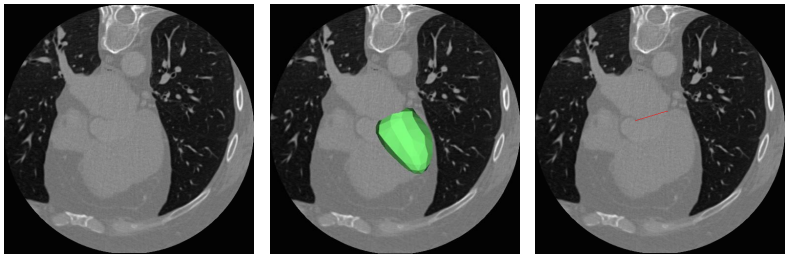
### 3.2 Fully Convolutional Network

We denote the 3D raw CT volume as $I$. For our two class problem, the set of possible labels is $L = \{0, 1\}$. Foreground LA voxels are marked with label 1 while other background voxels are marked with 0. For each voxel $v$, we define a variable $y_v \in L$ that denotes the assigned label. Given the image $I$, the FCN calculates the probability of assigning label $k$ to $v$, described by $P(y_v = k|I)$.

The first step in the training process, we trained FCN-32s network. FCN-32s networks give coarse segmentation results, much local information of the input is dropped while passing convolution and pooling layers. Hence we added links form lower layers to the final layer, and the whole network turns into a directed acyclic graph (DAG). Compared with simple linear networks, DAG-style networks combine the global structure with local information, thus the results became more sophisticated. Combining the second-to-last and then the third-to-last pooling layers, FCN-16s and FCN-8s structure were generated in succession.

Our dataset included 1200 slices, which were extracted from 10 CT volumes. These slices were divided into two parts: 1000 slices as training set, and the other 200 as validation set.

Due to the fact that FCN was trained on slices, some 3D structure information was lost unavoidably, continuity and smoothness cannot be ensured for the final 3D model. Some tiny tissues were misclassified as foreground. Therefore, we extract the maximum connected component from the FCN segmentation result, and apply hole filling algorithm to the maximum connected component. By doing this, most false positives can be excluded, then segmentation result become a smooth and solid model.



(a) raw image slice (b) LV segmentation (c) LA/LV boundary

**Fig. 2.** Segment LV with ASM based method, to get the LA/LV boundary

### 3.3 Shape Constraints

The knotty problem for LA segmentation task is determination for the fuzzy boundary of LA and LV. These two chambers are connected in structure, and

**Table 1.** LA segmentation performance comparison by three evaluation metrics. Acronyms of methods are from [13].

| Method | evaluation criterion | | |
|---|---|---|---|
| | DC | time(s) | train dataset size |
| BECHAR [3] | 0.66 | 900 | 10 |
| INRIA [8] | 0.82 | 1500 | 10 |
| LTSI_VRG [6] | 0.88 | 4700 | 10 |
| SIE_PMB [14] | 0.94 | 3 | 457 |
| UCL_1C [15] | 0.93 | 4200 | 30 |
| Proposed | **0.93** | **32** | **30** |

may have similar gray-scale for some images (see Fig. 2 (a)). Therefore, many machine learning methods [8] have shortcoming of deciding the boundary of LA and LV, as well as FCN method. These methods mainly focus on LA segmentation, while LV is ignored.

Some previous work [1] adopted 3D CRF as postprocessing to improve the segmentation results. Dense CRF mainly focuses on prior and posterior distribution of pixel classes, rather than global shape constraints. Furthermore, CRF method is time consuming, for both offline learning and online testing.

For fuzzy boundary problem, shape models with constraints are shown to be more robust [13]. Compared with LA wall, LV epicardium is much thicker, and the clearly visible posterior borders of the LV and LA are flat. We establish a statistic shape model to segment LV epicardium with active shape model (ASM) method [2]. The top of LV is regarded as the boundary of these two chambers. These two chambers can therefore be clearly distinguished (see Fig. 2).

## 4 Experiments

We trained networks on the benchmark CT dataset for Left Atrial Segmentation Challenge (LASC) carried out at the STACOM'13 workshop [13]. Ten CT volumes were provided with expert manual segmentations, and the other twenty volumes were used for algorithm evaluation. For statistical shape model method, we randomly selected other twenty volumes from patients who underwent a CTA examination using a Philips Brilliance iCT256 scanner. The volumes in the whole dataset set contain 210 to 455 slices while the size of all slices is of $512 \times 512$ pixels. The resolution inside each slice is isotropic but varies between 0.314 mm and 0.508 mm for different volumes, and the slice thickness varies between 0.450 mm and 0.510 mm.

