

Left Atrial Segmentation In a Few Seconds Using Fully Convolutional Network and Transfer Learning

Élodie Puybareau¹, Zhao Zhou², Younes Khoudli¹, Edwin Carlinet¹,
Yongchao Xu^{1,2}, Jérôme Lacotte³, Thierry Géraud¹

¹ EPITA Research and Development Laboratory (LRDE), France

² Huazhong University of Science and Technology, China

³ Institut Cardiovasculaire Paris Sud, Hôpital Privé Jacques Cartier, France

elodie.puybareau@lrde.epita.fr

Abstract. In this paper, we propose a fast automatic method that segments left atrial cavity from 3D GE-MRIs without any manual assistance, using a fully convolutional network (FCN) and transfer learning. This FCN is the base network of VGG-16, pre-trained on ImageNet for natural image classification, and fine tuned with the training dataset of the MICCAI 2018 Atrial Segmentation Challenge. It relies on the "pseudo-3D" method published at ICIP 2017, which allows for segmenting objects from 2D color images which contain 3D information of MRI volumes. For each n^{th} slice of the volume to segment, we consider three images, corresponding to the $(n - 1)^{\text{th}}$, n^{th} , and $(n + 1)^{\text{th}}$ slices of the original volume. These three gray-level 2D images are assembled to form a 2D RGB color image (one image per channel). This image is the input of the FCN to obtain a 2D segmentation of the n^{th} slice. We process all slices, then stack the results to form the 3D output segmentation. With such a technique, the segmentation of the left atrial cavity on a 3D volume takes only a few seconds. We obtain a Dice score of 0.92 both on the training set in our experiments before the challenge, and on the test set of the challenge.

Keywords: 3D heart MRI · atrial segmentation · fully convolutional network.

1 Introduction

Motivation. Atrial fibrillation (AF) is the most common heart rhythm disease, corresponding with the activation of an electrical substrate within the atrial myocardium. AF is already an endemic disease, and its prevalence is soaring, due to both an increasing incidence of the arrhythmia and an age-related increase in its prevalence [3, 10, 18]. Indeed, 1–2% of the population suffer from AF at present, and the number of affected individuals is expected to double or triple within the next two to three decades both in Europe and in the USA [5].

Due to the limited effects of anti-arrhythmic drugs, AF can only be cured by percutaneous radiofrequency catheter ablation (CA) targeting triggers and

critical areas responsible for AF perpetuation in left atrium (LA). Identification and quantification of AF electrical substrate prior to AF ablation remains an unsolved issue as the number of targets remains unpredictable using clinical criterias. AF CA is still a challenging intervention requiring a perioperative 3D mapping to identify AF substrate to select the best ablation strategy [5].

Exploration of LA substrate has suggested that AF may be a self-perpetuating disease with a voltage or electrogram (EGM) amplitude reduction which is an indicator of the severity of tissue corresponding with collagen deposition in the myocardial interstitial space. Non-invasive assessment of myocardial fibrosis has proved useful as a diagnostic, prognostic, and therapeutic tool to understand and reverse AF [3, 18]. Visualization and quantification of gadolinium in late gadolinium-enhanced cardiac magnetic resonance (LGE-CMR) sequences estimate the extracellular matrix volume and have been used as a LA fibrosis surrogate, and over the last years, several groups tested the ability of LGE-CMR to detect pre-existing fibrosis [1, 10].

Although these reports suggested that the extent of fibrosis may predict recurrences after ablation procedures, the lack of 3D automated LA reconstruction, the lack of reference values for normality has prompted the publication of several image acquisition and post-processing protocols and thresholds to identify fibrosis, eventually limiting the external validation and reproducibility of this technique [9, 11, 13]. Last, let us note that, because of these technical limits, the assessment of LA fibrosis has not yet been widely adopted in the clinical practice [2].

Context. The method presented in this paper has been developed in the context of the MICCAI 2018 Atrial Segmentation Challenge⁴. The aim was to provide a fully automated pipeline for the segmentation of the LA cavity from 3D GE-MRIs without any manual assistance.

Despite the relevance of LA segmentation on GE-MRIs, this task remains manual, tedious and user-dependant. This segmentation is challenging due to the low contrast between atrial tissue and background. The development of an algorithm that can perform fully automatic atrial segmentation for the LA cavity, reconstruct and visualize the atrial structure for clinical usage would be an important improvement for patients and practitioners.

