Performance Comparison of Deep Learning Approaches for Left Atrium Segmentation From LGE-MRI Data

Davide Borra[1], Daniela Portas[1], Alice Andalò[1], Claudio Fabbri[1], Cristiana Corsi[1]

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

Abstract

Quantification of viable left atrial (LA) tissue is a reliable information which should be used to support therapy selection in atrial fibrillation (AF) patients. Late gadolinium-enhanced magnetic resonance imaging (LGE-MRI) is employed for the non-invasive assessment of LA fibrotic tissue. Unfortunately, the analysis of LGE-MRI relies on manual tracing of LA boundaries. This task is time-consuming and prone to high inter-observer variability. Therefore, an automatic approach for LA wall detection would be very helpful. In this study, we compared the performance of different deep architectures – U-Net and attention U-Net (AttnU-Net) – and different loss functions - Dice loss (DL) and focal Tversky loss (FTL) to automatically detect LA boundaries from LGE-MRI data. In addition, AttnU-Net was trained without deep supervision (DS) and multi-scale inputs (MI), with DS and with DS+MI. No statistically significant differences were found training the networks with DL or FTL. U-Net was the best-performing algorithm overall, outperforming significantly AttnU-Net with a Dice Coefficient of 0.9015±0.0308 (mean ± standard deviation). However, no significant differences were found between U-Net and AttnU-Net DS/DS+MI. Based on these results, using a DL or FTL does not affect the performance and U-Net was the best-performing solution.

1. Introduction

Atrial fibrillation (AF) is the most common arrhythmia in the western world [1]. Consequences of AF could lead to a notable reduction in quality of life and, mainly, an increase of stroke risk by five-fold [2].

Radio frequency ablation (RFA) of the left atrium (LA) represents the clinical therapy for AF patients in which antiarrhythmic drugs and direct current cardioversion do not provide improvements. However, despite strong improvements for the targeting and the delivery of AF RFA, the long-term restoration of sinus rhythm is achieved only in a limited percentage of AF patients [3]. These results suggest that there is room for improvements in RFA treatment.

Information related to scarred and non-scarred atrial tissue may be of great importance to select the best AF treatment as well as to predict AF recurrence. Magnetic resonance imaging (MRI) can differentiate between scarred and non-scarred atrial wall by using late gadolinium enhancement (LGE) imaging. Unfortunately, the analysis of LGE-MRI is based on a time-consuming procedure of manual tracing of LA wall and PVs [5], thus reducing its use in clinical practice. In addition, results are affected by high variability among experts and low reproducibility in multicenter studies.

Deep learning is a branch of machine learning that received particular attention in computer vision applications, especially when handling images and time series [6]. Recently, these techniques were largely applied not only to physiological signals (such as electroencephalography, electromyography and electrocardiography) [7-9] but also to MRI data [10] to design detection, classification, reconstruction and segmentation algorithms. Convolutional neural networks (CNNs) were successfully applied to automatically segment biological structures. A successful CNN for semantic segmentation is U-Net [11], a fully convolutional architecture composed by a contracting and an expansion paths with multi-scale skip connections that has become the de facto standard for image segmentation. Class-imbalance is an important aspect in image segmentation to deal with. The Tversky loss function was proposed to alleviate class-imbalance but it struggles to balance precision and recall due to small regions-of-interest (ROI) in medical images. To overcome this limitation, the focal Tversky loss function was designed by Abraham et al. [12] and tested on lesion segmentation. Furthermore, efforts were made to segment small ROI by designing more discriminative models such as CNNs with attention gates (AGs) [13] allowing the model to focus on the target region related to the task. Among the approaches proposed in literature to automatic segment LA boundaries from LGE-MRI, U-Net is commonly used achieving outstanding results [14-16].

The aim of this study was to delve into three main methodological aspects when designing an automatic algorithm based on a fully convolutional CNN for LA

 ISSN: 2325-887X DOI: 10.22489/CinC.2020.306

segmentation from LGE-MRI. These were: i) compare U-Net with its variant including AGs [13]; ii) compare the dice loss, a traditional loss function used for image segmentation, with the focal Tversky loss proposed in [12] to improve precision and recall balance; iii) study the effect of deep supervision (DS) and multi-scale inputs (MS) on the performance of AttnU-Net as done in [12] in a different objective task.

2. Methods

2.1. Dataset

Experiments were conducted on the 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 ground truth segmentations obtained by manually tracing the LA endocardial wall. The resolution is 0.625x0.625x0.625 mm^3 and the images are composed by 88 axial slices with in-plane size of 576x576 or 640x640 pixels. Five-fold cross-validation was performed and early stopping was applied using a validation set of 10% extracted from the training set in each fold.

