An Automated Approach Based on a Convolutional Neural Network for Left Atrium Segmentation From Late Gadolinium Enhanced Magnetic Resonance Imaging

Davide Borra[1], Claudio Fabbri[1], Alessandro Masci[1], Lorena Esposito[1], Alice Andalò[1], Cristiana Corsi[1]

[1]DEI, University of Bologna, Cesena, Italy

Abstract

Late Gadolinium Enhanced (LGE) Magnetic Resonance Imaging (MRI) is a new emerging non-invasive technique which might be employed for the non-invasive quantification of left atrium (LA) myocardial fibrotic tissue in patients affected by atrial fibrillation. Nowadays, the analysis of LGE MRI relies on manual tracing of LA boundaries. An automated LA segmentation approach for the quantification of scar tissue would be highly desirable. This study focuses on the design of a fully automated LGE MRI segmentation pipeline which includes a convolutional neural network (CNN) based on the successful architecture U-Net. The CNN was trained with the data available from the Statistical Atlases and Computational Modelling of the Heart 2018 Atrial Segmentation Challenge (100 cardiac data) with two different approaches: using both stacks of 2-D axial slices and using 3-D data. Mean Dice coefficients on the test set were 0.896 and 0.914 by using the 2-D and 3-D approaches, respectively. Contour accuracy was highly variable along the LA longitudinal axis showing poorest results in correspondence of the pulmonary veins. These results suggest that, despite the increase of the number of trainable parameters, the proposed 3-D CNN learns better features leading to higher performance, feasible for a real clinical application.

1.   Introduction

Atrial fibrillation (AF) is the most common arrythmia in the western world with an incidence of about 0.4% in men and 0.6% in women. It is known that the prevalence of AF in US is about 2.2 million including paroxysmal or persistent AF [1]. Consequences of AF could lead to a notable reduction in quality of life, poor mental health, disability, dementia and, mainly, an increase of stroke risk by five-fold [2].

Radio frequency ablation (RFA) of the left atrium (LA) for pulmonary vein isolation (PVI) represents the standard clinical therapy for AF patients in which anti-arrhythmic drugs and direct current cardioversion do not provide improvements for patient's health. However, despite strong improvements for the targeting and the delivery of RFA, the long-term restoration of sinus rhythm is achieved only in a limited percentage of AF patients: AF-free rates after a single ablation vary between 30 and 50% at 5 years follow-up [5]. These results suggest there is room for improvements in RFA treatment and underline a lack of understanding of mechanisms sustaining AF.

Magnetic resonance imaging (MRI) can differentiate between scarred and non-scarred atrial wall by using late gadolinium enhancement (LGE) imaging. Unfortunately, in clinical practice, LGE MRI is rarely available since a standard acquisition protocol is not available [4]. In addition, even if studies on atrial structure segmentation applied to LGE MRI have shown promising results, most of them were based on manual tracing of LA boundaries and PVs [5]. This procedure is time-consuming and prone to high inter-observer variability given the different degrees of observers' experience, LA wall thickness and data resolution.

Different approaches for LA segmentation have been proposed but they are based on different MRI data [6,7]. Only very few studies were recently proposed to segment LA chamber from LGE MRI [8,9]. Unfortunately, these approaches require substantial computation time due to the extensive computation of the multi-atlas-based registration [8] or rely on complex convolutional neural networks (CNNs) involving a large amount of trainable parameters and thus requiring a large number of cardiac data for training as well as an extensive hyper-parameter's tuning procedure to optimize segmentation results since some hyper-parameter settings are not obvious for such complex CNNs [9].