We received 100 volumes with associated masks to develop our method. A set of 25 new volumes, released 2 weeks before the end of the challenge, has been used to evaluate our method by comparing our segmentation results with the manual segmented masks (these masks will not be released). This challenge will establish a fair comparison to state-of-the-art methods, while providing the community with a large dataset of 3D GE-MRI heart images.

Main related work. As machine learning really improved the results of some segmentation tasks, the use of such strategy seems meaningful in the context of medical image segmentation. In a work published in the IEEE Intl. Conf.

⁴ <http://atriaseg2018.cardiacatlas.org/>

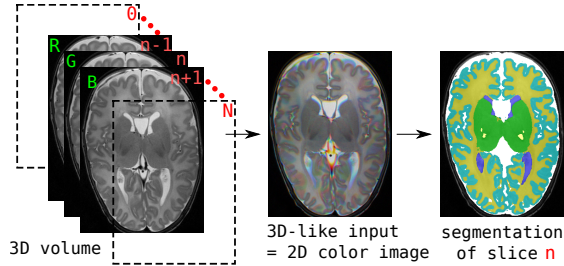


Fig. 1. Illustration of the main idea used in [15]: a segmentation of a 3D medical image is performed on a set of 2D color images. For the n^{th} slice, a 2D color image (middle) is formed from the triplet of slices $n-1, n, n+1$ (left) of the original volume, and the segmentation (right) is performed on this 2D color image. Actually, such a **2D** image contains **3D** information thanks to the 3 color channels (R, G, B); so, this approach is called “**pseudo-3D**”.

on Image Processing (ICIP) in 2017 [15], 3D brain MR volumes are segmented using fully convolutional network (FCN) and transfer learning. The network used for transfer learning is VGG (Visual Geometry Group) [14], pre-trained on the ImageNet dataset. It takes as input a 2D color image, formed from 3 consecutive slices of the 3D volume (see Fig. 1). Thus, we have a 2D image containing 3D information; this approach is called “pseudo-3D”. This method only relies on a single modality, and obtains good results for brain segmentation. For the left atrial segmentation challenge, we use the same approach, leveraging the power of a fully convolutional network pre-trained on a large dataset and later fine-tuned on the training set provided for the challenge.

2 Description of the Proposed Method

An overview of the proposed method is given in Fig. 2. The method is fully automatic, and takes 2D color images as input. The method is very computationally efficient, as it processes each complete scan volume in less than 2 seconds.

2.1 Pre-Processing

The study of the histograms of the training volumes shows a high variability among the different volumes, as depicted in Figure 3(a). A clue to improve the network results is to normalize the input volumes according to their histograms. To that aim, for each volume, we compute the histogram of the central sub-volume ($1/4$ of the pixels in all three dimensions, so containing 4^3 times less voxels than the whole volume; see Figure 4(c)). Let m be the maximum gray-level having a non-null histogram value. We requantize all voxel values using a linear function so that the gray-level range $[0, m]$ is mapped to $[0, 255]$. For our application, the question amounts to how to prepare appropriate inputs, RGB input images, given that a heart MR image is a 3D volume. To that aim, we

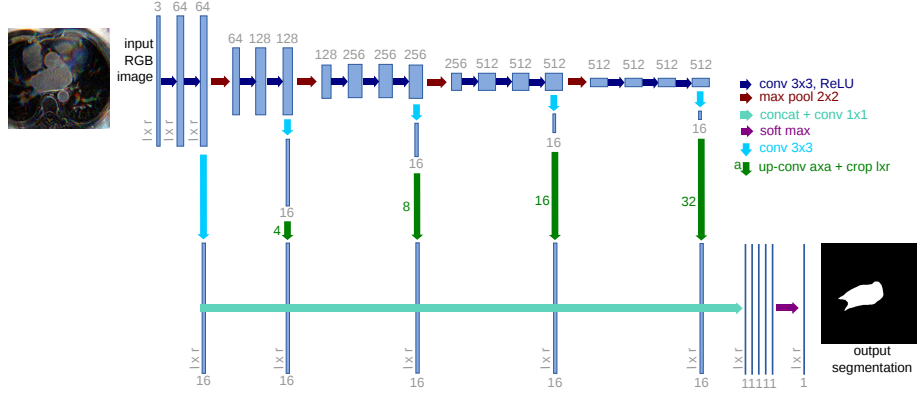


Fig. 2. Architecture of the proposed network. We fine tune it and combine linearly fine to coarse feature maps of the pre-trained VGG network [14]. Note that each input 2D color image is built from the slice n and its neighbouring slices $n - 1$ and $n + 1$.

propose to stack successive 2D slices. Precisely, to form an input artificial color image for the pre-trained network to segment the n^{th} slice, we use the slice n of the volume, its predecessor, the $(n - 1)^{\text{th}}$ slice, and its successor, the $(n + 1)^{\text{th}}$ slice, as respectively the green, red and blue channels (note that the order is not relevant). This process is depicted in Fig. 2 (left). Each 2D color image thus forms a representation of a part (a small volume) of the MR volume.

