BACHELOR'S THESIS / BIOMEDICAL ENGINEERING 2021

3D Reconstruction of the Left Atrial Appendage from Multiple 2D Echocardiographic Data

Daniel Cañadas Gómez

upf. Universitat Pompeu Fabra Barcelona

3D Reconstruction of the Left Atrial Appendage from Multiple 2D Echocardiographic Data

Daniel Cañadas Gómez

Bachelor's thesis UPF 2020/2021

Thesis supervisor(s):

Pr. Oscar Camara,

Department of Information and Communications Technologies (UPF)

Acknowledgments

First of all, I must acknowledge my supervisor Pr. Oscar Camara, for proposing me this interesting and challenging bachelor's thesis topic, and supporting and giving me advice during the development of the project. Also, thanks to his research group, Physense, who were listening to my periodic presentations and giving me the necessary support.

To Abdel Hakim Moustafa (Hospital de la Santa Creu i Sant Pau) for his time helping me in the obtention and processing of all the data.

Finally, thanks to my family and friends for their patience and daily support.

Abstract

Clot formation in the left atrial appendage (LAA) is a very common problem in patients with previous pathologies, such as atrial fibrillation (AF), which can lead to fatal consequences. In these patients, computer simulations can help predicting possible risks and assessing clinicians' decisions. However, to perform LAA fluid simulations, volumetric images are necessary, from which the 3D structure can be segmented. Computed tomography (CT) scans are useful for this, but due to their high cost, high dose of radiation or simply because they are not always available in every hospital, they are performed in a minimum number of patients. On the other hand, echocardiography is a common technique, which is done routinely in AF patients.

The objective of this project is to use deep learning (DL) techniques to reconstruct a 3D model of the left atrial appendage from single multi-view echocardiography images. The hypothesis to be tested is that a correctly trained DL algorithm can learn to reconstruct a 3D structure from ultrasound images, while keeping the most important details from each of the input views, thus obtaining a realistic enough model.

A modification of classic variational autoencoders was considered an optimal approach for the task of 3D reconstruction in echocardiography images. The performance of the network was evaluated by computing the Dice Score and Hausdorff distance, and comparing fluid simulation results between ground-truth and reconstructions. The results demonstrate the potential of DL and variational autoencoders in reconstruction problems and open the possibility of cheaper and more efficient stratification method for patients with AF.

Keywords

Deep learning, computer vision, variational autoencoder, ultrasound, left atrial appendage.

Preface

"Adopting a continuous learning mindset is and will continue to be needed to adjust to the changing healthcare landscape. Staff will need to possess and continually refine competencies in IT in order to contribute and compete. For our patients, who will be expected to handle more self-care through IT capabilities, the need to develop comfort with wearables and using self-monitoring, AI and care management will be essential. In short, the biggest challenge facing healthcare will be in embracing the embeddedness of IT in all facets of the health care and the delivery system."

Linda C. Lombardi, PhD, chief strategy officer and chief experience officer at NYC Health + Hospitals/Bellevue

This thesis was developed in collaboration with the Sensing in Physiology and Biomedicine Research Group in UPF, and with clinicians at Hospital de la Santa Creu i Sant Pau, Barcelona.

Index

List of figures

List of tables

1. INTRODUCTION

1.1 Clinical background and importance of the project

Atrial fibrillation (AF) is a common form of arrythmia with an overall prevalence in the general population of around 0.9% [1][15]. Also, its incidence increases with age, approximately a 0.2% per year, resulting in a prevalence of about 5% over the age of 40 [1][4]. This prevalence of AF has been increasing during the recent years, mainly due to some factors such as the population ageing or the increasing occurrence of some risk factors. Some of these important aspects increasing the probabilities to suffer AF include obesity, hypertension, diabetes, pulmonary disease, congenital cardiovascular diseases, cardiac insufficiency, etc.

In addition to the common symptoms, including palpitations, chest pain, seizure, reduced ability to exercise, fatigue, etc, this affection can lead to fatal consequences such as stroke. AF is associated with major systemic thromboembolism. In fact, about a third of patients with ischemic stroke have been found to have either clinical or subclinical AF due to high prevalence of left atrial thrombosis [2]. Thus, when having episodes of AF, blood can form clots inside the atrial cavities; subsequently clots can go out of the heart chambers and block the circulation to the brain causing an ischemic stroke, which can be fatal. After these thromboembolic events, 40% of patients are left with moderate functional impairments, while 15% to 30% are left with severe disability [32]. Moreover, strokes related to AF are associated with higher mortality, greater disability, longer hospital stays, and lower chance of being discharged home than strokes unrelated to AF [3][31]. The number of patients currently presenting AF is about 33 millions of patients around the world, with an annual incidence of stroke estimated to be around 5%-7% [4].

During Atrial Fibrillation (AF) more than 99% of the left atrial thrombi responsible for ischemic stroke events originate in the left atrial appendage (LAA) [16]. LAA is a small tubular sac in the muscle wall of the left atrium, formed during the fourth week of embryonic development. Due to its small size and its shape and location, during atrial arrhythmicity, blood can collect there and form clots.

Computer-based simulations have opened the possibility to predict which of the patients presenting pathologies such as AF are in a high risk of suffering from clot formation in the LAA, with its fatal consequences. Some studies (e.g, Bosi et al. [5] or Mill et al. [17]) aimed at predicting the thrombosis risk in patients with atrial fibrillation, by analysing the hemodynamic behaviour in four typical LAA morphologies: “Chicken wing”, “Cactus”, “Windsock” and “Cauliflower”. The objective was to identify potential relationships between the different shapes and the risk of thrombotic events. These studies highlighted the major role played by the LAA morphology on the hemodynamics, thus demonstrating the importance of computer simulations on patient risk stratification. Other studies, in contrast, use computer simulation tools with the aim of assessing clinicians to select the optimum size, shape and orientation of the implantable LAA occluder device (LAAO), which in patients in which the treatment with anticoagulants is not recommended, is useful to avoid clot formation inside the LAA cavity. For instance, Aguado et al. [6] used

in-silico models to identify the LAAO configurations associated to a lower risk of thrombus formation and concluded with the potential of this type of models during the pre-implant planning of the LAAO therapy to reduce clot formation due to sub-optimal configurations of the device.

However, to be able to run reliable enough computational simulations, the 3D geometry of the LAA should be segmented from volumetric images. Computed tomography scans (CTs) are very useful for the segmentation process, as they facilitate the extraction of volumetric information of the structure of the heart. Hence, the segmentation of the LAA becomes a direct and rapid process, generating the geometry on which a mesh is optimized to be able to run numerical simulations. However, the obtention of CTs is not always possible for all patients having AF, so the feasibility of an in-silico-based risk stratification for all these patients decreases. The cost of carrying out CTs is elevated with respect to other diagnostic imaging tools and patients are exposed to a high dose of radiation. The dimensions and the cost of a CT scanner are also factors making impossible to carry out this type of diagnostic tests in all hospitals. On the other hand, echocardiography or ultrasound (US) are done routinely in patients with atrial fibrillation; unfortunately, echocardiographic 2D acquisitions mainly provide information in the form of 2D images. 3D US is also a medical imaging technique to consider, which shows an image in volumetric form, but in a limited field of view. Figure 1 shows visually the differences between the 2D and 3D US images, and also between both of them and cardiac CTs.

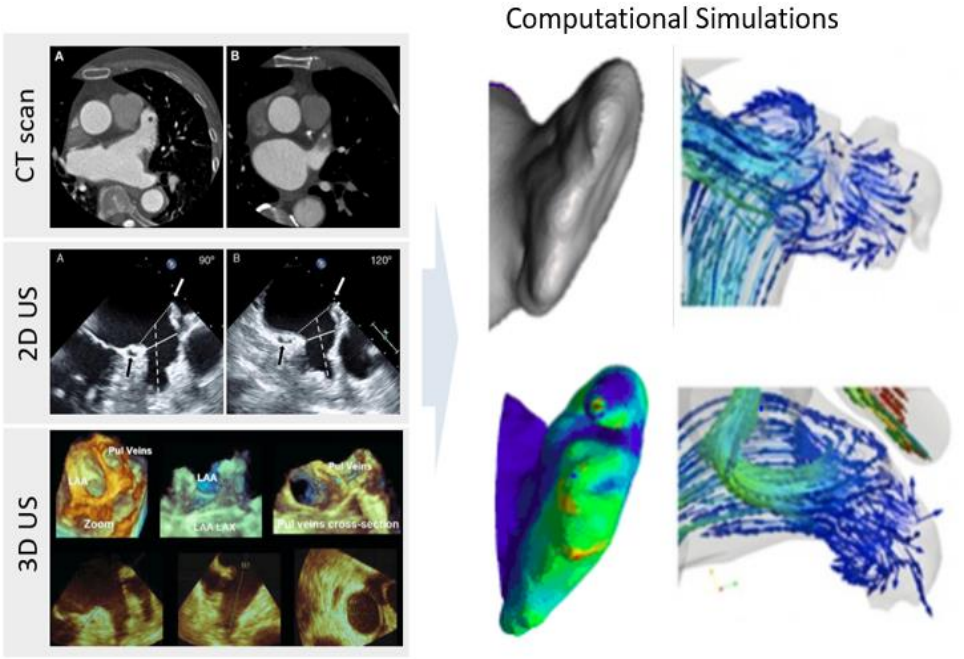

Figure 1. Comparison of the different imaging modalities exposed (Computed Tomography (CT) scans, 2D and 3D ultrasounds (US)). The objective is to obtain from them a reliable enough segmentation of the 3D geometry of the Left Atrial Appendage (LAA), in which computer simulations can be run. Images are taken from: [20][21].

For these reasons, the purpose of this bachelor's thesis is to develop deep learning algorithms which, starting from 2D images of transoesophageal LAA US images (the modality of US imaging used to visualize the LAA of the patient), are capable to reconstruct a reliable enough 3D geometry of the LAA, in which computer simulations can be run. As above-mentioned, echocardiography images are taken routinely in patients with atrial fibrillation, so being able to reconstruct the 3D geometry of the LAA from multiple 2D US views would be interesting to achieve an easier, cheaper and rapid risk prediction of the patients with AF to suffer from strokes. It would also represent an important simplification of the process of acquisition of the input images to obtain the whole reconstruction, as 2D transoesophageal (TEE) US images can be obtained with a simple micro-TEE transducer, which can be introduced through the patient's oesophagus with a minimal sedation process preceding it (see Figure 2 and Annex 3 for more information on TEE probes).

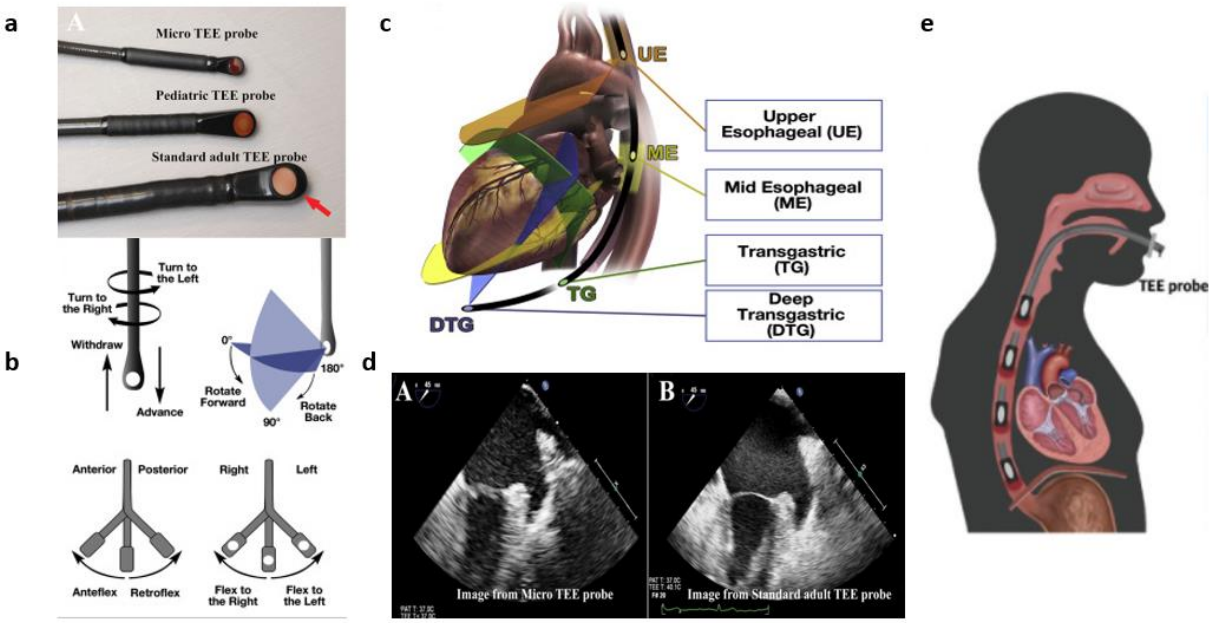

Figure 2. a. Different transoesophageal (TEE) transducer probes used in the acquisition of TEE ultrasound (US) images. b. Functioning and movements of the TEE probe. c. Placements of the probes inside of the oesophagus for the obtention of images from different angles. d. Examples of images from TEE US, showing the left atrium (LA), the left atrial appendage (LAA), and the left ventricle (LV). e. Collocation of the TEE probe. Images are taken from: [28][29][30]

1.2 State-of-the-art on 3D object reconstruction from 2D inputs

Several articles in the literature propose non-DL methods that allow creating correspondences between the 2D initial images and the 3D model. For instance, González Ballester et al. [18] presented a method based on a non-rigid 2D point matching process, to establish correspondences between 2D X-Ray images and a 3D model of the bone surface.

In recent years, the increasing knowledge on deep learning (DL) has offered new ways of analysing and processing medical images. Indeed, DL has achieved, in some medical imaging applications such as segmentation of certain biological structures, a performance comparable to human experts. Advances in deep learning have also allowed researchers to come up with more complex algorithms, such as generative adversarial networks (GAN), which are capable, for instance, of transforming Magnetic Resonance Imaging (MRI) images to CT space, obtaining the visual information of a common CT without exposing the patient to such high radiation [7].

Related to the objectives of this bachelor's thesis, some computer vision DL-based algorithms, outside the medical domain, have been developed to reconstruct 3D volumetric meshes or surfaces starting from single or multiple 2D images. Wang et al. [8] developed "Pixel2Mesh", a network based on a fully convolutional model and a graph neural network, transforming a single 2D input into a volumetric mesh, capturing the main characteristics of the introduced 2D image. This procedure was improved by Wen et al. in their algorithm named "Pixel2Mesh++" [9]. This improved version was capable to capture features from different points of view of the geometry of interest.

Another method from Pumarola et al. (2020) [13] was based on flow-based generative models. The training of a GAN algorithm to perform a 3D reconstruction could be very unstable and requires a careful tuning of the hyperparameters. In this paper, Pumarola et al. introduced a novel mechanism to condition flow-based generative models. The principal benefits of this method for our objective are the direct generation of a 3D point cloud from a single image, in comparison with other algorithms such as PointFlow [14], a generative based model in which instead, point clouds are generated one point at a time, incrementing the inference time.

Although the first two methods obtained good results, both were trained and evaluated on ShapeNet images, having characteristics and resolutions completely different to the echocardiographic images we use. Also, the effect of noise must be taken into account in US images, while in ShapeNet objects it is not a factor consider.

Other studies, in contrast, such as the one carried out by Tóthová et al. (2020) [10], used probabilistic approaches to reconstruct a 3D surface of cardiac ventricles from 3 quasi-orthogonal MRI data, while capturing at the same time the uncertainty. The above-mentioned method is built on the principles of probabilistic principal component analysis (PCA). This second method could be particularly interesting for the purpose of reconstructing the LAA from multiple views of echocardiographic images. As the data will be very conditioned by the resolution of ultrasound modality, which lacks resolution, details and is significatively more noisy than other typical modalities, such as MRI or CT, a model which captures the uncertainty and is able to consider more views in the reconstruction of the geometry could help to obtain more reliable results.

Vlontzos et al. (2020) [11] also tried to address the problem of the lack of volumetric information in 2D X-Ray images. Two methods were tested in this paper and evaluated

in more than 117 CT scans. Both methods are adaptations of segmentation deep learning algorithms: a typical 3D U-Net and PhiSeg, by Baumgartner et al. [12], an Open-Source segmentation algorithm able to capture the typical interpersonal uncertainty of a manual expert segmentation. Both models are adapted to augment the information contained in an X-Ray image and achieved a Dice score in the evaluation process of 0.91 and 0.90, respectively. However, they both failed to reconstruct very fine structures, for instance the ribs, with Dice scores below 0.5.

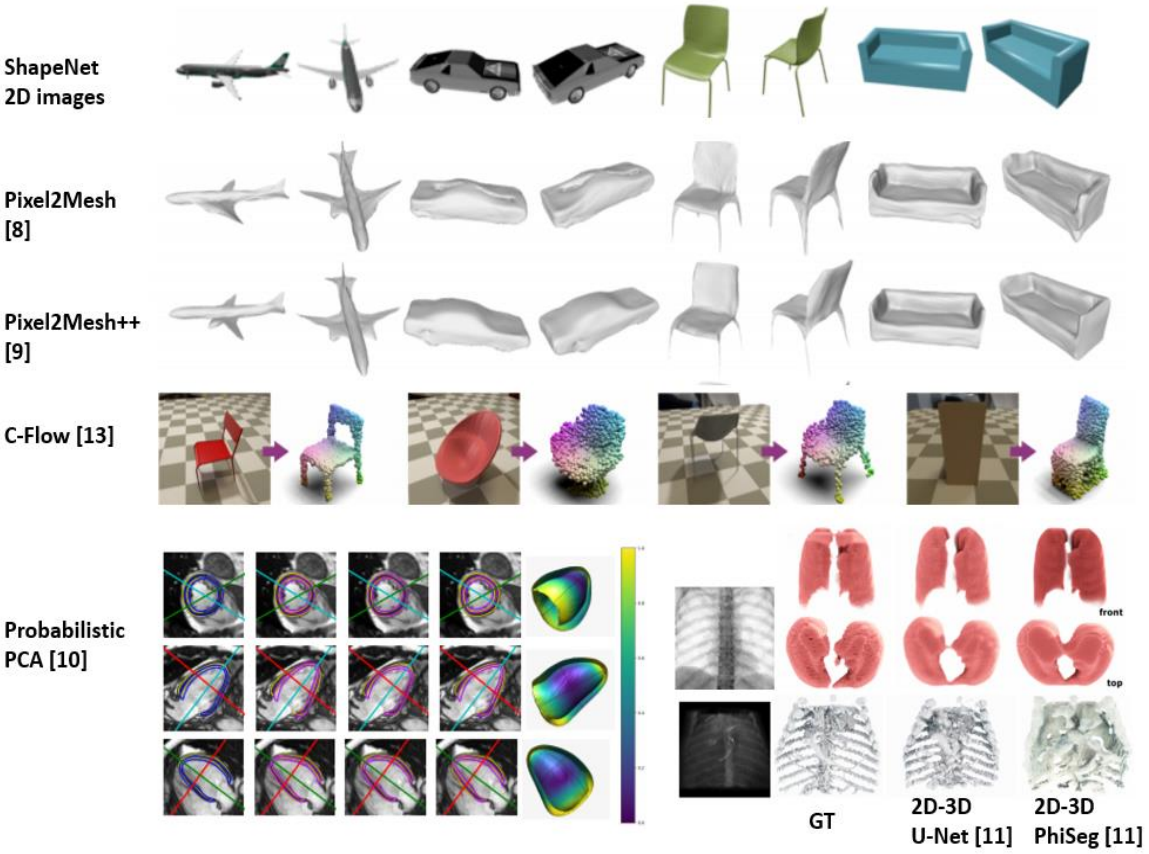

Figure 3. Comparison of the different deep learning (DL) methods exposed in the state-of-the-art, achieving 3D reconstructions from single or multiple 2D views.