The aim of this study was to design and test a simple fully automatic pipeline for LA segmentation from LGE MRI. The proposed approach was based on the successful U-Net and was tested on both 3-D data using as input the 2-D slices or the 3-D volume to provide a direct comparison between these two approaches and test if shape constraints improve segmentation performances.
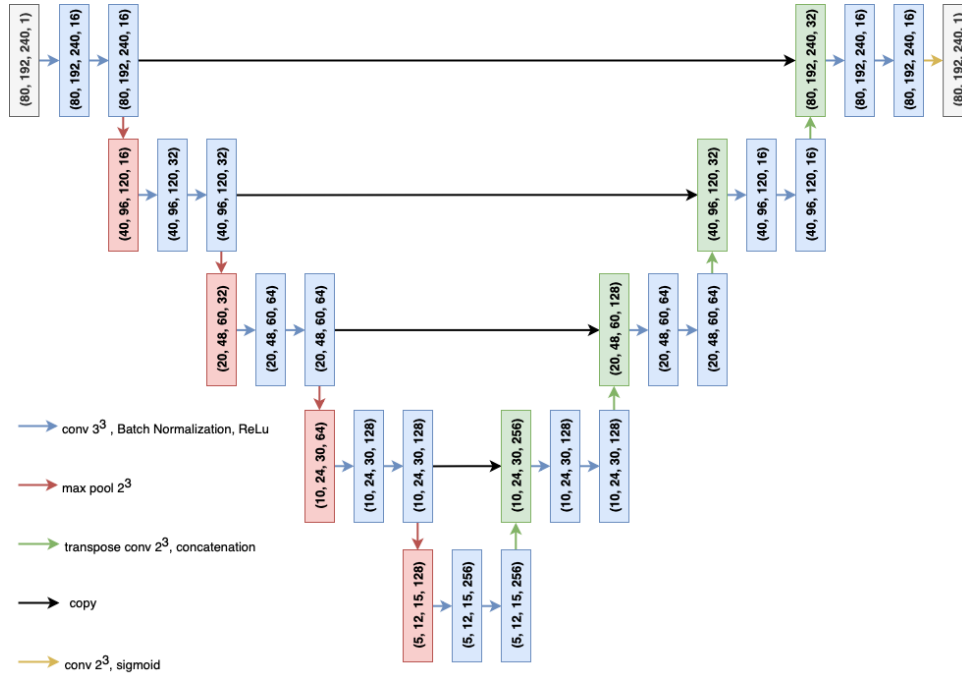
Fig. 1: 3-D CNN architecture. The 2-D architecture shares the same hyper-parameters and is easily obtainable from the 3-D architecture. Each item specifies the tensor shape for each of the represented layers of the neural network and the arrow colors encode different operators as explained in the figure legend. The two gray boxes represent the input and output tensors; the blue boxes are the outputs of the convolutional layers; red boxes represent the max pooled activations and the green boxes the concatenation between the activations of the transposed convolutional layers and the corresponding activations in the encoder module.

2.   Materials and Methods

2.1.   Clinical Data

The developed approach was applied to data from the Statistical Atlases and Computational Modelling of the Heart 2018 Atrial Segmentation Challenge (http://atriaseg2018.cardiacatlas.org/), which includes 100 LGE MRI 3-D cardiac data with the related 3-D ground truth segmentations obtained manually tracing LA endocardial wall.

Data resolution was 0.625x0.625x0.625 mm^3 and 3-D cardiac data were composed by 88 axial slices with in-plane size of 576x576 or 640x640 pixels. To train the neural network in both the proposed approaches, the dataset was randomly split into a training set (80%, 80 cardiac data) and a test set (20%, 20 cardiac data). To perform early stopping in the first training run, a validation set of 10% (8 cardiac data) of the training set was selected.

2.2.   Data Analysis

LGE MRI data were first pre-processed to automatically reduce the number of pixels or voxels from which the CNN extracts information. To this aim, the patient-specific LA position was automatically assessed using a rough LA segmentation based on Otsu's algorithm applied to the central axial slice of each 3-D cardiac image. Once the binary image resulting from Otsu's segmentation was obtained, the centroid of the region located in the center of the image was automatically extracted; the limits of the region of interest in the x direction were automatically computed and a 3-D crop of fixed size of 88x320x384 pixels was extracted from the original cardiac data. This procedure allowed the correct crop of the LA region in all LGE MRI data.