2.2 Deep FCN for Left Atrial Segmentation

Fully convolutional network (FCN) and transfer learning has proved their efficiency for natural image segmentation [7]. The paper [15] proposed to rely on a FCN and transfer learning to segment 3D brain MR images, although those images are very different from natural images. As it was a success, we adapted it to LA segmentation. We rely on the 16-layer VGG network [14], which was pre-trained on millions of natural images of ImageNet for image classification [6]. For our application, we keep only the 4 stages of convolutional parts called “*base network*”, and we discard the fully connected layers at the end of VGG network. This base network is mainly composed of convolutional layers: $z_i = w_i \times x + b_i$, Rectified Linear Unit (ReLU) layers for non-linear activation function: $f(z_i) = \max(0, z_i)$, and max-pooling layers between two successive stages, where x is the input of each convolutional layer, w_i is the convolution parameter, and b_i is the bias term. The three max-pooling layers divide the base network into four stages of fine to coarse feature maps. Inspired by the work in [7, 8], we add specialized convolutional layers (with a 3×3 kernel size) with K (e.g. $K = 16$) feature maps after the convolutional layers at the end of each stage. All the specialized layers are then rescaled to the original image size, and concatenated together. We add a last convolutional layer with kernel size 1×1 at the end. This last layer combine linearly the fine to coarse feature maps in

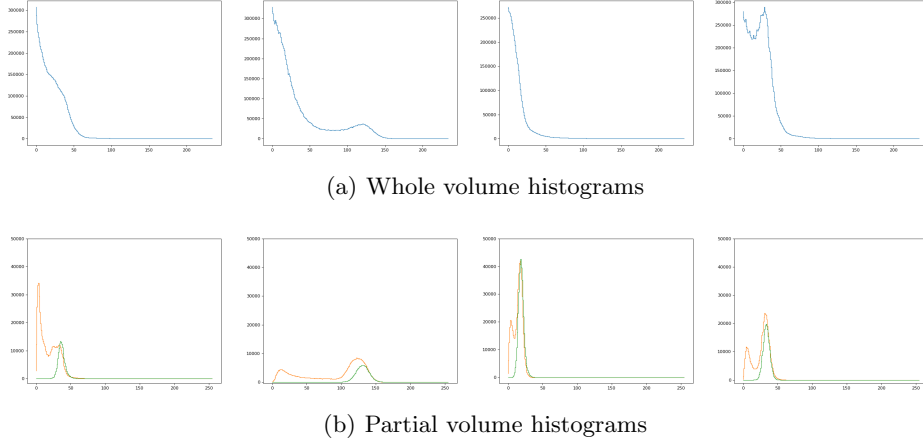


Fig. 3. (a): The histograms of the original volumes have various shapes; (b): to normalize the gray-level scale of each volume, we consider the histogram of their central sub-volume (in orange; see also Fig. 4(c)), which has the same dynamic than the one of the left atrial region given by the ground-truth (in green).

the concatenated specialized layers, and provide the final segmentation result. The proposed network architecture is schematized in Fig. 2.

The architecture described above is very similar with the one used in [8] for retinal image analysis, where the retinal images are already 2D color images. Using such a 2D representation avoids the expensive computational and memory requirements of fully 3D FCN.

For the training phase, we use the multinomial logistic loss function for a one-of-many classification task, passing real-valued predictions through a softmax to get a probability distribution over classes. We rely on the ADAM optimization procedure [4] (AMSGrad variant [12]) to minimize the loss of the network. The relevant parameters of the methods are the following: the learning rate is set to 0.002 (we did not use learning rate decay), the beta_1 and beta_2 are respectively set to 0.9 and 0.999, and we use a fuzz factor (epsilon) of 0.001. Each batch is composed of 12 centered images (for each channel we subtract 127 to values to ensure that input values are within the $[-127, 127]$ range). We trained the network for 10 epochs.

At test time, after having pre-processed the 3D volume (requantization), we prepare the set of 2D color images. Then we subtract 127 for each channel, and pass every image through the network.

