

Master thesis on Computational Biomedical Engineering
Universitat Pompeu Fabra

Deep Learning Surrogate of Computational Fluid Dynamics for Thrombus Formation Risk in the Left Atrial Appendage

Xabier Morales Ferez

Supervisor: Oscar Camara Rey

July 2019



**Universitat
Pompeu Fabra**
Barcelona

Master thesis on Computational Biomedical Engineering
Universitat Pompeu Fabra

Deep Learning Surrogate of Computational Fluid Dynamics for Thrombus Formation Risk in the Left Atrial Appendage

Xabier Morales Ferez

Supervisor: Oscar Camara Rey

July 2019



Contents

1	Introduction	1
1.1	Atrial Fibrillation	1
1.2	Left Atrial Appendage (LAA)	2
1.3	Atrial fibrillation treatment	3
1.4	Hemodynamics and TIA/CVA risk	5
1.5	Computational fluid dynamics for hemodynamics characterization . . .	6
1.6	Deep neural networks and computational fluid dynamics	8
1.7	Objectives	9
2	Methods	10
2.1	Geometry	11
2.1.1	LAA Statistical Shape Model (SSM)	12
2.1.2	Oval representation of the LA	14
2.2	Preliminary simulations	15
2.2.1	Geometry Assemble	17
2.3	Generation of risk indices for thrombus formation - ECAP	19
2.3.1	CFD simulations and boundary conditions	19
2.3.2	ECAP definition	21
2.3.3	Simulation post-processing	22
2.4	Deep Learning Model	23
2.4.1	Simple fully connected (SFC) network	24
2.4.2	PCA dimensionality reduction network	25

2.4.3	Training of the model	28
2.4.4	Evaluation of the model	28
3	Results	30
3.1	Deep Learning model	30
3.1.1	Performance	30
3.1.2	Detection of potentially high risk thrombotic areas	35
4	Discussion and conclusions	38
4.1	Oval geometry and dynamic mesh	39
4.2	DL Performance	40
4.3	DL network model comparison	42
4.4	High ECAP values	44
4.5	Study limitations	45
4.6	Conclusions	46
	List of Figures	49
	List of Tables	53
	Bibliography	54
A	First Appendix	61
A.1	Prediction figures - Front	62
A.2	Prediction figures - Back	64
A.3	SFC - PCA difference	67
A.4	GT - Prediction difference back-face	68
A.5	FP back-face	69
B	Second Appendix	70
B.1	Total amount of nodes in each ECAP range	71
B.2	Mean Absolute Error of simulations H-U	71

Acknowledgement

I would like to express my most sincere gratitude to everyone involved on the project. In particular, I would like to start thanking Kristine A. Juhl for providing the statistical shape model crucial on the success of this thesis. I would also like to emphasize Guillermo Jimenez-Perez's support in the development of the DL model. Likewise, the assistance given by other members of the COMPILAAO project such as Jordi Mill or Andy L. Olivares has been invaluable. The help provided by the fellow master student, Pablo Martí as well as my friends Antton Lamarka and Ekain Arrieta was also key in overcoming the most challenging moments. Finally, I would like to stress the support given by the thesis supervisor, Oscar Camara, throughout all year.

Abstract

Atrial fibrillation (AF) is the most common clinically significant arrhythmia, often severely disrupting cardiac hemodynamics and drastically increasing the risk of thromboembolic events. Around 90% of such intracardiac thrombus formation in AF patients takes place in the left atrial appendage (LAA). Such thrombus have been related to blood stasis, which at the moment, can be only assessed through noisy imaging data from transesophageal echocardiography (TEE) at one single point in space and time, vastly oversimplifying the characterization of the complex 4D nature of blood flow patterns. Alternatively, attempts have been made to relate LAA morphology to the risk of thrombi formation, some studies suggesting reduced risk of thrombosis on chicken-wing morphologies. Nonetheless, such classification of the LAA morphology has been found to be highly inconsistent and subjective, excluding as well, several fundamental morphological parameters such as the ostium size or the pulmonary vein (PV) orientation among others.

More recently, computational fluid dynamics (CFD) have been employed on the left atrium (LA), seeking to assess the risk of thrombogenesis more quantitatively. CFD has proven to be an invaluable tool in establishing a mechanistic relation between patient-specific organ morphology and its characteristic hemodynamics. In fact, it has long been implemented in other human tissues, such as the coronary arteries, cerebral aneurysms and the aorta with unparalleled success, enabling early diagnosis and risk assessment of various cardiovascular diseases. Nevertheless, traditional CFD methods are renowned for their large memory requirements and long computing times, which severely hinders its suitability for time-sensitive clinical applications.

Hence, this thesis seeks to harness the immense potential of deep learning (DL) by developing a deep neural network (DNN), with the objective of generating a fast and accurate surrogate of CFD, capable of instantaneously evaluating the risk of thrombus formation in the LAA. Already having revolutionized fields such as data processing, it has only recently begun to employ DNNs in high-dimensional, complex dynamical systems such as fluid dynamics. In fact to our knowledge, this study

represents the first successful implementation of a DL surrogate of CFD analysis in a structure as complex as the LAA, which had only been previously attempted in the aorta. For that purpose, two DL architectures have been successfully designed and trained, which receive the specific LAA geometry as an input, and accurately predict its corresponding endothelial cell activation potential (ECAP) map, parameter linked to the risk of thrombosis. The first approach, is based on a simple fully-connected feedforward network, while the latter, also embeds unsupervised learning. An statistical shape model (SSM) of the LAA was created to generate the training dataset, encompassing 210 virtual shapes, on which CFD simulations were performed to attain the ground truth ECAP mappings. Once trained, the final DL networks have accurately predicted the ECAP distributions resulting in an average error of 4.72% for the simple fully-connected network and 5.75% for the unsupervised learning model. Most importantly, the obtention of the ECAP predictions was quasi-instantaneous, orders of magnitude faster than conventional CFD. Therefore, this study is one of the first to demonstrate, the feasibility and unparalleled potential of DL models as accurate and substantially faster surrogates of CFD, potentially enabling future real-time assessment of thrombogenesis risk on the LAA.

Keywords: Deep Learning; Thrombus Formation; Left Atrial Appendage; Computational Fluid Dynamics; Hemodynamics

Chapter 1

Introduction

1.1 Atrial Fibrillation

Atrial fibrillation (AF) is the most common arrhythmia of clinical significance, with an estimated lifetime development risk of around 22% to 26% and affecting approximately 0.4% to 1% of the global population. Aging is the predominant demographic cause of this type of arrhythmia, multiplying AF incidence by a factor of 15 to 20 between the ages of 35 and 85 [1, 2]. As a consequence, the number of patients with AF is likely to increase 2.5-fold by 2050, reflecting the growing proportion of elderly individuals [3]. AF is defined by a quivering or irregular heartbeat (arrhythmia), which prevents effective flow of the blood to the ventricles [4]. This leads to increased morbidity of heart failure, blood clots, ischemic stroke and other hemodynamic disturbances [5, 6]. AF may be asymptomatic and as a result, is often only diagnosed when complications such as stroke arise. In addition, these strokes tend to be more severe and deadly than those originating from other aetiologies [7, 8, 9].

AF is often classified in two subgroups: valvular and non-valvular. The first encompasses only those patients suffering from mitral valve (MV) stenosis or that have undergone surgery for artificial heart valve implantation or MV reparation. As for the latter, presence of non-valvular AF increases the risk of cerebrovascular accidents, such as transient ischemic attack and ischemic stroke, by allowing thrombus

formation within the left atrium (LA); if dislodged, such thrombi can occlude the cerebral circulation, causing a thromboembolic (ischemic) stroke [10]. As a matter of fact, AF is responsible for 15 to 20% of all cardioembolic ischemic strokes, 90% of which are caused by thrombus formed in the left atrial appendage (LAA), a finger-like cavity located in the LA [11, 12].

1.2 Left Atrial Appendage (LAA)

The LAA, also known as auricle, is a long, tubular, hooked structure derived from the anterior wall of the LA. It is a remnant of the original fetal LA, mainly formed by the adsorption of the primordial pulmonary veins and their branches. Its junction to the LA is well defined by a narrowing of the appendage's orifice [2].

For a long period of time the LAA has been considered as a mere atrial outgrowth, but its direct relationship with stroke and the increase in LAAO interventions has renewed the interest in the possible influence of LAA over LA hemodynamics. No longer considered a static chamber, LAA is thought to be a highly dynamic structure, working both as a contractile reservoir and decompression chamber, by producing suction during ventricular systole and as a conduit during diastole [13]. Besides its role in LA hemodynamics, LAA seems to have an endocrine function by releasing atrial and brain natriuretic peptides (BNP). Patients suffering from AF, especially those suffering from non-valvular AF, produce high levels of BNP in the LAA, which has been directly related with a pro-thrombotic state [14].