Finally, Biffi et al. [19] used an adaptation of variational autoencoders (VAEs), conditional variational autoencoders (CVAEs), with the potential of reconstructing 3D models of the ventricles of the heart, conditioned by three single 2D MRI images (two long-axis and one short-axis). Their approach achieved an 87.92 of Dice Score in the test set. Figure 3 shows some of the results obtained in the different articles in the state-of the-art, using DL in the task of 3D reconstruction from 2D inputs.

1.3 Contribution of this project

As previously mentioned, the main purpose of this bachelor's thesis is to perform a reliable enough 3D reconstruction of the LAA from echocardiographic images, in which computer simulations can be run. To achieve this, once all these previously exposed state-of-the-art approaches are carefully investigated in order to see the possible benefits or weak points that they could have for our objective, they were be trained and evaluated for the 3D reconstruction on 2D echocardiographic images, in which the structure of interest, the LAA was previously segmented with a common U-Net. The objectives were to check their performance on this imaging modality, see which of the approaches outperforms the rest and improve its performance. Our approach also had to deal with the difficulties in the task of adapting these existing methods to our main objective, not only because of the low resolution and presence of noise in echocardiography, but also due to the complexity of the LAA geometry. This anatomy is also one of the causes why, as previously explained, more than 90% of the thrombi related to cerebrovascular accidents are generated there. The reconstruction of an anatomy of these characteristics is a challenge itself, as 2D images will lack of many important details. This project has been performed in collaboration with clinicians at the Hospital de la Santa Creu i Sant Pau de Barcelona, Spain.

2. METHODS

2.0 General Methodology

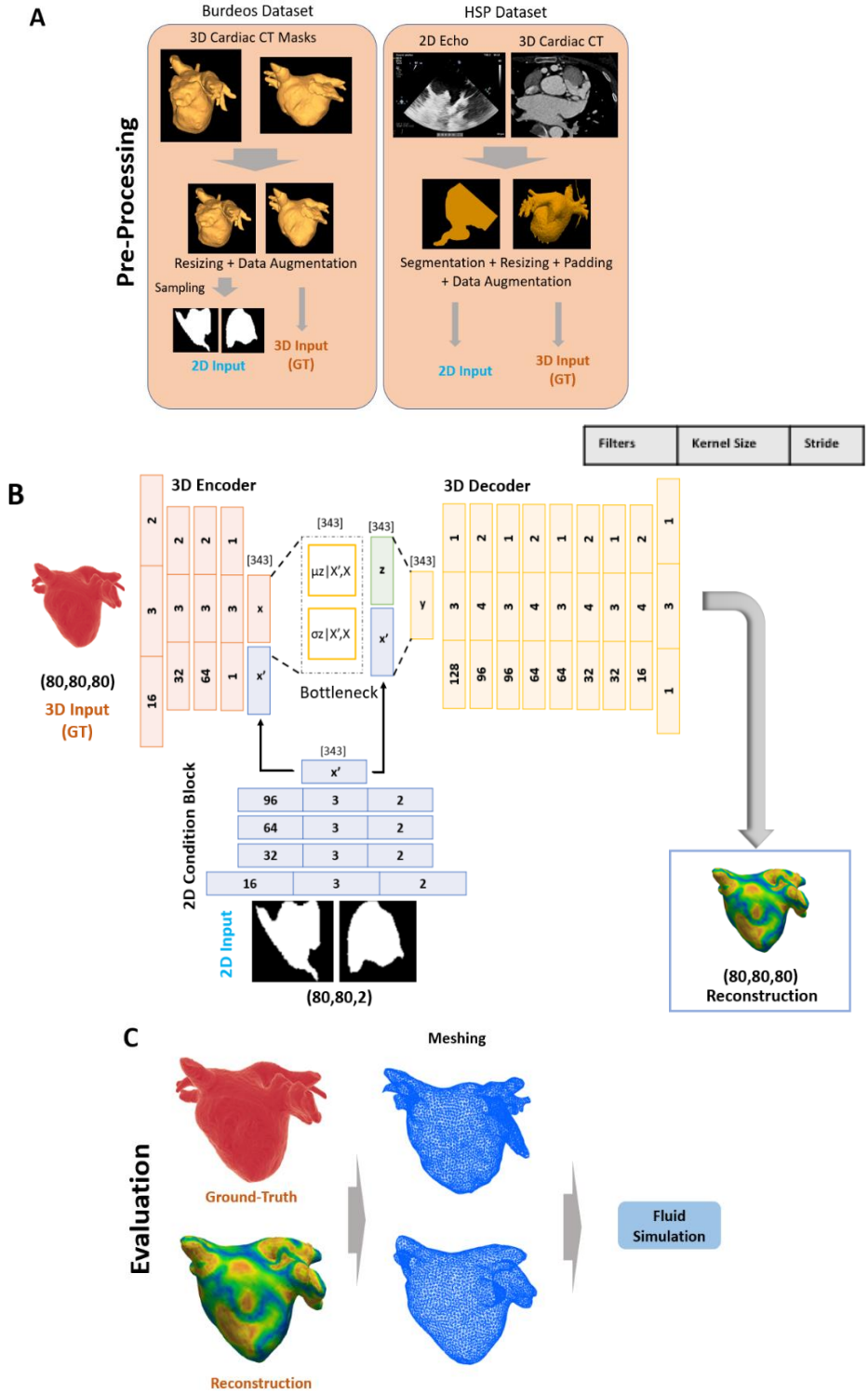

Figure 4. Scheme summarizing the whole methodology used to perform the 3D reconstruction from 2D views of echocardiography. A. Pre-processing. B. Model for the 3D reconstruction from 2D inputs. C. Fluid simulation-based evaluation of the results.

In this part of the report, a full detailed explanation of the main activities and procedures carried out in this project will be done. The final objective is to design a deep learning pipeline which could be followed to obtain a trained neural network capable to perform 3D reconstructions from single 2D views. Figure 4 shows a diagram summarizing the whole methodology progress, proposed in this bachelor's thesis. The different sections of this pipeline will be fully explained in the next sections, and include the acquisition of the data in Hospital de la Santa Creu i Sant Pau, the pre-processing, an explanation of the architecture and training process of the model, the performed experiments with their corresponding results, the discussions, and a final section summarizing the main conclusions of this study.

2.1 Materials

To accomplish the objectives of the thesis and obtain the expected 3D reconstructions, left atrial appendage (LAA) transoesophageal echocardiography (TEE) 2D images, from at least 2 different angles were used. In this work, as it is especially interesting to have also a complete 3D structure of the LAA of these patients, a cohort of patients in which CT scans were also acquired was used. This way, for each of the subjects, we had 2 or more different 2D views of ultrasound modality and a 3D segmented structure from a CT scan. This dataset, which will be referred to from now as HSP, was collected from Hospital de la Santa Creu i Sant Pau (Barcelona, Spain), and it includes CT and TEE images from patients suffering from different pathologies, such as atrial fibrillation or mitral regurgitation, giving a final database size of 50 subjects. The CT scans were acquired in a Philips iCT 256 Elite scanner (voxel size: 0.67x0.67x0.67 mm). The US images were acquired in a Philips Epiq 7 and in a Philips iE33 Ultrasound Machine. Some examples of these cardiac CT and TEE US images can be found in Figure 5.

Also, in this project, a second database, which we will refer to from now as Burdeos, was used consisting of left atrium CT masks. This was done, not only to compare the performance of the deep learning method in the task of reconstructing a 3D geometry from 2D ultrasound images against 2D CT masks, but also to perform other experiments, including transfer learning, training the architecture firstly for CT reconstruction and then re-training it for the TEE US reconstruction task. This second database includes cardiac CT images from 194 patients, which were already segmented and refined.

2.2 Experimental Setup

2.2.1 Pre-Processing of the Data

Before using the data in deep learning algorithms to obtain a full 3D reconstruction, data was previously pre-processed, to match the required size and format, as well as to reduce excessive computational costs.

A. HSP CT scans

The 50 CT scans from the HSP database were segmented in Hospital de Sant Pau using Mimics Research 23.0 from Materialise[1], and were validated by a clinician. This process included a segmentation of the heart cavities using a threshold, the separation of the different anatomical parts, and the smoothing of the left atrial surface. In some of the cases, the left atrial appendage had to be segmented by hand, as the noise produced by the occluder device (LAAO) difficulted an automatic threshold-based segmentation of the whole structure. The complete process of segmentation using Mimics Research 23.0 is fully detailed in Annex 1 (See Annex 1.1). The obtained 3D masks were downsampled to a size of 80x80x80 pixels. As previously exposed, these 3D structures will serve as ground-truth.

B. HSP Transoesophageal Echocardiography Images

In parallel, the corresponding 2D TEE images from HSP were manually segmented using 3D Slicer[2]. These 2D slices will be the ones introduced as 2D input for this first database. The objective is that the model learns to reconstruct the 3D geometry from them. See Annex 1.2 for a fully detailed explanation on this process.

C. Left atrium 3D CT masks

As explained in the materials section, a second database only including left atrium segmentations from 194 patients was also used (Burdeos). These left atrium masks from CT images were cropped, in order to get rid of empty slices, and downsampled to a size of 80x80x80 pixels, as in the HSP 3D masks.

Once we have these true 3D geometries, 2-4 different 2D samples (views) were selected, corresponding to these structures. The idea was to train an algorithm which, from these views, is able to reconstruct the whole geometry of the atrium. For this reason, 2-4 slices in total were randomly selected between the 30th and 50th slice of each of the anatomical axis.

[1] Materialise: https://www.materialise.com/en/medical/mimics-innovation-suite/mimics

[2] 3D Slicer: https://www.slicer.org/

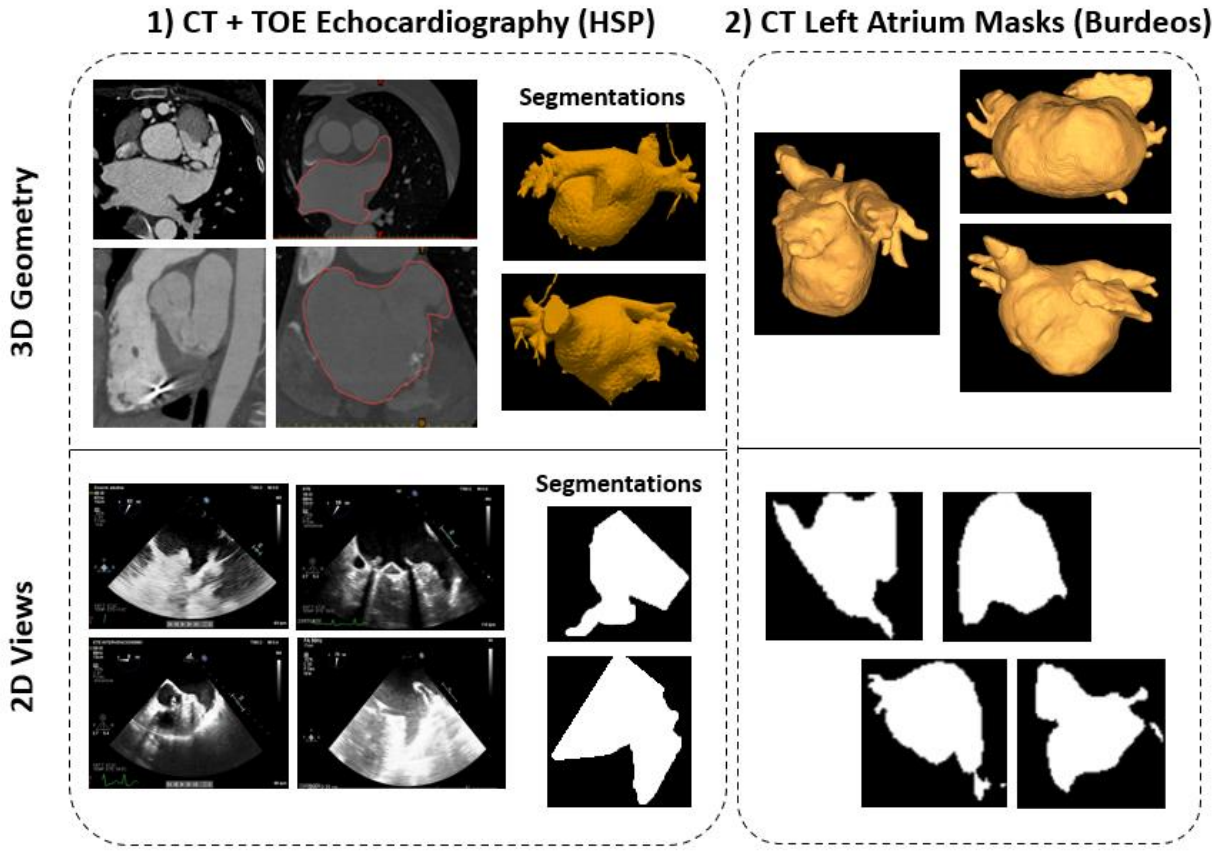

Figure 5. Databases used in the project. On the left, database collected from Hospital de la Santa Creu i Sant Pau (HSP), showing the 2D views from echocardiography with the respective segmentations, and the 3D CT segmentations. On the right, the second database (Burdeos) including only 3D CT masks. The 2D views in this case are samples from the 3D structure.

2.2.2. Data Augmentation

To be able to train the model with more data, data augmentation was applied. This data augmentation consisted of rotations and binary erosions and dilations, as they are an efficient way to increment the dataset volume without altering the typical morphology of a left atrium. In other words, it is required to have more data to train a more efficient algorithm, but this data augmentation must be carried out in a way we are preserving the anatomy and the structural relationships of the elements of the geometry. If anatomically incorrect structures are formed by data augmentation, the training of the network will be affected, harming the overall reliability of the reconstructions obtained for the test set. These rotations consisted of:

- In the case of Burdeos dataset, for each of the 3D geometries, 3 different random rotations were applied (between 1º and 359º). This made possible to increment the volume of the data to 776 subjects.

- In the case of the database including CTs and the corresponding 2D echocardiography images (HSP), 30 rotations were performed per subject. The reason behind this, is that we cannot know exactly the correspondence of the 2D TEE samples in the real 3D space, so performing a considerable number of rotations per each case could help the network to not take into account this difference in location between the 3D structure and the 2D samples. Also, as it is a more limited database, more augmentation is required to ensure an optimal convergence of the network.
- To increment the number of volumes to train the data, some morphological transformations were also applied to HSP data, including erosions and dilations, both with kernel size in the range of 2-6.

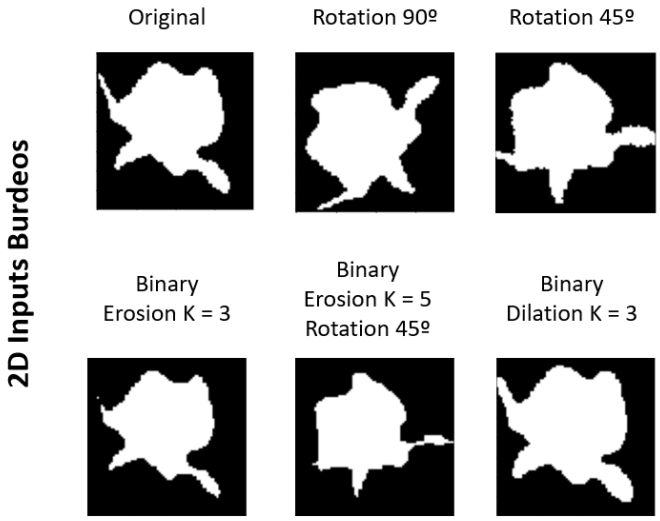

Figure 6. Examples of the data augmentation using rotations, erosions and dilations in 2D samples of Burdeos dataset.

2.2.3. Conditional variational autoencoders in reconstruction problems

As mentioned in the first sections of the document, in recent years, many researchers have been trying to obtain volumetric information from 2D inputs using different methodologies, including some based on deep learning. In this project, we are trying to deal with two main challenges: the complexity of the geometry of interest and the lack of details of the ultrasound imaging modality, which hampers significatively the task of reconstruction, as explained in the introduction. To face these challenges, the method used in Biffi et al. [19], using conditional variational autoencoders, could be particularly interesting, as previously justified.

In simplified way, autoencoders (AE) are generative models consisting of an encoder and a decoder, which are designed to learn the essential features to produce their own inputs,

thus reducing the inputs to a lower dimensional latent space (encoder), from which the decoder can produce the original structure. On the other hand, variational autoencoders (VAE), instead of encoding the inputs into features which represent the initial data, they encode distributions with a certain mean and variance [22-24]. This means that, in the case of introducing as inputs a dataset of LA geometries, the network would capture the variability that exists between the different geometries. In other words, new geometries could be generated knowing this variability. To sum up, this is the main advantage of variational autoencoders with respect to simple autoencoders: the capability to address variability [25]. This solves a new problem and is that we want to reconstruct 3D structures from 2D images, which are new, so we do not know the original geometry.

But, as we can imagine, if we sample at random from the latent space created by an encoder of a VAE, a random 3D geometry would be produced. What we really want is that this 3D structure matches with a 2D input, in this case a 2D segmentation from a TEE image. This means that we obtain a reconstruction consistent with the 2D information we are introducing as input. This problem could be solved with conditional variational autoencoders (CVAE). These models are characterized by an extra input, the condition. Now, the decoder will not generate a random reconstruction from a latent space, but an output matching the features of the 2D input.

In conclusion, the idea is to train the model to capture the variability of different left atrial appendages thanks to the encoder-decoder architecture. In other words, the network is learning the features that an anatomically correct LAA must have, and it is able to generate from the created latent space, new geometries corresponding to new LAA. By the addition of a conditional layer, we achieve that the model learns to sample adequately from the latent space produced by the encoder to decode a 3D geometry consistent with the 2D data of interest.

2.2.4. Architecture of the CVAE

As explained, a typical VAE consists of an encoder and a decoder. In this case, the encoder was composed of a series of four 3D convolutional layers, with a stride s = 2 (downsampling the initial data in each of the layers), and kernel size k=3, and a final convolutional layer with s = 1. The output of this final layer is flattened, and a dense layer gives a vector of 343 values. The encoder is then basically reducing the input image $X \in \mathbb{R}^3$ into an encoded lower dimensional space $\mathbf{z} \in \mathbb{R}^{343}$, which has a regularized prior distribution p(z) (standard Gaussian distribution). This can be defined as:

$$z = Encoder(X) \sim q(z|X), \quad \text{(Equation 1)}$$

$$p(z) = \mathcal{N}(0,I), \quad \text{(Equation 2)}$$

where X is the input 3D image, and q(z|X) could be translated as the probability of variable z to pertain to X.

