Deep Learning-Based Left Atrial Segmentation and Hemodynamics Evaluation using 4D Flow MRI

Jonas Leite[1], Louis Parker[1], Marie Shannon Soulez[1], Tom Da-Silva Faria[1], Khaoula Bouazizi[1,2], Perrine Marsac[1], Elie Mousseaux[3], Moussa Gueda Moussa[1], Nicolas Badenco[2,4], Estelle Gandjbakhch[2,4], Alban Redheuil[1,2,4], Mikaël Laredo[1,2,4], Emilie Bollache[1], Nadjia Kachenoura[1]

[1] Laboratoire d'Imagerie Biomédicale (LIB), Sorbonne Université, INSERM, CNRS, Paris, France
[2] Institute of Cardiometabolism and Nutrition (ICAN), Paris, France
[3] European Hospital Georges Pompidou, Assistance Publique - Hôpitaux de Paris (AP-HP)/PARCC, Université Paris-Cité, INSERM, Paris, France
[4] Institut de Cardiologie, Pitié-Salpêtrière Hospital, Paris, France

Abstract

We designed a deep learning (DL) pipeline for automated left atrial (LA) segmentation, which is crucial to evaluate function and hemodynamics from 4D flow MRI. We included 59 patients (1.5T) with atrial fibrillation (AF) and 27 healthy controls (3T) who underwent MRI. We implemented and trained a dual 3D ResUnet architecture to crop and segment the LA. Results on the testing set (N=16 patients/9 controls) revealed good DL performances against ground truth (GT) expert masks on the time frame used for GT initialization (mean Dice score, DSC=0.85±0.06; Hausdorff distance, HD=4.15±1.41 mm), with higher performances for the more represented AF patients (DSC=0.87±0.03, HD=3.73±0.89 mm) than controls (DSC=0.82±0.07, HD=4.80±1.76 mm). A drop in DSC was observed across all time frames on which GT was registered (all data: DSC=0.81±0.07, patients: DSC=0.84±0.03, controls: DSC=0.77±0.08), mainly attributed to limited 3T data and external time frame in training. Finally, extracted LA volume and hemodynamic indices showed good agreement with expert-derived values (LA volume: r=0.85, bias μ=6.06±8.67%, main vortex eccentricity: r=0.88, μ=0.55±9.23%, vortex amplitude evaluated through λ_2 and Q criteria: r = 0.94, μ=-12.47±9.25% and r=0.91, μ=-16.27±10.95%, respectively).

1. Introduction

Left atrial (LA) cardiomyopathy underlies atrial fibrillation (AF), which is a cardiac arrhythmia and considered to be a major contributor of thrombi formation and stroke. Indeed, AF involves alterations in atrial wall dynamics, tissue characteristics (including fibrosis and adipose tissue infiltration), as well as inner flow hemodynamics.

Therefore, the accurate assessment of LA function and hemodynamics is crucial to decipher such complex phenomena. MRI is considered as the reference technique for the evaluation of cardiac chamber volumes and function, as well as myocardial mass. Furthermore, three-dimensional (3D) cine (time-resolved) phase-contrast MRI with three-directional velocity encoding, referred to as "4D flow MRI", has proven useful to provide non-invasive comprehensive visualization and quantification of intra-cardiac blood flow hemodynamics.

However, the analysis of 4D flow MRI data, which generate thousands of images per exam, and particularly the segmentation step is complex, labor-intensive and time-consuming. As such, most image segmentation methodologies previously reported in the 4D flow MRI literature based either on conventional image processing methods [1], [2] or deep learning (DL) [3], [4], focused on the aorta, which has a large and tubular shape with straight segments aligned along its longitudinal axis inducing a laminar flow. On the other hand, geometry and hemodynamic patterns of left heart cavities and especially the LA are more complex, including low and time-varying velocity magnitude and direction, resulting in spatially heterogeneous signal throughout the cardiac cycle.