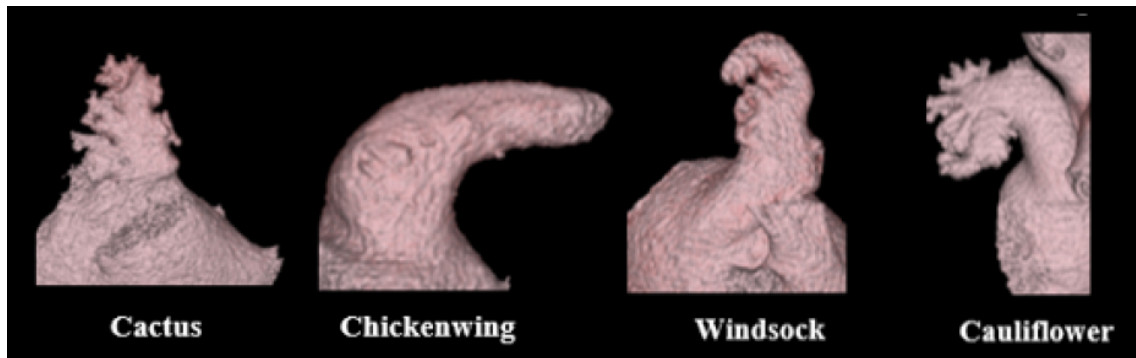


Figure 1: Main classification of the LAA morphology (from Di Biase *et al.*[12])

Moreover, as aforementioned, the structure of the LAA is highly heterogeneous presenting a tremendous inter-patient variability in its size and shape. The wide range in which the different features used to characterize the LAA vary, represent a prime example of the high heterogeneity of the LAA structure: volume can range from 0.7 ml to 19.2 ml; ostium minimum diameter can go from 5 to 27 mm and its maximum diameter ranges from 10 to 40 mm. Finally, there can be also great variation in the LAA orifice length, from 16 up to 51 mm [15]. Furthermore, LAA typically presents several lobes that protrude from their main body. Based on the heterogeneity of the shape, number of lobes, trabeculations and overall dimensions of the LAA, it is commonly classified into four distinct morphological categories as shown in Figure 1.

Several attempts have been made to establish a relationship between LAA morphology and the risk of thromboembolic events. In fact, latest research suggests that patients with chicken wing LAA morphology are less likely to develop thromboembolic events than those with non-chicken wing morphology [12, 16]. Nevertheless, inconsistent LAA morphology classification coupled with low interobserver and intraobserver agreement, suggests caution when directly linking LAA morphology to the risk of stroke. Indeed, in a recent study by Wu *et al.* [17], CT images of 2264 AF patients were handed to 3 experts for independent judgement of their morphology. Of all the cases studied, only in 655 was consensus achieved between the 3 reviewers, which means that less than 30% of the patients were classified to be part of the same morphology group. This highlights the need for a clinically feasible practice, to directly evaluate and link the patient-specific mechanic structure of the LA and its appendage with the risk of intracardiac thrombosis.

1.3 Atrial fibrillation treatment

AF patients, often suffer from chaotic electric activity, which leads to irregular contraction and wall rigidity of the LA, altering the hemodynamics and decelerating the blood. Together with blood stagnation in the LAA, AF patients often present endothelial or endocardial dysfunction, fibrinolysis, abnormal hemostasis and platelet

count, effectively fulfilling the Virchow triad criteria and explaining the prothrombotic state caused by AF [18].

To avoid such deleterious outcome, specially among the elderly, the most common approach consists in the prescription of oral anticoagulation therapy. Nevertheless, attaining perfect equilibrium between thrombus prevention and avoidance of hemorrhagic events is complex, requiring delicate personalized dose monitoring.



Figure 2: LAA occlusion devices (a) Watchman device (b) Amplatzer device [19]

Left Atrial Appendage Occlusion (LAAO) devices have risen as response for those on high hemorrhagic risk. The task of isolating the LAA is extremely complex due to its highly heterogenic structure morphology [20]. These innovative devices are implanted percutaneously through the femoral vein, under fluoroscopic and echocardiographic guidance to avoid the blood stasis in the LAA. Despite the relative novelty of these therapies, models such as the Watchman Device (Boston Scientific Corp., Natick, Massachusetts) or the Amplatzer Cardiac Plug (ACP) (St. Jude Medical, Inc., St. Paul, Minnesota) shown on Figure 2, have proven to successfully occlude the LAA in several randomized trials and multiple published cohorts, leading to reduced thrombus formation and consequently avoiding stroke events [21, 19].

Despite low invasiveness and relative success of LAAO device implantations, there have been documented cases of air embolism, LAA tear/rupture and stroke related

to the trans-septal puncture and implantation of the device. One of the most common causes of such undesired outcomes, is the improper sizing and placement of the device, which unfortunately leads to a substantial percentage of thrombi formation post-LAAO implantation [22, 23, 24]. In order to avoid such post-surgical complications, the device is properly sized using a chart after meticulously measuring the LAA structure through contrast X-Ray Computer Tomography (CT) or 3D Rotational Angiography (3DRA). Nonetheless, this pre-operative procedure does not consider the way in which the blood flow patterns are affected post-LAAO implantation, which is crucial to ensure that the risk of thrombus formation no longer exists. Thus, an innovative procedure allowing instantaneous analysis of the behavior of the local hemodynamic to changes in the position and size of the LAAO devices would be of great help in avoiding these setbacks.

1.4 Hemodynamics and TIA/CVA risk

Thrombi formation within the LAA is associated to a combination of several hemodynamic conditions. A number of studies have pointed out that blood stasis is good risk predictor of blood clot formation. As a matter of fact, blood flow velocities below 40 cm/s, have been associated to increased stroke risk and those below 20 cm/s, with the formation of clots in the LAA [2]. This emphasizes the great potential of LA hemodynamics analysis as a useful tool to determine the risk of thrombi formation.

With this in mind, there have been several attempts at characterizing LAA hemodynamics precisely, though at the moment, it can be only assessed through noisy imaging data from transesophageal echocardiography (TEE). Unfortunately, TEE can only provide blood velocity at one single point in space and time, vastly oversimplifying the characterization of the complex 4D nature of blood flow patterns. In this regard, 4D-flow magnetic resonance imaging is said to revolutionize the field in the near future, as it is able to capture the LA blood flow on the 3 dimensions of space through time, allowing for comprehensive evaluation of complex blood flow patterns. In the meantime, given that 4D MRI is still under development [25], Computational

fluid dynamics (CFD) arises as a promising alternative to fully characterize the complex hemodynamics and blood flow patterns in the different patient-specific LA and LAA morphologies.

1.5 Computational fluid dynamics for hemodynamics characterization

Computational fluid dynamics (CFD) is a branch of fluid mechanics that deals with numerically solving and analyzing fluid flow. For years, it has been a vital tool for computer simulation in such disparate fields such as aerodynamics, aerospace industry, combustion analysis and weather simulation. More recently, it has also been successfully employed and become widespread in the field of biomedical engineering. It has been of special interest in the study of the cardiovascular system being employed in the design of blood pumps as early as 1990 [26].

Advances in imaging techniques have made it possible to study the physiological mechanisms of human tissues and organs, allowing the development of patient-specific treatment strategies. As aforementioned, there have been several attempts to employ CFD in blood flow analysis of human LA and LAA. Otani *et al.* performed personalized blood flow analysis with 3D LA endocardial surface motion estimated from patient-specific cardiac CT images [27]. Garcia-Isla *et al.* performed a sensitivity analysis to identify the most relevant LA and LAA morphological parameters in atrial hemodynamics, pointing out the relevance of pulmonary vein configuration on LAA blood flow patterns [28]. In addition, Bosi *et al.* [29] attempted patient stratification under thrombogenesis risk, based on the conventional 4 group LAA classification previously shown in Figure 1, while Masci *et al.* sought to link blood stasis to a set of relevant geometrical descriptors and morphological features [30]. Finally, the work by Aguado *et al.* [31], focused on clinical translation by developing a web-based 3D interactive virtual implantation platform designated as VIDAA. The aim is to help clinicians in selecting the most appropriate LAAO configurations for the given patient-specific LAA morphology and export them for posterior

computation of CFD simulations.

Hence, CFD rises as an invaluable tool to establish a mechanistic relationship between the patient-specific LAA morphology and intracardiac thrombosis by characterizing the LA hemodynamics. Nevertheless, despite its huge potential and high-accuracy results, conventional CFD methods are known for their high computational cost, vast memory requirements and long computing times, often taking several hours or even days for the simulations to converge. In many studies, CFD analysis becomes one of the most intensive and time-consuming processes in the workflow, specially if numerical convergence issues emerge [32, 33]. Moreover, the generation of the mesh can also turn out to be a very time consuming process and can lead to multiple problems if not completed properly.

This represents a major drawback of CFD models and prevents its embedding in the clinical workflow of time-sensitive applications, which require fast diagnosis and treatment planning [34, 35]. Effectively skipping the computation of CFD would provide clinicians with really powerful tools to enable widespread implementation of patient-specific hemodynamic characterization in the clinical environment and instantaneously assess the risk of thrombosis. This would suit perfectly the aforementioned application consisting in the pre-operative evaluation of the optimal positioning and sizing of LAAO devices in the LAA, aiming to minimize future risk of thrombus formation. At the moment, it is not feasible to perform a high-accuracy CFD analysis to prompt quasi-instantaneous feedback about the LAAO device positioning to clinicians, due to obvious time constraints on daily clinical practice. This would be an invaluable tool on optimizing the positioning of devices such as the Amplatzer or the Watchman which could potentially avoid several post-surgical complications.