Here it enters the conditional block, composed by 5 layers of 2D convolutions with k=3 and s=2. The output gives a vector of 343 values as well, which is concatenated with the

output of the encoder. From this concatenation, a new dense layer produces a new vector of 343 values, the new latent dimension vector z. Half of its values correspond to means ($\mu z|X',X$) and the other half to standard deviations ($\sigma z|X',X$), as a VAE do not codify for features, but distributions of these features. By doing this concatenation, the latent dimension vector z is transformed from Equation 1 into:

$$z = q(z \mid X, X'), \text{ (Equation 3)}$$

where X' is the 2D input from the conditional block. This means that now the network is learning to create a latent variable z with a standard Gaussian distribution, which can match both 3D input geometries, and the corresponding 2D inputs. This part of the network is also known as bottleneck, as we have forced the model to learn to represent the input space in $\mathbb{R}^3$ by a vector of 343 values representing features, thus ensuring that it only learns the most necessary features which better represent the input data.

From this point, a sample z is taken from this latent space and it is again concatenated with the output vector ($x' \in \mathbb{R}^{343}$) of the conditional layer. A dense layer produces the input vector of size 343 for the decoder. This last part of the model consists of an alternation of 3D transposed convolutions and 3D convolutions, and its function is to decodify the sample z taken from the latent space to a new 3D geometry upsampling again to a shape of 80x80x80, which matches the original 3D shape. We can model this equation as:

$$Y = p(X \mid z, X'), \text{ (Equation 4)}$$

Where Y is the 3D output of the decoder, z is the latent variable and X' is the 2D condition input. With an optimal performance of the model, Y should match the original geometry X. So basically, what we are doing is to train the model with 2D samples (conditional block) and the true 3D geometry (encoder) to learn how, parting from this 2D input the decoder has to sample the adequate z from the latent space and reconstruct the geometry that we are putting as an input for the encoder. Once the network has been appropriately trained, during the test, the encoder will be disabled, so we will only introduce 2D data to the conditional layer, and the decoder will be able to predict the original shape. The architecture of the employed model is schematized in Figure 7.

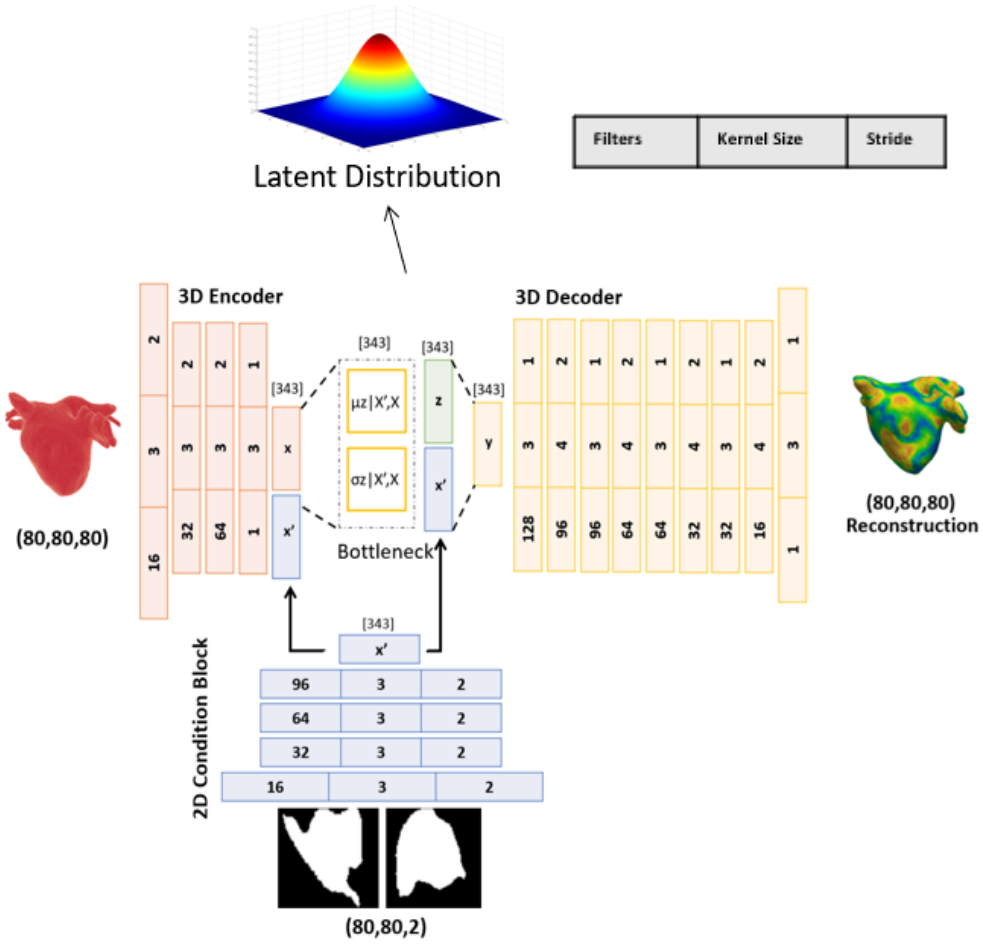

Figure 7. Scheme summarizing the deep learning model used in this 3D reconstruction problem.

2.2.5. Metrics and losses

The loss function to minimize is composed of two terms (see Equation 1). The first one is a reconstruction loss, a binary-cross entropy between the output of the decoder Y and the input of the encoder X, as the produced geometry should match the one we are inputting. The second term is called Kullback-Leibler (KL) Divergence. It is a regularization term which basically, acts as a force keeping the different clusters in the latent space together, leading the mean (μ) to 0 and the standard deviation (σ) to 1.

In a very simplified and intuitively way, by only imposing the network to create distributions in the latent space instead of encoding features just as classic autoencoders, it is not enough to ensure continuity and completeness. Without a regularisation term, the model, in order to maximize the reconstruction loss, could either return distributions with tiny variance (punctual distributions) or with very different means (points very far from each other in the latent space). This would cause overfitting, and our objective would be

impossible, as the network would not be capable to produce new geometries if they are not part of the training set. For this reason, we introduce the second term of the loss function, the KL divergence, which, when minimized, would lead the variances to 1 and approximate the means to 0, thus ensuring a distribution close to a standard normal distribution, as it can be seen in Equation 5.

$$Loss = BCE + \beta \sum_{i=1}^{d} (Var(z) - \log[Var(z)] - 1 + E(z)^2) \quad \textit{(Equation 5)},$$

where Var(z) are the variances and E(z) are the means for each value of the d-dimensional latent space vector. On the other hand, β is an arbitrary parameter which weights this second term of the loss function. In this project, β was chosen to be 1. For $\beta < 1$, the reconstruction term gained too much weight, prioritizing the reconstruction accuracy and preventing the model to approximate to normal standard distributions in the latent space, thus producing overfitting to the input data. For $\beta > 1$ the reconstruction term was strongly penalized, ensuring strict normal standard distributions, but making difficult the convergence of this first term.

In conclusion, with the absence of this last part of the equation, the autoencoder would tend to separate the clusters in the latent space, because of the difference in geometries of the different LAA, and this would result in a situation in which the model cannot predict new geometries from the latent space if it has not seen them during the training process.

Finally, a Dice Score (DS) metric was applied, so as in each iteration it was be possible to see how the DS improves and tune the hyperparameters to the most adequate ones (Equation 6).

$$DS = \frac{2 \cdot |X \cap Y|}{|X| + |Y|} \qquad \textit{(Equation 6)}$$

2.2.6. Evaluation of the performance in the test set

Once the networks are trained, their performance was evaluated on the corresponding test sets for each of the experiments. In this final part, the 3D encoder must be disabled, as we are only introducing 2D samples, and we do not have a true 3D geometry to encode. The 2D samples from the test set were provided to the conditional layer of the CT, and the latent variables z were set to 0. The obtained reconstructions were evaluated in 2 different ways: by computing the Dice Score (Equation 6) and Hausdorff Distance (HD) (Equation 7) and by performing a very simple static fluid simulation both in the true geometry and in its reconstruction, in order to check if the behaviour is similar.

$$\mathrm{HD} = \max\left(\max_{y \in Y}\left(\min_{x \in X}(d(x,y))\right), \max_{x \in X}\left(\min_{y \in Y}(d(x,y))\right)\right) \quad \textit{(Equation 7)}$$

Hausdorff Distances were computed in Meshlab[3]. The meshes from both ground-truth and reconstructed geometries were loaded. First of all, the GT geometry was glued, and the reconstructed shape was registered using landmarks. Once aligned, the filter "Hausdorff Distance" was used, which computes the HD (Equation 7) in each of the nodes of the selected mesh (reconstruction) to the reference mesh (GT).

The main objective of this thesis, as previously justified, is to develop a DL algorithm which is capable, parting from 2D data of echocardiography, to recreate an accurate enough 3D structure of the left atrium, in which computer simulations can be run. This DL network must then also ensure that the reconstructed structure will have a similar behaviour when doing computational fluid simulations. For this reason, we decided that it was necessary to check the performance of the network in a third way: doing simple fluid simulations both in the ground-truth and reconstructed geometries and compare the obtained patterns, as included in the scheme of Figure 3.

To perform these fluid simulations, both reconstruction and ground-truth meshes were created and processed in Meshmixer[4]. The inlets and outlets, which were decided to be the pulmonary veins and the mitral valve, respectively, were flattened, so as to impose the boundary conditions on them. Final meshes had between 40000 and 50000 elements. The final volumetric meshes were generated in Gmsh[5], and had between 300000 and 450000 elements.

Volumetric meshes were opened with Ansys Fluent[6], to set up the physical properties and boundary conditions, and perform the fluid simulations. These properties and boundary condition data is summarized in Table 1. The simulations were run for 200 iterations, and the results for the velocities at the mitral valve, and the absolute pressures were visually and quantitatively compared for both ground-truth and reconstructed geometries in Paraview[7].

Attribute	**Value**
Blood Density (kg/m^3)	1060
Blood Viscosity (Pa/s)	0.0035
Pulmonary Veins: Velocity inlet (m/s)	0.4
Mitral Valve: Pressure Outlet (mmHg)	8
Iterations	200

Table 1. Physical properties and boundary conditions applied to all the fluid simulations for both ground-truth and reconstructions.

[3] Meshlab: https://www.meshlab.net/

[4] Autodesk Meshmixer: https://www.meshmixer.com/

[5] Gmsh v.4.8.4: https://gmsh.info/

[6] Ansys Fluent: https://www.ansys.com/products/fluids/ansys-fluent

[7] Paraview: https://www.paraview.org

2.2.7 Experiments

The performed experiments were carried out with the main objective of obtaining a 3D reconstruction from 2D views of CT and echocardiography. These experiments are:

1. **CVAE applied to Burdeos data**. The 3D inputs for this experiment are 3D segmentations from CT scans of the left atrium, while the 2D inputs are 2-4 slices from approximately the center of the 3 anatomical parts. Data augmentation was as well performed, as explained in the previous sections.
2. **CVAE applied to HSP data**. The 3D inputs in this case are 3D segmentations of the left atrium from CT scans, while the 2D inputs are 2 planes including the left atrial appendage, manually segmented from echocardiography.
3. **CVAE applied to HSP data with data augmentation**. The distribution of inputs is the same as in Experiment 2, but in this case, data augmentation was firstly performed, as explained in the previous sections.
4. **CVAE applied both to Burdeos and HSP data, with data augmentation.** The inputs of the network are indistintively inputs from Experiment 1 and 3.
5. **Transfer Learning: pre-trained network in Burdeos applied in HSP.** The network was firstly trained as in Experiment 1, for Burdeos data (see Figure 8). The network was again compiled with the pre-trained weights and re-trained in HSP data, with data augmentation.
6. **CVAE applied to the reconstruction of the LAA in HSP data with data augmentation.** In this 6th experiment, the 3rd experiment was replicated in HSP dataset, but using only the LAA geometry, discarding the rest of the LA.
7. **Transfer Learning in the reconstruction of the LAA.** The pre-trained network in Burdeos dataset was then used in HSP dataset as in Experiment 5, but using only the LAA geometry and discarding the rest of the LA. Data augmentation was also performed.

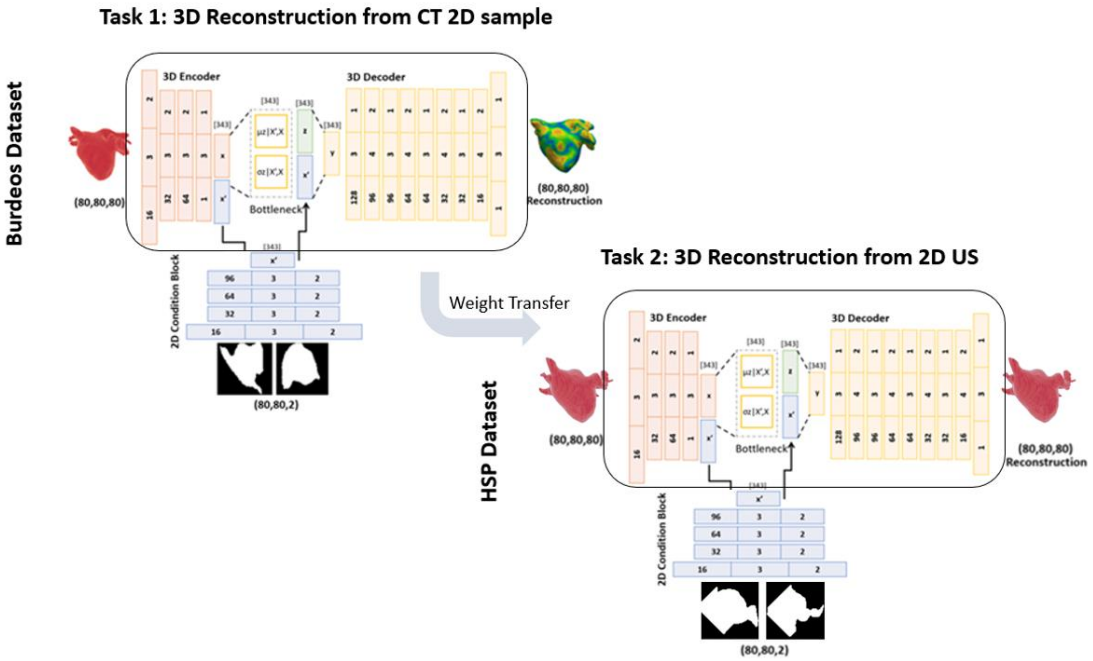

Figure 8. Schema summarizing the methodology for Experiment 5 (Transfer Learning). The network is pre-trained in Burdeos dataset. Once done, the weights for the encoder and the decoder are used, and the model is re-trained on HSP dataset.

For the experiments including the database HSP, the whole left atrium was tried to be reconstructed, as well as the left atrial appendage separately. This was done as many of the 3D CT ground-truth geometries were incomplete, so it could affect the network training and finally its performance as well.

2.2.8 Training the model

To train the model, as previously mentioned, we divided our final datasets including 3D geometries with their corresponding 2D views, used in each of the experiments, into 85-90% training and 10-15% test. Out of this first 90%, a 10% was used for validation (Table 2).

Adam optimizer was selected as optimization algorithm, with a learning rate between $1x10^{-4}$ and $1x10^{-5}$, for 100-200 epochs or iterations, with a batch size of 8. The callback 'ReduceLROnPlateau' was used with a patience of 16 epochs. Basically, when validation loss is not increasing for a certain number of iterations, the learning rate will be decreased to 70% of the original value, so as to optimise the convergence of the network. All these parameters were adjusted by trial and error observing the convergence curve of the model during the iterations.

	Exp. 1	Exp. 2	Exp. 3	Exp.4	Exp.5	Exp.6	Exp.7
Training	174	45	45	218	174+45*	45	174+45
Training (Data Augmentation)	591	-	1215	1877	591+1215	1215	591+1215
Validation	65	5	135	209	65+209	135	65+209
Test	29	5	5	5	29+5	5	29+4

Table 2. Dataset distribution. * The sum indicates that the model is first trained in one dataset, then the trained weights are used in a second training process with a new dataset. Thus, the first terms refer to Burdeos dataset elements, while the second terms to HSP elements.

3. RESULTS

3.1 Results in Burdeos dataset

The conditional variational autoencoder (CVAE) developed in this project was trained, as a first experiment in Burdeos dataset, consisting of 3D segmentations of the left atrium from CTs with their respective 2D planes. The network was trained in 591 cases, including data with rotations as data augmentation, and validated in 65 cases, for 100 epochs. The total training time was around 11.3 hours. Once the training process had finished, the performance was checked in a test set including 29 new volumes. Each prediction is created in < 2 sec.

	Average Dice Score	**Average Hausdorff Distance (mm)**
Total	0.824 ± 0.093	3.08 ± 0.7

Table 3. Quantitative results for the test set subjects of Burdeos database (First experiment).

Each reconstruction is obtained in around 2 seconds. Each of these predictions were compared with their respective ground-truth geometry by computing the Dice Score and the Hausdorff Distance (Equations 6, 7). Table 3 collects these quantitative values evaluating the 3D reconstruction by the CVAE in Burdeos dataset. Figure 9 shows the training and validation curves along the 100 epochs, and two confusion matrices obtained by computing the DS and the HD between all the ground-truth geometries and reconstructions. The higher values in the diagonal show that each of the reconstructions correspond to their respective ground-truth geometries. Finally, Figure 10 shows some of the results obtained for five of the patients of the test set. The rest of results for this part of the project can be found in Annex 2.

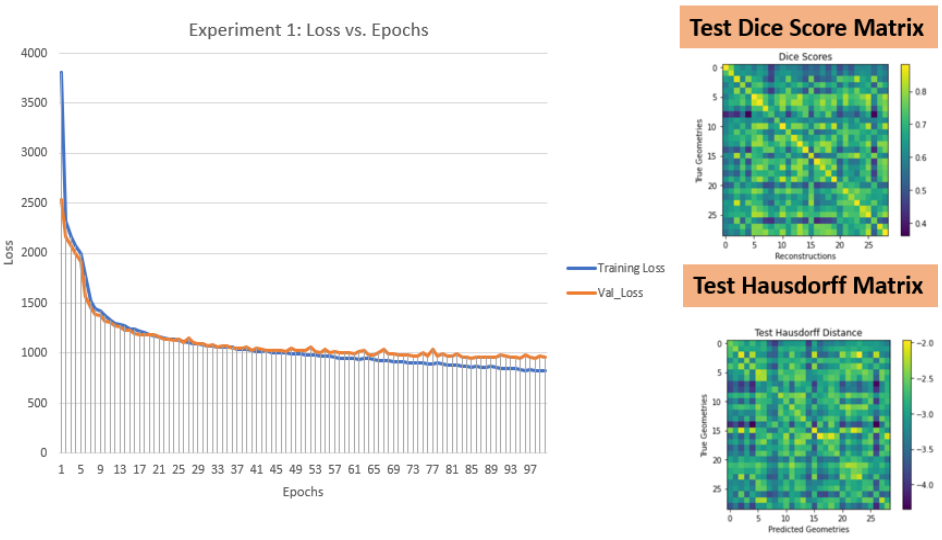

Figure 9. Left: Evolution of training and validation curves for experiment 1. Right: Matrices of correlation computing the DS (top) and the Hausdorff Distance between all the reconstructions an ground-truth geometries in the test set. Higher values in the diagonal indicate that the reconstructions correspond to their respective ground-truth geometries.