2.2. Two-stage segmentation approach

All the approaches were based on a two-stage segmentation. The first stage was devoted to reducing the total number of pixels and therefore, the computational cost of the proposed algorithms. This was used also in our past studies [15,16] and was accomplished by applying the Otsu's algorithm to the central slice of each 3-D LGE-MRI data. Once the binary image resulting from this stage was obtained, the centroid of the region located in the centre of the image was automatically extracted; the limits of the region of interest were automatically computed and a 3-D crop centred in the LA of fixed size of 88x320x384 pixels was extracted. Each 3-D LGE-MRI image was then then subsampled to 192x240 in the axial plane to further reduce the computational cost. Within each fold, the axial slices were collected and used to train, test and validate the 2-D models using 6336, 1760 and 704 examples, respectively.

The second stage was devoted to providing the fine segmentation from the LA-centred crops as obtained from the previous stage. This was performed using U-Net and

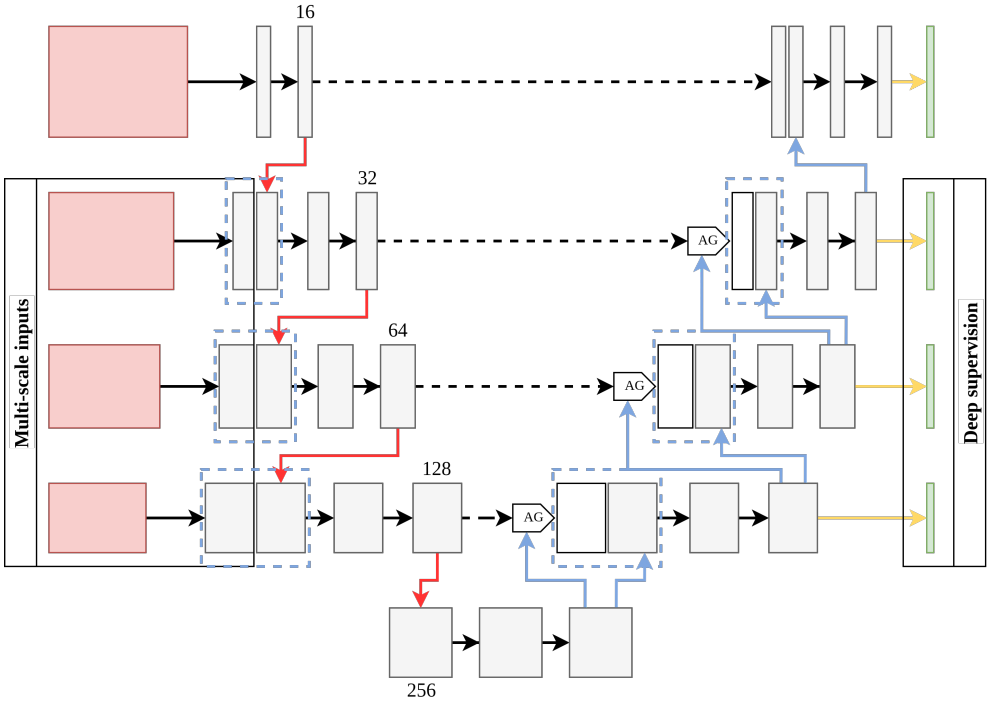

Figure 1: Schematization of the architectures. Feature maps are represented as gray and white boxes (AG outputs). Blue dashed boxes represent concatenation operators. Black arrows denote 3x3 convolution + batch normalization + ReLU non-linearity, red arrows 2x2 max pooling, blue arrows 2x2 transposed convolution, black dashed arrows skip connection and yellow arrows 1x1 convolution + sigmoid non-linearity. From the scheme reported in the figure, the U-Net architecture result without considering MS, DS and AGs, AttnU-Net without considering MS and DS and AttnU-Net+DS without considering MS. Lastly, the whole scheme corresponds to AttnU-Net+DS+MI.

