EasyChair Preprint

№ 11574

Deep Learning Models for Left Atrial Segmentation in MRI

Meriem Triki and Mohammed Ammar

December 19, 2023

Deep Learning Models for Left Atrial Segmentation in MRI

Meriem Triki
Engineering Systems and Telecommunication Laboratory
M'Hamed Bougara University
Boumerdes, Algeria
m.triki@univ-boumerdes.dz

Mohammed Ammar
Engineering Systems and Telecommunication Laboratory
University M'Hamed Bougara
Boumerdes, Algeria
m.ammar@univ-boumerdes.dz

***Abstract*— The left atrium (LA) segmentation is a crucial and essential procedure in the field of cardiac imaging due to its significance in cardiovascular health and its role in diagnosing heart conditions. Magnetic resonance imaging (MRI) is a non-invasive medical imaging technique that has become indispensable in the cardiovascular field, especially for visualizing and evaluating atrial myopathy. In recent years, deep learning has emerged as the approach with the highest adoption rate for segmenting cardiac images. For the purpose of segmenting the left atrium from MRI images, we applied three Unet variation architectures to compare the optimal one, and we experimented with dice and cross-entropy loss.**

Keywords*— left atrial, segmentation , cardiac MRI, deep learning, U-net.

I. Introduction

The left atrium is a vital component of the heart's structure and holds significant importance in the fields of cardiology and cardiac health. Due to the intricate variations in its structure among patients, accurately understanding it becomes challenging, often resulting in misdiagnosis and ineffective treatment [1].

Left atrial segmentation identifies and delineates the boundaries of the left atrium in medical imaging, typically using techniques from image analysis and computer vision. This segmentation task holds significant importance in cardiac imaging, as it allows precise quantification of the left atrium's size and shape, which can be critical in diagnosing and monitoring various heart conditions including atrial fibrillation, mitral valve regurgitation, and other cardiac diseases.

Noninvasive medical imaging became essential in cardiovascular medicine. The exceptional measurement precision offered by MRI positions it as a perfect tool for monitoring the advancement and treatment outcomes of cardiovascular diseases [1].

Existing models of left atrial segmentation are based on encoder-decoder architectures like FCN [2] and U-Net [3]. For instance, [4] segmented the left atrial from magnetic resonance (MR) images with two consecutive 3D U-Net architectures. The initial 3D U-Net aimed to identify and coarsely extract the region of interest, while the second network focused on performing detailed segmentation from the cropped images. [5] devised an automated left atrial segmentation technique utilizing a FCN to enhance the accuracy of left atrial structural delineation. This method integrates a dual-path, multi-scaled architecture capturing both local atrial tissue geometry and the broader positional information of the left atrium. [6] proposed a deep 2D U-Net, a derivation of the original U-Net, was introduced for automatic left atrium segmentation from GE-MRIs. This method employed multi-task learning by augmenting the network depth and incorporating an additional classification branch.[7] modified the 3D U-Net [4] by incorporating a hierarchical aggregation unit (HAU) serving as a trunk branch and an attention unit (AU) in the encoder path for 3D left atrial segmentation. To address the segmentation issues in GE-MRI for the left atrium, [8] created two image segmentation networks using FCN and U-Net architectures. They improved the segmentation accuracy by modifying the dice loss function, aiming to reduce the impact of imbalances between positive /negative samples. [9] constructed a modified 3D U-Net by integrating dilated convolutions at the deepest layer of the encoder and incorporating residual connections, aiming to merge and capture local and global information. [10] proposed a two-stage method for LA segmentation from LGE MRI. The method involves an Otsu-based localization phase followed by fine segmentation, utilizing both 2-D and 3-D pipelines which is based on the original U-Net architecture. [1] developed a method for automatic left atrial region segmentation in MRI scans of patients with LAE. This method comprises two parts: the first one, a U-Net architecture with Gaussian blur and channel weight neural network (GCW-UNet) segments the left atrial region. The subsequent part reconstructs the 2D segmented left atria into a 3D model.

The achievements of these studies demonstrate that designed U-Net-based architectures can perform satisfactorily on tasks like left atrial segmentation. There are, however, several U-Net versions, including Residual U-Net [11], Dense U-Net [12], V-Net [13], Attention U-Net

[14], U-Net++ [15], SegResNetVAE [16], to mention a handful. Due to the various U-Net architectures available, selecting the optimal one is challenging.

In this study, we compare the U-net variation for left atrial segmentation. We tested 3DU-Net [17], Attention U-Net [14], and SegResNet [16] variants and experimented with the Dice and Cross Entropy loss function.