For all these reasons, there is an ever-growing body of work to create reduced order or surrogate models of fluid flow simulations that can be evaluated with significantly less computational resources [33]. Towards this end, deep neural networks (DNN) provide an innovative approach that both compresses the computation time and reduces the memory usage of CFD simulations.

1.6 Deep neural networks and computational fluid dynamics

It was only a matter of time before deep learning (DL) made its way into fluid flow modelling. DNNs have been widely applied on several fields such as image segmentation with outstanding success, several studies claiming to outperform human clinicians, not without controversy [36, 37]. Nonetheless, it has not been until recently that DL has made its way into the field of high-dimensional, complex dynamical systems. First applied by Ling *et al.* in 2016, modern DNNs have achieved performance improvement over competing state of the art methods such as finite element analysis, which suggests DNNs may play a critically enabling role in the future of modelling complex flows [38].

It hasn't been long before DNNs have been also implemented in the study of tissue biomechanics. Liang *et al.* [35] proposed a DNN based CFD surrogate that directly estimated the aortic wall shear stress (WSS) distributions given the patient-specific geometry of the aorta, which to the best of our knowledge represents the first time this kind of approach has been successfully attempted in biological fluid flow [35]. The DL model was designed and trained to take as an input the finite element mesh of the aorta and directly output the aortic WSS distribution. The novel approach managed to reproduce the CFD results with an average discrepancy of 0.492% and simultaneously reducing computation time from hours to mere seconds. This unbelievable time performance improvement over current methods without the need to sacrifice accuracy is the main reason making DNNs so compelling. In this regard, DNNs could pave the way for development of instantaneous fluid dynamic analysis that could be embedded in time-sensitive clinical applications.

1.7 Objectives

Despite the immense escalation of computational power and the increased availability of commercial software in the recent decades, high fidelity CFD model solving, still requires an impressive amount of time and computational resources that severely hinders its embedding in time-sensitive clinical applications. Hence, this master thesis seeks to develop a deep learning surrogate of CFD, with the goal of directly retrieving the endothelial cell activation potential (ECAP) in the LAA, which has been linked to potential risk of thrombogenesis, solely from patient-specific LA and LAA morphologies, effectively skipping the numerical computation of the CFD model. With this in mind, two distinct DNNs have been developed, one being based upon a simple fully-connected (SFC) feedforward network, while the other performs dimensionality reduction of the input and output data by means of principal component analysis (PCA), as a previous step to the non-linear regression on the multilayer neural network. In the future, the aim is to embed the created CFD surrogate on the aforementioned VIDAA web platform, as an additional tool to speed up the hemodynamic characterization of the LA by orders of magnitude.

For that purpose, the hosting research group, BCN MedTech, has collaborated with several hospitals and research groups for the obtention of an artificial population of LA and LAA 3D meshes. More specifically, researchers from the Technical University of Denmark (DTU) have collaborated by providing the tools to generate sets of virtual LAA representations [39].

Chapter 2

Methods

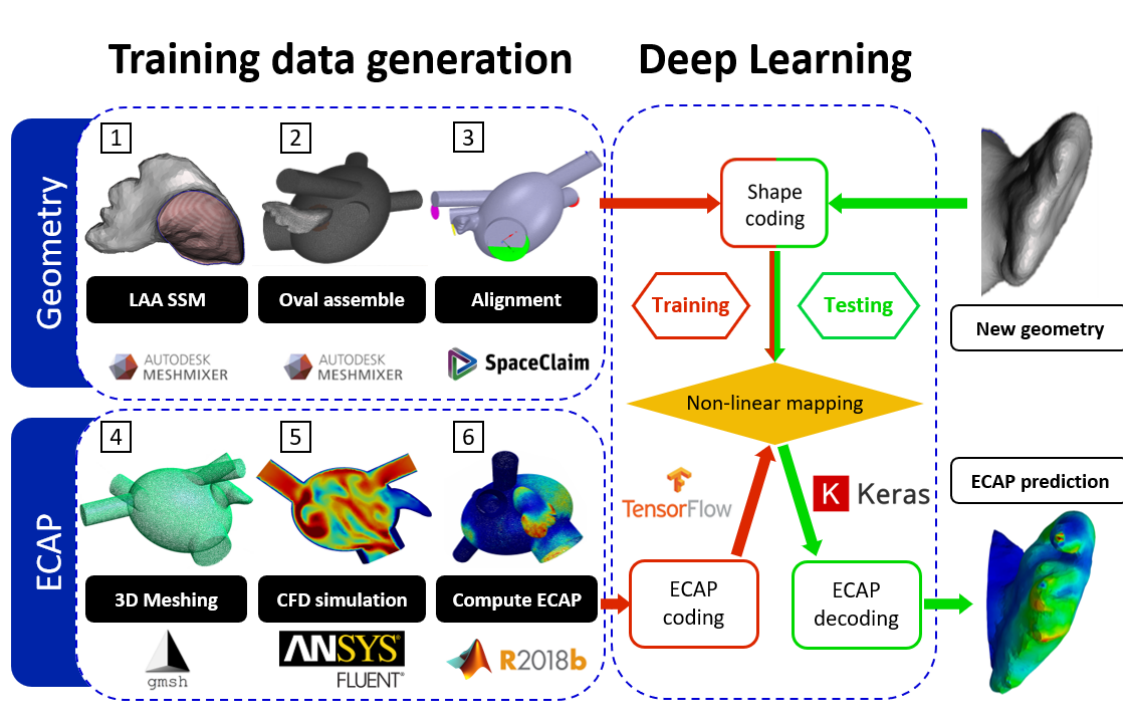


Figure 3: General pipeline of the project. Following the enumeration, the geometries are generated from the SSM [1], assembled [2] and aligned [3]. Afterwards, the 3D volumetric mesh is generated [4] and CFD simulations are carried out [5] before calculating the ECAP maps [6]. Once the training data has been generated, the deep neural networks are trained (in red). As the final step the accuracy of the networks is tested with a set of new geometries (in green). LAA: Left Atrial Appendage, SSM: Statistical Shape Model, CFD: Computational Fluid Dynamics, ECAP: Endothelial Cell Activation Potential.

This section seeks to delve into the procedures applied throughout the development of this thesis. As shown in Figure 3, the workflow is divided into two major blocks. First, a breakdown of the procedures or the generation of the training data is carried out. Likewise, the deep neural networks employed for the machine learning analysis are described.

The first objective of the project consisted on ensuring that the same connectivity and number of elements was shared across all geometries. That meant establishing correspondence between all the different morphologies. Several assumptions had to be made in order to accelerate the working pipeline and keep the model simple. For the LA, the geometry was to be approximated as an ovoid geometry. As for the LAA, an statistical shape model (SSM) was constructed, allowing the generation of an unlimited population of varying geometries sharing the same connectivity and number of nodes. Secondly, the development of the DL architecture took place. The network needed to be trained both, with the LAA geometries and the ECAP maps attained through CFD analysis. Two approaches were proposed: The first entailing the dimensionality reduction of the input tensors through truncated principal component analysis (PCA). Then a multilayer DNN would relate the low dimensional representation of the shape and the endothelial cell activation potential (ECAP) which would then be mapped back to the LAA geometry. For the latter, a simple fully-connected feedforward network was employed to perform a direct non-linear mapping. Finally, the effectiveness of the DL surrogate model was be validated against the ECAP distributions predicted by the recently trained DNNs through a Monte Carlo cross-validation.

2.1 Geometry

Due to the complexity and high variability of the anatomy of both the LA and the LAA, a decision was taken to represent the LA cavity by an ovoid structure based on the work by García-Isla *et. al.* [28]. The idea behind this choice was to disregard the effect of LA heterogeneity on the atrial hemodynamics and focus on the effects caused by the complex morphology of the LAA. In addition, it was decided that

the input to the neural network would consist on the spatial (x, y, z) coordinates of each of the nodes that make up the mesh of the LAA geometries. Therefore, correspondence between LAA meshes was set as a requirement, that is to say, each of the cases had to have the same number of nodes and maintain the same inter-nodal connectivity. To fulfill such demands an statistical shape model (SSM) of the LAA was created, allowing the creation of a big enough training dataset while establishing node correspondence between each geometry. As a result, the final set of virtual geometries, represented by a unstructured triangular 3D mesh, all shared the same oval LA structure and cavity. Furthermore, inter-nodal connectivity and correspondence was imposed on the LAA morphologies that vary according to an SSM, so that changes in hemodynamics depend solely on the LAA geometry.

We have made use of several software to carry out the task. The assemble of the LAA meshes was performed on Meshmixer [40] and MeshLab [41] open-source software, while the alignment process was carried out on Space Claim 19.2 by ANSYS (Student License Version) [42]. Subsequently, the generation of the tetrahedral volumetric mesh was completed on gmsh 4.3.0 [43]. The CFD simulations were all automatized with the AAS toolbox on MATLAB R2018b (Academic License) [44] that allowed to complete all the simulations on ANSYS Fluent 19.2 [45] without the need for supervision. During the simulations, pressure values, the velocity vector, the vorticity magnitude, the wall shear stress and the shear strain were stored at each time step. The ECAP parameter was calculated using MATLAB scripts, later visualizing the results on ParaView 5.6.0-RC3 (open-source) [46].

