# Automatically Segmenting the Left Atrium from Cardiac Images Using Successive 3D U-Nets and a Contour Loss

Shuman Jia, Antoine Despinasse, Zihao Wang, Hervé Delingette, Xavier Pennec, Pierre Jaïs, Hubert Cochet, Maxime Sermesant

**HAL Id: hal-01860285**
**https://inria.hal.science/hal-01860285v3**

Submitted on 4 Oct 2018 (v3), last revised 1 Nov 2018 (v4)

# Automatically Segmenting the Left Atrium from Cardiac Images Using Successive 3D U-Nets and a Contour Loss

Shuman Jia[1], Antoine Despinasse[1], Zihao Wang[1], Hervé Delingette[1], Xavier Pennec[1], Pierre Jaïs[2], Hubert Cochet[2], and Maxime Sermesant[1]

[1] Université Côte d'Azur, Epione Research Group, Inria, Sophia Antipolis, France
[2] IHU Liryc, University of Bordeaux, Pessac, France

**Abstract.** Radiological imaging offers effective measurement of anatomy, which is useful in disease diagnosis and assessment. Previous study [1] has shown that the left atrial wall remodeling can provide information to predict treatment outcome in atrial fibrillation. Nevertheless, the segmentation of the left atrial structures from medical images is still very time-consuming. Current advances in neural network may help creating automatic segmentation models that reduce the workload for clinicians. In this preliminary study, we propose automated, two-stage, three-dimensional U-Nets with convolutional neural network, for the challenging task of left atrial segmentation. Unlike previous two-dimensional image segmentation methods, we use 3D U-Nets to obtain the heart cavity directly in 3D. The dual 3D U-Net structure consists of, a first U-Net to coarsely segment and locate the left atrium, and a second U-Net to accurately segment the left atrium under higher resolution. In addition, we introduce a loss function based on additional distance information to adjust the final segmentation. We randomly split the data into training datasets (80 subjects) and validation datasets (20 subjects). Experiments show that the average Dice coefficient for validation datasets is 0.924, the sensitivity is 0.92 and the specificity 0.99. We also compare the prediction results using the two loss functions, Dice coefficient loss and proposed weighted Dice/Contour distance loss, in a 3D view.

**Keywords:** 3D U-Net, segmentation, left atrium, loss function, contour loss, distance map

## 1 Introduction

Atrial fibrillation is the most frequently encountered arrhythmia in clinical practice, especially in aged population [2,3]. It is characterized by uncoordinated electrical activation and disorganized contraction of the atria. This condition is associated with life-threatening consequences, such as stroke and heart failure. AF also leads to increased public resource utilization and expense on health care.

With evolving imaging technologies, the analysis of cardiovascular diseases and computer-aided interventions have been developing rapidly. Imaging of the

heart is routinely performed in some hospital centers when managing AF and prior to atrial ablation therapy, an invasive treatment to establish transmural lesions and block the propagation of arrhythmia. Automated segmentation from cardiac images will benefit the studies of left atrial anatomy, tissues and structures, and provide tools for AF patient management and ablation guidance.

In recent years, with the continuous development of deep learning, neural network models have shown significant advantages in different visual and image processing problems [4]. In medical image analysis, automatic segmentation of 3D volumes from medical images by deep neural network also attracts increasing attention in the research community [5, 6].

In this study, we utilize 3D U-Nets with convolutional neural network. CNN shows clear advantages compared with traditional feature extraction algorithms [7, 8]. Based on that, Ronneberger et al. proposed the original U-Net structure [9] with CNN. Traditional 2D U-Net has achieved good results in the field of medical image segmentation [9, 10]. However, it performs convolution on the 2D slices of the images and cannot capture the spatial relationship between slices. Its 3D extension [11], expands the filter operator into 3D space. This extracts image features in 3D, and hence takes into account the spatial continuity between slices in medical imaging. This may better reflect shape features of the corresponding anatomy.

Previously, Tran et al. used 3D CNNs to extract temporal and spatial features [12]. They experimented with different sets of data. Hou et al. used 3D CNN to detect and segment pedestrians in a video sequence [13]. The previous studies show that 3D CNN outperformed 2D CNN when dealing with sequences issues.