We run the train and test phases on an NVIDIA GPU card. The training phase lasts about 50min, while the testing lasts less than 2 seconds, 1.78 second in average. These reasonable computational times allow a re-training if new scans are available, and allow clinical use. To this running time, 1 second is added for pre- and post-processing.

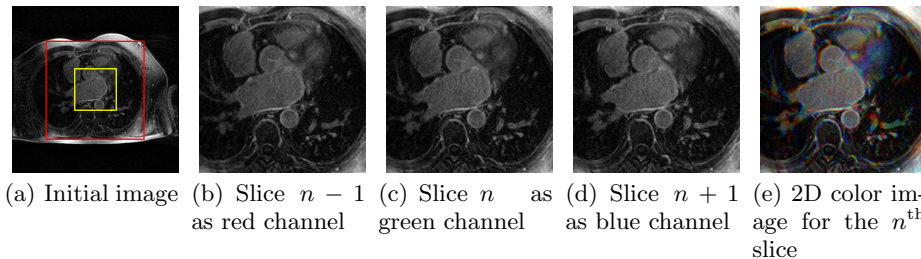


Fig. 4. The yellow box in (a) depicts the sub-volume used for the gray-level normalization stage, and the red box depicts the boundary of the cropped input image. Three successive cropped slices (b-d) are used to build a 2D color image (e).

2.3 Post-processing

The output of the network for one slice during the inference phase is a 2D segmented slice. After treating all the slices of the volume, all the segmented slices are stacked to recover a 3D volume with the same shape as the initial volume, and containing only the segmented lesions. We clean up the segmentation keeping only the largest component. We regularized the 3D segmentation result using a 5x5 median filter in the sagittal axis. This step smoothes the segmentation and erases potential inconsistencies.

3 Experiments and Results

3.1 Experiments and results for the training phase

To experiment, we only relied on the 100 training scans provided by the challenge.

Our model has been trained on 80 scans, chosen randomly, using the parameters described in the previous section. Then we have tested our model on the 20 remaining scans. This procedure have been repeated 5 times to cross-validate the results.

Our preliminary results on the training dataset without any normalization showed a dice score of 0.86 after 10 epochs. We performed two 5-fold. With the help of normalization, the dice score reaches 0.9 ± 0.02 (under cross-validation), after 10 epochs, and with an average of 0.92. Note that these results were obtained without any training data augmentation.

The evolution of the dice score according to the number of epochs (Fig. 5) shows that the network converges quickly and, actually, 4 epochs are enough to reach a dice score of 0.90.

Some qualitative results can be found in Fig. 6.

3.2 Results of the challenge

Our method reaches a top rank during the challenge, with a Dice score of 0.923. The best score was 0.932 and the second best 0.926.

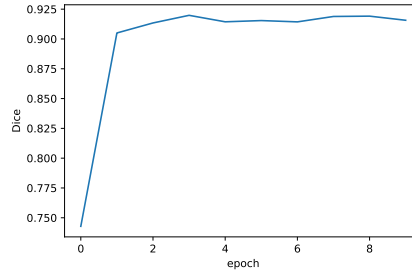


Fig. 5. Evolution of the dice score with the number of epochs.

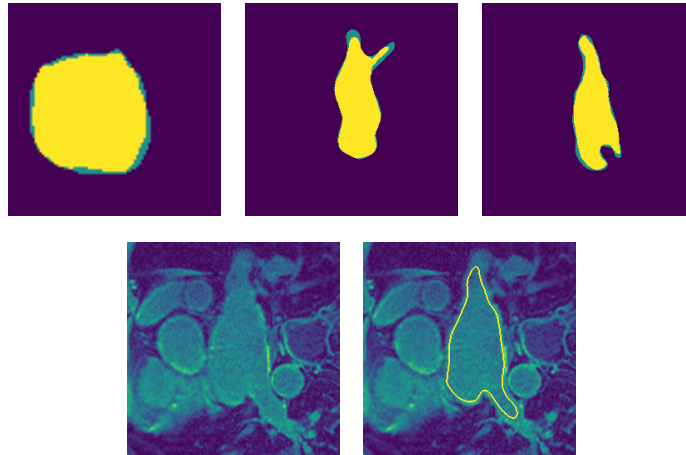


Fig. 6. Examples of segmentation results. First row: superposition between our result and GT (wrongly labeled pixels are in green); second row: an MRI slice (left) and our segmentation superimposed over it (right).