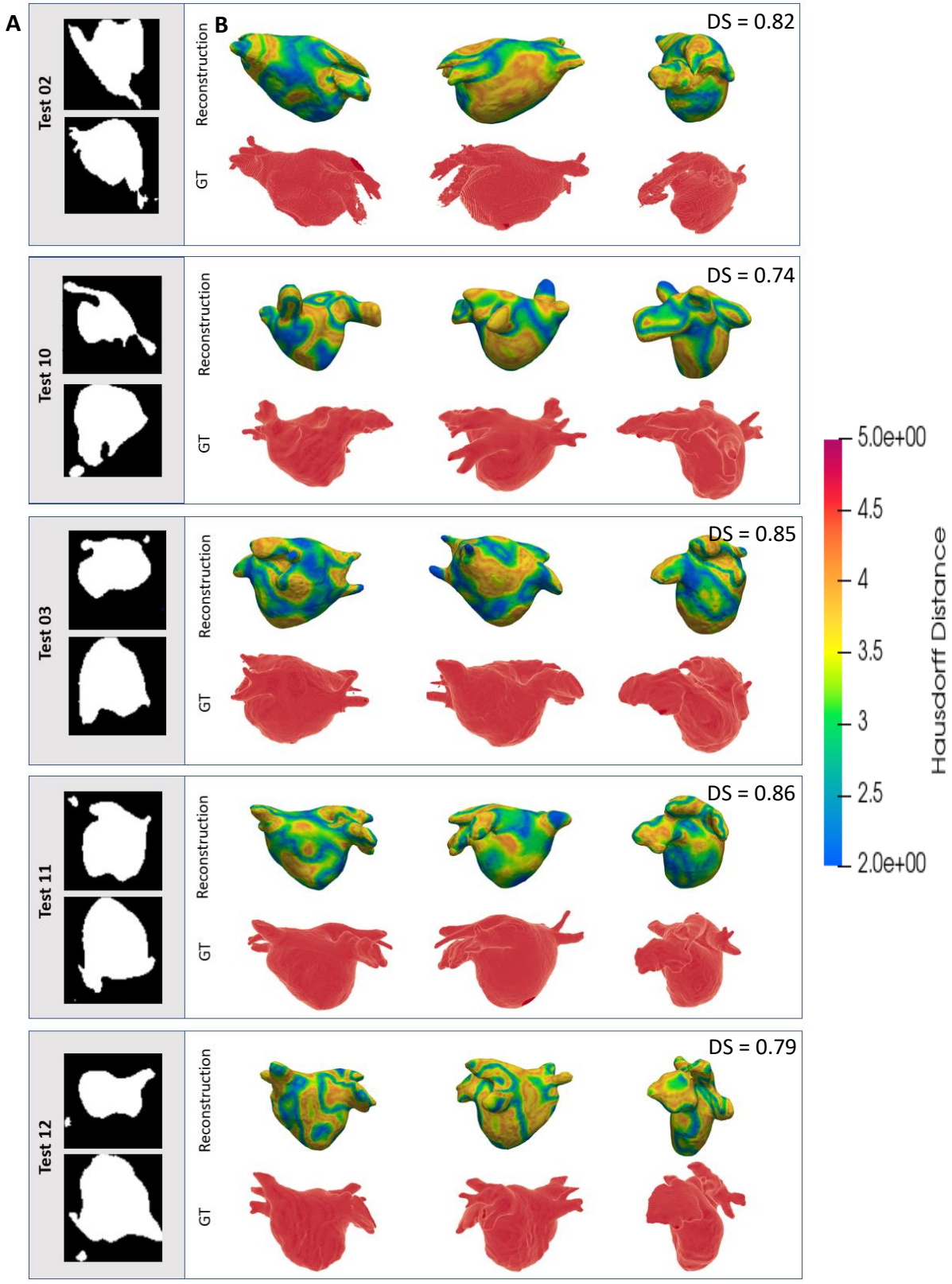

Figure 10. Qualitative results for five subjects from the test set. A. Subject's number with their corresponding two 2D input views. B. Top: Reconstructions performed by the CVAE parting from the input 2D views. The heatmap indicates in red the regions with biggest Hausdorff Distance (mm), while blue regions correspond to lower HD zones. Bottom: Ground-truth geometries.

Fluid Simulation Results

As explained in Methods section, once the reconstructions were obtained, fluid simulations were run with the aim of comparing the behaviour in both ground-truth and predicted geometries. Simulations were running for 200 iterations, and the results for the wall shear stress were analysed in Paraview. Figure 11 shows some of these results for test cases 10 and 12.

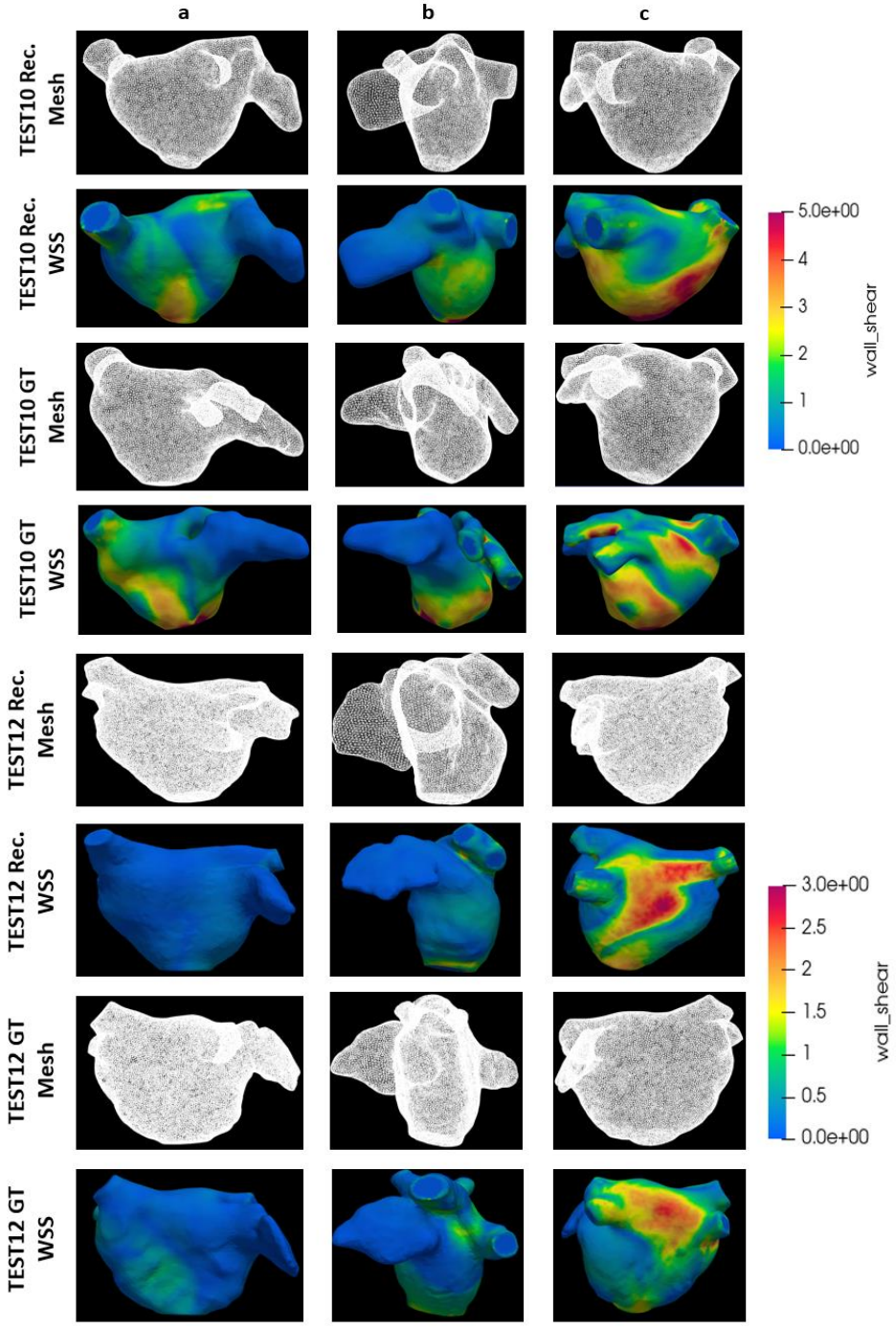

Figure 11. Fluid simulation results for the wall shear stress (WSS), for both ground-truth and predicted geometries (test cases number 10 and 12). a, b and c are different views of the 3D GT left atriums and their reconstructions.

3.2 Results in HSP dataset

Once the network had been trained and was capable to reconstruct 3D geometries from single 2D inputs of CT, it was adapted to perform such task parting from 2D images of echocardiography. This was tried in four different ways, as explained in methods:

- **Experiment 2:** CVAE applied to HSP data.
- **Experiment 3:** CVAE applied to HSP data with data augmentation.
- **Experiment 4:** CVAE applied both to Burdeos and HSP data.
- **Experiment 5:** Transfer Learning.

The best results were obtained for experiment number 3 and 4, as for the other experiments the training curves indicated overfitting or low reconstruction accuracy. The results for experiment 5 can be seen in Figure 12.

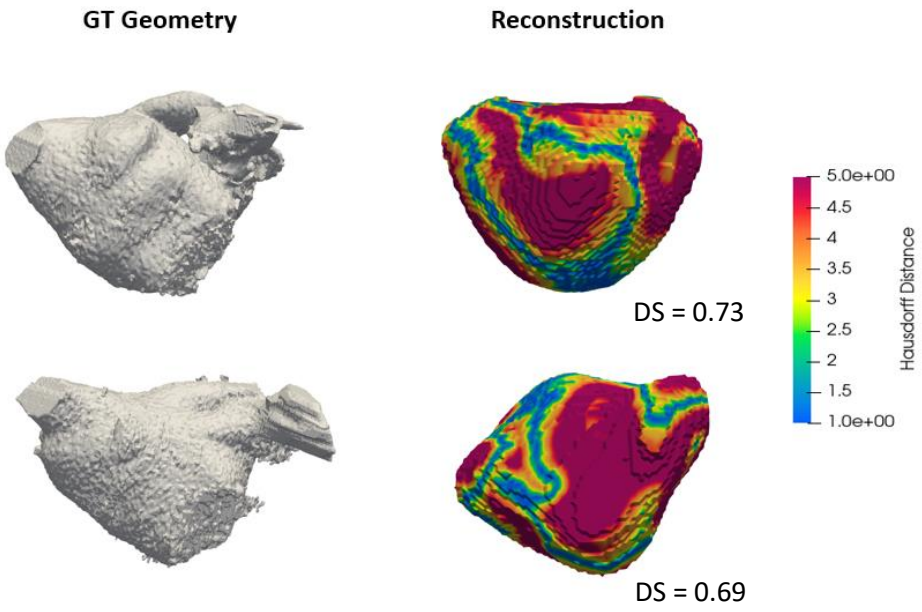

Figure 12. Reconstructions obtained for two test subjects, for experiment 5 (Transfer Learning). The surface of the LA is reconstructed with minimal details.

After these experiments evaluating the performance of the model in the 3D reconstruction of left atriums from 2D US, two more experiments were carried out, aiming to reconstruct the left atrial appendage separately:

- **Experiment 6:** CVAE applied to HSP LAA data with data augmentation.
- **Experiment 7:** Transfer Learning applied to LAA reconstruction.

The results for this last experiment (Experiment 7) indicate a certain better accuracy in terms of the DS. Also, the visual results shown in Paraview (See Figure 13) indicate a better reconstruction of the surface of the LAA. See Table 4.

	Dice Score	**Hausdorff Distance Mean (mm)**
Test_Subject_01	0.79	3.72
Test_Subject_02	0.65	4.5
Test_Subject_03	0.6	4.7
Test_Subject_04	0.78	3.31
Total	**0.70 ± 0.1**	**4.05 ± 0.74**

Table 4. Quantitative results for the test set subjects of HSP database (7^{th} experiment).

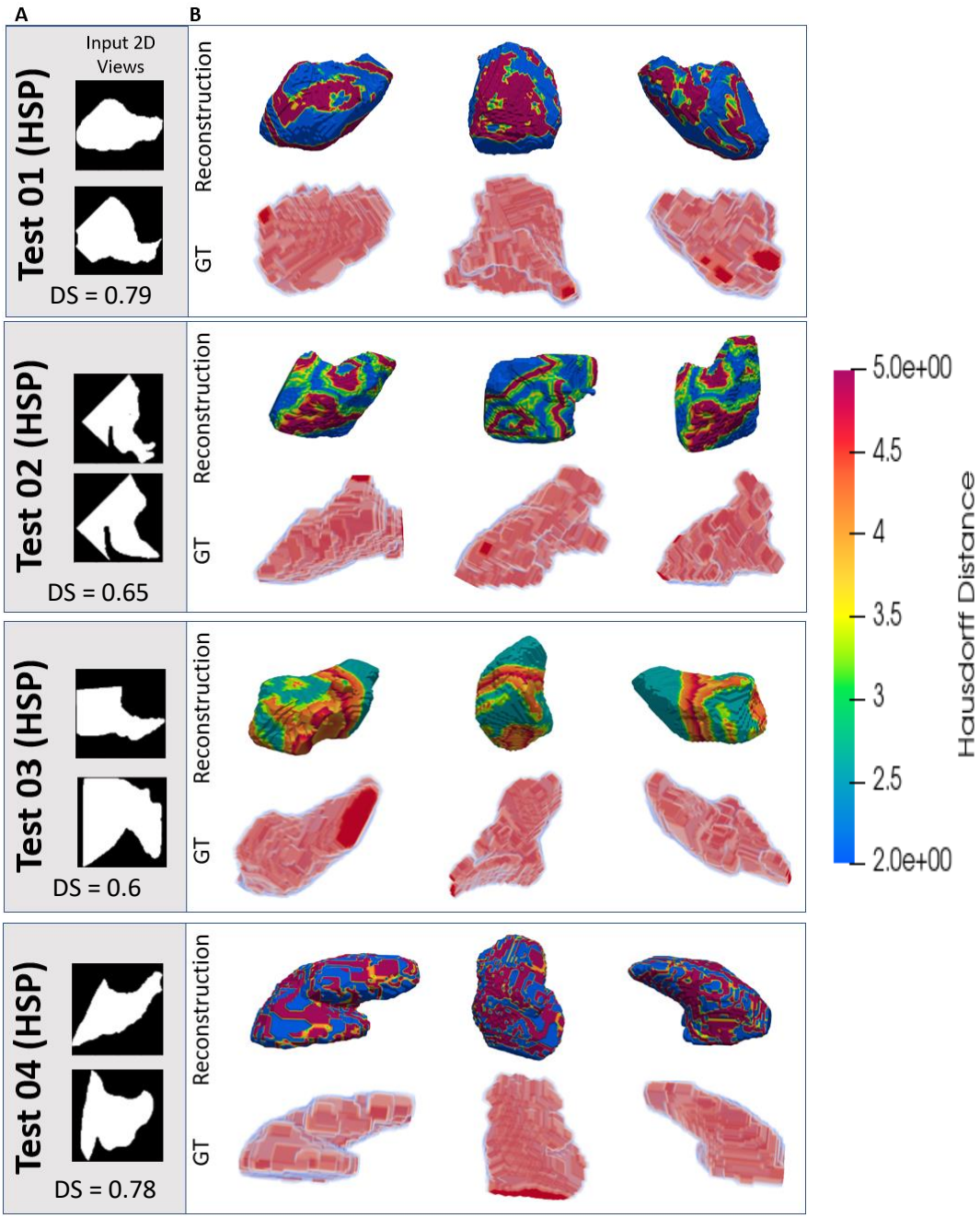

Figure 13. Qualitative results for five subjects from the test set (7th Experiment, HSP dataset). A. Subject's number with their 2 corresponding 2D input views from US. B. Top: Reconstructions performed by the CVAE parting from the input 2D views. The heatmap indicates in red the regions with biggest Hausdorff Distance, while blue regions correspond to lower HD zones. Bottom: Ground-truth geometries.

4. DISCUSSIONS

This project demonstrates the feasibility of a 3D reconstruction problem starting from 2D input data, using deep learning models. The experiments carried out in this work show how starting from a very reduced dataset of CT and ultrasound images of the left atrium, accurate reconstructions of the true 3D geometry can be obtained in a few seconds. This has the potential to be used in clinical environments to run computational simulations in an easy and faster way, to assess clinicians not only in the risk assessment of patients with atrial fibrillation, but also in other cases in which approximate 3D geometries of anatomical structures are needed to be obtained.

As exposed in the introduction and state of the art sections, in the recent years, researchers have been trying to develop new deep learning methodologies for the task of 3D reconstruction. However, the task of reconstructing a 3D geometry from 2D ultrasound images, in which fluid simulations can be performed is still a challenge itself, and not many investigations have been carried out on it, despite the potential benefits that a model being able to perform this action could provide. In the current work, the implementation of the conditional variational autoencoder proposed in Biffi et al. [19] was adapted to this task and optimized by applying some modifications and proposing different pre-processing ways.

These 3D reconstructions were achieved, on one hand, starting from 2D slices of 3D CT images, in a dataset including 193 left atrium 3D segmentations from CT, Burdeos dataset. The obtained reconstructions demonstrated a promising performance of the algorithm chosen in this project to perform the 3D reconstructions, not only because of the high DS and low HD obtained comparing ground-truth and predictions, but also because of their similar behaviour when performing fluid simulations. This, at the same time, suggests that our first objective would be accomplished: obtaining a reconstruction in which fluid simulations could be run, giving similar results to the expected ones, thus being able to trust the model predictions when needing to do patient risk assessment. Observing the results obtained for this first step of the project, we can say that the model is capable of obtaining an approximate enough reconstruction of the whole left atrium structure, with only introducing the information contained in two 2D slices of this whole geometry. The body of the left atrium is accurately reconstructed, while some lack of details can be found if we look at the pulmonary veins' reconstruction (see Figure 10). This behaviour was already expected, as we are not introducing enough information in these 2D samples of these pulmonary veins, thus the model just does an approximation of them. By increasing the number of views introduced to train the model to perform this reconstruction, the accuracy of the reconstruction of the pulmonary veins also increased. However, this was not our main goal, and in practice, with echocardiography images, it will not be optimal to introduce more than 2-4 views including the LAA. Moreover, when performing the fluid simulations to compare the behaviour between both ground-truth and reconstructed geometries, similar patterns for the WSS were found in the body of the left atrium, in the same range of values, and with some minimal differences, suggesting that

the reconstruction from 2 2D views is enough to run predictive computational simulations.

On the other hand, the results on HSP showed that very approximated reconstructions can be obtained from 2D echocardiography segmentations, as happened in the case in which we tried to reconstruct the LAA geometry training the model with the 3D true geometries, and ultrasound 2D segmentations. However, if we look at the reconstructions performed when trying to reconstruct the whole LA from 2D ultrasound images, we can observe that the performance of the model is decreased. This was somehow expected for three different reasons: the reduced database used in this second experiment, the diversity of geometries of the left atrium (being most of them incomplete geometries, not showing the entire left atrium) and the partial information of the left atrium present in the 2D transoesophageal ultrasound images. For these reasons, the performance of the network was affected, as too much variability was found by the autoencoder-based structure for the left atrium geometries, thus making difficult to learn the probabilistic distribution of features of an anatomically correct left atrium. Also, when trying to reconstruct the whole LA, the reconstruction of the LAA was also harmed.