II. Method

A. Dataset

The dataset employed in this work from Left Atrial Segmentation Challenge Dataset (LASC2013) [18], sourced from STACOM'13 in MICCAI'13. This dataset comprises 30 MRI images in NIFTI format, with 20 images designated as the training set, accompanied by manually labeled segmentation images. The remaining images constitute the test set, lacking manual segmentation labels. Each sample contained 100–130 slices. Therefore, we confined the dataset to the 20 MRI images with labels, partitioned into 15 images for training purposes and 5 images for validation.

B. Data Preprocessing

As an initial step, all files were stacked in such a way that each example has the shape (320×320×110) and the same-sized segmentation mask. Then, on the borders of each volume, redundant background voxels are cropped. This has no effect on the dataset because it provides no significant information, which the network could ignore, but it does reduce its size and hence its computation burden.

C. Model Architectures

To determine the optimal U-Net variation architecture, we studied some models, including 3D-UNet [17], Attention-UNet [14], and SegResNet [16]. Here, we present brief descriptions of each of these models.

3D U-Net [17] (shown in Fig. 1) is composed of two essential stages: a contracting encoder, whose primary purpose is to examine the entire image, and a full-resolution segmentation produced by a successively expanding decoder. A rectified linear unit (ReLu) and two 3×3×3 convolutions are present in every encoder layer. Following each of them, a max pooling layer of 2×2×2 is added. In a similar manner, two 2×2×2 up-convolutions, two 3×3×3 convolutions, and ReLu comprise the decoder stage. To suit the amount of labels, the final convolution layer 1×1×1 decreases the output channels [17].

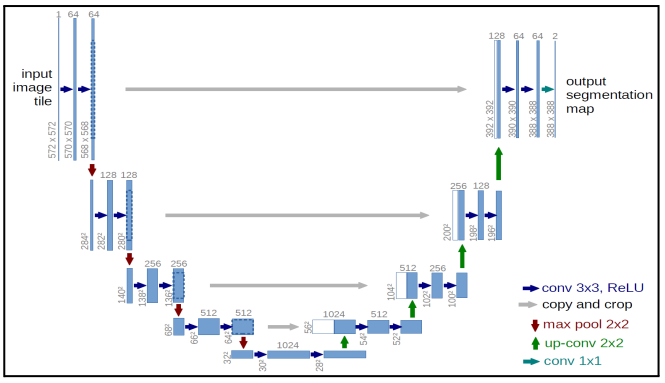

Fig. 1. The 3D U-Net architecture.

Attention U-Net [14] expands the basic U-Net by incorporating an attention gate (seen in Fig. 2) in the decoder section to emphasize important features that move through the skip connections. Prior to the concatenation process, the attention gate modifies the encoder's feature map to merge only the relevant activations. It figures out which parts of the encoder's feature map are most significant, taking cues from the contextual information provided by the previous decoder block's feature map. This is accomplished by multiplying the weight values determined by the attention gate by the encoder feature map. These weight values, which are limited to the interval (0, 1), indicate the degree of focus that the neural network is applying to a certain pixel [14].

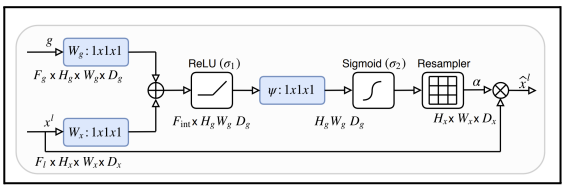

Fig. 2. The attention gate's architecture. Attention weights (α) get multiplied by input features (xl). To calculate α, the feature map from the associated encoder level and the input features (xl) are first processed by 1x1x1 convolution and then added together. After that, a further 1x1x1 convolution and ReLU activation are used. Finally, trilinear interpolation is used to upsample attention weights [14].

SegResNet is based on **[16]** but the variational autoencoder (VAE) mentioned in the article is not included in the module. The method relies on a CNN architecture featuring an encoder-decoder design. It employs an asymmetrical larger encoder to capture image features and a smaller decoder for reconstructing the segmentation mask (shown in Fig. 3). In the encoder, ResNet is implemented as blocks. Each block comprises two 3x3x3 convolutions accompanied by normalization and ReLU activation, followed by an additive identity skip connection. The decoder follows a similar structure to the encoder, but it contains a single block for each spatial level. Each decoder level commences

by decreasing the feature count by a factor of two using 1x1x1 convolutions while simultaneously increasing the spatial dimension utilizing 3D bilinear operations, then adding encoder output from the same spatial level. The resulting decoder output matches the original input feature size both in terms of spatial dimensions and feature size after 1x1x1 convolution and a sigmoid function [16]

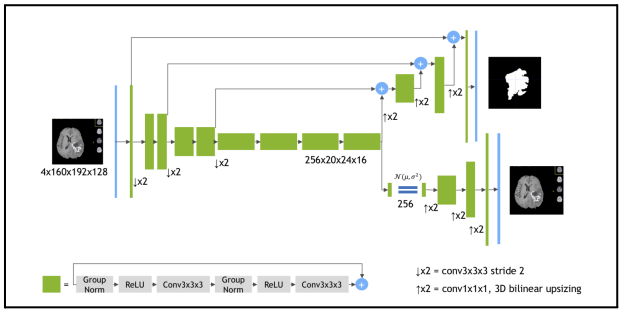

Fig.3. The architecture of SegResNet. Every green block represents a ResNet-like block incorporating group normalization [16].