Data were then subsampled to 88x192x240 for the 2-D pipeline and to 80x192x240 for the 3-D pipeline. This last resizing of the images along the third dimension (from 88 to 80) allowed a match between the dimensions of the tensors in the concatenation layers of the neural network in the 3-D approach. Only these subsampled crops containing the LA chamber were used for the training procedure.

The neural network was then trained using these subsampled cardiac data with the 2-D axial slices extracted from the 3-D volumes in the 2-D pipeline and with the 3-D volumes in the 3-D pipeline. In the 2-D pipeline the number of training, test and validation examples was 6336, 1760, and 704, respectively, while in the 3-D pipeline the number of training, test and validation examples was 72, 20, and 8, respectively.

The deep learning approach proposed for both the 2-D and 3-D pipelines was based on the successful U-Net
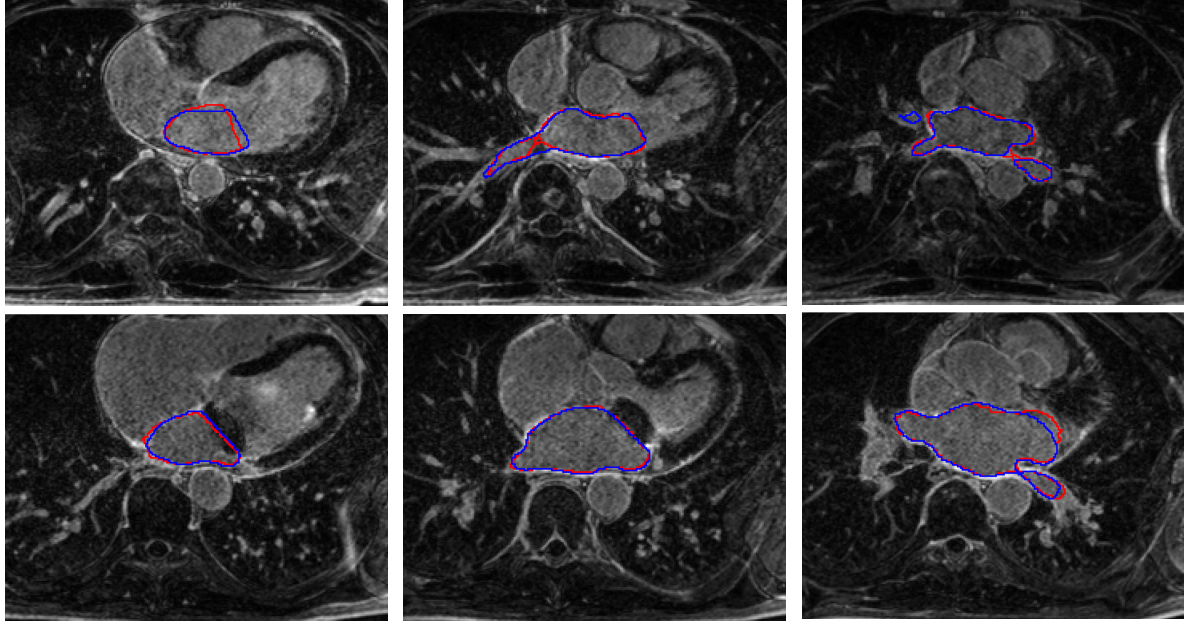
Fig. 2: Contours obtained applying our 3-D approach (red, top panels) and the 2-D approach (red, bottom panels) are compared with contours from manual tracing (blue) for the worst 2-D (D=0.855) and 3-D (D=0.888) predictions. Images were chosen at different levels of the LA (20%, 40% and 60% of the LA longitudinal axis in the 1st, 2nd and 3rd column, respectively).

architecture (Figure 1). The main hyper-parameters were chosen following the original U-Net architecture [11] while the number of convolutional filters and the learning rate were chosen empirically during an early evaluation stage.