2.1.1 LAA Statistical Shape Model (SSM)

The complex morphology of the LAA coupled with the vast amount of samples required by DNNs to be effectively trained, made fulfilling the aforementioned requirements a non-trivial task. Consequently, generating a dataset of artificial LAA geometries was regarded as the most sensible approach. For that purpose, researchers from the Technical University of Denmark (DTU) provided us with datasets obtained from a Statistical Shape Model (SSM) of the LAA developed by Slipsager *et*

al. [39]. In the study, non-rigid volumetric registration of signed distance fields was applied to 103 patient LAA surfaces, ensuring that all geometries shared the same inter-nodal connectivity and number of nodes. The LAA surfaces were reconstructed from CT images, provided by the Department of Radiology, Rigshospitalet, University of Copenhagen. Afterwards, the SSM was generated based on PCA, which represented the shape probability distribution by a mean shape and modes of shape variation through a vector of spatial point coordinates. As a result, the generated SSM allowed to create as many artificial geometries as desired by varying the most meaningful eigenvalues as shown in the equation,

$$x = \bar{x} + Pb \quad (2.1)$$

where \bar{x} is the mean shape of the dataset, $P = (p_1, \dots, p_n)$ is the set of n retained eigenvectors and $B = (b_1, \dots, b_n)$ is the set of n weights for each eigenvector defined as follows:

$$-3\sqrt{\lambda_k} \leq b_k \leq 3\sqrt{\lambda_k} \quad (2.2)$$

The set of weights for the selected eigenvectors was chosen following a normal distribution $N(0, 1)$. In order to preserve the variance of the SSM, we attempted to keep as many eigenvectors as possible. Nevertheless, as the number of eigenvalues included in the model increased, some kind of “noise” appeared in the mesh surface. This leads to defects such as self-intersecting faces that might cause the volumetric meshing to crash and the CFD simulations to diverge. For that matter, as shown on Figure 4, datasets with a number of principal components (PC) equal to $PC = 15$, retained 92% of the variance but the defects observed in its mesh rendered the dataset useless causing continuous convergence problems. On the other hand, the dataset with $n=5$ showed a nice, smooth surface mesh but only retained 70% of the variability. Consequently, $PC=10$ was chosen as an acceptable trade-off between preserving variability (87% of the variance) and surface mesh quality.

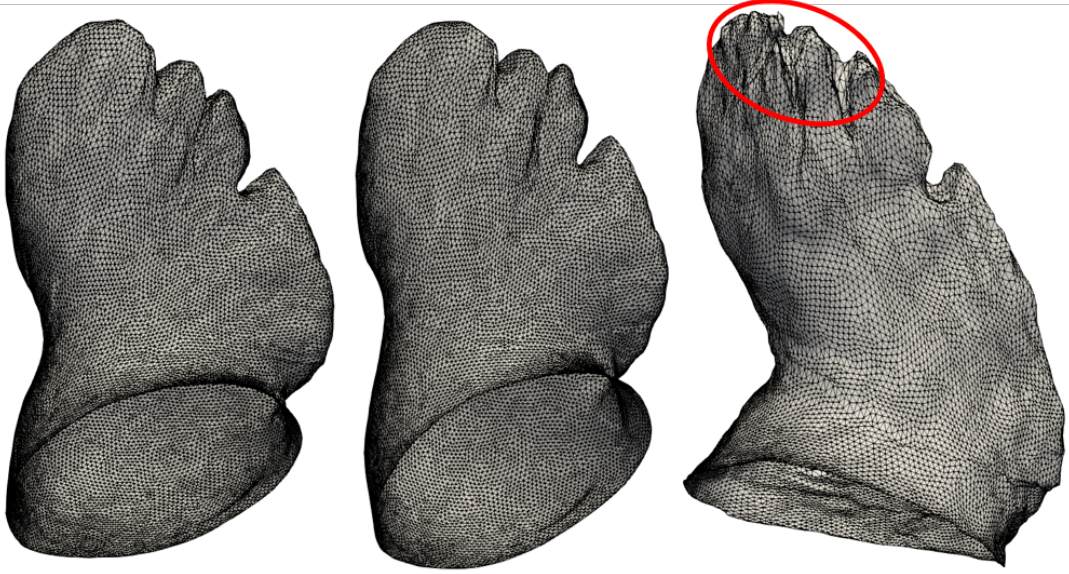


Figure 4: SSM generated with varying number of retained principal components (PC) A) $PC = 5$, B) $PC = 10$, C) $PC = 15$. In red a set of self-intersecting faces responsible for the divergence of the CFD simulations

Additionally, the number of nodes in the artificial LAA meshes was also studied. Considering the dependence of the fineness of the 3D volumetric mesh from gmsh on the surface grid, the number of nodes was minimized to avoid long computational times. Therefore, a geometry with 5000 elements was chosen as the template to which all the generated geometries from the SSM were registered.

With all of this in mind, a final dataset of 300 virtual LAA geometries was generated, preserving 10 PCs with a total number of 5000 faces and 2536 nodes.

2.1.2 Oval representation of the LA

The LA, similarly to its appendage, also shows an extraordinary interpatient variability, such as a varying number of pulmonary veins with several possible orientations and diverse cavity volumes. Consequently, aiming to reduce the complexity of the model and accelerating the workflow, the LA was approximated to an ovoid shape based on the work by García-Isla *et al.* [28]. As reported on Figure 5 the oval LAA consists of 4 cylindrical PVs and a cylindrical MV. Doing so, the effect of LA morphology on atrial hemodynamics is disregarded.

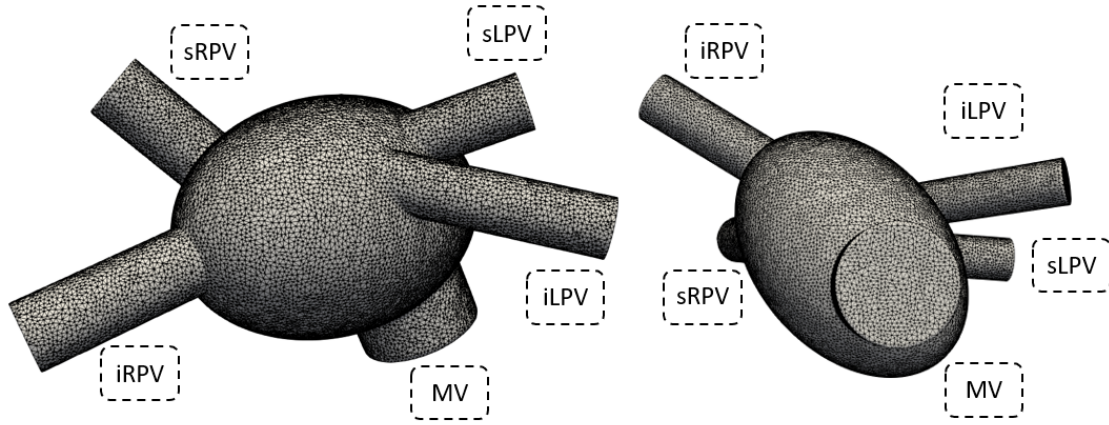


Figure 5: Oval representation of the LA cavity. sRPV: Superior right pulmonary vein, iRPV: Inferior right pulmonary vein, sLPV: Superior left pulmonary vein, iLPV: Inferior left pulmonary vein, MV: Mitral valve.

It must be emphasized, that contrary to the LAA, no correspondence is imposed in the LA and therefore, the total number of nodes and the connectivity varies from one case to another. Only the nodes belonging to the LAA on which correspondence has actually been imposed comprise the input to the DNN. The values for the volume and area of the LA chamber, veins and mitral valve shown on Table 1, were set by computing the average of 11 real study cases [28]. After several numerical convergence problems the fineness of the oval mesh was reduced to match that of the SSM generated LAAs. This resulted in an oval geometry with a surface mesh mesh composed of approximately 20000 elements.