D. *Loss function*

When handling medical images, it is usual for the anatomy segment to occupy minor areas of the image, which could result in a significant bias in the background; thus, the foreground regions are typically underrepresented or missing in the subsequent predicted segmentations. To solve this class unbalance, we applied DiceCross-Entropy loss, which is a hybrid of dice loss [13] and cross-entropy loss.

III. Experiments And Results

A. *Augmentation*

The quantity of training data has a significant impact on their excellent segmentation task performance. The data augmentation technique can have a big impact when these networks are applied to tiny datasets. It is used to reduce the issue of overfitting by increasing the dataset during training, thereby improving the performance of network architecture.

During the training, we apply to each image/mask pair a composition of image transformations such as RandomFlip, CropForeground, and Random Scale and Shift Intensity, which provides an expanded dataset that offers a wider range of image variations. As a result, during training, random transformations prevent the network from fixating on specific features within its perceptive field by constantly shifting these features across the field.

B. *Evaluation Metric*

We have evaluated our results using the Dice Coefficient score, referred to as the overlapping index, which measures the extent of overlap between the ground truth and the predicted output.

For a given class, Gi and Pi indicate the ground truth and predicted values for voxel i. The Dice score metric is defined by the following formula:

$$Dice(G, P) = \frac{2\sum_{i=1}^{I} G_i P_i}{\sum_{i=1}^{I} G_i + \sum_{i=1}^{I} P'_i} \qquad (1)$$

C. *Experimental Results*

To compare the segmentation models of various U-Net variations as described earlier, we implemented our work in PyTorch1 and MONAI2. We set the learning rate at 1*10-5 and employed the Adam optimizer.

According to our experiments (the detailed results are shown in Table 1), the 3D U-Net achieves the highest average Dice score. SegResNet has a similar score as of Attention U-Net, but their training time is two times greater than that of a 3D U-Net.

Table1. Averaged Dice scores of the data comparing the 3D U-Net, SegResNet, Attention-Unet models.

Model	3D U-Net	SegResNet	AttentionU-net
DSC (Train)	0.9294	0.8240	0.8027
DSC (Test)	0.9005	0.8098	0.7774

Fig.4 below presents the visual comparison of segmentation effects achieved by the previous approaches using predicted segmentation images from one dataset example, contrasted with manual segmentation images. The first row displays the MRI images earmarked for segmentation, while the second row shows the manual labeling of the left atrial structures by an expert. Commencing from the third row are the segmentation outcomes for various models. Regarding the results obtained, the 3D-UNet showed the best performance, successfully segmenting the left atrium along with its appendages (LAA). On the other hand, the remaining models achieved decent segmentation but failed to encompass the left atrium's appendages. An inherent limitation across all three models, notably visible in the Attention U-Net, was the inability to emphasize key features in cardiac MRI images, primarily due to sequence connections and low contrast between the left atrium and its surrounding tissues.

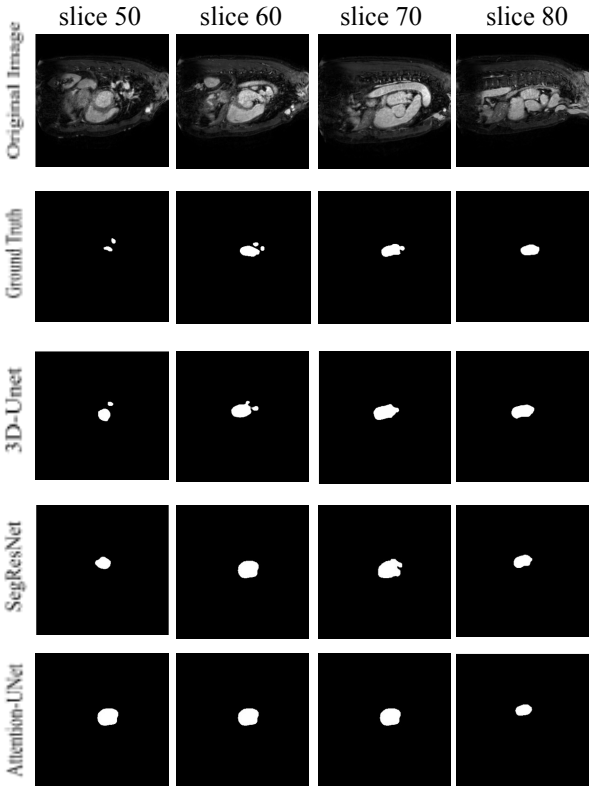

Fig.4. Segmentation results for different networks.