3D U-Net was used in [11] to realize semi-automatic segmentation of volumetric images. Oktay et al. used 3D U-Net ventricle segmentation from magnetic resonance images. They introduced an anatomical regularization factor into the model [14], while we choose to use loss function at pixel level.

In the following sections, we will present the two-stage network to segment the left atrium from MR images. The network consists of two successive 3D U-Nets. The first U-Net is used to locate the segmentation target. The second U-Net performs detailed segmentation from cropped region of interest. We introduced a new loss function, Contour distance for the second U-Net. Results will be shown in Section 3.

## 2 Method

### 2.1 Dual 3D U-Nets - cropping and segmenting

U-Net is a typical encoder-decoder neural network structure. The images are encoded by the CNN layers in the encoder. The output characteristics and the feature maps at different feature levels of the encoder serve as input of the decoder. The decoder is an inverse layer-by-layer decoding process. Such a codec structure can effectively extract image features of different levels so as to analyze the images in each dimension. The 3D U-Net used in this paper is a specialization of 3D U-Net proposed by Çiçek et al [11]. The implementation of U-Net

follows the work of Isensee et al [15]. We propose a successive dual 3D U-Net architecture, illustrated in Fig. 1.

The first 3D U-Net locates and coarsely extracts the region of interest. Its input is MR images normalized and re-sized to $(128, 128, 128)$. Its output is preliminary predicted masks of the left atrium. We keep the largest connected component in the masks, and compute the spatial location and size of the left atrium. Then, we crop the MR images and ground truth masks with a cuboid, whose dimension in each direction equals to 1.5 times the size of the predicted left atrium in this direction.

The second network performs a secondary processing of the cropped images using the full resolution. Because the higher is the resolution, the larger is the needed memory, we keep only the region around the left atrium, so as to preserve information that is essential for left atrial segmentation. But also, this allows to put a higher resolution on the region of interest with the same amount of memory resource. The input for the second U-Net is MR images cropped and and re-sized to $(128, 128, 128)$. Its output is our prediction for the left atrial segmentation. We train the second U-Net with two ground truth channels, binary segmentation masks $M$ and euclidean distance maps $D(M)$, as shown in Fig 2 (top). Here, we introduce a weighted loss function based on Dice coefficient and Contour distance.
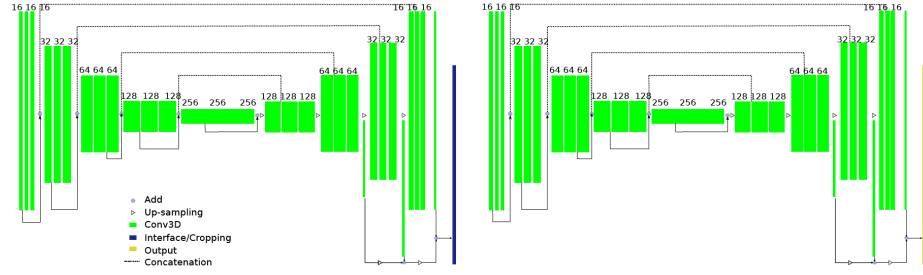


**Fig. 1.** Proposed Dual 3D U-Net Structure. Green blocks represent 3D features; Dark blue refers to interface operation to crop the region of interest based on first U-Net prediction.

## 2.2 Loss Functions

Using Dice coefficient as loss function can reach high accuracy. However, experiments show that, because the inside of the left atrial body accounts for most pixels, the network would stop to optimize when it finds satisfying segmentation of the left atrial body. Instead of the volume inside, the contour is what we want to obtain accurately for the segmentation. It is challenging to segment accurately especially the region around the pulmonary veins and the appendage.

To improve the performance, we introduce the contour distance into the loss function.

