

# Automatic Segmentation of the Left Atrium from LGE-MRI Based on U-Net and Bidirectional Convolutional LSTM

Ze Zhang<sup>1\*</sup>, Kuanquan Wang<sup>1</sup>, Qince Li<sup>1,2</sup>, Yashu Liu<sup>1</sup>, Yongfeng Yuan<sup>1</sup>,  
Yacong Li<sup>1</sup>, Henggui Zhang<sup>2,3</sup>

<sup>1</sup>School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

<sup>2</sup>Peng Cheng Laboratory, Shenzhen, China

<sup>3</sup>School of Physics and Astronomy, The University of Manchester, Manchester, UK

## Abstract

*Atrial fibrillation (AF) is the most common cardiac arrhythmia causing morbidity and mortality. The segmentation of the left atrium (LA) is very important for image-guided ablation of AF and quantification of left atrial fibrosis. However, manual segmentation is labor-intensive and highly subjective. Therefore, the automatic segmentation of the left atrium is of great significance. In this study, we developed a U-Net based network for automatic segmentation of the left atrium. Due to the high computational cost and GPU memory consumption of 3D deep learning networks, a 2D network was used for learning. However, the 2D network only learned the features in 2D slices. The spatial context information of the images was not considered and used, so bidirectional convolutional long short-term memory (LSTM) was combined to obtain the context information in the z-axis direction. This study aimed to design a two-step method based on U-Net and bidirectional convolutional LSTM for the automatic segmentation of the left atrium from LGE-MRI. The model was trained and tested on the dataset of the 2018 Atrial Segmentation Challenge. The dice coefficient obtained by the method was 0.906. By combining the context information between image slices, the segmentation results were optimized.*

## 1. Introduction

Atrial fibrillation is the most common persistent arrhythmia, which has a great risk of causing heart failure, vascular embolism, and even sudden death [1]. Its early diagnosis and treatment can help improve the function of the heart and the prognosis of patients, and effectively reduce the mortality of patients. In order to understand the anatomical structure of the atrium better, assist in interventional surgery and postoperative observation, the segmentation of the atrium is a key step. Images from late gadolinium enhancement magnetic resonance imaging

(LGE-MRI) are widely used to evaluate the structure and function of the heart because of its advantages such as no radiation damage, the ability to perform heart and blood vessel imaging without contrast agents and can be used to detect and quantify the scar tissue in the atrial wall. Although it is possible to reconstruct the atrium and explore its structure by manually segmenting the images, it often requires professional domain knowledge and lots of labor costs [2]. Therefore, the research on the automatic segmentation method of the left atrium has important scientific significance and application value.

There is considerable variability across subjects' anatomical structures of LA, that is, highly individualized differences in atrial shape and size [3]. The image noise and the dynamic movement of the heart will also cause great difficulty in image segmentation. At the same time, the volume of the atrium is smaller than that of the ventricle, and the atrial myocardium is thinner. The left atrium segmentation is much more challenging. The segmentation methods of medical images can be divided into image-driven methods based on no prior knowledge or weak prior knowledge, model-driven methods based on strong prior knowledge, and deep learning methods [4]. For image-driven methods, there is thresholding, region growing [5], clustering, etc. For model-driven methods, there are atlas-guided methods [6,7], statistical shape model based methods like ASM [8] and AAM, etc. Deep learning networks such as CNN are also often used for image segmentation tasks and have achieved good results. Among them, the research of the left atrium segmentation method has also made good progress. In [9], the author summarized the research status of the deep learning method in the field of the left atrium segmentation from LGE-MRI.

In this study, we used a deep learning method for the automatic segmentation of the left atrium from LGE-MRI. The method combined the characteristics of U-Net network and bidirectional convolutional LSTM, and realized effective segmentation of the left atrium by a two-step learning.

## 2. Methods

### 2.1. Preprocessing

Due to the inconsistent image size, we cropped the images and adjusted them to the same size (256×256) suitable for input to the network. The ground truth was processed so that the grayscale value of the left atrium is 1, and the grayscale value of the rest area is 0. We performed z-score standardization on the data. The formula is as follows:

$$X^* = \frac{X - \bar{X}}{\sigma} \quad (1)$$

where  $X$  is the original data,  $\bar{X}$  is the mean value of the original data, and  $\sigma$  is the standard deviation of the original data. The z-score standardization operation scales the original data to a specific interval. The distribution of the processed data  $X^*$  satisfies the normal distribution  $N(0,1)$ . The mean value of the data is 0, and the variance is 1.