For the training process, the mean validation pixel-wise accuracy of FCN-32s, 16s, and 8s is 86%, 92%, and 95%, respectively. In general, 32s networks give coarse results, with very low false positive rate. Furthermore, 16s networks give higher accuracy, but high false positive rate as well. Moreover, 8s networks give obviously better results than the others.

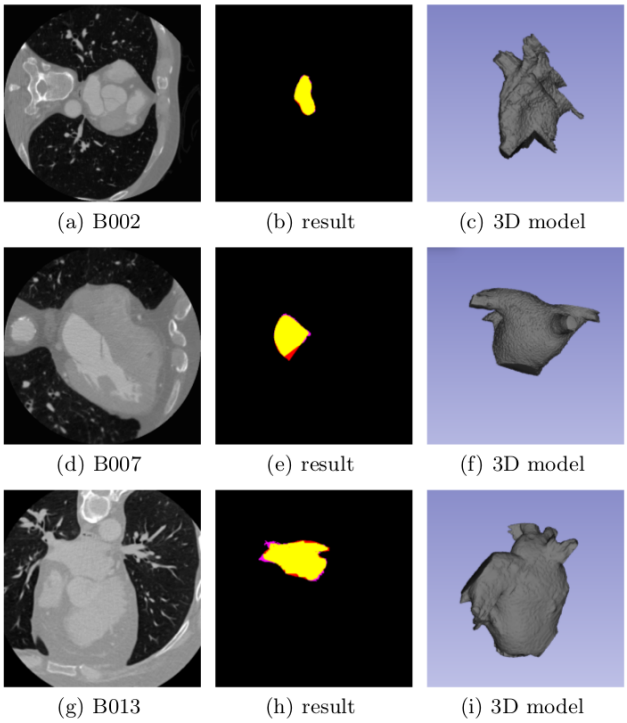



**Fig. 3.** Segmentation examples for B002, B007, B013 data. (a)(d)(g): raw image, (b)(e)(h): segmentation result, and (c)(f)(i): segmented 3D model. Yellow region indicates true positive, red indicates false negative, and magenta indicates false positive.

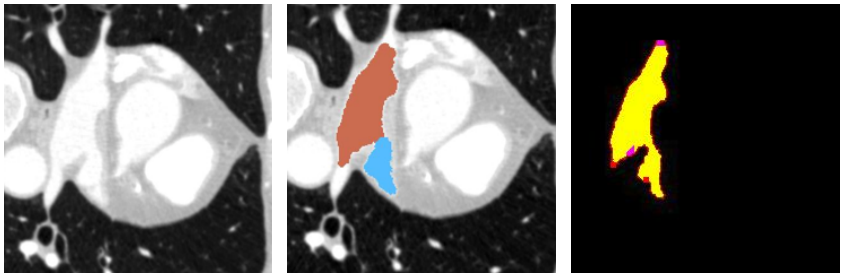
The test data set includes 20 volumes. We choose Dice coefficient as evaluation criterion, which is defined by:

$$DC = \frac{2|V_{GT} \bigcap V_{SEG}|}{|V_{GT}| + |V_{SEG}|}, \tag{1}$$

where we denote the ground truth volume mask by $V_{GT}$, and the segment result mask by $V_{SEG}$.

Table 1 shows the results of some different methods. Segmentation accuracy is the main consideration for algorithm evaluation, while the other two aspects affect the applicability.

As a comparison experiment, we choose dense 3D CRF [7] as postprocessing method, the result shows that ASM based LV segmentation gives simpler yet robust LA/LV boundary. For the other comparison experiment, we establish LA statistical shape model, and apply ASM method for both LA and LV, to compare FCN with ASM method for LA segmentation. Table 2 shows the results of these comparison experiments.