The distance maps $D(M)$ illustrate how far is each pixel from the contour of the left atrium. The value is negative inside the left atrium and positive outside. We compute the signed distance based on the euclidean distance transform algorithm implemented in scipy. The definition of a Hausdorff distance is symmetric between two point sets. But to make it easy to be implemented in neural networks, we do not compute the distance map of the changing prediction $P$ in the training process, and use an unitary distance:

$$d = \sum_{p \in P} \min_{m \in M} \|p - m\|_2 = -\sum (D(M) \circ P), \tag{1}$$

where $\circ$ performs element-wise multiplication. $M$ is the first ground truth channel, the binary segmentation, $D(M)$ is the second ground truth channel, the signed distance map, and $P$ is the prediction of the U-Net, after sigmoid activation of the output layer. Then, the loss function is given by a weighted combination of Dice coefficient loss and Contour distance loss,

$$Loss_{Dice} : \ -\frac{\sum (M \circ P)}{\sum M + \sum P}; \tag{2}$$

$$Loss_{Contour} : \frac{\sum (D(M) \circ P)}{\sum \|D(M)\|}. \tag{3}$$

We normalize the second loss by $\sum \|D(M)\|$. To compute the distance loss function only from pixels on contour, we can either clip the value of $P$ in the range of $[0.1, 0.9]$; or apply 3D Sobel filters on $P$.

Contour distance provides a spatial information for the segmentation, while Dice coefficient measures the volumes inside the contours. We assign a bigger weight to Dice so that it controls the convergence in the beginning, and when it stops to improve, Contour loss begins to alter the the segmentation results. The loss function is still differentiable and converges in the training.

## 2.3 Experimental setting

In this study, we use 1 as batch size. The initial learning rate of our neural network was $5e^{-4}$. For the first U-Net, it takes around $300 - 350$ epochs to reach early convergence. For the second U-Net, it takes around $250 - 300$ epochs to reach early convergence. The early convergence is defined as no improvement after 50 epochs.

Using large initial learning rate reduces the time to find the minimum and may also overstep the local minima to converge closer to the global minimum of the loss function. However, the accuracy of segmentation also relies on when we stop the training to avoid over-fitting. Besides, to obtain an optimal weight between the two loss functions, experiments are needed. We present the results of weighted Contour distance loss and Dice loss as 0.05:0.95. With Dice coefficient loss alone, the average Dice index for predicted segmentations in validation

datasets is around 0.91. We choose this ratio so that Contour distance loss plays a fine-tuning role.

## 3  Evaluation on Clinical Data

### 3.1  Materials

A total of 100 3D GE-MRIs from patients with AF are provided by the STACOM 2018 Atrial Segmentation Challenge. The original resolution of the data is $0.625\times 0.625 \times 0.625$ $mm^3$. 3D MR images were acquired using a clinical whole-body MRI scanner and the corresponding ground truths of the left atrial masks were manually segmented by experts in the field.

We performed normalization on the raw MR images. The datasets were split into 80 training sets and 20 validation sets.

### 3.2  Results

We visualize the predicted segmentation results of validation datasets in Fig. 2&3. Case 1 and Case 2 are randomly selected to be presented in a 3D view (see Fig. 3).

The proposed method closely segmented the atrial body compared with manual segmentation, with both loss functions. As the manual segmentations were probably performed slice-by-slice, there is sometimes discontinuity between two axial slices. While our segmentation is based on 3D operators, the segmented surface is smooth in 3D. The use of 3D U-Net to segment the cardiac MR image enables full use of the spatial information of the 3D image as compared to segmentation on a 2D image.

**Evaluation Metrics** We assess the segmentation results of our method by confusion matrix and Dice coefficient, as shown in Table 1. The Dice index for predicted segmentation in validation datasets attained 0.92. The sensitivity of prediction was 0.92 and the specificity 0.99, both with smaller standard derivation, compared with traditional Dice coefficient loss. Based on previous studies on the left atrial segmentation and other cardiac cavities segmentation, our network obtained good results.

**Comparison of Loss Functions** We compare the prediction segmentation using Dice loss alone and weighted loss, as featured in Fig. 3.

For the two loss functions, differences lay in the boundary, the region close to the pulmonary veins and septum. The prediction results from the weighted loss is closer to the manual segmentation in terms of global shape. Moreover, the prediction based on Dice coefficient loss alone contains disconnected spots not belonging to the left atrium in Case 1 and a irregular bump on the septum in Case 2, while, with the fine-tuning of Contour distance loss, these effects are reduced.