Consequently, only few studies have proposed thresholding techniques, atlas-based approaches [5], or DL [6] to segment cardiac chambers from 4D flow images. However, these studies focused only on the most contrast-enhanced phase of the cardiac cycle. As such, there is a notable gap in dynamic segmentation methods for cardiac 4D flow MRI images, which can be attributed to the complexity of these images, the lack of databases since this sequence is primarily used in research, and the scarcity of annotations.

Accordingly, the purpose of our study was to propose an automated DL-based LA segmentation and hemodynamics evaluation, using 4D flow MRI. Our ultimate goal is to investigate LA hemodynamic patterns in patients with AF.

2. Materials and methods

2.1. Population and acquisitions

We included the first 59 patients (65.3% male, 64 ± 10 years) with AF from the two CATS-AF (NCT05565183) and CT-AF (NCT04281329) ongoing prospective studies. Each patient underwent, within an interval of 51 ± 98 days: 1) a multi-phase cardiac volume CT scan , 2) a 1.5 T cardiac MRI (Siemens), including 2D+t cine SSFP anatomical images in short-axis, as well as 2- and 4-chamber views, and a 4D flow acquisition covering the entire heart. We also retrospectively included 4D flow data from 27 healthy volunteers (71.4% male, 51 ± 18 years) acquired on a General Electric 3T MRI at a second site (NCT02537041) as controls. Acquisition parameters are summarized in Table 1.

Table 1. Data acquisition parameters.

	3D CT scan	2D+t cine SSFP MRI	4D flow MRI (patients)	4D flow MRI (controls)
Acquisition matrix	512 x 512 x [141-748]	128 x 256	160 x [93-102] x [40-60]	256 x [96-134] x [120-148]
Time phases (n)	1-2	25-50	25	25-50
Spatial resolution (mm²)	0.31 x 0.75	[1.40-1.56] x [1.40-1.56]	1.82 x 2.37	2.93 x 3.51
Slice thickness (mm)	0.4	6-8	3	1
Encoding velocity (cm/s)	-	-	150 - 200	250

2.2. Dataset annotation

To provide a training dataset for DL models we used the 3D Slicer software to generate and then register the annotations first performed on highly-resolved CT volumes in systole and/or diastole onto 4D flow MRI anatomical modulus volumes, for each patient. To achieve this, we first isolated the MRI temporal phase closest to that used to annotate the CT data, using trigger time from the DICOM headers. If the patient had two annotated phases on the CT data, we selected the atrial systolic phase, corresponding to maximal LA dilation, as the reference. We then manually registered CT and MRI volumes to align and overlay the LA, using prominent anatomical structures such as the pulmonary veins as landmarks. For this registration, we used a T_M 3D affine transformation matrix (CT→MRI), allowing for translation, rotation, scaling, and reflection. Finally, a manual correction was applied to each registered mask. Finally, we excluded pulmonary veins and LA appendage from 4D flow MRI volumes to isolate LA cavity only.

In the second step, all temporal phases of 4D flow MRI anatomical modulus images were registered to the initialization phase for which CT annotation was available. This registration was carried out using the Elastix registration algorithm [7], available on 3D Slicer. Elastix functionalities include the following: 1) the algorithm uses an initial transformation to estimate a basic alignment between images, typically based on prominent anatomical features; 2) an optimization method is then applied to refine such transformation; 3) finally, rigid, affine, or non-rigid (B-splines) transformations are employed to estimate the deformations necessary for precise anatomical alignment. The resulting deformation fields are then applied to the LA mask from the initialization image to obtain LA masks aligned to all time phases of the cardiac cycle (Figure 1). Visual check is finally performed to assess the overall quality of the deformation. For 3T MRI datasets, manual segmentation was performed on systolic frame since CT was not acquired in healthy individuals, and time registration step using Elastix did not perform well due to less contrast. Thus, we only used the initial single frame where manual annotations were made. Finally, LA was manually segmented on all 4D flow volume time frames using 3D Slicer in 16 patients and 9 controls for DL model evaluation purposes.