In addition to the quantity and quality of data and the different pre-processing ways applied in this work to optimize the training of the model, one of the factors that have been playing a key role in the 3D reconstruction process was the number of 2D views. Different number of views were introduced in different experiments in the training of the model, both in Burdeos reconstruction (from 2D CT slices) and in HSP reconstructions (from 2D echocardiography data). Although it seems obvious that with more number of views we are incrementing the information that has the model from the real structure, thus optimizing the performance of it, it is not exactly like that. It is true that in Burdeos dataset, we achieved more accuracy in the reconstruction process, due to an improvement of the accuracy of reconstruction for the pulmonary veins. However, as exposed, this is not our main goal. Also, most of the echocardiography studies only included at most two different views showing the LAA. As the improvements in terms of the accuracy are not relevant for a number of views above 2, this was selected as the optimum number. Here it is also important to remember how this would be applied into clinical practice, using a micro-transoesophageal ultrasound probe, thus ensuring a process which is faster, efficient and more comfortable for the patient, with minimal sedation. In contrast, a more complex procedure could be applied, using more advanced TEE ultrasound probes, with the cost of having to sedate the patient, but being able to obtain more views of the left atrium, which might increase the performance of the model.

Finally, our methodology achieved to solve the problems associated with the 3D reconstruction of a complex structure such as the LA and the LAA, with partial information. Although more investigation is needed to perform the procedure in more extensive datasets, our preliminary results show that our methodology based on variational autoencoders can successfully learn the distribution of features that an anatomically correct LA must have. This is essential, so as when we want to obtain a full reconstruction of the real geometry from the minimal information that 2D views can provide, the model can recreate a LA which is plausible anatomically, and at the same time, corresponds to the information contained in these views.

Further Work

The preliminary results of this work have shown the potential of conditional variational autoencoders in the task of obtaining 3D reliable geometries of the LAA starting from the partial information contained in 2D views from CT slices or 2D echocardiography segmentations. However, more investigation should be carried out in this ambit in order to be able to apply the methodology presented in this work in clinical practice, obtaining a 3D geometry in a few seconds after the obtention of the 2D echocardiography images and being able to run reliable enough fluid simulations on them.

As it is obvious after the exposition of results and the discussions, one of the main limitations of this project has been the quantity and quality of data, overall, in HSP dataset, thus affecting the results obtained for the reconstruction from echocardiography. Further investigations should focus principally on the obtention of the data for this second dataset of the project. It is not enough to obtain more cardiac CT images with their corresponding 2D TEE ultrasound images, but also making enough emphasis on the quality of them. In the present job, the variability of the 3D geometries of the LA, have constituted a major obstacle in the learning process of the model of the geometry of interest. Some of the obtained ground-truth geometries of the LA were incomplete, as only 17 of the 50 ground-truth geometries contained information about the lower part of the LA, including the mitral valve. This is due to the fact that most of the patients included in this work were patients with atrial fibrillation who needed an implantation of the LAAO device so clinicians were interested in the upper part of the LA (including the LAA). To reduce the radioactivity dose, and the cost of the CT scan, the acquired CT images did only include this upper part of the LA. This is beneficial in clinical practice, but it difficulted us the task of finding complete geometries of the LA, which would have been ideal to train the model with the full shape and avoiding this source of variability.

On the other hand, one of the problems of echocardiography which make it difficult the objective of this project, is that we do not know exactly the correspondence of these 2D images in the real 3D space. To face this problem, we were designing a strategy to co-registrate the 2D segmentations of echocardiography with the 3D CT segmentations of the left atrium, which consisted of a re-formatting of the CT scan to obtain a 2D plain which matches the 2D segmentation of ultrasound, then performing a co-register between these 2 2D images. This way, the 2D condition block of the model would be transformed into a 3D condition block, to which the 2D ultrasound segmentation inputs would be introduced in the form of a 3D matrix with their real orientation. However, it is a complex process, and the lack of time made impossible to perform this last task, which could give our network more clues about the real coordinates of the 2D inputs, thus improving the performance.

Also, in the last part of the project, when trying to perform the reconstructions in HSP dataset, we have observed that the LAA is better reconstructed when it is separated from the LA, training the model exclusively for this task. We have also seen how the information of the LA in 2D TEE images including the LAA is limited, and this is one of the reasons why the model is not capable to reconstruct properly the LA surface. To solve this, further investigations could be focused on training two different models separately. The first of them should be dedicated to be trained with 2D TEE images of the LAA, and with 3D geometries of the LAA, as we did in Experiments 6 and 7. The second one instead, should be trained with 2D TEE images of different views of the LA, and be trained to reconstruct the body of the LA as good as possible. Both reconstructions (LA and LAA) should be joined at the end.

Last of all, more research on the post-processing of the obtained segmentations should be carried out in order to return the predicted reconstructions to the most accurate real dimensions, so as to obtain the most realistic fluid simulation results.

5. CONCLUSIONS

The present work has demonstrated the potential of deep learning and variational autoencoders in a 3D reconstruction task of the LA and LAA from 2D data with a high dice score. Being able to improve these results also with echocardiography images, would help clinicians to optimise the process of risk stratification of patients with pathologies such as AF, by just acquiring a pair of 2D echocardiography images of the LA, obtaining a 3D reconstruction in a few seconds, and doing computational simulations on them. This would mean, avoiding the necessity of having to perform CT scans to obtain a 3D segmentation of the LA, thus reducing the radioactivity dose for the patient, saving time and costs.

The preliminary results obtained in this project, show how variational autoencoders can successfully learn the anatomical shape of a LA and a LAA, and relate these 3D shapes with 2D information in the latent space. Despite this, further investigation is needed, and should focus on the obtention of more data, and with more quality, so as the trained model would be capable to generate an anatomically correct 3D structure and relate it with the 2D information provided by the echocardiography. Other investigation lines could also focus on other methodologies, such as performing a previous co-registration between the 2D echocardiographic image and the CT scan, or trying new types of data augmentation. This would probably allow to move closer to the objective of obtaining a cheaper, safe, and more efficient method for risk stratification of patients with AF thanks to US imaging and DL.

6. BIBLIOGRAPHY

[1] Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population. Am J Cardiol. 2013 Oct 15;112(8):1142-7. doi: 10.1016/j.amjcard.2013.05.063

[2] Freedman B, Potpara TS, Lip GY. Stroke prevention in atrial fibrillation. Lancet. 2016 Aug 20;388(10046):806-17. doi: 10.1016/S0140-6736(16)31257-0. PMID: 27560276.

[3] Lip GY, Lane DA. Stroke prevention in atrial fibrillation: a systematic review. JAMA. 2015 May 19;313(19):1950-62. doi: 10.1001/jama.2015.4369

[4] Oladiran O, Nwosu I. Stroke risk stratification in atrial fibrillation: a review of common risk factors. J Community Hosp Intern Med Perspect. 2019 Apr 12;9(2):113-120. doi: 10.1080/20009666.2019.1593781

[5] Bosi GM, Cook A, Rai R, Menezes LJ, Schievano S, Torii R, Burriesci G Burriesci. Computational Fluid Dynamic Analysis of the Left Atrial Appendage to Predict Thrombosis Risk. Front Cardiovasc Med. 2018 Apr 4;5:34. doi: 10.3389/fcvm.2018.00034.

[6] Aguado AM, Olivares AL, Yagüe C, Silva E, Nuñez-García M, Fernandez-Quilez Á, Mill J, Genua I, Arzamendi D, De Potter T, Freixa X, Camara O. *In silico* Optimization of Left Atrial Appendage Occluder Implantation Using Interactive and Modeling Tools. Front Physiol. 2019 Mar 22;10:237. doi: 10.3389/fphys.2019.00237.

[7] MRI to CT Translation with GANs. Bodo Kaiser et al. (2019) arXiv:1901.05259

[8] Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images. Nanyang Wang et al. (2018) ECCV2018 arXiv:1804.01654

[9] Pixel2Mesh++: Multi-View 3D Mesh Generation via Deformation. Chao Wen et al. (2019) ICCV2019. arXiv:1908.01491

[10] Probabilistic 3D surface reconstruction from sparse MRI information. Katarína Tóthová et al. (2020) MICCAI 2020 pp 813-823|

[11] Vlontzos, Athanasios & Budd, Samuel & Hou, Benjamin & Rueckert, Daniel & Kainz, Bernhard. (2020). 3D Probabilistic Segmentation and Volumetry from 2D Projection Images. 10.1007/978-3-030-62469-9_5.

[12] PHiSeg: Capturing Uncertainty in Medical Image Segmentation. Baumgartner, C.F et al. MICCAI pp. 1–14 (2019) arXiv:1906.04045

[13] C-Flow: Conditional Generative Flow Models for Images and 3D Point Clouds. Albert Pumarola et al. Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 7949-7958 arXiv:1912.07009

[14] PointFlow: 3D Point Cloud Generation with Continuous Normalizing Flows. Guandao Yang et al. 2019 IEEE/CVF International Conference on Computer Vision (ICCV) arXiv:1906.12320

[15] Feinberg WM, Blackshear JL, Laupacis A, Kronmal R, Hart RG. Prevalence, age distribution, and gender of patients with atrial fibrillation. Analysis and implications. Arch Intern Med. 1995 Mar 13;155(5):469-73.

[16] Cresti A, García-Fernández MA, Sievert H, Mazzone P, Baratta P, Solari M, Geyer A, De Sensi F, Limbruno U. Prevalence of extra-appendage thrombosis in non-valvular atrial

fibrillation and atrial flutter in patients undergoing cardioversion: a large transoesophageal echo study. EuroIntervention. 2019 Jun 12;15(3):e225-e230. doi: 10.4244/EIJ-D-19-00128.

[17] Genua I. et al. (2019) Centreline-Based Shape Descriptors of the Left Atrial Appendage in Relation with Thrombus Formation. In: Pop M. et al. (eds) Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges. STACOM 2018. Lecture Notes in Computer Science, vol 11395. Springer, Cham. https://doi.org/10.1007/978-3-030-12029-0_22

[18] Zheng G, Gollmer S, Schumann S, Dong X, Feilkas T, González Ballester MA. A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated X-ray images. Med Image Anal. 2009 Dec;13(6):883-99. doi: 10.1016/j.media.2008.12.003.

[19] Biffi, Carlo & Cerrolaza, Juan & Tarroni, Giacomo & de Marvao, Antonio & Cook, Stuart & O'Regan, Declan & Rueckert, Daniel. (2019). 3D High-Resolution Cardiac Segmentation Reconstruction from 2D Views using Conditional Variational Autoencoders.

[20] Morales, Xabier & Mill, Jordi & Juhl, Kristine & Olivares, Andy & Jimenez-Perez, Guillermo & Paulsen, Rasmus & Camara, Oscar. (2020). Deep Learning Surrogate of Computational Fluid Dynamics for Thrombus Formation Risk in the Left Atrial Appendage. 10.1007/978-3-030-39074-7_17.

[21] Aguado AM, Olivares AL, Yagüe C, Silva E, Nuñez-García M, Fernandez-Quilez Á, Mill J, Genua I, Arzamendi D, De Potter T, Freixa X, Camara O. *In silico* Optimization of Left Atrial Appendage Occluder Implantation Using Interactive and Modeling Tools. Front

Physiol. 2019 Mar 22;10:237. doi: 10.3389/fphys.2019.00237. PMID: 30967786; PMCID: PMC6440369.

[22] An Introduction to Variational Autoencoders. Diederik P. Kingma, Max Welling. Foundations and Trends in Machine Learning: Vol. 12 (2019): No. 4, pp 307-392. DOI: 10.1561/2200000056

[23] Bank, Dor & Koenigstein, Noam & Giryes, Raja. (2020). Autoencoders.

[24] Doersch, Carl. (2016). Tutorial on Variational Autoencoders.

[25] Jinwon An et al. 2015 Variational Autoencoder based Anomaly Detection using Reconstruction Probability. - 2 Special Lecture on IE

[26] S. Suh, D. H. Chae, H. Kang and S. Choi, "Echo-state conditional variational autoencoder for anomaly detection," 2016 International Joint Conference on Neural Networks (IJCNN), 2016, pp. 1015-1022, doi: 10.1109/IJCNN.2016.7727309.

[27] Cerrolaza J.J. et al. (2018) 3D Fetal Skull Reconstruction from 2DUS via Deep Conditional Generative Networks. In: Frangi A., Schnabel J., Davatzikos C., Alberola-López C., Fichtinger G. (eds) Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. MICCAI 2018. Lecture Notes in Computer Science, vol 11070. Springer, Cham. https://doi.org/10.1007/978-3-030-00928-1_44

[28] Wunderlich N.C.,Beigel R.,Swaans M.J.,Ho S.Y., Siegel R.J. Percutaneous interventions for left atrial appendage exclusion: options, assessment, and imaging using 2D and 3D echocardiography. *JACC Cardiovasc Imaging*. 2015; **8**: 472-488

[29] Bin Wang, MD et al. Transnasal Transesophageal Echocardiography Guidance for Percutaneous Left Atrial Appendage Closure VOLUME 108, ISSUE 3, E161-E164, SEPTEMBER 01, 2019:https://doi.org/10.1016/j.athoracsur.2019.01.039

[30] Afonso B. Freitas-Ferraz et al. Safety of Transesophageal Echocardiography to Guide Structural Cardiac Interventions, Journal of the American College of Cardiology, Volume 75, Issue 25, 2020, Pages 3164-3173, ISSN 0735-1097, https://doi.org/10.1016/j.jacc.2020.04.069.

[31] Candelise L, Pinardi G, Morabito A. Mortality in acute stroke with atrial fibrillation. The Italian Acute Stroke Study Group. Stroke. 1991 Feb;22(2):169-74. doi: 10.1161/01.str.22.2.169

[32] Duncan PW, Zorowitz R, Bates B, Choi JY, Glasberg JJ, Graham GD, Katz RC, Lamberty K, Reker D. Management of Adult Stroke Rehabilitation Care: a clinical practice guideline. Stroke. 2005 Sep;36(9):e100-43. doi: 10.1161/01.STR.0000180861.54180.FF

[33] Transesophageal Echocardiography for Congenital Heart Disease. Hui, Ling and Sahn, David J. (2014). Springer London. doi: 10.1007/978-1-84800-064-3_2

[34] Esophageal echocardiography. L Frazin, J V Talano, L Stephanides, H S Loeb, L Kopel, and R M Gunnar. Jul 1976. Circulation, Vol. 54 No. 1, doi: https://doi.org/10.1161/01.CIR.54.1.102

[35] Feasibility of pediatric probes in transesophageal echocardiography guiding of left atrial appendage closure in adults. M Maarse (Nieuwegein,NL), MN Klaver (Nieuwegein,NL), MJ Swaans (Nieuwegein,NL), LVA Boersma (Nieuwegein,NL). EuroIntervention 2021;17:93-96. DOI: 10.4244/EIJ-D-19-00943

[36] Sebastian Stec, Beata Zaborska, Małgorzata Sikora-Frąc, Tomasz Kryński, Piotr Kułakowski, First experience with microprobe transoesophageal echocardiography in non-sedated adults undergoing atrial fibrillation ablation: feasibility study and comparison with intracardiac echocardiography, EP Europace, Volume 13, Issue 1, January 2011, Pages 51–56, https://doi.org/10.1093/europace/euq349

ANNEX 1: Additional Methods

1. CT Segmentation (HSP Database) using Mimics Research 23.0

In this section, a fully detailed explanation on the segmentation process of the CTs from HSP database will be done. The final objective is to create a smoothed CT mask of the whole left atrium, prepared to be introduced as 3D input (ground-truth) for the CVAE model, performing the reconstruction.

Once all the necessary CTs, from patients also having TEE images, are exported, DICOMS must be opened with Mimics Research 23.0 (Materialise).

1. Select the 35-40% phase of the cardiac CT scan. It corresponds to the ventricular systole phase, so atriums would be in diastolic phase.
2. If it is possible, select a studio in which contrast is used. This way, the segmentation of the atrium using thresholds will be much easier.

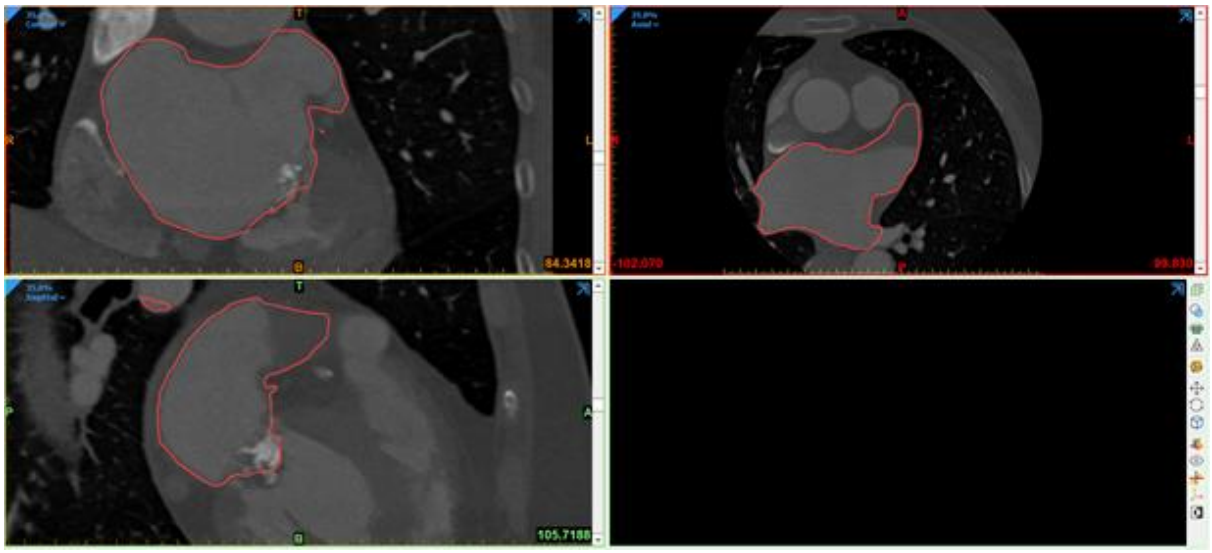

Figure 14. Visualization of case 1 after being opened with Mimics Research 23.0. The red line surrounding the left atrium structure indicates that this particular case has been already segmented, as the study being used as tutorial was already processed.

3. In the section “Segment”, select the Threshold tool. Adjust the lower threshold in a way that all the heart cavities are effectively marked, eliminating the other tissues from the segmentation the best as possible. Adjust also the upper threshold.

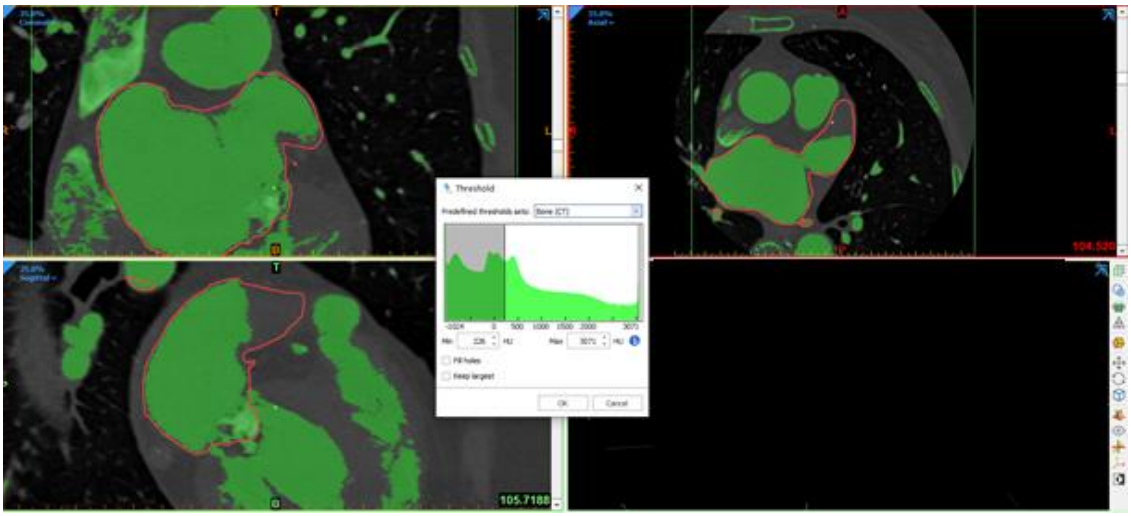

Figure 15. Visualization of the window of the Threshold tool. In the background, we can see the result of the threshold applied, segmenting all the heart cavities, as well as other tissues.