With the weighted loss function, we can better maintain the smoothness of the contour, and shape consistency.
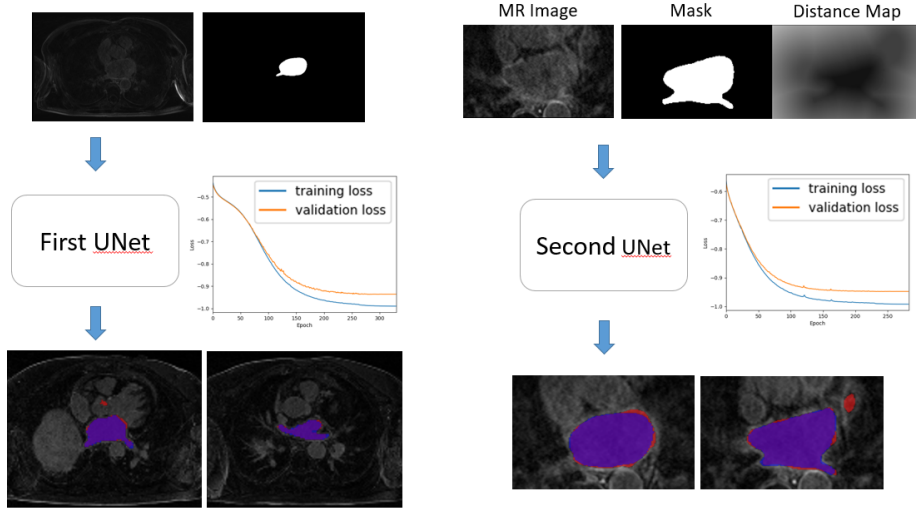
**Fig. 2.** Input (top) and output (bottom) of the successive U-Nets. For prediction results, we show axial slices of MR images, overlapped with manual segmentation of the left atrial wall in blue, our segmentation in red, intersection of the two in purple.
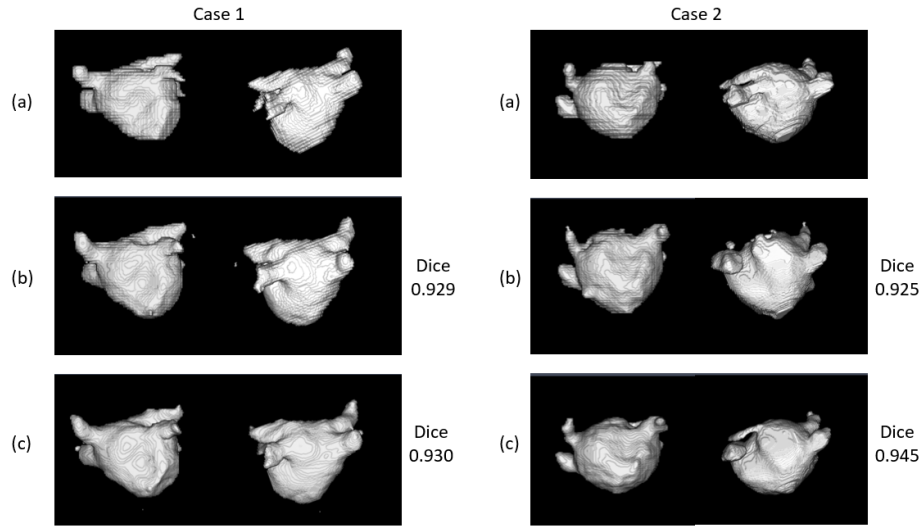


**Fig. 3.** Comparison of two loss functions. (a) manually segmented; (b) predicted with Dice coefficient loss; (c) predicted with weighted Loss of Dice/Contour distance.

**Table 1.** Validation results using two loss functions: Dice coefficient loss (top) and weighted loss of Dice/Contour distance (bottom).

| Confusion Matrix (%) | | Evaluation Metrics | |
|---|---|---|---|
| True positive | False positive | Sensitivity | $0.91 \pm 0.04$ |
| $7.81 \pm 1.70$ | $0.77 \pm 0.37$ | Specificity | $0.99 \pm 0.00$ |
| False negative | True negative | | |
| $0.79 \pm 0.42$ | $90.64 \pm 1.85$ | Dice | $0.907 \pm 0.026$ |

| Confusion Matrix (%) | | Evaluation Metrics | |
|---|---|---|---|
| True positive | False positive | Sensitivity | $0.92 \pm 0.03$ |
| $7.78 \pm 1.68$ | $0.61 \pm 0.23$ | Specificity | $0.99 \pm 0.00$ |
| False negative | True negative | | |
| $0.66 \pm 0.35$ | $90.96 \pm 1.87$ | Dice | $0.924 \pm 0.020$ |