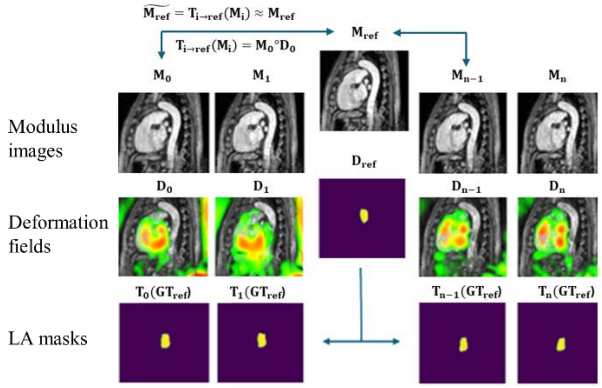

Figure1. Temporal registration of 3D volumes from a 4D flow MRI sequence to the segmented reference phase from CT, and propagation of the reference segmented LA masks across the entire cardiac cycle.

2.3. Deep learning models, training strategies, loss function and implementation

We implemented and evaluated a DL architecture for LA segmentation, based on ResUnet [8]. The pipeline consists of two main steps: the first stage is a 3D ResUnet with three levels, performing an initial binary segmentation of the heart. This network identifies the region of interest and applies a cropping step, reducing background noise. The second stage is also a 3D ResUnet with four levels, to segment the LA from the background. By leveraging deeper architecture, this model captures finer anatomical details and enhances spatial coherence of the predicted masks. 3D architecture, trained across spatial frames, ensures morphological consistency throughout the volume.

We trained our model using a database including 43 AF patients and 18 controls. To ensure high quality annotations, we selected 7 volumes near the initialization

phase (initial phase ± 3 time frames) for patients and only the initial frame for controls, resulting in N = 638 training volumes (319 + augmentation using rotation, noise and blur).
In the evaluation, we used the fully annotated database including 16 patients and 9 controls, resulting in N = 625 evaluation volumes.

For both networks, we used a combination of Binary Cross-Entropy (BCE) and Dice Loss where the BCE component provides stable gradients, while the Dice Loss encourages better global shape alignment between predictions and ground truth. Networks were implemented in PyTorch and initialized via a truncated normal distribution centered on 0. The two networks were trained with the Adam optimizer and a training rate of 0.001 and for 50 and 100 epochs, respectively, with a batch size of 32 and 64, respectively. Our model was trained on a dedicated computing hub: dual Intel® Xeon® Gold 6226R 3.90 GHz (16 hearts), 512 GB ddr4 RAM, two NVIDIA A6000 GPU (48 GB of dedicated memory each).

2.4. Prediction evaluation and quantitative indices

We evaluated our model performances using the Dice similarity score (DSC), which measures the overlap between predicted and ground truth segmented masks, and the Hausdorff distance (HD), which quantifies the maximum distance between masks, both on the initial time frame (systolic time frame) of the reference label and across the full cardiac cycle. To achieve a more comprehensive assessment over the full cycle, we also evaluated the model's performance in terms of volume estimation by comparing predicted LA volumes to reference LA volumes estimated from cine SSFP MRI images. These reference volumes were derived from 2- and 4-chamber views using the biplane area-length formula as follows:

$$V = 10 \times \left(\frac{8}{3\pi}\right) \times \frac{A_{2C} \times A_{4C}}{(L_{2C} + L_{4C})/2}, \qquad (1)$$

where A_{2C} and A_{4C} represent the surface area and L_{2C} and L_{4C} correspond to the length measured between mitral plane and the most distant LA apical point, in the 2- and 4-chamber views, respectively.

In order to investigate abnormal flow patterns, we also evaluated our model by analyzing the main vortex magnitude within the LA using λ_2 and Q criteria [9], where λ_2 is the second highest eigenvalue of the $S^2+\Omega^2$ matrix (derived from Navier-Stokes equations) and when negative indicates organized rotational flow (vortex), and Q identifies areas where the vorticity magnitude is greater than the magnitude of the rate of strain as $\frac{1}{2}(\|\Omega\|^2 - \|S\|^2)$. λ_2 and Q values were averaged throughout the LA volume and the cardiac cycle. We also assessed vortex eccentricity, defined as the distance between LA center of mass and the vortex center, normalized by LA radius.