In the convolutional layers, kernel size of 3x3x3 (3-D approach) or 3x3 (2-D approach), stride size of 1x1x1 or 1x1 were applied; Rectified Linear Units (ReLUs) activation functions in the hidden layers or sigmoidal activation function in the output layer were used. In the max pooling layers, a pooling size of 2x2x2 and stride size of 2x2x2 or 2x2, halving the shape of hidden activations, were employed. Lastly, in the transposed convolutional layers, kernel size of 2x2x2 or 2x2 and stride size of 2x2x2 or 2x2 were applied. For both convolutional and transposed convolutional layers, padding size was such that the output shape of the layer was the same of the input shape. Furthermore, biases and weights were randomly initialized from a truncated normal distribution and using the initialization scheme proposed in [12] for ReLUs, respectively.

In addition to the original version of the U-Net, after each convolutional layer and before the activation function, a batch normalization layer was included.

The neural network training was driven by a Soft-Dice (SD) loss function proposed to introduce a balancing between foreground and background voxels (or pixels in the 2-D approach). The training process was subdivided into 2 runs: in the first run the neural network was trained until the validation SD reaches its minimum (with a maximum number of epochs of 220). The training loss recorded at this minimum was the target threshold loss to be reached during the second run. In the second run the training continued including the validation set in the training set until the validation set loss matched the threshold loss recorded during the first run (with a maximum number of epochs of 100).

3. Results

The overall number of parameters for the 2-D approach was 1946705 (1943761 trainable parameters), while for the 3-D approach it was 5650801 (5647857 trainable parameters).

In both our pipelines the validation loss reached the desired threshold within the maximum numbers of epochs. The first run took up to 320 s/epoch in the 3-D pipeline and up to 115 s/epoch in the 2-D pipeline, while the second run took a few more seconds due to the slight increase of the number of training samples.

The LA surface model was obtained in few seconds: 1.01 s for the 3-D pipeline and 0.02 s for the 2-D pipeline, considering only the forward propagation time of a sample through the deep neural network.

The 2-D examples of the LA contours in the worst predictions obtained applying the 3-D approach (top panels) and the 2-D approach (bottom panels) are reported in Figure 2.

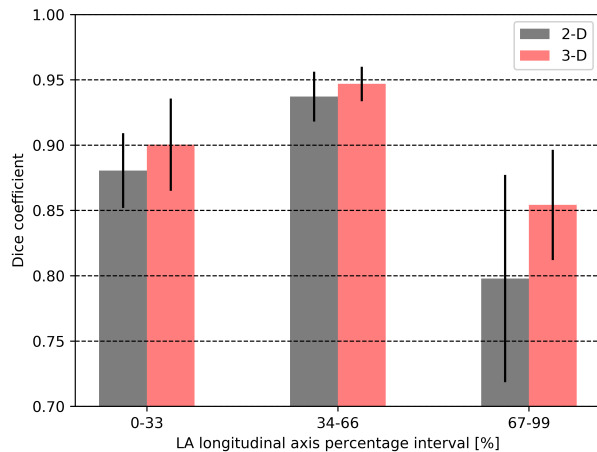In the 2-D approach the test Dice coefficient was

Fig. 3: Dice coefficient trend along the LA longitudinal axis. Dice coefficients were computed in the sub-volumes identified by the 0-33%, 34-66% and 67-99% of the longitudinal LA axis ($p < 0:05$ in each specific sub-volume between dice coefficients obtained applying the 2-D pipeline in gray and the 3-D pipeline in red).

0.896±0.025, while in the 3-D approach it resulted 0.914±0.015.