4 Conclusion

In this article, we proposed a method to accurately segment the left atrial cavity in few seconds based on transfer learning from VGG-16, a pre-trained network used to classify natural images. This method takes the advantage of keeping 3D information of the MRI volume and the speed of processing only 2D images, thanks to the pseudo-3D concept. An entire 3D volume is processed in less than 3 seconds. We were ranked 3^{rd} during the challenge.

This method can also deal with multi-modality, and can be applied to other segmentation problems, such as in [17], where a similar method is proposed to segment white matter hyperintensities, but pseudo-3D has been replaced by an association of multimodality and mathematical morphology pre-processing to improve the detection of small lesions. We might also try to modify our inputs

thanks to some highly non-linear filtering to help the network segment the LA, more specifically using some mathematical morphology operators [16].

The strength of this method is its modularity and its simplicity. It is easy to implement, fast, and does not need a huge amount of annotated data for training (in our work on brain segmentation [15], only 2 or 4 images were used for training).

Acknowledgments The authors want to thank the organizers of the Atrial Segmentation Challenge. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Quadro P6000 GPU used for this research.

References

1. Ambale-Venkatesh, B., Lima, J.: Cardiac MRI: a central prognostic tool in myocardial fibrosis. *Nat Rev Cardiol* 12, 18–29 (2015)
2. Benito, E.M., et al.: Left atrial fibrosis quantification by late gadolinium-enhanced magnetic resonance: a new method to standardize the thresholds for reproducibility. *EP Europace* 19(8), 1272–1279 (2016)
3. Hansen, B.J., et al.: Atrial fibrillation driven by micro-anatomic intramural re-entry revealed by simultaneous sub-epicardial and sub-endocardial optical mapping in explanted human hearts. *European Heart Journal* 36(35), 2390–2401 (2015)
4. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. *CoRR* abs/1412.6980 (2014)
5. Kirchhof, P., et al.: ESC guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 37, 2893–2962 (2016)
6. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. pp. 1097–1105 (2012)
7. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*. pp. 3431–3440 (2015)
8. Maninis, K.K., Pont-Tuset, J., Arbeláez, P., Gool, L.V.: Deep retinal image understanding. In: *Proceedings of International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Part II. Lecture Notes in Computer Science*, vol. 9901, pp. 140–148. Springer (2016)
9. Marrouche, N., et al.: Association of atrial tissue fibrosis identified by delayed enhancement MRI and atrial fibrillation catheter ablation: the DECAAF study. *JAMA* 311, 498–506 (2014)
10. McGann, C., et al.: Atrial fibrillation ablation outcome is predicted by left atrial remodeling on MRI. *Circulation: Arrhythmia and Electrophysiology* pp. CIRCEP–113 (2013)
11. Oakes, R.S., et al.: Detection and quantification of left atrial structural remodeling with delayed-enhancement magnetic resonance imaging in patients with atrial fibrillation. *Circulation* 119(13), 1758–1767 (March 2009)
12. Reddi, S.J., Kale, S., Kumar, S.: On the convergence of adam and beyond. In: *International Conference on Learning Representations* (2018)
13. Seitz, J., Horvilleur, J., Lacotte, J., et al.: Correlation between AF substrate ablation difficulty and left atrial fibrosis quantified by delayed-enhancement cardiac magnetic resonance. *Pacing Clin Electrophysiol* 34(10), 1267–1277 (2011)

14. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR abs/1409.1556 (2014)
15. Xu, Y., Géraud, T., Bloch, I.: From neonatal to adult brain MR image segmentation in a few seconds using 3D-like fully convolutional network and transfer learning. In: Proceedings of the 23rd IEEE International Conference on Image Processing (ICIP). pp. 4417–4421. Beijing, China (September 2017)
16. Xu, Y., Géraud, T., Najman, L.: Connected filtering on tree-based shape-spaces. IEEE Transactions on Pattern Analysis and Machine Intelligence 38(6), 1126–1140 (June 2016)
17. Xu, Y., Géraud, T., Puybureau, E., Bloch, I., Chazalon, J.: White matter hyperintensities segmentation in a few seconds using fully convolutional network and transfer learning. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries— 3rd International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Revised Selected Papers, Lecture Notes in Computer Science, vol. 10670, pp. 501–514. Springer (2018)
18. Zhao, J., et al.: Three dimensional integrated functional, structural, and computational mapping to define the structural ”fingerprints” of heart specific atrial fibrillation drivers in human heart ex vivo. Journal of the American Heart Association 6(8), 5922 (2017)