4. Select the tool “Split Mask”. With this tool, we are going to separate the left atrium from the other cavities, as it is our region of interest. We are going to use two regions: region A, and B. For Region A, in red, we are going to mark in some slices the left atrium, as it can be seen in Figure 16. For Region B, in blue, we are going to mark all the tissues which have been segmented in the last step, but are not part of the left atrium.

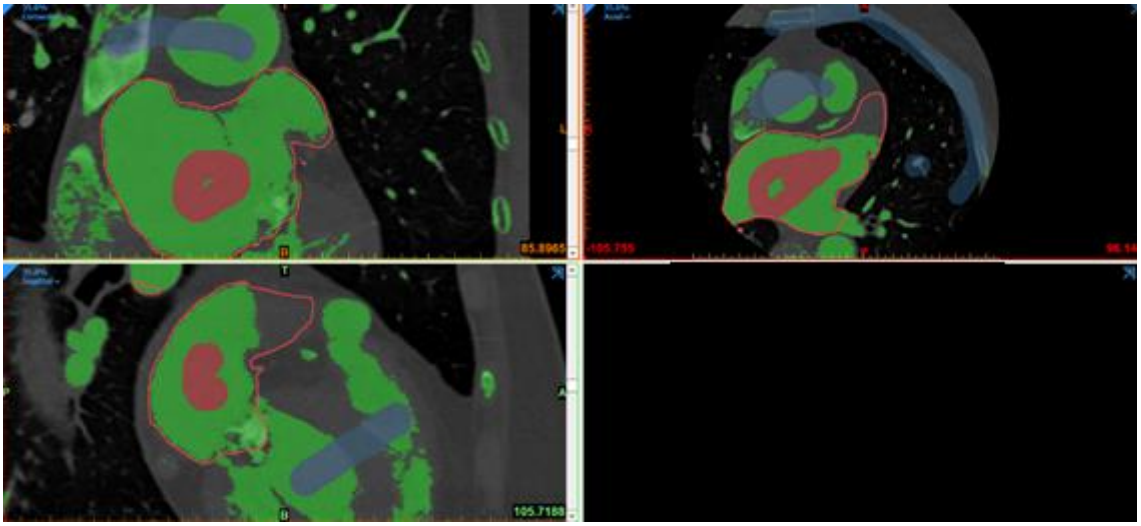

Figure 16. Visual example of the procedure that must be followed with the tool “Split Mask”, separating the left atrium and the other tissues.

5. Once the 4th step is done, it should look like in Figure 17. As we can see, from the selection we have done for regions A and B, the program has automatically doe a region growing, separating the left atrium and the other anatomical parts. We can discard region B.

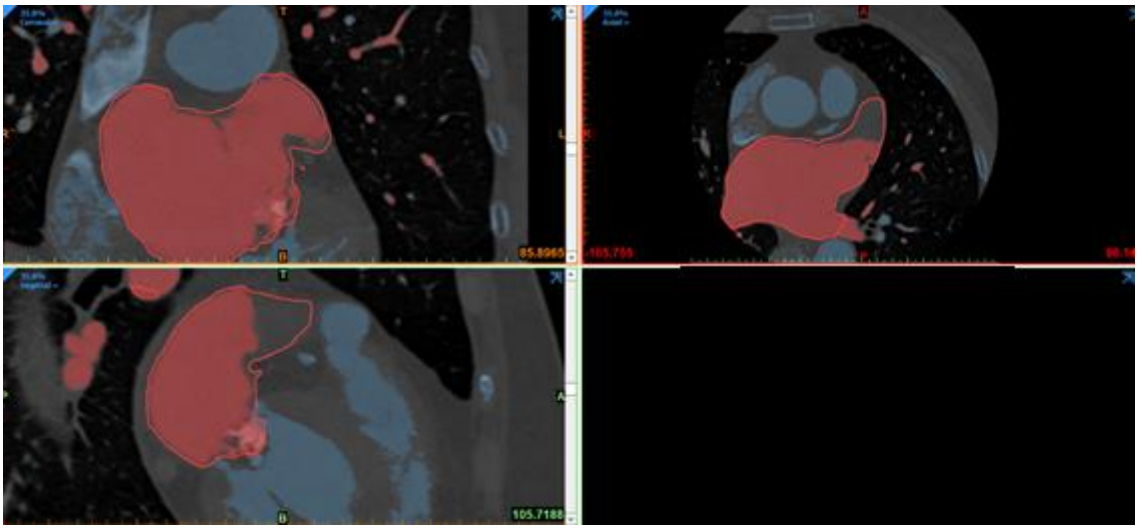

Figure 17. Results after applying the tool "Mask Split". In red, the segmented LA, in blue the other tissues and cavities.

6. For some of the patients, overall the ones who have the LAAO implanted, it is very difficult to segment the left atrial appendage with a simple threshold. The reason behind this is that blood do not circulate properly through the LAA because of the LAAO, or because of the clot formation in this cavity. Here we have two solutions:

   - Manual segmentation of the left atrial appendage.
   - Apply as in step 3 a new threshold which can segment the LAA.

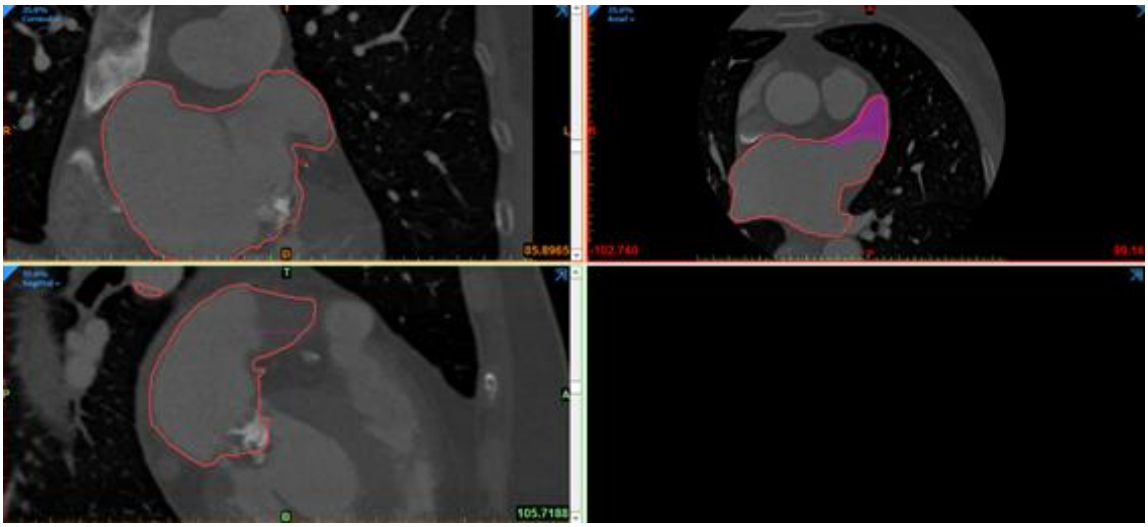

Figure 18. Application of the threshold to segment the LAA.

7. Now the whole left atrium is segmented. Use the tool “smooth” to post-process the resulting surface. The result for this case can be observed in Figure 19.

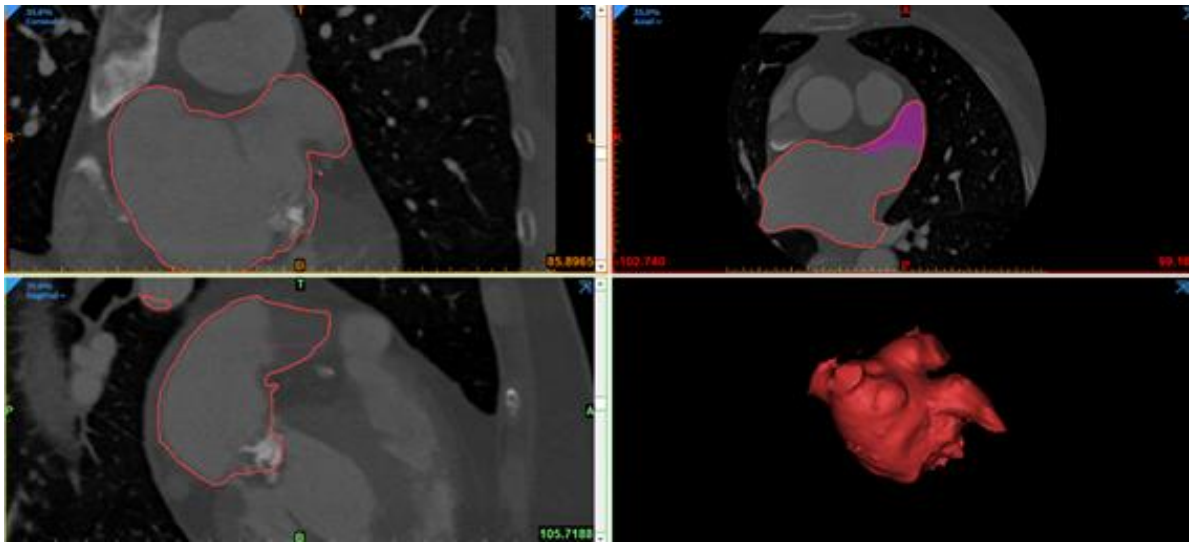

Figure 19. Final segmentation of the left atrium after the whole process described in this Annex 1.

8. Finally, the last part, is to export the segmentation as a .stl.

2. Segmentation of the 2D TEE US Images Using 3D Slicer 4.11

1. Run 3D Slicer 4.11 and load the DICOM file that needs to be segmented. The selected image should include the LAA, as in Figure 20.

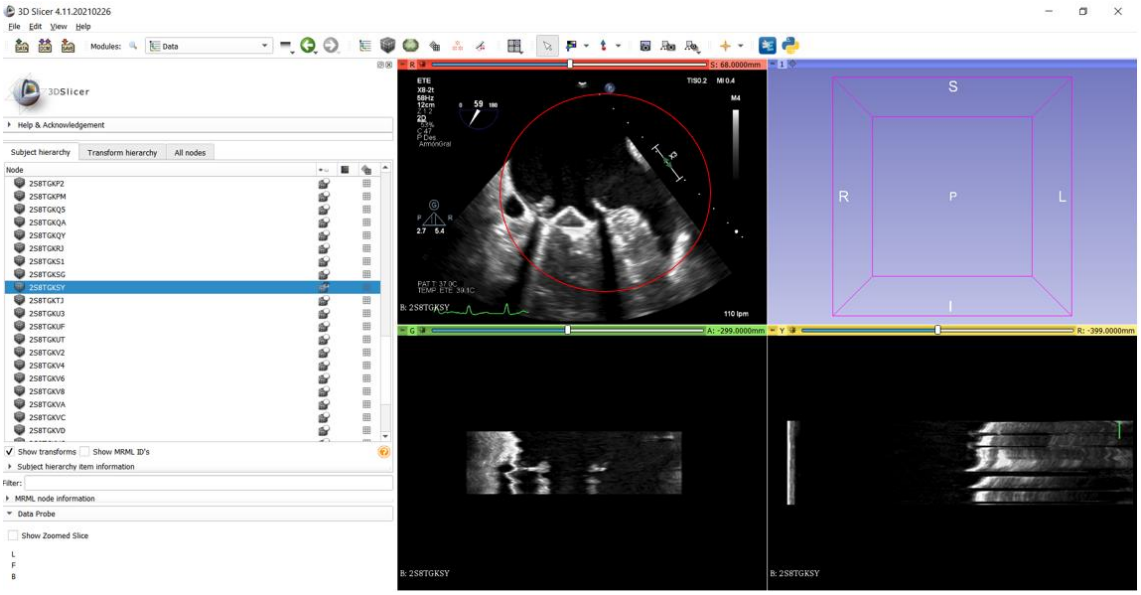

Figure 20. Visualization of the chosen frame to segment.

2. Select “Segment Editor” in the menu, and press “Add”.

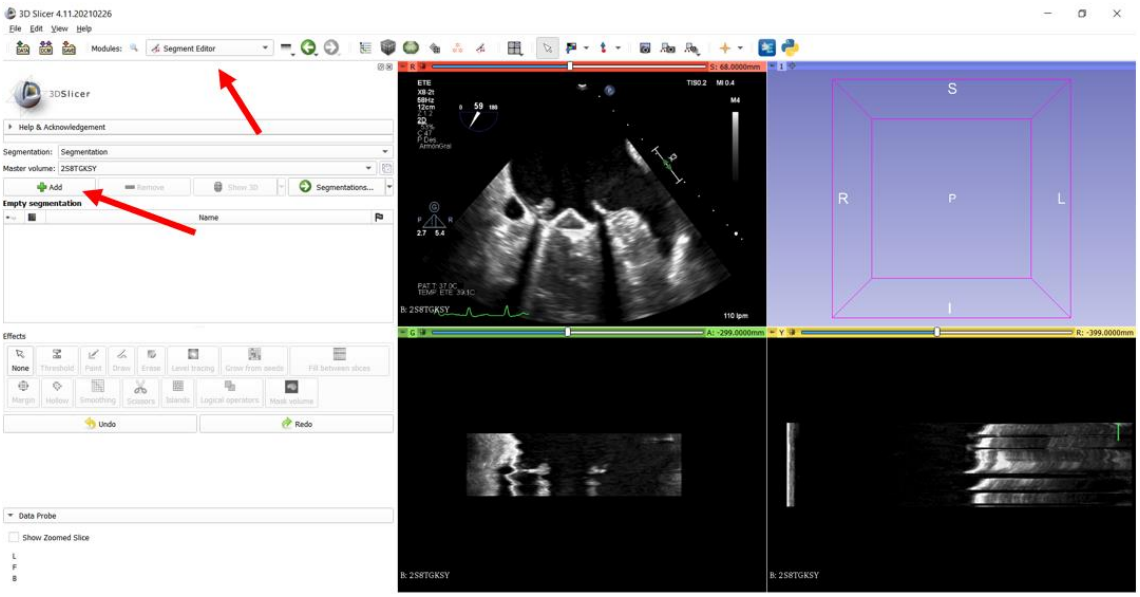

Figure 21. Procedure followed to add a segmentation on the echocardiography frame.

3. Choose the “Threshold” tool, and adjust the upper and lower thresholds, so as the darkest parts of the image are selected, thus discarding the tissue part. Press “Apply”.

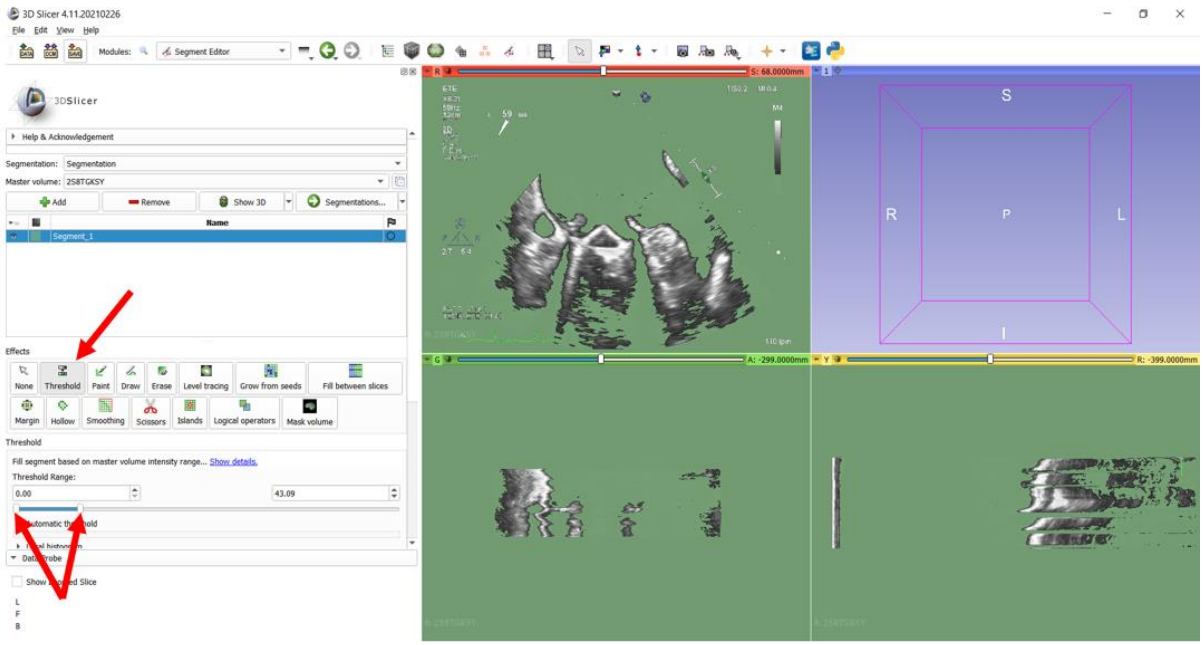

Figure 22. The tool threshold is selected, and the upper and lower thresholds are adjusted so as the whole cavity is included in the segmentation.

4. With the tools “Scissors” and “Erase” eliminate the other tissues, not belonging to the LA, that have been oversegmented with the threshold.

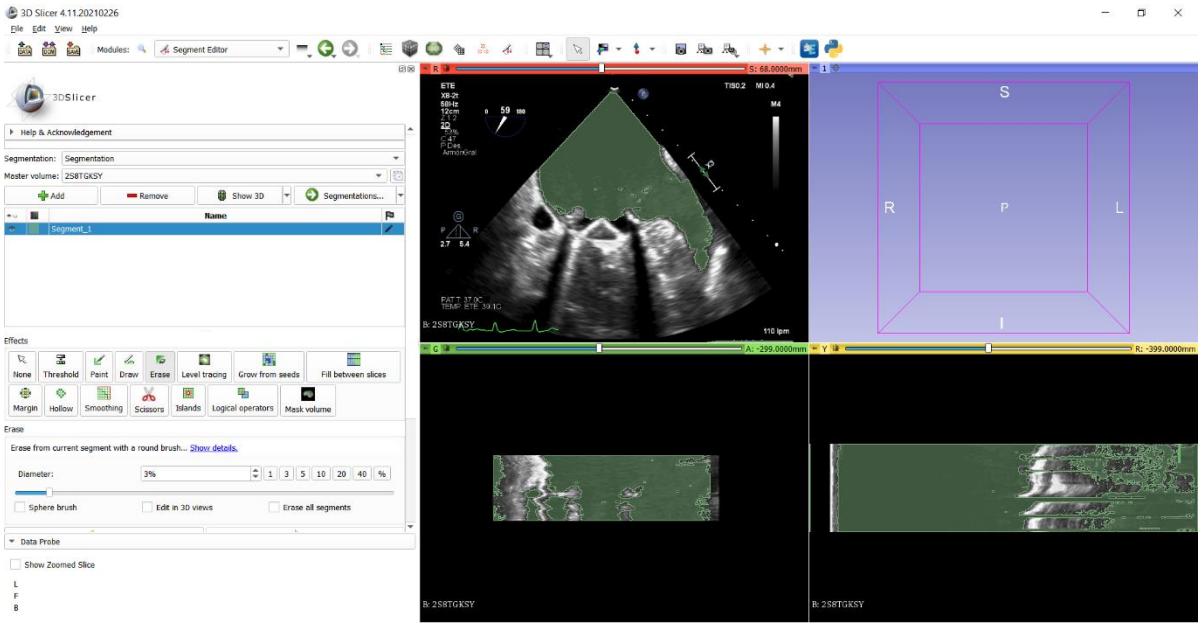

Figure 23. With the tool “Scissors” and “Erase”, the oversegmentation is eliminated.