Table 1: Ovoid LAA dimensions, VLA: Volume Left Atrium, VLAA: Volume Left Atrial Appendage, ALPV: Area Left Pulmonary Veins, ARPV: Area Right Pulmonary Veins, AMV: Area Mitral Valve

VLA [mm ³]	VLAA [mm ³]	ALPV [mm ²]	ARPV [mm ²]	AMV [mm ²]
170.63	11.59	165.59	224.76	855.83

2.2 Preliminary simulations

Several preliminary CFD simulations were carried out to ensure that the oval approximation managed to generate sufficiently “realistical” ECAP values compared to

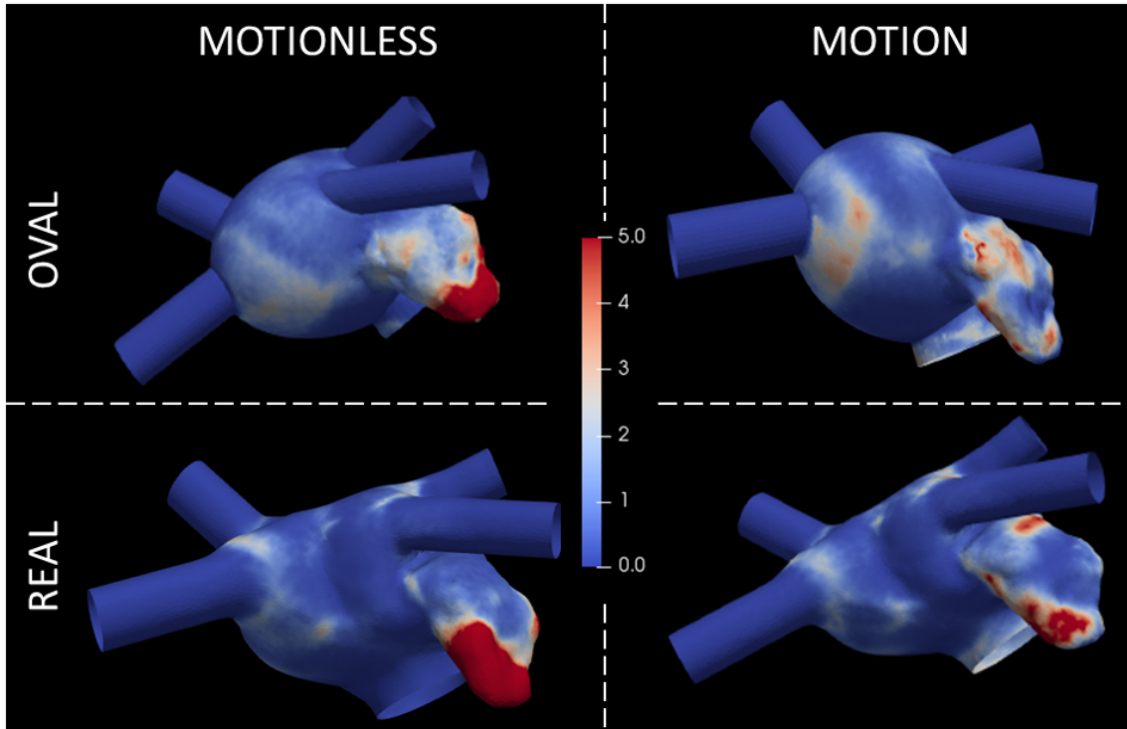


Figure 6: ECAP maps for both real patient and SSM generated LA surfaces. On the left the LA cavity is supposed to be motionless, while on the right a dynamic mesh is implemented on the MV annulus.

those obtained from CFD simulations on real patient LA geometries. Furthermore, a dynamic mesh method [47] was implemented on the MV to simulate the physiological motion of the LA cavity during the heart cycle and avoiding mass imbalance [48]. More details about its implementation can be found on Section 2.3.1.

Figure 6 showcases the ECAP maps obtained from real patient geometries and its oval counterparts. Both the realistic and artificial mesh of the LA, shared the volume and area shown on Table 1. The simulations with motion, implemented the dynamic mesh, while the motionless examples did not.

Both motionless simulations attained significantly higher maximum ECAP values than its counterparts, reaching peak values of 12 on the oval geometry and 21 on the patient geometry. In contrast, the dynamic simulation resulted on a maximum of 6.7 on the ovoid while the realistic geometry had a ECAP score of 8.3. The 90th quartile for both the motionless simulations were 15 and 9 respectively, well above its counterparts, which just reached values of 4 and 6.

2.2.1 Geometry Assemble

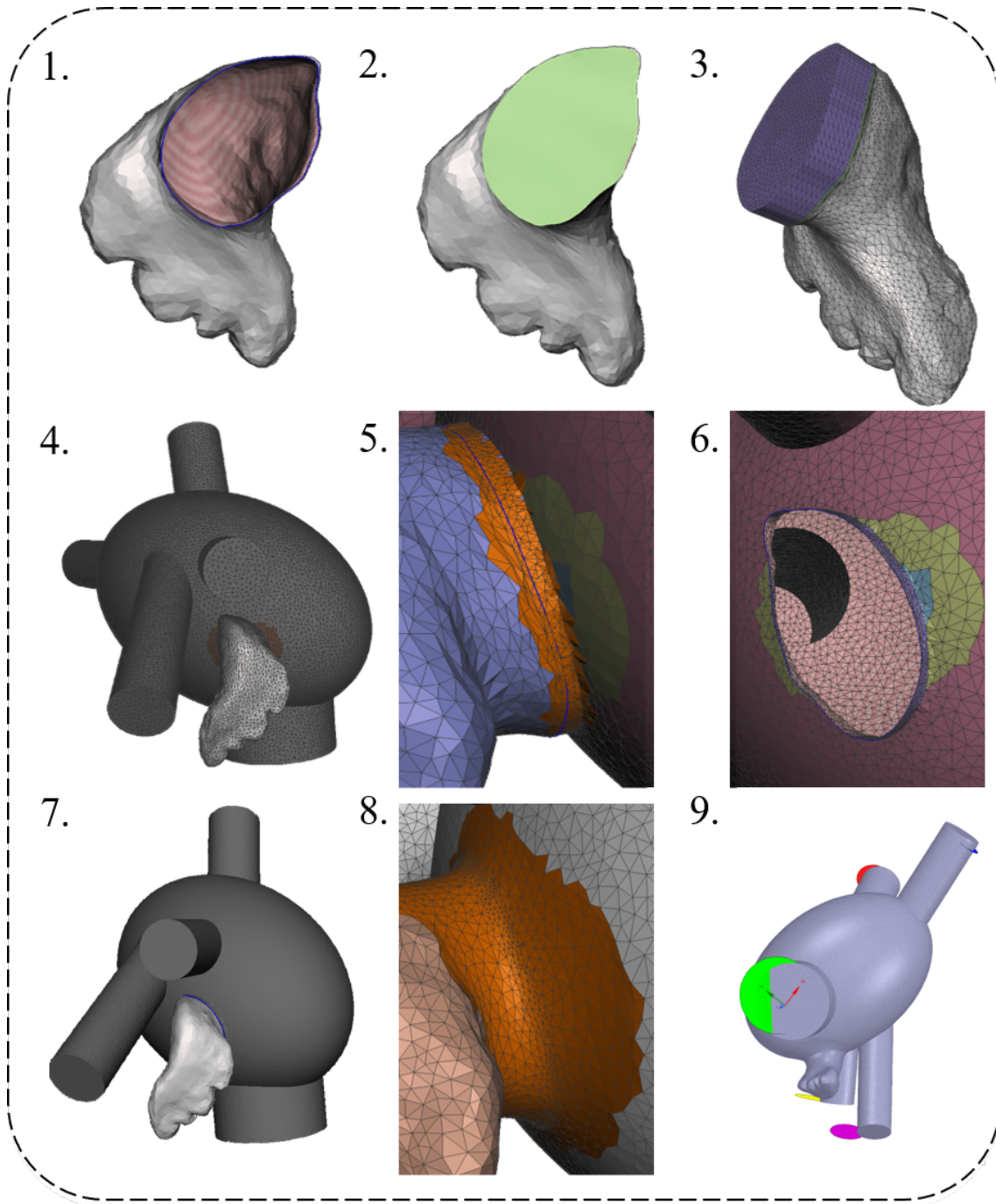


Figure 7: Assembly procedure of the geometry. 1. LAA shape obtained from the SSM 2. Seal the ostium 3. Extrude twice the ostium cover 4. Assemble with the Oval 5. Separate the second extrude and perform a boolean union with the Oval 6. Combine back the oval and second extrude to the LAA and the first extrude, eliminate LAA 7. Load the LAA again 8. Seal and smooth the elements 9. Align the MV on the z-axis.

Due to the need to maintain unaltered the inter-nodal connectivity, this process turned out to be one of the most delicate in the whole computational pipeline. It was completed in its entirety in Meshmixer. The different steps involved on the assembly of the final geometry are shown in Figure 7.

First of all, the original LAA SSM shapes, stored at first on vtk files, were converted to ply format both, to be able to work with them on Meshmixer and to track that the connectivity of the mesh was not altered throughout the process. The aim of steps 2-6 is to allow the LAA geometry to be attached to the oval cavity without altering the LAA surface mesh in anyway. With that in mind, the ostium was sealed (2) and later on, extruded twice (3). Then, the oval LA was imported, and slightly rotated to align it to the LAA (4). To aid in this endeavour, the oval was color coded so that all the appendages were attached nearly on the same position. Afterwards, the second extrude was separated from the LAA and joined with the LA ovoid through a boolean union. If done directly on the LAA, this process would have completely remeshed the geometry of interest, ruining the mesh correspondence. Once completed, the first extrude was joined back to the second one (5) using the "Combine" functionality in Meshmixer, having created a mesh that perfectly matches our LAA geometry.

Before the ensemble was completed, the LAA had to be removed and loaded again, to preserve the order of the nodes on the ply files (6-7). This time as the first extrude matches exactly the LAA they can be joined without altering the mesh connectivity. Later on, the smoothing tool allowed to eliminate defects on the LAA surface by rearranging all the elements as evenly as possible without altering the connectivity, seeking to avoid numerical convergence problems (8). This process was completed manually through visual inspection of the mesh.