AttnU-Net with the hyper-parameters set as in [12,16]. At the deepest encoding level of the contractive path, the model retains the richest feature representation. However, the cascade of convolutions and non-linearities is detrimental for the spatial resolution leading to wrong detection when small objects with high morphology variability are processed. Attention gates mitigate this issue, identifying relevant spatial information from low-level feature maps and propagating them up to the decoding stage. The structure of the AGs adopted in this study was the one proposed by Oktay et al. [13]. The coefficients α_i produced by the AGs for each pixel i-th, scale the input feature maps to output semantically relevant features. Lastly, the design of AttnU-Net as proposed in [12] was modified only in its number of feature maps learned at each scale, in order to match the one used in our previous study [16] using U-Net and the main hyper-parameters are summarized in Figure 1. In addition, the supervision of AttnU-Net was modified by including the combination of deep supervision, computing the loss also from 2-D probability distributions at lower spatial scales, and multiple-scale inputs, providing the input images at the different spatial scales to the encoder, as this was found to be beneficial when segmenting small ROIs [12]. Thus, this resulted in 2 additional architectures, namely AttnU-Net+DS and AttnU-Net+DS+MI.

2.2. Optimization

The optimization of the three architectures was driven alternatively by 2 different loss functions. These were the dice loss (DL) and the focal Tversky loss (FTL) function.

The dice coefficient (DC) is an overlap index widely used to evaluate segmentation maps. The 2-class DC can be computed as (Equation 1):

$$DC_c = \frac{2\sum_{i=1}^{N} p_{ic} g_{ic}}{\sum_{i=1}^{N} p_{ic} + g_{ic} + \epsilon} \quad (1)$$

where c is the c-th class, the index i runs over all pixels (N in total), $g_{ic} \in \{0,1\}$ is the ground truth value, $p_{ic} \in [0,1]$ is the prediction value and ϵ is added for numerical stability. Thus, the DL can be computed as $DL = \sum_c 1 - DC_c$. The DC has two main limitations: i) it equally weights false positive (FP) and false negative (FN) detections, corresponding to predicted segmentations with high precision and low recall; ii) it struggles to segment small ROIs due to their small contribution to the loss function.

The Tversky similarity index (TI) is a generalization of the DC that enables flexibility balancing FPs and FNs (Equation 2):

$$TI_c = \frac{\sum_{i=1}^{N} p_{ic} g_{ic}}{\sum_{i=1}^{N} p_{ic} g_{ic} + \alpha \sum_{i=1}^{N} p_{i\bar{c}} g_{ic} + \beta \sum_{i=1}^{N} p_{ic} g_{i\bar{c}} + \epsilon} \quad (2)$$

where $\bar{c}$ indicates the complementary class in the 2-class scenario. The hyper-parameters α and β can be tuned to change the recall improvement in case of large class-imbalance. If $\alpha = \beta = 0.5$, TI corresponds to DC; in this study, we adopted $\alpha = 0.7$ and $\beta = 0.3$ as in [12]. Thus, the Tversky loss function can be computed as $TL = \sum_c 1 - TI_c$. In order to overcome the point ii), Abraham et al. [12] proposed the FTL, which is the TL parametrized by $\gamma \in [1,3]$, $FTL = \sum_c (1 - TI_c)^{1/\gamma}$. In this study, we used $\gamma = 4/3$ as this value was found optimal in [12].

Lastly, Adam was used as optimizer with a learning rate of 10^{-3}, a maximum number of training epochs of 100 and a batch size of 32.

3. Results

All the experiments were conducted using Keras as framework to build the deep neural networks and were accelerated using the free resources offered by the Google Colaboratory project.

An example of the detected LA boundaries is shown in Figure 2; the gold standard (blue) and the predicted (red) LA boundaries as obtained with the 4 architectures are reported in Figure 2 for the same representative example.

On average, training CNNs using FTL or DL resulted in a DC of 0.8897 and 0.8912, respectively (p>0.05, Wilcoxon signed-rank test). In addition, paired Wilcoxon signed-rank test were performed between U-Net and its variants based on AGs, applying Bonferroni correction for multiple comparison. Results are presented in Table1.

Table 1. DC scored with the 4 different architectures trained with the FTL and DL. *p<0.05, Wilcoxon signed-rank test corrected for multiple comparison.

Architecture	DL	FTL
U-Net	**0.9015±0.0308**	**0.8941±0.0444**
AttnU-Net	0.8906±0.0521*	0.8855±0.0386*
AttnU-Net+DS	0.8985±0.0615	0.8923±0.0389
AttnU-Net+DS+MI	0.8945±0.0323	0.8931±0.0312

3. Discussion and conclusion

All the approaches investigated provided fast joint segmentations of LA and PVs in patients with AF, exploiting a dual-stage segmentation algorithm with a Otsu-based localization stage and a CNN-based fine segmentation stage. Despite the variability of the LA morphology, all the approaches scored accurate prediction of LA boundaries. Prospectively, these solutions could be useful to support ablation therapy in terms of (1) making available an accurate patient specific