Due to the scarcity of data and the high cost of labelling, before training the network, we performed the data augmentation on the training data. Data augmentation such as rotation, translation and zoom was performed on the images in a random combination of data transformation, so that different epochs can obtain different training data and increasing noise data, which can improve the robustness and generalization ability of the model.

### 2.2. U-Net

The medical image dataset tends to only have a small amount of labelled data. The boundary of the target structure is often blurred, and the gradient is complex. The U-Net network [10] was first presented at MICCAI 2015. For the learning of the small dataset, a great segmentation result can also be obtained. Simultaneously, it can supplement the contextual semantic information of the segmentation target at the high level of the network and provide more detailed information such as gradient for segmentation at the low level of the network through skip connection and concatenation. Therefore, it is very suitable for the segmentation of medical images. In this study, our basic structure adopted the U-Net network.

The U-Net network adopts a U-shaped structure. The first half of the network is an encoder process with four coding blocks, and each coding block contains two 3×3 convolutions with stride 1. After each convolutional layer, the rectified linear unit (ReLU) is used as the activation function. Each coding block uses a 2×2 max-pooling layer for down-sampling. The coding process performs four down-sampling. In the decoder part, the resolution of the features is restored to the resolution of the original

images block by block through up-sampling. Correspondingly, there are four decoding blocks. Each of them contains two 3×3 convolutions with stride 1 and similarly uses ReLU as the activation function.

The skip connection structure of the network connects the shallow feature map output by each encoding block in the encoder path and the deep feature map input by the corresponding decoding block in the decoder path. The batch normalization layers were added to the network. The data can be normalized so that the training process can be speed up, and it increased the generalization ability of the model to a certain extent. Data augmentation and dropout regularization were used to prevent overfitting. All layers except the last layer used the ReLU activation function. The last layer of the network used the sigmoid activation function. Adam optimizer with a learning rate of 0.001 was selected for the learning.

### 2.3. Bidirectional Convolutional LSTM

Since only 2D slices were used for learning, the inter-layer information in the z-axis direction of the images was not considered and used. The convolutional LSTM [11] can combine the spatial relationship of the input image sequence into the process of learning, and use context information to make up for those poor segmentation results. Therefore, we added the convolutional LSTM layer in the network to combine the information of the z-axis direction of the 2D slices, that is, the spatial correlation of the 2D slice sequence.

The essence of convolutional LSTM is the same as LSTM. The difference is that convolutional LSTM adds a convolution operation. It can not only get the timing relationship, but also extract spatial features like a convolutional layer. The convolutional LSTM is defined as follows:

$$\begin{aligned} i_t &= \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i) \\ f_t &= \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f) \\ C_t &= f_t \circ C_{t-1} + i_t \circ \text{Relu} \left( \begin{matrix} W_{xc} * X_t \\ + W_{hc} * H_{t-1} + b_c \end{matrix} \right) \\ o_t &= \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ C_t + b_o) \\ H_t &= o_t \circ \text{Relu}(C_t) \end{aligned} \quad (2)$$

where  $X$  is input,  $C$  is cell output,  $H$  is the hidden state,  $i$  is the input gate,  $f$  is the forget gate,  $o$  is the output gate,  $\sigma$  is the logistic function,  $*$  is convolution operator and  $\circ$  is Hadamard product operator.  $W$  denotes the weights.  $b$  denotes the biases. Since the 2D image is related to the slices of both its upper and lower layers, here we use bidirectional LSTM to obtain the spatial information of the images. The sketch map of bidirectional LSTM is shown in Figure 1.

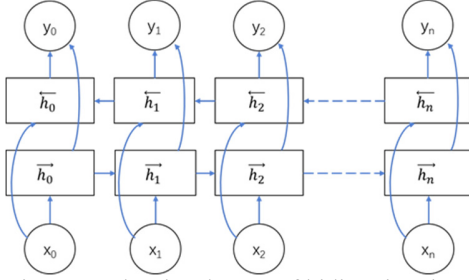


Figure 1. The sketch map of bidirectional LSTM.

In this study, we used the bidirectional convolutional LSTM layer after the last up-sampling of the decoder part to combine the spatial relationship between high-resolution features in adjacent 2D slices. The structure of the network is shown in Figure 2.