IV. CONCLUSION

The study is limited to a small number of samples, the changes in scale, and the high imbalance between the left atrial structure and its surrounding areas. These factors made challenges in accurately segmenting the left atrial, making it difficult to achieve precise delineation. These limitations might affect future research general findings. To improve this, future studies could explore and compare more U-Net variant models, helping to better understand how to segment the left atrial effectively.

REFERENCES

[1] WONG, Kelvin KL, ZHANG, An, YANG, Ke, *et al.* GCW-UNet segmentation of cardiac magnetic resonance images for evaluation of left atrial enlargement. *Computer Methods and Programs in Biomedicine*, 2022, vol. 221, p. 106915.

[2] LONG, Jonathan, SHELHAMER, Evan, et DARRELL, Trevor. Fully convolutional networks for semantic segmentation. In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015. p. 3431-3440.

[3] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation.

[4] JIA, Shuman, DESPINASSE, Antoine, WANG, Zihao, *et al.* Automatically segmenting the left atrium from cardiac images using successive 3D U-nets and a contour loss. Springer International Publishing, 2019. p. 221-229.

[5] Xiong, Z.; Fedorov, V.V.; Fu, X.; Cheng, E.; Macleod, R.; Zhao, J. Fully Automatic Left Atrium Segmentation From Late Gadolinium Enhanced Magnetic Resonance Imaging Using a Dual Fully Convolutional Neural Network. IEEE Trans. Med. Imaging 2019, 38, 515–524.

[6] Chen, Chen, BAI, Wenjia, et RUECKERT, Daniel. Multi-task learning for left atrial segmentation on GE-MRI. Springer International Publishing, 2019. p. 292-301.

[7] LI, Caizi, TONG, Qianqian, LIAO, Xiangyun, *et al.* Attention based hierarchical aggregation network for 3D left atrial segmentation. Springer International Publishing, 2019. p. 255-264.

[8] LIU, Yashu, DAI, Yangyang, YAN, Cong, *et al.* Deep learning based method for left atrial segmentation in GE-MRI. Springer International Publishing, 2019. p. 311-318.

[9] Vesal, S.; Ravikumar, N.; Maier, A. Dilated Convolutions in Neural Networks for Left Atrial Segmentation in 3D Gadolinium Enhanced-MRI. In International Workshop on Statistical Atlases and Computational Models of the Heart; Springer: Berlin/Heidelberg, Germany, 2019.

[10] BORRA, Davide, ANDALÒ, Alice, PACI, Michelangelo, *et al.* A fully automated left atrium segmentation approach from late gadolinium enhanced magnetic resonance imaging based on a convolutional neural network. *Quantitative Imaging in Medicine and Surgery*, 2020, vol. 10, no 10, p. 1894.

[11] HE, Kaiming, ZHANG, Xiangyu, REN, Shaoqing, *et al.* Deep residual learning for image recognition. In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016. p. 770-778.

[12] HUANG, Gao, LIU, Zhuang, VAN DER MAATEN, Laurens, *et al.* Densely connected convolutional networks. In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. p. 4700-4708.

[13] MILLETARI, Fausto, NAVAB, Nassir, et AHMADI, Seyed-Ahmad. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In : *2016 fourth international conference on 3D vision (3DV)*. Ieee, 2016. p. 565-571.

[14] Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B., Rueckert, D.:

Attention u-net: Learning where to look for the pancreas (2018)

[15] ZHOU, Zongwei, RAHMAN SIDDIQUEE, Md Mahfuzur, TAJBAKHSH, Nima, *et al.* Unet++: A nested u-net architecture for medical image segmentation. Springer International Publishing, 2018. p. 3-11.

[16] MYRONENKO, Andriy. 3D MRI brain tumor segmentation using autoencoder regularization. Springer International Publishing, 2019. p. 311-320.

[17] ÇIÇEK, Özgün, ABDULKADIR, Ahmed, LIENKAMP, Soeren S., *et al.* 3D U-Net: learning dense volumetric segmentation from sparse annotation. Springer International Publishing, 2016. p. 424-432.

[18] Tobon-Gomez, C.; Geers, A.J.; Peters, J.; Weese, J.; Pinto, K.; Karim, R.; Ammar, M.; Daoudi, A.; Margeta, J.; Sandoval, Z.; et al. Benchmark for Algorithms Segmenting the Left Atrium From 3D CT and MRI Datasets. IEEE Trans. Med. Imaging 2015, 34, 1460–1473.

[19] ZHANG, Zhilu et SABUNCU, Mert. Generalized cross entropy loss for training deep neural networks with noisy labels. *Advances in neural information processing systems*, 2018, vol. 31.