The mean Dice coefficient was computed for the three sub-volumes of interest near the mitral valve (0-33% of the LA longitudinal axis; 2-D: 0.881±0.029 and 3-D: 0.901±0.035), in the mid LA region (34-66%, 2-D: 0.937±0.019 and 3-D: 0.947±0.013) and in the region containing the PVs (67-99%, 2-D: 0.798±0.079 and 3-D: 0.854±0.042) applying both pipelines. Results are shown in Figure 3.

4. Discussion and Conclusion

The proposed dual stage segmentation approach produces a joint fast segmentation of the LA and PVs in AF patients exploiting an Otsu-based localization stage and a fine segmentation stage based on a deep neural network trained end-to-end from scratch in 2-D and 3-D. Despite the high variability of the LA anatomy, the model allows an accurate prediction that could be useful for supporting ablation therapy in both the deepened pipelines.

Our approach was completely automatic and feasible in all study subjects. Dice coefficients obtained applying the two approaches designed and implemented in this study showed a good and reliable accuracy for the LA segmentation, especially for the 3-D pipeline. We hypothesize that, applying the 3-D approach, the CNN was able to consider shape constraints which improved segmentation performances leading to Dice coefficients up to 0.939.

Future developments include the design and development of a new custom loss function and the separation of the PVs structures from the joint segmentation of LA and PVs, evaluating the performance metric solely of the LA chamber without the PVs.

References

[1] S.S. Chugh, J.L. Blackshear, W.K. Shen, et al. Epidemiology and natural history of atrial fibrillation: clinical implications, Journal of the American College of Cardiology 2001;37(2), 371-378.

[2] P.A. Wolf, R.D. Abbott, W.B. Kannel. Atrial fibrillation as an independent risk factor for stroke: the Framingham study. Stroke 1991;22(8):983-988.

[3] R. Weerasooriya, P. Khairy, J. Litalien, et al. Catheter ablation for atrial fibrillation: are results maintained at 5 years of follow-up? Journal of the American College of Cardiology 2011;57(2):160-166.

[4] A. Giannakidis, E. Nyktari, J. Keegan, et al. Rapid automatic segmentation of abnormal tissue in late gadolinium enhancement cardiovascular magnetic resonance images for improved management of long-standing persistent atrial fibrillation. Biomedical engineering online 2014;14(1):88-

[5] N.F Marrouche, D. Wilber, G. Hindricks, et al. Association of atrial tissue brosis identified by delayed enhancement mri and atrial fibrillation catheter ablation: the decaaf study. JAMA 2014;311(5):498-506.

[6] M. Valinoti, C. Fabbri, D. Turco, et al. 3D patient-specic models for left atrium characterization to support ablation in atrial fibrillation patients. Magnetic resonance imaging 2018;45:51-57.

[7] G. Yang, X. Zhuang, H. Khan, et al. Fully automatic segmentation and objective assessment of atrial scars for long-standing persistent atrial fibrillation patients using late gadolinium-enhanced MRI. Medical Physics 2018;45(4): 1562-1576.

[8] Q.Tao, E.G. Ipek, R. Shahzad, et al. Fully automatic segmentation of left atrium and pulmonary veins in late gadolinium-enhanced mri: Towards objective atrial scar assessment. Journal of Magnetic Resonance Imaging 2016;44(2):346-354.

[9] Z. Xiong, V. Fedorov, X. Fu, et al. Fully automatic left atrium segmentation from late gadolinium enhanced magnetic resonance imaging using a dual fully convolutional neural network. IEEE Transactions on Medical Imaging 2018;38(2):515-524.

[10] O. Ronneberger, P. Fischer, T. Brox. U-net: Convolutional networks for biomedical image segmentation. CoRR abs/1505.04597 (2015).

[11] K. He, X. Zhang, S. Ren, J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. CoRR abs/1502.01852 (2015).

Address for correspondence.

Davide Borra
DEI, University of Bologna
Via dell'Università 50, 47521 Cesena, Italy
E-mail: davide.borra2@unibo.it