5. Finally, go to “Segmentations” in the menu, select your segmentation and export to the desired folder in the format .stl.

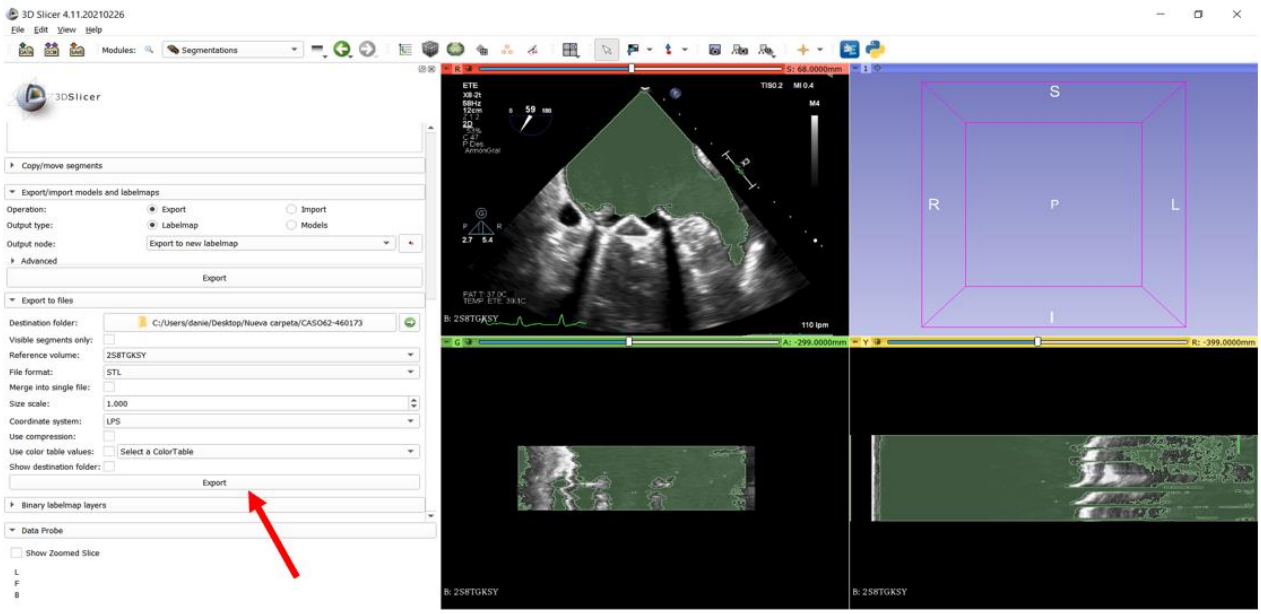

Figure 24. Exporting the echocardiography segmentation in format .stl.

ANNEX 2: ADDITIONAL RESULTS FOR THE RECONSTRUCTIONS FROM 2D SLICES

1. CT Reconstructions From 2D Slices

In this second Annex the different 29 ground-truth geometries for the test set with their respective reconstructions are shown (Burdeos Dataset).

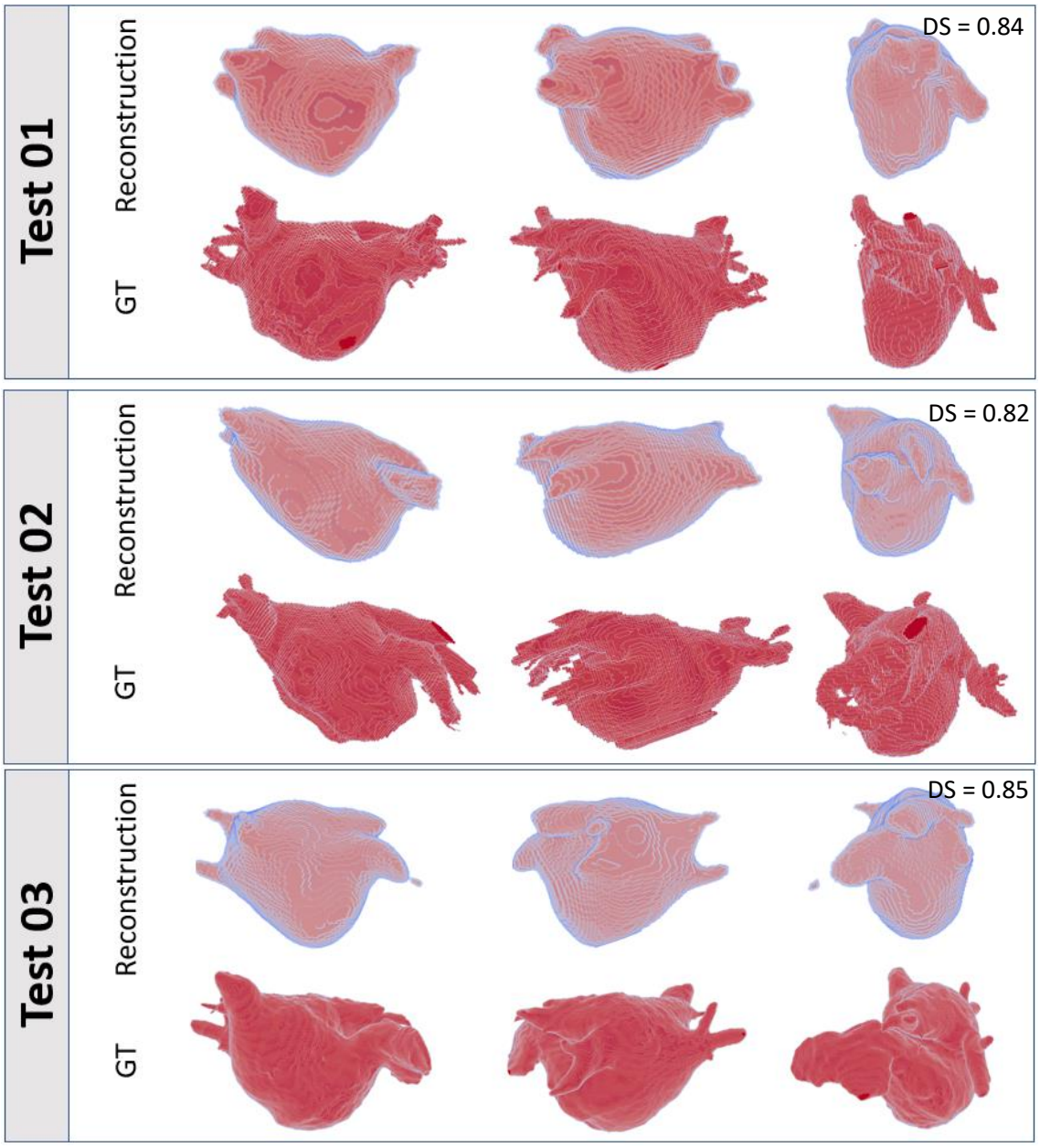

Figure 25. Ground-truth and respective reconstructions for test subjects 1-3.

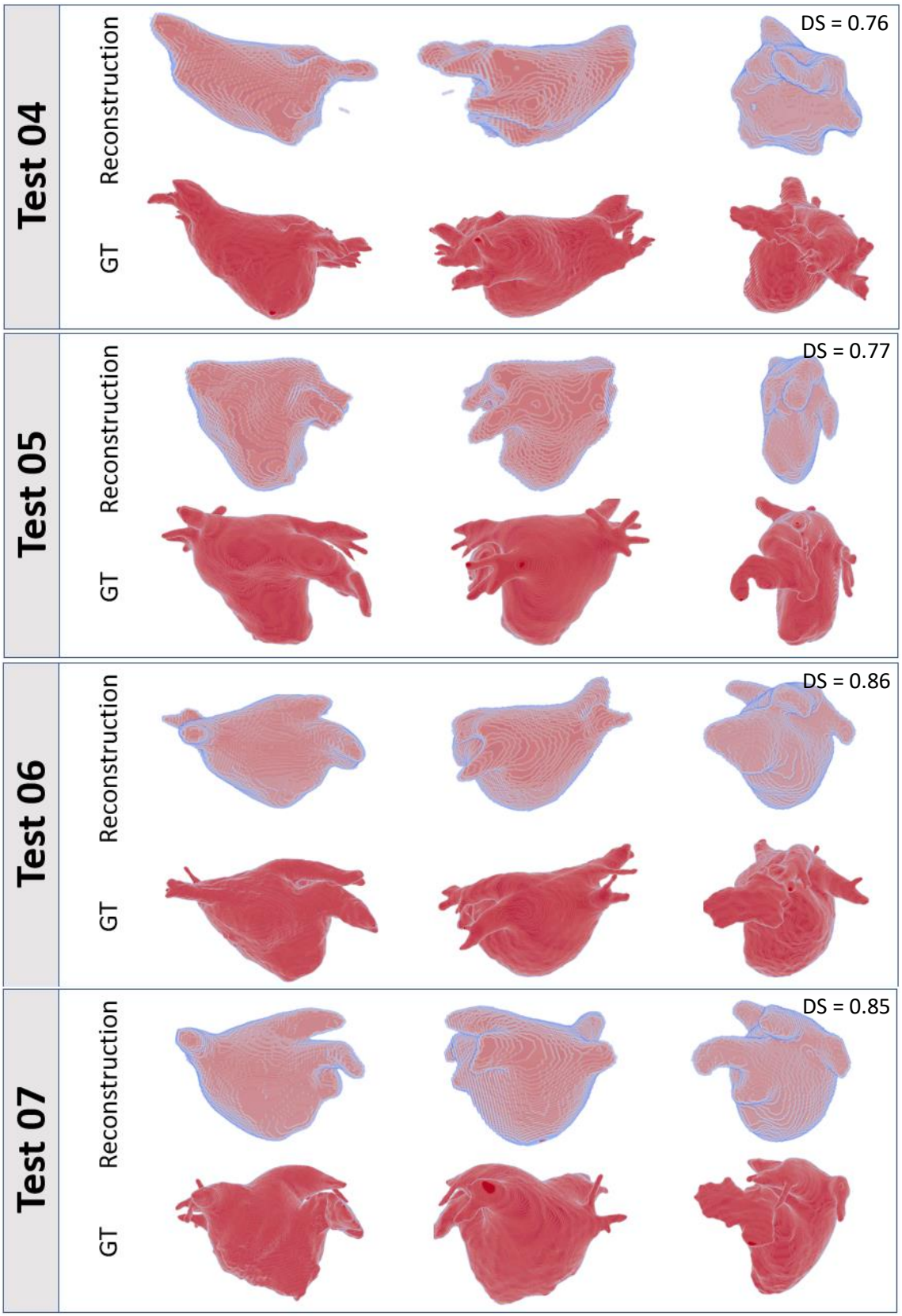

Figure 26. Ground-truth and respective reconstructions for test subjects 4-7.

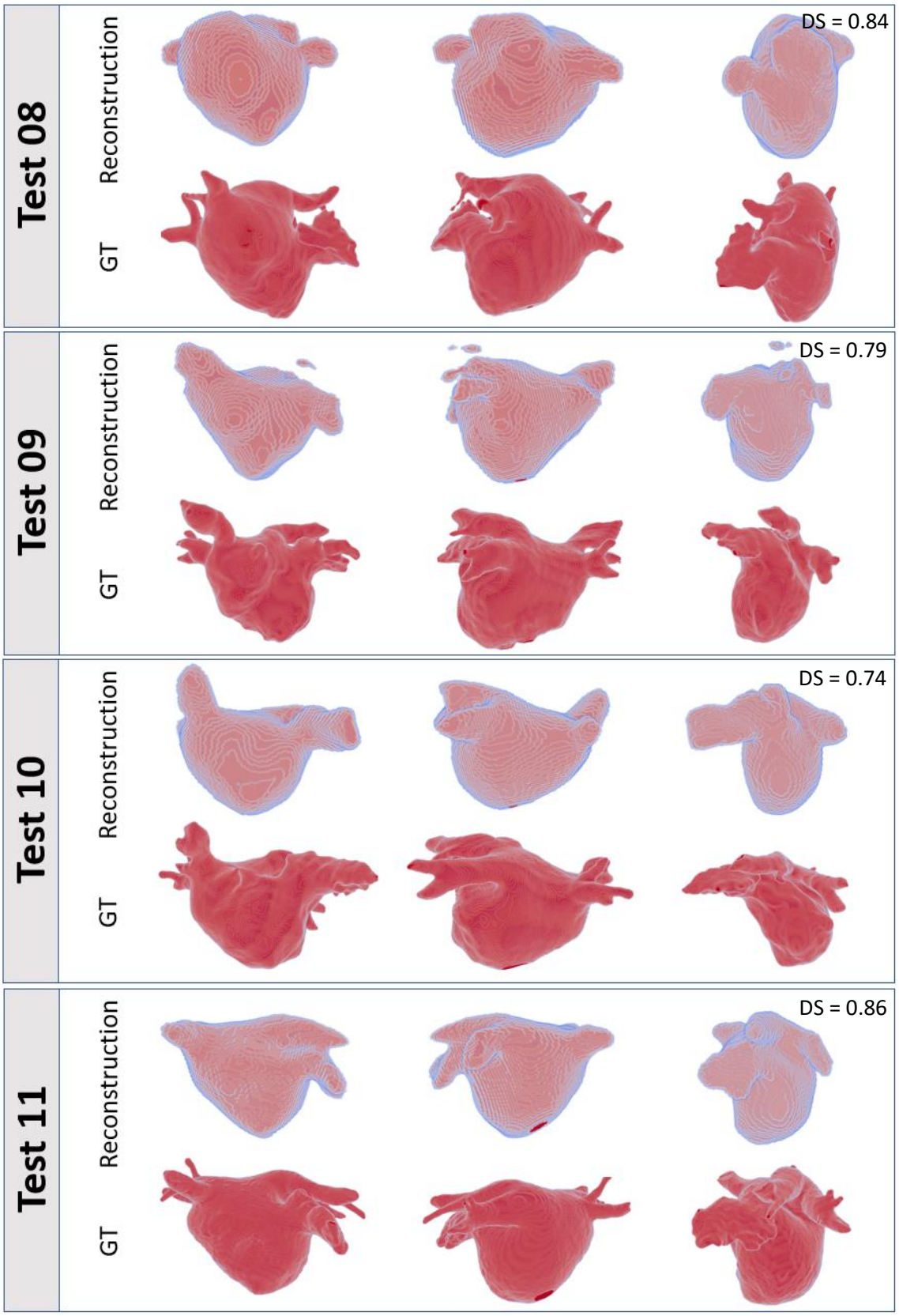

Figure 27. Ground-truth and respective reconstructions for test subjects 8-11.

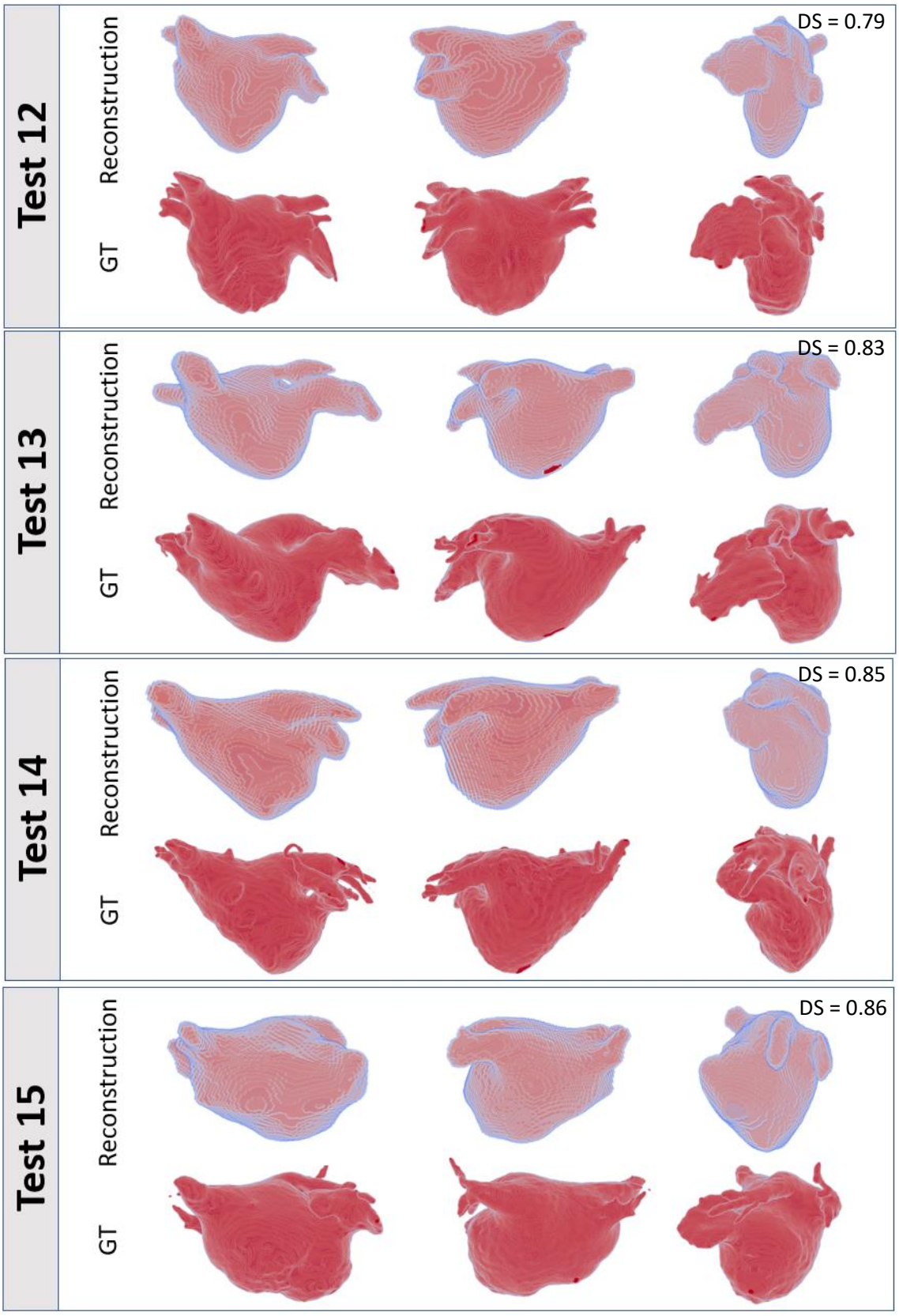

Figure 28. Ground-truth and respective reconstructions for test subjects 12-15.