3. Results

The proposed model was evaluated for LA segmentation from 4D flow MRI across the full cardiac cycle. Multiphase performance results are summarized in Table 2, reporting both DSC and HD metrics across the entire dataset, as well as separately for patients and controls. On the initial time frame, the model achieved a mean DSC of 0.85 ± 0.06 and HD of 4.15 ± 1.41 mm, with improved scores for patient data (DSC = 0.87 ± 0.03, HD = 3.73 ± 0.89 mm) compared to healthy controls (DSC = 0.82 ± 0.07, HD = 4.80 ± 1.76 mm). We observed same trends on all time frame evaluations. This reduction in performances might be due to the less representativity of 3T control data in the training.

Table 2. Results of the DL pipeline evaluation for LA segmentation in terms of Dice score and Hausdorff distance against ground truth on both initial and all time-frames.

	Initial time frame N=25		**All time frames N=625**	
	DSC	**HD (mm)**	**DSC**	**HD (mm)**
All data	0.85 ± 0.06	4.15 ± 1.41	0.81 ± 0.07	8.50 ± 4.94
Patients	0.87 ± 0.03	3.73 ± 0.89	0.84 ± 0.03	8.43 ± 5.76
Controls	0.82 ± 0.07	4.80 ± 1.76	0.77 ± 0.08	8.60 ± 3.35

LA volume and hemodynamic biomarker predictions were compared to expert-derived ground truth values, as detailed in Table 3 and Figure 2. LA volume was accurately estimated, with a correlation R = 0.85 and a bias of 6.06% (limits of agreement: -2.60 to 14.73%). For vortex coherence, λ_2 values were well approximated (r = 0.94, with a bias of -12.47 [-5.79; -19.16] %). Q also showed strong association (r = 0.91) with a moderate underestimation (-16.27 [-27,21; 5.31] %). Finally, eccentricity predictions were consistent with the reference correlation r = 0.88 and a bias of 0.55 [-8.68; 9.78]%.

Table 3. Comparison of LA volume and hemodynamic indices obtained using DL against ground truth (GT). p values for comparison between variables, Pearson correlation coefficients (R), and biases (μ) along with limits of agreement (loa) expressed in percentage of reference GT value, are provided.

	GT	**Prediction**	**p**	**R**	μ**(loa) (%)**
Volume (ml)	71.75±33.07	76.09±28.50	>0.05	0.85	6.06 (-2.60; 14.73)
λ_2 (s^{-1})	-41.80±24.43	-36.59±19.35	>0.05	0.94	-12.47 (-5.79; -19.16)
Q (s^{-1})	40.83±24.03	34.19±18.47	>0.05	0.91	-16.27 (-27,21; 5.31)
Eccentricity	0.045±0.02	0.045±0.02	>0.05	0.88	0.55 (-8.68; 9.78)

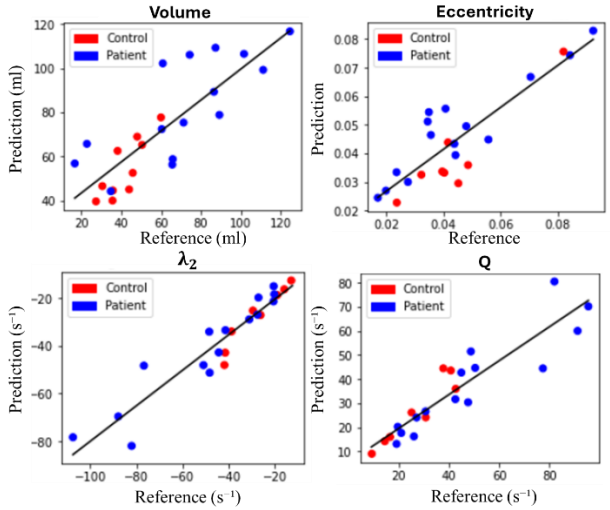

Figure 2: Linear regressions for comparison between predicted and reference LA volume and hemodynamic-related vortex eccentricity and amplitude as provided by λ_2 and Q criteria.