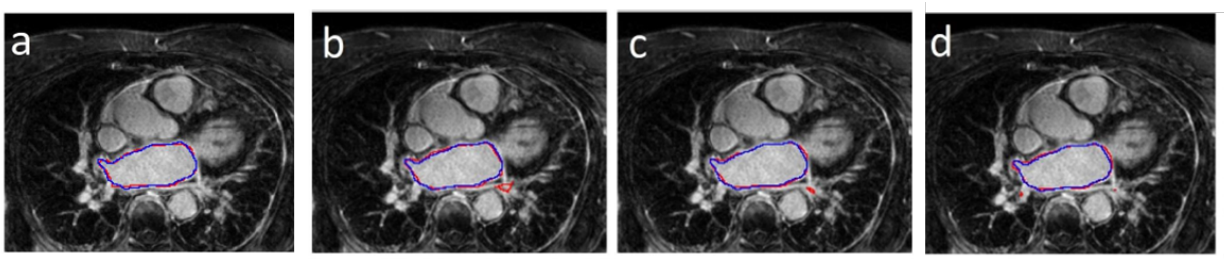

Figure 2: Example of the detected LA boundaries in one mid LA slice applying the different architectures (blue: reference contours; a: U-net; b: AttnU-Net; c: AttnU-Net+DS; d: AttnU-Net+DS+MI).

anatomical model and (2) as a first step for fibrosis quantification on the LA wall. From the results obtained in the performed experiments, the use of the improved TL did not change significantly the performance in our target decoding task. U-Net resulted the best-performing architecture both with FTL and DL, but statistical significance was found only respect to the baseline AttnU-Net architecture. The variants AttnU-Net+DS and AttnU-Net+DS+MI showed comparable performance with U-Net and this can be also observed in the predicted LA boundaries reported in Figure 2 for a representative input. In the future, the comparison of these approaches will be extended using also other metrics and using 3-D architectures.

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 Fibrosis Identified by Delayed Enhancement MRI and Atrial Fibrillation Catheter Ablation: the Decaaf Study. JAMA 2014; 311(5):498-506.

[6] Y. LeCun, Y. Bengio, G. Hinton. Deep Learning. Nature 2015; 521:436-444.

[7] D. Borra, S. Fantozzi, E. Magosso. Interpretable and Lightweight Convolutional Neural Network for EEG Decoding: Application to Movement Execution and Imagination. Neural Networks 2020; 129:55-74.

[8] M. Simoes, D. Borra, E. Santamaria-Vázquez, et al. BCIAUT-P300: a Multi-Session and Multi-Subject Benchmark Dataset on Autism for P300-based Brain-Computer-Interfaces. Frontiers in Neuroscience 2020; 14.

[9] O. Faust, Y. Hagiwara, T. J. Hong, et al. Deep Learning for Healthcare Applications Based on Physiological Signals: A Review. Computer Methods and Programs in Biomedicine 2018; 161:1-13.

[10] M. A. Mazurowski, M. Buda, A. Saha, et al. Deep learning in radiology: An Overview of the Concepts and a Survey of the State of the Art with Focus on MRI. J Magn Reson Imaging 2019; 49(4):939-954.

[11] O. Ronneberger, P. Fischer, T. Brox. U-net: Convolutional Networks for Biomedical Image Segmentation. arXiv preprint 2015: 1505.04597.

[12] N. Abraham, N. M. Khan. A Novel Focal Tversky Loss Function with Improved Attention U-Net for Lesion Segmentation. IEEE 16th International Symposium on Biomedical Imaging 2019; 683-687.

[13] O. Oktay, J. Schlemper, L. L. Folgoc, et al. Attention U-net: Learning Where to Look for the Pancreas. arXiv preprint 2018: 1804.03999.

[14] Z. Xiong, Q. Xia, Z. Hu, et al. A Global Benchmark of Algorithms for Segmenting the Left Atrium from Late Gadolinium-Enhanced Cardiac Magnetic Resonance Imaging. Medical Image Analysis 2020; In Press.

[15] D. Borra, A. Masci, E. Lorena, et al. A Semantic-Wise Convolutional Neural Network Approach for 3-D Left Atrium Segmentation From Late Gadolinium Enhanced Magnetic Resonance Imaging. Springer Lecture Notes in Computer Science 2019; 11395:329-338.

[16] D. Borra, A. Andalò, M. Paci, et al. A Fully Automated Left Atrium Segmentation Approach From Late Gadolinium Enhanced Magnetic Resonance Imaging Based on a Convolutional Neural Network. Quant Imaging Med Surg 2020; 10(10): 1894-1907.

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