(a) raw image (b) ground truth (c) result

**Fig. 4.** LAA segmentation example. Note that LA body ground truth is from the original STACOM' 13 dataset, and LAA ground truth was labelled by ourselves.: (a) raw image, (b) ground truth, LAA is labelled with blue, and (c) segmentation result, the meaning for each color is the same as Fig. 3

**Table 2.** Segmentation performance comparison by different LA/LV processing

| Method | evaluation criterion | | |
|---|---|---|---|
| | DC | time(s) | train dataset size |
| FCN + CRF [7] | 0.87 | 120 | 30 |
| LA ASM + LV ASM | 0.85 | 15 | 30 |
| FCN + LV ASM (Proposed) | **0.93** | **32** | **30** |

As an example, Fig. 3 shows segmentation results for CT volume B002, B007, and B013. The result and ground truth are compared in Fig. 3(b), (e), (h).

STACOM' 13 challenge ignored LAA while evaluating each method. However, atrial fibrillation usually leads to LAAs emptying obstruction, blood stasis and induces thrombosis [11]. An additional advantage of our method, also not reflected in Table 1, is that it can also segments LAA accurately (see Fig. 4), which is hard for ASM based methods.

## 5 Conclusion

In this paper, we propose a fully convolutional network based automated approach for left atrium segmentation, and adopt statistical shape models to make up for the lack of constraints. The experiments showed that our proposed method is comparable with state-of-the-art methods while using only a small training set. Furthermore, we preliminarily attempted to combine neural networks with statistical shape models.

The proposed method can be improved from many perspectives. For reliability and simplicity, we prefer to learn LA/LV fuzzy boundary in FCN model itself. Neural networks and shape models have respective advantages, next step we will attempt to better unify the two into the same framework. As the dataset increases, the FCN models are expected to be trained better in the future.

## References


1. Christ, P.F., Elshaer, M.E.A., et al.: Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields. In: Proc. of International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 415–423 (2016)
2. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models-their training and application. Computer Vision and Image Understanding 61(1), 38–59 (1995)
3. Daoudi, A., Mahmoudi, S., Chikh, M.A.: Automatic segmentation of the left atrium on CT images. In: Proc. of International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 14–23 (2013)
4. Everingham, M., Eslami, S.M.A., Gool, L.V., Williams, C.K.I., Winn, J., Zisserman, A.: The PASCAL visual object classes challenge: A retrospective. International Journal of Computer Vision 111(1), 98–136 (2015)
5. Heimann, T., Meinzer, H.: Statistical shape models for 3D medical image segmentation: A review. Medical Image Analysis 13(4), 543–563 (2009)
6. Iglesias, J.E., Sabuncu, M.R.: Multi-atlas segmentation of biomedical images: A survey. Medical Image Analysis 24(1), 205–219 (2015)
7. Koltun, V.: Efficient inference in fully connected CRFs with gaussian edge potentials. Adv. Neural Information Processing Systems 2(3), 4 (2011)
8. Margeta, J., McLeod, K., Criminisi, A., Ayache, N.: Decision forests for segmentation of the left atrium from 3D MRI. In: Proc. of International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 49–56 (2013)
9. Marrouche, N.F., Wilber, D.J., Hindricks, G., et al.: Association of atrial tissue fibrosis identified by delayed enhancement MRI and atrial fibrillation catheter ablation: The DECAAF study. Journal of the American Medical Association 311(5), 498–506 (2014)
10. Noh, H., Hong, S., Han, B.: Learning deconvolution network for semantic segmentation. In: Proc. of the IEEE International Conference on Computer Vision. pp. 1520–1528 (2015)
11. Patti, G., Pengo, V., Marcucci, R., et al.: The left atrial appendage: from embryology to prevention of thromboembolism. European Heart Journal p. ehw159 (2016)
12. Shelhamer, E., Long, J., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3431–3440 (2015)
13. Tobongomez, C., Geers, A.J., Peters, J., et al.: Benchmark for algorithms segmenting the left atrium from 3D CT and MRI datasets. IEEE Transactions on Medical Imaging 34(7), 1460–1473 (2015)
14. Zheng, Y., Wang, T., John, M., Zhou, S.K., Boese, J., Comaniciu, D.: Multi-part left atrium modeling and segmentation in C-arm CT volumes for atrial fibrillation ablation. In: Proc. of International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 487–495. Springer (2011)
15. Zuluaga, M., Cardoso, M.J., Modat, M., et al.: Multi-atlas propagation whole heart segmentation from MRI and CTA using a local normalised correlation coefficient criterion. In: Proc. of International Conference on Functional Imaging and Modeling of the Heart. pp. 174–181 (2013)