Finally, the pulmonary veins and mitral valve of all the geometries were aligned to a template on Space Claim (9). The MV of the template was previously aligned on the Z-axis direction, which is essential for the correct implementation of the dynamic mesh as it will be explained on the next section.

Following this pipeline, we managed to obtain 300 completely aligned virtual LAA geometries. From case to case, only the relative position of each node forming the LAA varies slightly. The tetrahedral 3D mesh was then generated on gmsh and saved on a bdf file, which resulted in volumetric meshes composed of approximately 350.000 elements each ($\mu = 352180$, $\delta = 11237$). It is worth mentioning that, while we imposed correspondence on the surface mesh, the same is not true for the volumetric grid. Nonetheless, since the parameter of interest is computed over the elements of the surface this fact is irrelevant as this remains untouched.

2.3 Generation of risk indices for thrombus formation - ECAP

The endothelial cell activation potential (ECAP) was the parameter chosen to evaluate the risk of thrombus generation following studies such as the one by Di Achille *et al.* [49]. This parameter is based on the wall shear stress, but it is averaged over time. Therefore, it greatly simplified the required DL architecture. The wall shear stress was obtained through conventional CFD simulations before computing the ECAP values which served as the ground truth for the posterior DL analysis.

2.3.1 CFD simulations and boundary conditions

As previously stated, all the simulations were automatically performed through the MATLAB - ANSYS Fluent interface. A script was developed for that purpose, which automatically switched the volumetric mesh and run all the simulations on a loop. Each of the completed simulations encompasses a full cardiac cycle. The systole lasted 0.40s , imposing an input blood flow through the PVs obtained from clinical observations made by Fernandez-Perez *et al.* [50]. Wall boundary conditions were applied to the MV during this cardiac phase, as it is closed during this period. On the other hand, the diastole was defined to last 0.65 imposing an outlet pressure of 1067 Pa on the MV, simulating an open valve. Thus, the total length of the cardiac cycle is 1.05 seconds.

To simulate the motion of the LA and to prevent breaking the law of mass conservation during the cardiac cycle [48], a diffusion-based smoothing dynamic mesh is applied in the MV annulus on the longitudinal direction (z-axis). Accordingly, the mesh motion is governed by the following diffusion equation [47],

$$\nabla \cdot (\gamma \nabla \vec{u}) = 0 \quad \gamma = \frac{1}{d^\alpha} \quad (2.3)$$

where \vec{u} is the mesh displacement velocity, d is a normalized boundary distance and γ is the diffusion coefficient, a function of the boundary distance. When using smoothing, the interior nodes of the mesh move and “absorb” the movement of the boundary, but the number of nodes and their connectivity does not change. At first several CFD simulations were attempted with $\gamma = [1, 2]$. After facing continuous problems on the convergence of the simulations due to negative cell volumes, a diffusion coefficient of $\gamma = 0$ was chosen as a last resort.

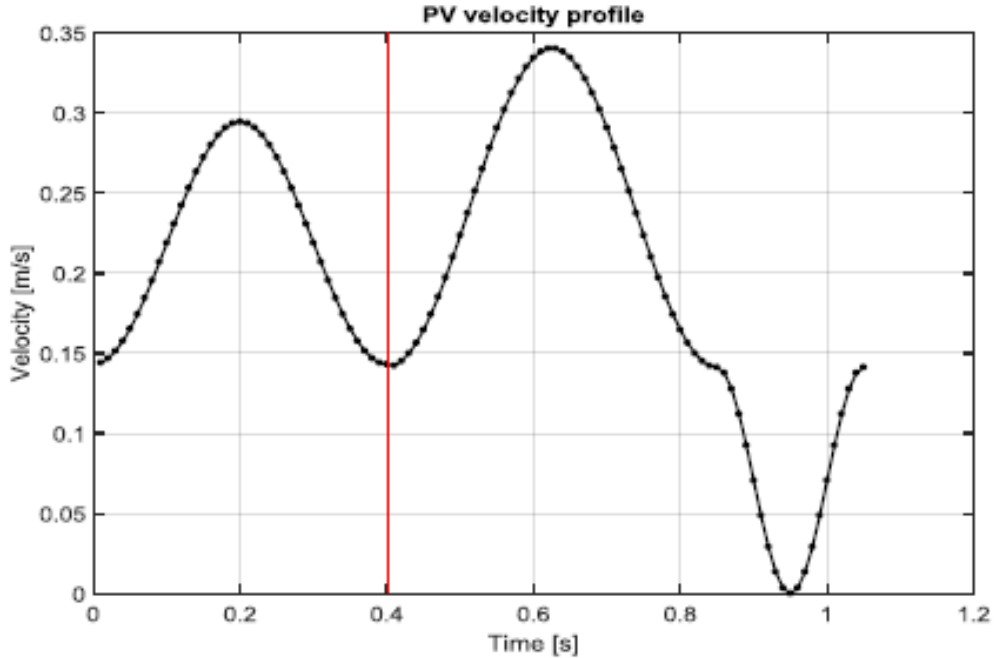


Figure 8: PVs inlet velocity profile throughout the whole cardiac cycle from systole 0 – 0.40s to diastole 0.41 – 1.05s. The red line indicates the transition between the two cardiac cycles.

Regarding the remeshing process, local cell remeshing was chosen, setting the length scale to the values obtained from the mesh scale info, a cell skewness of 0.9 and remeshing interval of 6 to avoid exceedingly long computation times. The interior and exterior zones were defined as deforming areas, the PVs as stationary and the MV as a rigid body with a cell height of 0.001.

The longitudinal motion is imposed through a transient-table oriented on the z-axis with data from real-time 3D echocardiography measurements performed by Veronesi *et al* [51]. Blood was modeled as an incompressible Newtonian fluid with density $\rho = 1060 \text{ kg/m}^3$ and dynamic viscosity of $\mu = 0.0035 \text{ Pa.s}$. Flow was assumed laminar with isothermal and non-gravitational conditions. Blood flow was simulated using the incompressible Navier-Stokes equations. For the equation's resolution, a time step of $\nabla t = 0.01s$, 200 maximum iterations and a second order implicit unsteady formulation was used, together with a standard partial discretization for the pressure with an under-relaxation factor of 0.3 and 0.7 respectively.

Each time step the density, the velocity field, the absolute and total pressure, the vorticity, the mass-imbalance and the wall-shear stress were recorded in an encase gold file, essential parameters for the ECAP calculation and that ensured the successful completion of the simulation.

2.3.2 ECAP definition

The endothelial cell activation potential (ECAP), defined by Di Achille *et al.* [49], unifies several parameters related with the localization of thrombogenic areas. The time-averaged wall shear stress (TAWSS) is defined as follows

$$TAWSS = \frac{1}{T} \int_0^T |\vec{\tau}_W| dt \quad (2.4)$$

, where $|\vec{\tau}_W|$ represents the instantaneous magnitude of the WSS $[\text{N/m}^2]$ and $T[s]$ the duration of the pulse period.

The oscillatory shear index (OSI) measures the differences between both parame-

ters exposing the WSS vector deflection from the main blood flow direction. It is computed by the following expression,

$$OSI = \frac{1}{2} \left(1 - \frac{|\int_0^T \vec{\tau}_W dt|}{\int_0^T |\vec{\tau}_W| dt} \right) \quad (2.5)$$

Its value range goes from 0, which represents the complete absence of deviation, to 0.5, that indicates a 180° WSS vector deflection. High values are associated to complex flows and complex shapes.

Finally, the ECAP measures the ratio of OSI to TAWSS with the aim of localizing regions with high clot susceptibility. High values are associated with endothelial susceptibility and risk of thrombus formation.

$$ECAP = \frac{OSI}{TAWSS} \quad (2.6)$$

2.3.3 Simulation post-processing

After the completion of a simulation for a given geometry, each one of the executed simulations was imported to ParaView [46] to extract all the required data for the ECAP computation. With the aim of accelerating the process, the ParaView-python interface, pvpython, was employed to run a trace file that automatically stored all the data necessary for the obtention of the parameters of interest. The information of each time-step stored in the encase file was converted to .csv files and the LA geometry was saved to an stl file. Finally, a MATLAB script was used to compute the ECAP.