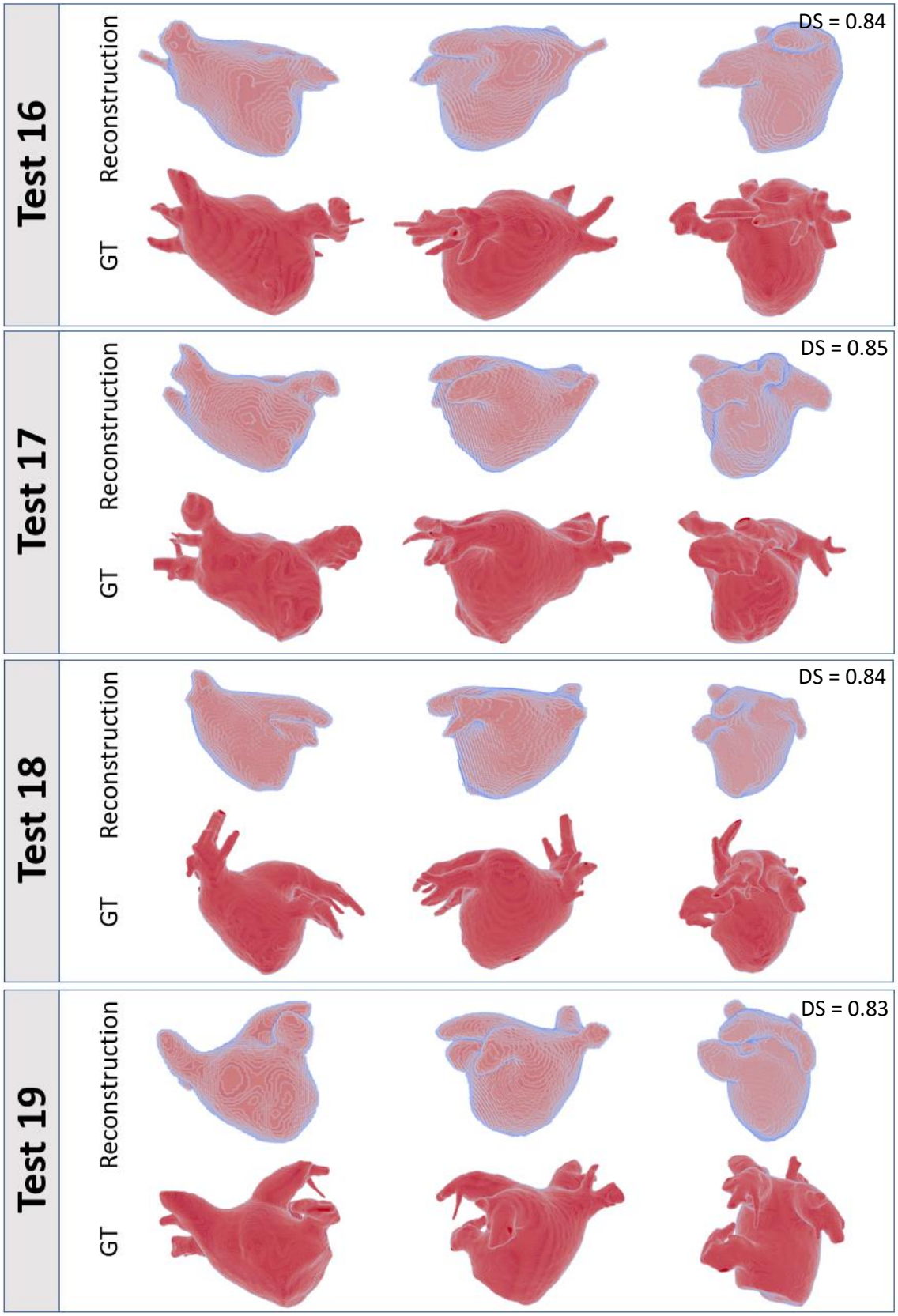

Figure 29. Ground-truth and respective reconstructions for test subjects 16-19.

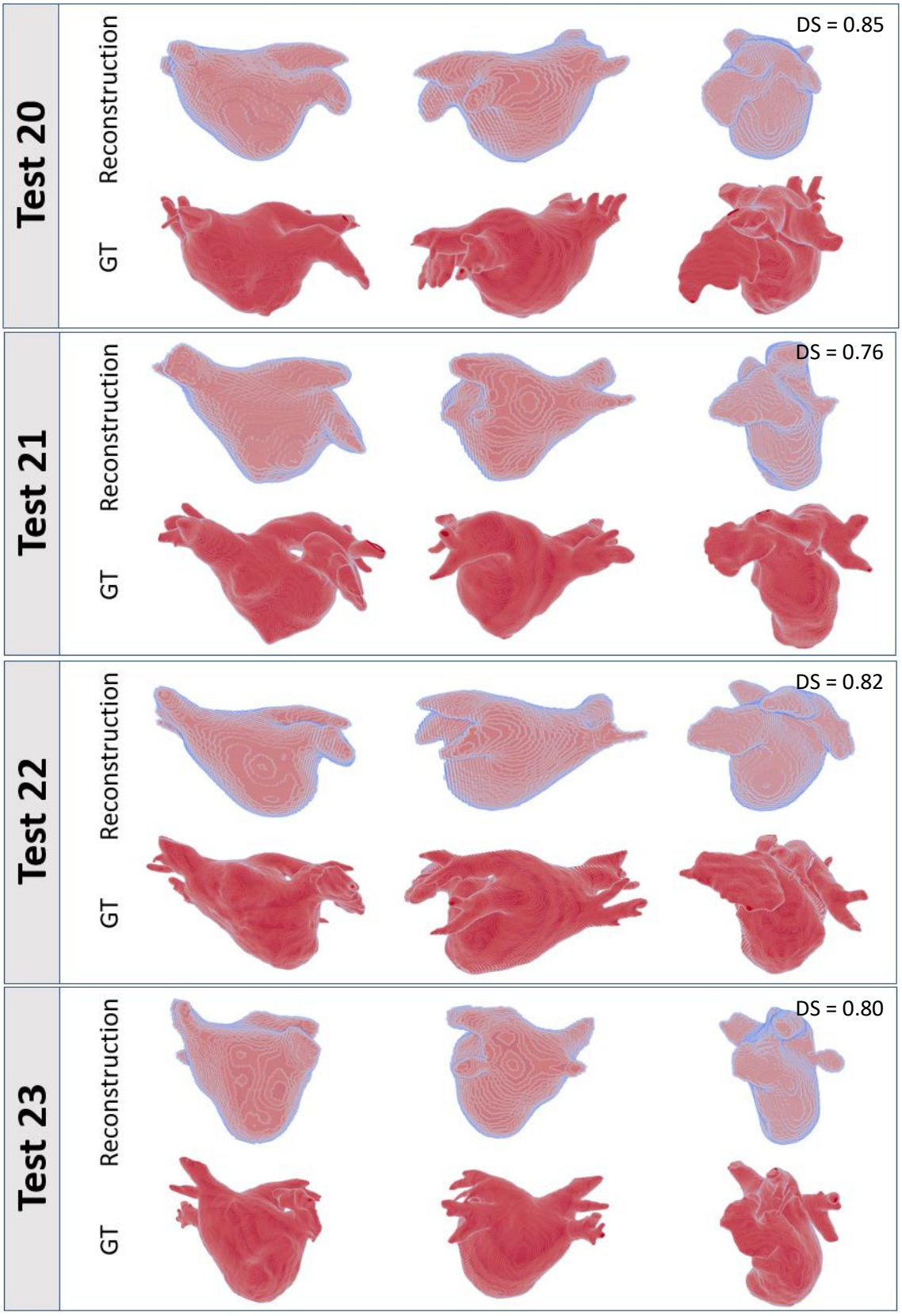

Figure 30. Ground-truth and respective reconstructions for test subjects 20-23.

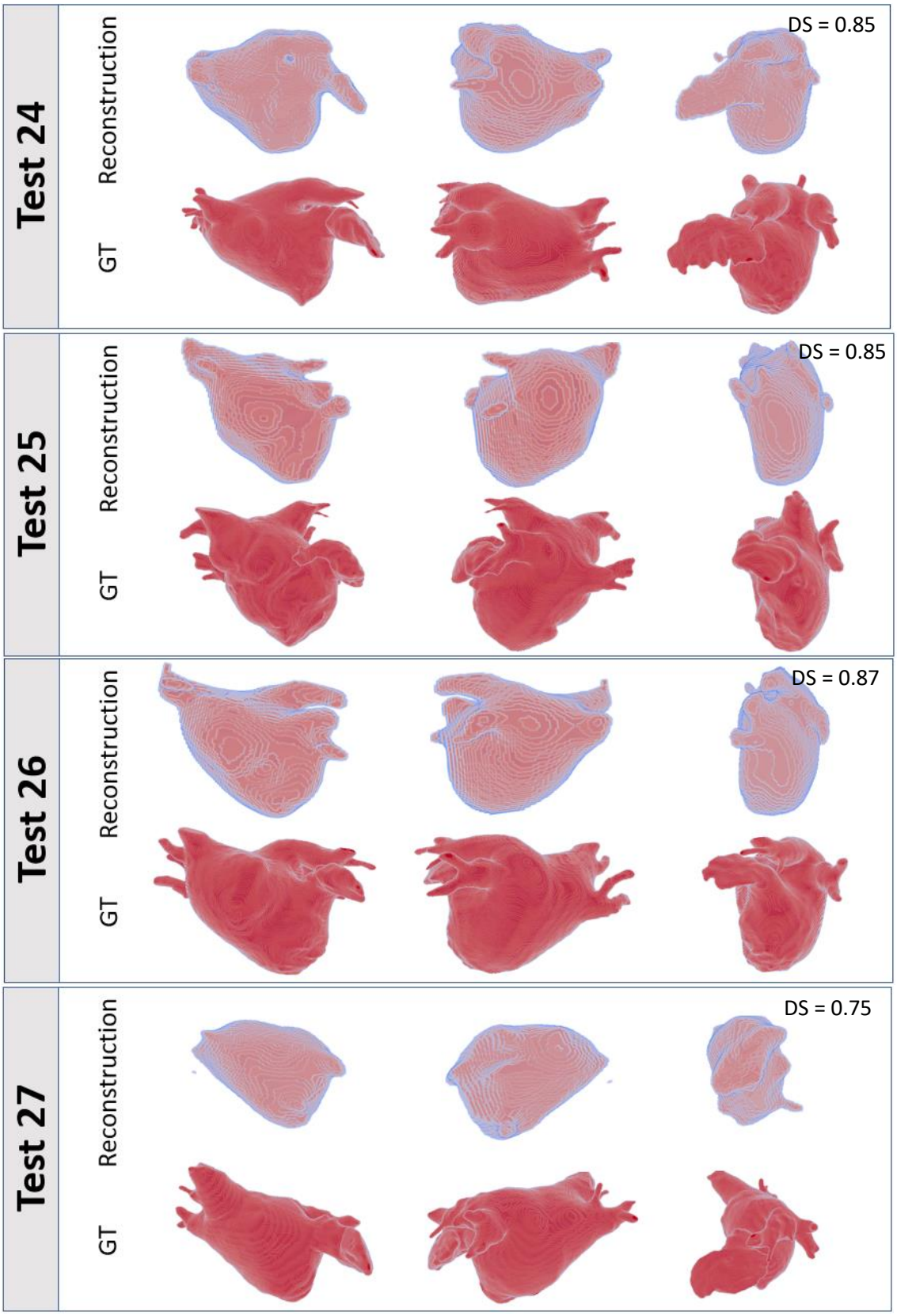

Figure 31. Ground-truth and respective reconstructions for test subjects 24-27.

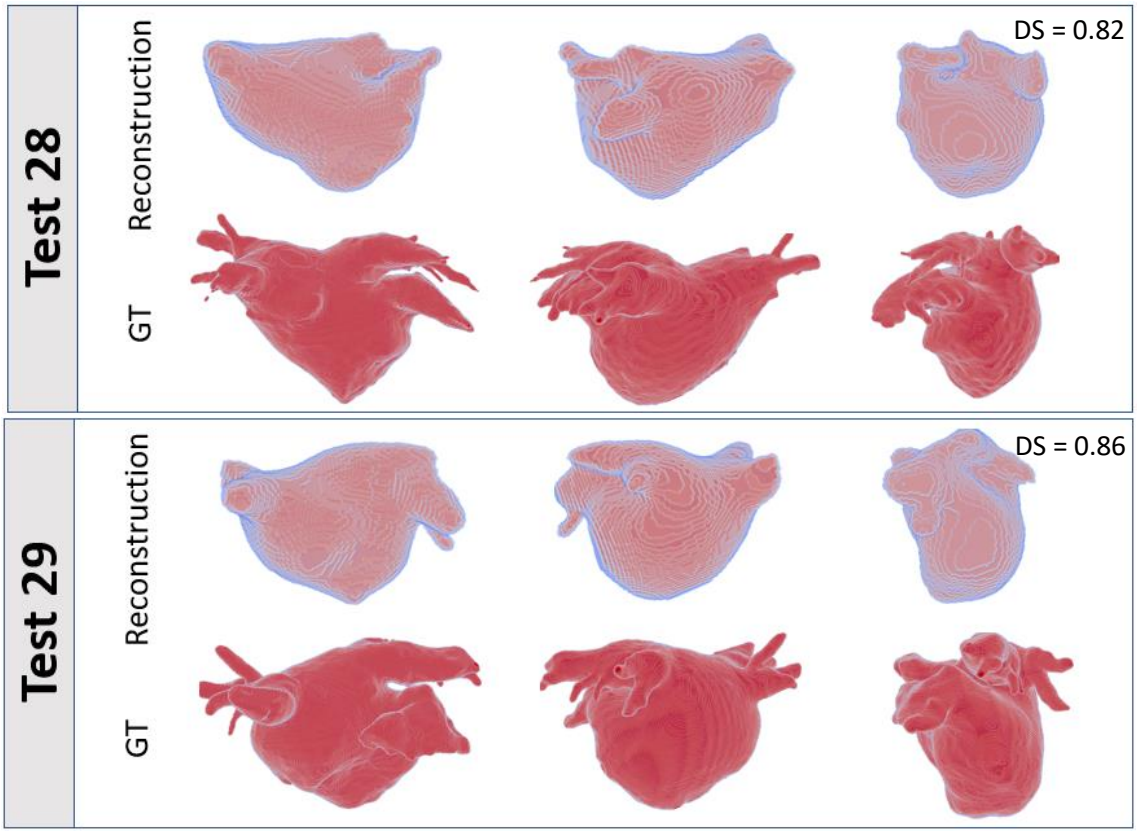

Figure 32. Ground-truth and respective reconstructions for test subjects 28-29.

ANNEX 3: Transesophageal Echocardiography

Transesophageal echocardiography (TEE) uses ultrasound waves to examine the heart and great vessels. TEE can provide valuable information quickly and at low risk, and therefore is a useful tool to improve outcomes and patient survival.

Since its first reported use to evaluate intracardiac flow in 1971 and to visualize cardiac structures in 1976, the TEE US modality has now been universally adopted for their use in patients with congenital heart disease, including applications in neonates and infants [33][34]. With advances in technology, multiplane TEE probes have gradually decreased in size from 16 to 14, to 13, 10, and even 7.5 mm. These devices include phased array technology, with the array on a rotating turntable at the tip of the probe engineered to steer the ultrasound-examining plane through a 180° arc. This technology effectively provides a 360° "panoramic" image of the heart[1].

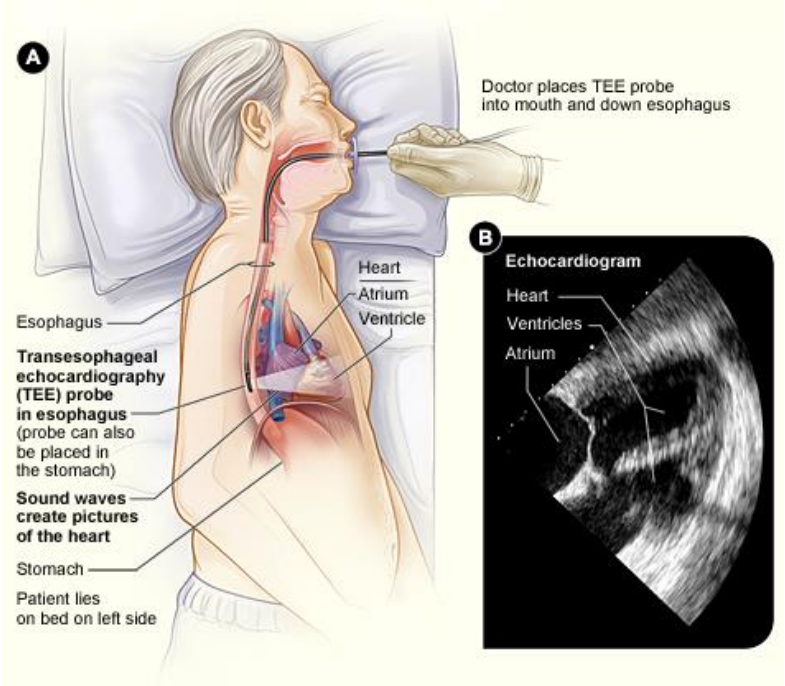

Figure 33. A. TEE procedure. B. Example of an US view from TEE modality.

[1] https://thoracickey.com/instrumentation-for-transesophageal-echocardiography/

TEE Probes

Today, all adult TEE probes as well as the smaller pediatric TEE probes are fully equipped with two-dimensional imaging, M-mode, and complete Doppler capabilities. Some of the currently used TEE probes, with some of their specifications can be found in Table 5.

	Tip dimensions (Width × Height × Length, mm)	Shaft dimensions (Width × Length, mm)	Number of elements	Imaging frequencies (MHz)
Philips (Hewlett Packard/Agilent)				
Pediatric biplane (21366A)	9.3 × 8.8 × 27	8 × 80	64	5.5–7.5
Pediatric mini-multiplane (T6207/S7–3t)	10.7 × 7.2 × 25.4	7.4 × 70	48	4.0–7.0
Pediatric micro-multiplane (S8–3t)	7.5 × 5.5 × 18.5	5.2 × 85	32	3.0–8.0
Adult multiplane (Omni III)	14.5 × 11.2 × 42	10.5 × 100	64	4.0–7.0
Adult 3D TEE (X7–2t)	16 × 12 × 40	10 × 106	2,500+ (Matrix)	2.0–7.0
Siemens (Acuson)				
Pediatric biplane (V7B)	9.5 × 8.5 × 31	8.5 × 85	48	5.0–8.0
Pediatric mini-multiplane (V7M)	10.7 × 7.2 × 25.4	7.4 × 70	48	4.0–8.0
Adult multiplane (V5M)	14.5 × 11.5 × 45	10.5 × 110	64	3.5–7.0

Table 5. Some of the TEE probes available in the market with some of their corresponding specifications. Information obtained from [1].

The pediatric TEE probes provide high frequency, high quality imaging. However, the smaller probe size and greater flexibility can lead to difficulty maintaining constant wall contact in a larger (adult-size) esophagus, and the higher probe frequencies can reduce the effective depth useful for two-dimensional and color flow Doppler imaging.

TEE is essential in the guidance of left atrial appendage closure (LAAC). Current practice often requires general anesthetics to endure a standard size probe during LAAC procedures. On the other hand, smaller probes for pediatric purposes, such as the Philips micro-multiplane (S8-3t), might be useful for the guidance of LAAC in adults, reducing the need of general sedation and also procedure times, and for the assessment of the LA and LAA. [35][36]. It could also be useful to extract the main planes needed in this thesis to obtain a 3D reconstruction of the LA to perform computational simulations.