4. Conclusion

In this study, we proposed a pipeline to register LA segmented masks from 3D CT scan to 4D flow MRI volumes, as well as through time to generate 3D+t consistent ground truth annotations. More importantly, we then tested a dual 3D ResUnet architecture to crop and segment the LA. This configuration achieved overall good performances, with DSC values above 0.80 on average , and demonstrated good to strong agreement with expert reference for LA volume, main vortex amplitude and eccentricity. The robustness of the model across dynamic sequences highlights its suitability for accurate, automated LA segmentation and functional biomarker extraction from 4D flow MRI.

Acknowledgment

We acknowledge the H2020 MAESTRIA project (965286) for funding JL, Marie Skłodowska-Curie grant agreement #101105768 for funding LP, “Fondation Coeur & Artères—FCA 21T1” for funding PM, as well as Djibouti Embassy in Paris (654596/AMB/BGDEF/20-23) for funding MG.

Compliance with Ethical Standards

Data were collected in protocols reviewed and approved by institutional review boards, and participants gave written informed consent (NCT05565183; NCT04281329; NCT02537041).

References

[1] R. V. Bergen, H.-Y. Lin, M. E. Alexander, et C. P. Bidinosti, « 4D MR phase and magnitude segmentations with GPU parallel computing », *Magn. Reson. Imaging*, vol. 33, nº 1, p. 134-145, janv. 2015, doi: 10.1016/j.mri.2014.08.019.

[2] P. Volonghi *et al.*, « Automatic extraction of three-dimensional thoracic aorta geometric model from phase contrast MRI for morphometric and hemodynamic characterization », *Magn. Reson. Med.*, vol. 75, nº 2, p. 873-882, 2016, doi: 10.1002/mrm.25630.

[3] T. Fujiwara *et al.*, « Segmentation of the aorta and pulmonary arteries based on 4D flow MRI in the pediatric setting using fully automated multi-site, multi-vendor, and multi-label Dense U-Net », *J. Magn. Reson. Imaging*, vol. 55, nº 6, p. 1666-1680, 2022, doi: 10.1002/jmri.27995.

[4] J. Aviles *et al.*, « Domain adaptation for automatic Aorta segmentation of 4d flow magnetic resonance imaging data from multiple vendor scanners », in *FIMH*, D. B. Ennis, L. E. Perotti, et V. Y. Wang, Éd., Cham: Springer International Publishing, 2021, p. 112-121. doi: 10.1007/978-3-030-78710-3_12.

[5] M. Bustamante, V. Gupta, D. Forsberg, C.-J. Carlhäll, J. Engvall, et T. Ebbers, « Automated multi-atlas segmentation of cardiac 4D flow MRI », *Med. Image Anal.*, vol. 49, p. 128-140, oct. 2018, doi: 10.1016/j.media.2018.08.003.

[6] M. Bustamante, F. Viola, J. Engvall, C.-J. Carlhäll, et T. Ebbers, « Automatic time-resolved cardiovascular segmentation of 4d flow mri using deep learning », *J. Magn. Reson. Imaging*, vol. 57, nº 1, p. 191-203, 2023, doi: 10.1002/jmri.28221.

[7] S. Klein, M. Staring, K. Murphy, M. A. Viergever, et J. P. W. Pluim, « elastix: a toolbox for intensity-based medical image registration », *IEEE Trans. Med. Imaging*, vol. 29, nº 1, p. 196-205, janv. 2010, doi: 10.1109/TMI.2009.2035616.

[8] Z. Zhang, Q. Liu, et Y. Wang, « Road extraction by deep residual U-Net », *IEEE Geosci. Remote Sens. Lett.*, vol. 15, nº 5, p. 749-753, mai 2018, doi: 10.1109/LGRS.2018.2802944.

[9] J. Jeong et F. Hussain, « On the identification of a vortex », *J. Fluid Mech.*, vol. 285, p. 69-94, févr. 1995, doi: 10.1017/S0022112095000462.

Address for correspondence:

Jonas Leite
15 rue de l’école de médecine, 75006,
Paris france
Jonas.leite@sorbonne-universite.fr;
nadjia.kachenoura@inserm.fr