Once all the data was gathered, the generated models and simulations were visually inspected to ensure the ground truths for the training data were correct. For that purpose, the connectivity of all geometries was verified to be the same. As shown on Figure 9, it was discovered that the identification number for the nodes in the first ring of the LAA (in red), area corresponding to the ostium, had been altered. As a consequence, this nodes had to be removed, since reestablishing correspondence

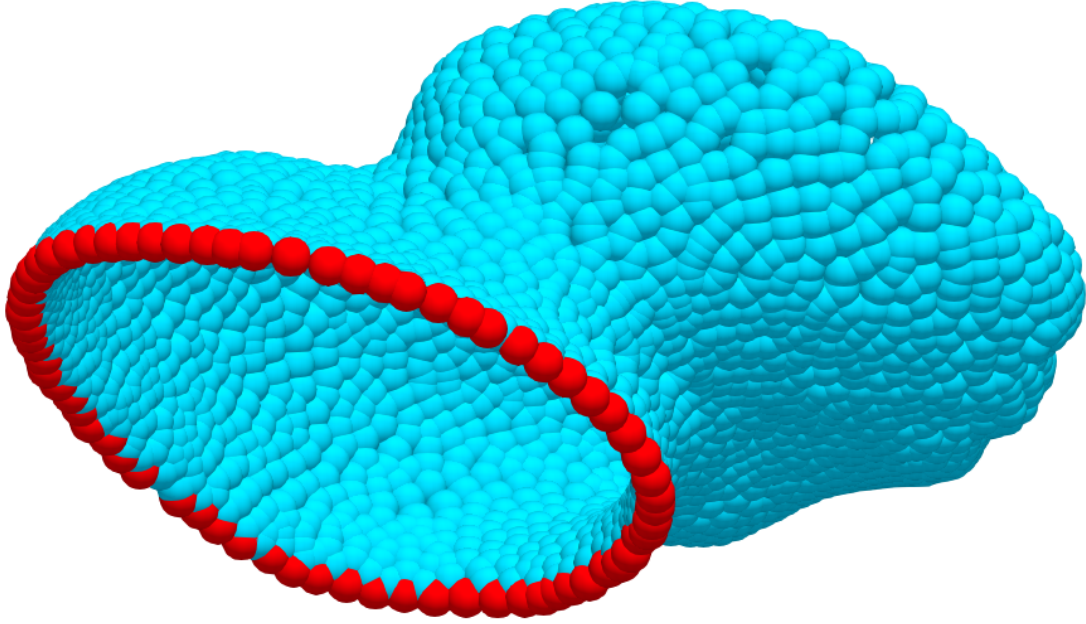


Figure 9: 70 nodes (in red) where eliminated on the outer ring of the LAA. The rest of the 2466 nodes (in blue) where retained for the deep learning analysis.

on all the geometries proved to be imposible. By removing this set of 70 nodes only 2466 remained for the posterior DL analysis.

Afterwards, the ECAP maps obtained for the SSM generated LAAs, were subjectively compared to those obtained on a small group of real patients LAA geometries in order to spot simulations in which the results were clearly unrealistic due to numerical problems during the simulation. After this screening, a final training dataset of 210 LAA geometries and corresponding ECAP maps was obtained.

2.4 Deep Learning Model

The main goal of the machine learning analysis was to develop a surrogate of the CFD simulations, enabling an accurate analysis of LAA thrombogenesis with unprecedented speed. Given an input LAA shape, the machine learning algorithm should be able to output the LAA ECAP maps while being consistent with the performed finite element analysis. Thanks to its innate capability to establish non-linear relationships between two sets of variables, DNNs suit perfectly the description

above, given that sufficient data is provided for an adequate training of the network.

With these objectives in mind, the workflow of Liang *et al.* [35] was taken as a basis, due to the obvious resemblance of the tasks in hand, and adapted to the specific requirements of this project. That being said, it is of utmost importance to emphasize the substantial differences between the aorta and the LAA, specially in terms of morphological complexity, which represents a far more daunting challenge. Moreover, two distinct DL models were tested on this work, further details about their implementation being provided on the following sections.

With regard to software, Keras 2.2.4 (open-source) [52] was chosen as the high-level neural network API in which to implement the DL model, due its highly modular networks and fast prototyping. It was run on top of TensorFlow 1.13.1 (open-source) [53], using Anaconda (Python 3.7) [54] as the python development environment. The code compilation was carried out through the Spyder IDE. The hardware available for this process was composed by a laptop with 16 GB RAM and an Intel Core i7-8565 quad-core processor. Due to the availability of a NVIDIA GTX 1050 graphic card with enough compute capability (6.1), it was possible to leverage TensorFlow's GPU support by installing the CUDA toolkit 1.10.1 [55] and its corresponding cuDNN 7.6.0 [56] library for accelerated parallel computing on the model training.

2.4.1 Simple fully connected (SFC) network

At first sight, having achieved complete nodal correspondence among all of the LAA meshes, performing a simple non-linear regression by implementing a simple fully-connected feed-forward network between the space coordinates and the ECAP maps seemed to be the most sensible approach. Accordingly, a multilayer neural network was designed to map the space coordinate triplets to their corresponding ECAP values. As shown on Figure 10, the final layout, is comprised of two equally sized hidden layers, counting a total of 5000 nodes each. The depth of the network was chosen following the recommendations by Heaton *et al.* [57] suggesting the use of no more than two layers in feedforward neural networks, which at the same time

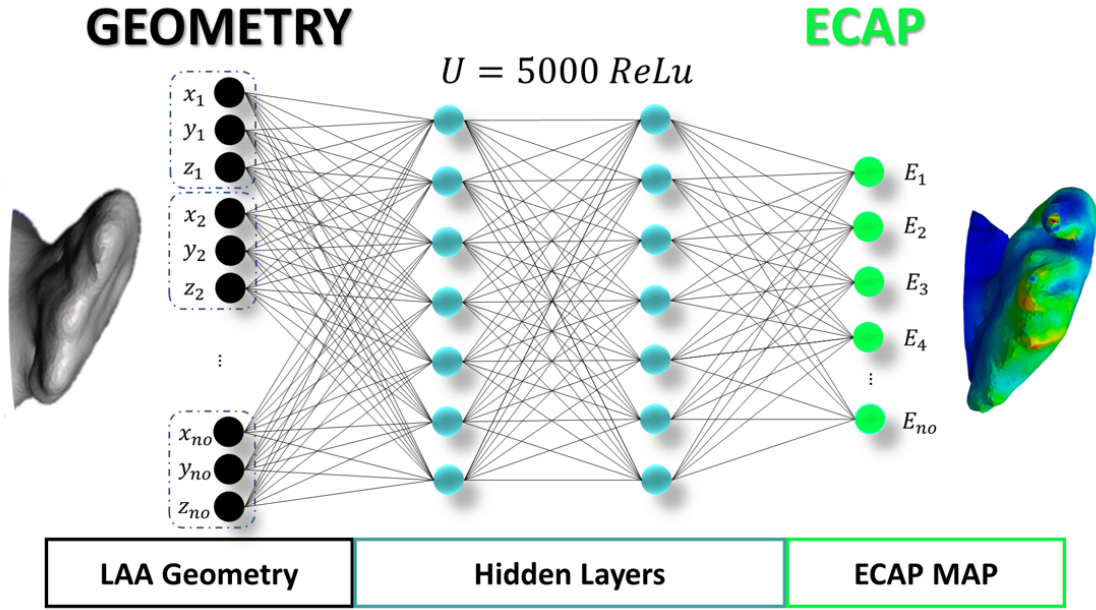


Figure 10: Network structure for the simple fully-connected feedforward network. no = Number of mesh nodes, U = Number of activation units.

happened to offer superior performance over deeper architectures. Following the same guidelines, the chosen number of hidden nodes was somewhere between the input (7398) and output (2466) unit number.

2.4.2 PCA dimensionality reduction network

Afterwards, an alternative approach was attempted, which more closely resembled that of Liang *et.al* [35]. As stated earlier, the input of the neural network would be the spatial (x, y, z) coordinates of the 2466 nodes forming the LAA geometries. The deep neural network would try to approximate the non-linear function relating the LAA shape defined by the spatial coordinates and its corresponding ECAP map. This means that the input and output of the neural network would be composed of 7398 and 2466 nodes, representing the surface of the mesh and the predicted ECAP values respectively.

This implies a exceedingly high input and output space dimensionality that could hinder the network training time cost, lead to undesirably high computational power consumption and potentially increase the likelihood of convergence problems of the

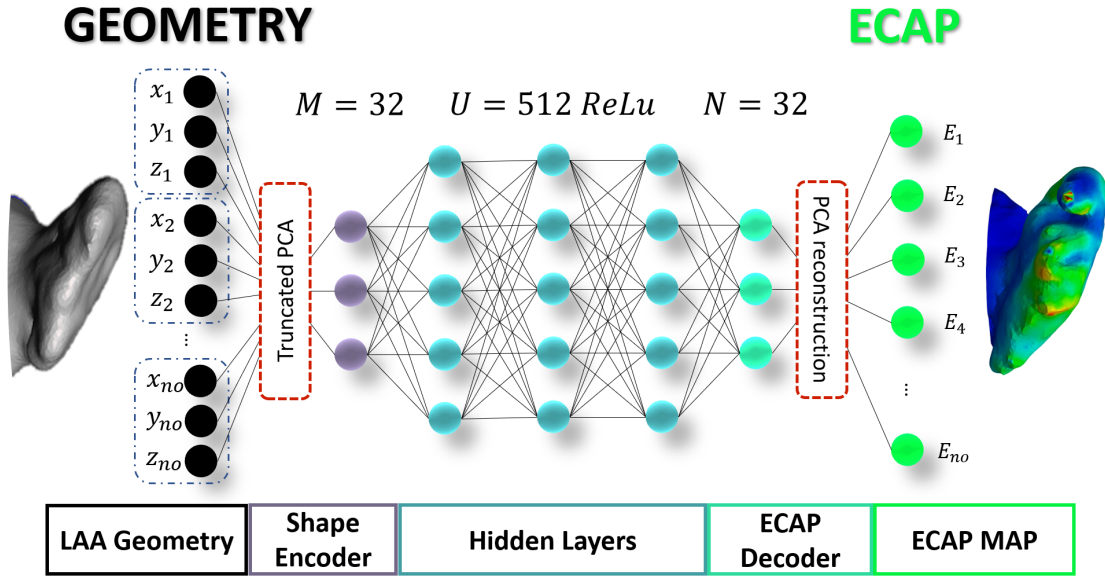


Figure 11: Network structure for the dimensionality reduction DL model. no = Number of mesh nodes, U = Number of activation units. Figure adapted from Liang *et al.* [35].

training algorithm [58]. For this reason, truncated PCA was performed in both the input feature space and the output ECAP variables as a previous step to carrying out the non-linear regression. The basic pipeline of this approach goes as follows:

1. Dimensionality reduction of both the spatial geometry and ECAP map data to a small set of scalar values through truncated PCA;
2. Training of the non-linear mapping by feeding through the set of low dimensional scalar values representing the shape and ECAP of the training dataset;
3. Predict the ECAP maps of the test geometric dataset by feeding its corresponding low dimensional representations to the net and reconstructing the ECAP values through matrix multiplication.