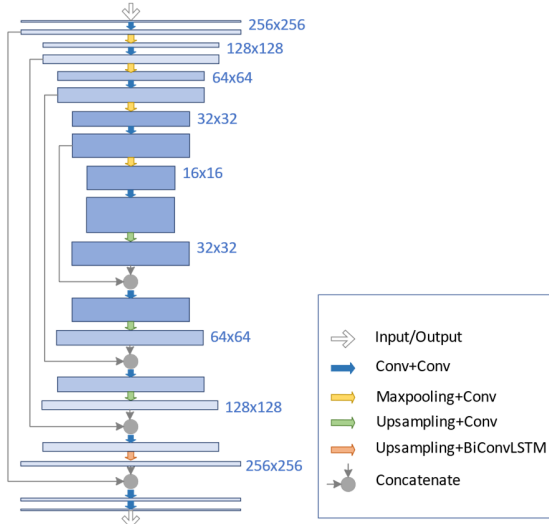


Figure 2. The structure of the network.

## 2.4. A Two-step Learning

The complete learning process was achieved through the following two steps:

Step1: The first learning step used all images containing the left atrium in the training data for network learning. The loss function used was binary cross-entropy. The formula is as follows:

$$Loss = -y * \log(\hat{y}) - (1 - y) * \log(1 - \hat{y}) \quad (3)$$

where  $y$  denotes the ground truth, and  $\hat{y}$  denotes the predicted labels of the images.

Step2: The second learning step used all images in the training data for network learning, in order to make the distribution close to the true distribution. The optimal parameters learned in Step1 were used as the initial parameters for the learning of Step2. The loss function used was the AC loss [12]. The AC loss can constrain the shape of the predicted labels, and the model parameters from Step1 can also make the learning of the AC loss

more stable. The formula of the AC loss is defined as follows:

$$Loss = Length + \lambda \cdot Region \quad (4)$$

in which,

$$Length = \int_C |\nabla u| ds \quad (5)$$

$$Region = \int_{\Omega} ((c_1 - v)^2 - (c_2 - v)^2) u dx$$

in which,

$$c_1 = \int v \cdot u dx / \int u dx$$

$$c_2 = \int v \cdot (1 - u) dx / \int (1 - u) dx \quad (6)$$

where  $u$  denotes the predicted labels, and  $v$  denotes the ground truth.  $c_1$  and  $c_2$  denote the internal and external energy respectively, that is, the energy of the foreground and the background correspondingly.

## 3. Dataset and Results

The network was trained and tested on the dataset of the 2018 Atrial Segmentation Challenge. The dataset contains 100 cases of 3D images from LGE-MRI with a spatial resolution of  $0.625\text{mm} \times 0.625\text{mm} \times 0.625\text{mm}$ . The raw MRIs are in grayscale and the segmentation labels are in binary. The grayscale value of the ground truth was processed as 0 and 1 before being input to the network.

The dataset was split such that 60-patient data were used for training, and 40-patient data were used for validation and testing (20 sets respectively). There are 88 slices in each data, and the size of each slice is either  $640 \times 640$  or  $576 \times 576$ . The dice coefficient was used to evaluate the results. The dice coefficient is essentially a measure of the overlap between the ground truth and the predicted label. The range of the value is between 0 and 1. That value is 1 means complete overlap. The formula of the dice coefficient is as follows:

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \quad (7)$$

where  $A$  is the ground truth,  $B$  is the predicted label,  $|A \cap B|$  is the intersection of  $A$  and  $B$ ,  $|A|$ ,  $|B|$  are the number of elements of  $A$  and  $B$  respectively.

The dice coefficient obtained by the original U-Net network we built was 0.897. The two-step segmentation method that combined the U-Net network and the bidirectional convolutional LSTM was executed. The parameters of the model were adjusted and optimized. The evaluation was then performed on the 20 sets of data. The dice coefficient obtained by the proposed method was 0.906. Figure 3 shows the ground truth and predicted labels of four slices. The first column is ground truth, the second column is the predicted labels using only the original U-Net network we built, and the third column is

the predicted labels of the proposed method. It can be seen in the figure that after combining the bidirectional convolutional LSTM, the segmentation result was optimized through the context.