## 4 Conclusion

In this paper, we proposed a deep neural network with dual 3D U-Net structure, to segment the left atrium from MR images. To take into consideration the shape characteristics of the left atrium, we propose to include Contour distance and create a weighted loss function. The entire model can accurately locate the segmentation target and clearly segment the target. Experiments show that the proposed method well captured the anatomy of atrial volume in 3D space from MR images. The new loss function achieved a fine-tuning of accuracy and shape consistency.

The neural network model can predict segmentations within seconds. This can largely reduce manual work load for clinicians and has promising applications in clinical practice. Future work includes integrating the segmentation model into a clinic-oriented AF management pipeline.

## References

1. Christopher McGann, Nazem Akoum, Amit Patel, Eugene Kholmovski, Patricia Revelo, Kavitha Damal, Brent Wilson, Josh Cates, Alexis Harrison, Ravi Ranjan, et al. Atrial fibrillation ablation outcome is predicted by left atrial remodeling on mri. *Circulation: Arrhythmia and Electrophysiology*, 7(1):23–30, 2014.
2. Massimo Zoni-Berisso, Fabrizio Lercari, Tiziana Carazza, Stefano Domenicucci, et al. Epidemiology of atrial fibrillation: European perspective. *Clin Epidemiol*, 6(213):e220, 2014.

3. Carlos A Morillo, Amitava Banerjee, Pablo Perel, David Wood, and Xavier Jouven. Atrial fibrillation: the current epidemic. *Journal of geriatric cardiology: JGC*, 14(3):195, 2017.

4. Carlo N De Cecco, Giuseppe Muscogiuri, Akos Varga-Szemes, and U Joseph Schoepf. Cutting edge clinical applications in cardiovascular magnetic resonance. *World Journal of Radiology*, 9(1):1, 2017.

5. Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A W M van der Laak, Bram van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, dec 2017.

6. Jin Liu, Yi Pan, Min Li, Ziyue Chen, Lu Tang, Chengqian Lu, and Jianxin Wang. Applications of deep learning to MRI images: A survey. *Big Data Mining and Analytics*, 1(1):1–18, 2018.

7. Matthew D. Zeiler and Rob Fergus. Visualizing and Understanding Convolutional Networks arXiv:1311.2901v3 [cs.CV] 28 Nov 2013. *Computer VisionECCV 2014*, 8689:818–833, 2014.

8. Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation Learning: A Review and New Perspectives. (1993):1–30, 2012.

9. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. pages 1–8, 2015.

10. Xiaomeng Li, Hao Chen, Xiaojuan Qi, Qi Dou, Chi-Wing Fu, and Pheng Ann Heng. H-DenseUNet: Hybrid Densely Connected UNet for Liver and Liver Tumor Segmentation from CT Volumes. (1):1–10, 2017.

11. Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3D U-net: Learning dense volumetric segmentation from sparse annotation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9901 LNCS:424–432, 2016.

12. D Tran, L D Bourdev, R Fergus, L Torresani, and M Paluri. Learning Spatiotemporal Features with 3D Convoutional Networks. *CoRR*, abs/1412.0, 2015.

13. Rui Hou, Chen Chen, and Mubarak Shah. An End-to-end 3D Convolutional Neural Network for Action Detection and Segmentation in Videos. 14(8):1–15, 2017.

14. Ozan Oktay, Enzo Ferrante, Konstantinos Kamnitsas, Mattias Heinrich, Wenjia Bai, Jose Caballero, Stuart Cook, Antonio De Marvao, Timothy Dawes, Declan O Regan, Bernhard Kainz, Ben Glocker, Daniel Rueckert, and C V Dec. Anatomically Constrained Neural Networks ( ACNN ): Application to Cardiac Image Enhancement and Segmentation. pages 1–13, 2017.

15. Fabian Isensee, Philipp Kickingereder, Wolfgang Wick, Martin Bendszus, and Klaus H Maier-Hein. Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge. In *International MICCAI Brainlesion Workshop*, pages 287–297. Springer, 2017.