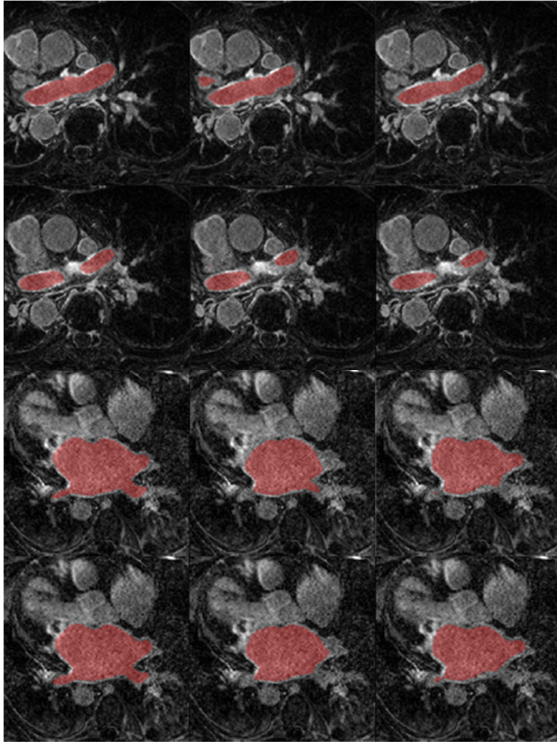


Figure 3. Ground truth and predicted labels of four slices. Each row is for one slice, the first column is ground truth, the second column is the predicted labels using only the original U-Net network we built, and the third column is the predicted labels of the proposed method.

#### 4. Conclusions

In this study, an automatic segmentation method of the left atrium from LGE-MRI was developed. The method performed a two-step learning combining U-Net and bidirectional convolutional LSTM to learn the segmentation model. At the same time, the AC loss was adopted to constrain the shape of the left atrium. The model was trained and tested on the dataset of the 2018 Atrial Segmentation Challenge. The result showed that the method can combine the spatial context information of the images and obtain a good segmentation result.

#### Acknowledgments

The work is supported by the National Natural Science Foundation of China (NSFC) under Grants No. 61572152, 61601143 and 81770328, the Science Technology and Innovation Commission of Shenzhen Municipality under Grants nos. JSGG20160229125049615 and JCYJ20151029173639477, China Postdoctoral Science

Foundation under Grant nos.2015M581448.

#### References

- [1] H. Calkins et al., "2012 HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: Recommendations for patient selection, procedural techniques, patient management and follow-up, definitions, endpoints, and research trial design," *Europace*, vol. 14, no. 4, pp. 528–606, Apr. 2012.
- [2] R. S. Oakes et al., "Detection and quantification of left atrial structural remodeling using delayed enhancement MRI in patients with atrial fibrillation," *Circulation*, vol. 119, no. 13, pp. 1758, 2009.
- [3] A. M. Maceira et al., "Reference left atrial dimensions and volumes by steady state free precession cardiovascular magnetic resonance," *Journal of cardiovascular magnetic resonance*, vol. 12, no. 1, pp. 65, 2010.
- [4] L. Wang et al., "Principles and methods for automatic and semi-automatic tissue segmentation in MRI data," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 29, no. 2, pp. 95–110, 2016.
- [5] L. Zhu et al., "Automatic segmentation of the left atrium from MR images via variational region growing with a moments-based shape prior," *IEEE Trans. Image Process*, vol. 22, no. 12, pp. 5111–5122, Dec. 2013.
- [6] I. Isgum et al., "Multi-Atlas-Based Segmentation with Local Decision Fusion—Application to Cardiac and Aortic Segmentation in CT Scans," *IEEE Transactions on Medical Imaging*, vol. 28, no. 7, pp. 1000–1010, 2009.
- [7] Z. Sandoval et al., "Multi-atlas-based segmentation of the left atrium and pulmonary veins," *International Workshop on Statistical Atlases and Computational Models of the Heart*, Springer, Berlin, Heidelberg, pp. 24–30, Sep. 2013.
- [8] H. C. van Assen et al., "Cardiac LV segmentation using a 3D active shape model driven by fuzzy inference," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Berlin, Heidelberg, pp. 533–540, Nov. 2003.
- [9] K. Jamart et al., "Mini Review: Deep Learning for Atrial Segmentation from Late Gadolinium-Enhanced MRIs," *Frontiers in Cardiovascular Medicine*, Jul. 2020.
- [10] O. Ronneberger et al., "U-net: Convolutional networks for biomedical image segmentation," *International Conference on Medical image computing and computer-assisted intervention*, Springer, Cham, pp. 234–241, Oct. 2015.
- [11] S. H. I. Xingjian et al., "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," *Advances in neural information processing systems*, pp. 802–810, 2015.
- [12] X. Chen et al., "Learning active contour models for medical image segmentation," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 11632–11640, 2019.

Address for correspondence:

Ze Zhang  
Building of Integrated Labs, Harbin Institute of Technology,  
Harbin, China, 150001.  
zhangze568@yeah.net