As aforementioned, truncated PCA was implemented on both the input and output spaces to partially reduce the complexity of the problem. Besides, it is well known that neural network training tends to converge faster if the input is whitened, that is, linearly transformed to have zero mean, unit variances, and being decorrelated

[59]. That is precisely what PCA does, allowing for dimensionality reduction of large datasets at the same time that the information loss is minimized by generating new uncorrelated variables that are linear functions of the original dataset. These new variables are known as principal components (PC). Following this definition we can describe a specific LAA shape as shown on the equation below [60]:

$$X = \bar{X} + \sum_{m=1}^M \alpha_m \sqrt{\lambda_m} W_m \quad (2.7)$$

, where \bar{X} is the mean shape and W_m and λ_m is the set of eigenvectors and eigenvalues of the covariance matrix for the retained number m of PCs. If the variability of the dataset is explained by a small set of PCs, that is, most of the variability is concentrated in a small number of PCs, the shape of the LAA can be accurately represented by a small set of scalar values as follows

$$\beta_m = \frac{W_m^T (X - \bar{X})}{\sqrt{\lambda_m}} \quad \epsilon_n = \frac{W_n^T (Y - \bar{Y})}{\sqrt{\lambda_n}} \quad (2.8)$$

The same can be employed on the ECAP distributions, effectively reducing the dimensionality of the input and the output. In our model the number of PCs retained was chosen by performing a grid search on the obtained prediction accuracy. In the final model, $m = n = 32$ PCs were kept for both the shape and the ECAP, which retained a total variability of 97,6% and 90,3% respectively.

Having performed the truncated PCA, a fully-connected feed-forward neural network was designed to perform a non-linear regression between the low dimensional scalars, β_m and ϵ_n , representing both the geometry and the ECAP, similarly to Section 2.4.1. The structure of the network is shown on Figure 11 which is composed of 3 hidden layers of 512 units, both the number of nodes and the layer depth being selected due to superior performance over other alternatives.

Once the deep neural network has been successfully trained, predicted ECAP values are reconstructed from its low-dimensional representations reversing Equation 2.8

$$Y = \bar{Y} + \sum_{n=1}^N \left(\epsilon_n \cdot \sqrt{\lambda_n} \right) W_n \quad (2.9)$$

2.4.3 Training of the model

After setting up the network structure, the loss function, optimizers and different hyperparameters were chosen in order to tune up the DL model. Most of the following parameters have been employed equally on both of the DL networks. First of all, Adam was selected as the compiler of choice, a modification of the classic stochastic gradient that has been shown to outperform similar algorithms while offering improved efficiency [61]. On the other hand, the mean square error (MSE) was selected as the loss function due to superior performance.

Likewise, among all the options available for non-linear regression, ReLU units were chosen over other nonlinear units, such as Leaky ReLU or Softplus, as they delivered better results. Furthermore, the number of epochs was also assessed to find the sweet-spot between under and overfitting. For the simple feedforward network, 200 epochs seemed optimum while for the PCA approach 300 epochs were employed. A batch size of 20 was used on both cases.

2.4.4 Evaluation of the model

The predictive capabilities of both DL models were evaluated through Monte Carlo cross-validation. This means that in each round of the cross-validation 90% of the dataset was randomly selected for training purposes, while the rest served to compute the performance of each neural network.

With this in mind, two different metrics were chosen: mean absolute error (MAE) and the root mean square error (RMSE), being some of the most common metrics to measure accuracy on continuous variables.

MAE measures the average magnitude of the errors in a set of predictions, where all observed individual differences between prediction and ground truth share equal weight. NMAE is its dimensionless variation, obtained by dividing MAE by the maximum absolute ground truth observation.

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \bar{y}_j| \quad (2.10)$$

RMSE is calculated as the square root of the average of squared differences between prediction and ground truth. It also expresses average model prediction error in units of the variable of interest, but it also gives more weight to large errors as the difference is squared.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \bar{y}_j)^2} \quad (2.11)$$

Aside from the aforementioned accuracy measures, a set of different metrics was used in order to evaluate the performance of the DL model on predicting the peak ECAP values that have been linked to potential risk of thrombogenesis. For that purpose, the absolute error (AE) was computed between the maximum predicted and ground truth values. Aiming to avoid this measurement to be conditioned by a small set of outlier nodes with exceedingly high values of ECAP, this metric was also calculated for the 99th and 90th percentile of the highest ECAP values.

$$AE = | \max(|y_j|) - \max(|\bar{y}_j|) | \quad (2.12)$$

In addition the normalized absolute error (NAE) was also calculated as follows

$$NMAE = \frac{MAE}{\max(|y_j|)} \cdot 100 \quad (2.13)$$

Chapter 3

Results

3.1 Deep Learning model

As previously stated, a final dataset of 210 artificial LAA surface meshes was assembled, each composed of 2466 nodes sharing the same inter-nodal connectivity. At the same time, the ECAP values considered as the ground truth were obtained through CFD simulations and later contrasted with the predicted ECAP maps from the DNNs.

Once all the required data was gathered both the DL networks were trained. During this process a significant difference on time performance was noticed between the two employed approaches. Each cross-validation iteration on the truncated PCA model required just 30 seconds to perform the unsupervised learning analysis, the non-linear mapping and the reconstruction matrix multiplications. In contrast, the training of simple fully-connected (SFC) network took about 25 minutes.

3.1.1 Performance

With the aim of quantitatively assessing each network's performance a Monte Carlo cross-validation of 100 folds was implemented. The details about the methods, data and performance metrics employed are described in depth on Section 2.4.4. The accuracy achieved by each of the networks is shown on Table 2. It is important to

Table 2: Performance of DL models on a 100 fold Monte Carlo cross-validation. MAE: Mean Absolute Error, RMSE: Root Mean Square Error, NMAE: Normalized Mean Absolute Error.

Network	MAE	RMSE	NMAE
PCA	0.6457 ± 0.0493	0.8905 ± 0.0856	$4.7197 \pm 1.4625\%$
SFC	0.6486 ± 0.0462	0.8852 ± 0.0848	$5.7558 \pm 1.523\%$

point out that the ECAP parameter is a pure positive scalar quantity and therefore, both the MAE and the RMSE have no physical dimension assigned. Both DL models achieve very similar accuracy, attaining an average MAE of 0.6457 ± 0.0493 for the dimensionality reduction model and 0.6486 ± 0.0462 for the SFC network.

Table 3: Absolute error (AE) error of peak ECAP values. Max: Maximum peak values, 99th: Upper 99th percentile, 90th: Upper 90th percentile.

Network	AE Max	AE 99th	AE 90th
PCA	1.4978 ± 0.1559	0.7869 ± 0.0893	0.5210 ± 0.0460
SFC	1.4759 ± 0.2112	0.7723 ± 0.1330	0.5157 ± 0.0815

Table 4: Normalized absolute error (NAE) of peak ECAP values. Max: Maximum peakvalues, 99th: Upper 99th percentile, 90th: Upper 90th percentile.

Network	NAE Max	NAE 99th	NAE 90th
PCA	$20.93 \pm 1.98\%$	$16.13 \pm 3.05\%$	$14.52 \pm 1.83\%$
SFC	$21.63 \pm 3.79\%$	$16.6 \pm 6.3\%$	$14.98 \pm 3.84\%$

Regarding peak ECAP values related to thrombosis risk, there seems to be a noticeable downward trend as we go from peak values to lower percentiles of the absolute error (AE), as seen on Table 3. This trend seems to be mirrored by the normalized absolute error (NAE) values on Table 4.

The predicted ECAP mappings alongside with the ground truth, can be visualized for 7 representative LAA shapes in Figure 12. These specific set of LAA shapes, belong to one single iteration of the Monte Carlo cross-validation, both networks being trained with same dataset composed of 189 instances. Once trained, a test dataset of 21 shapes was used to generate the predictions shown in the figure. The

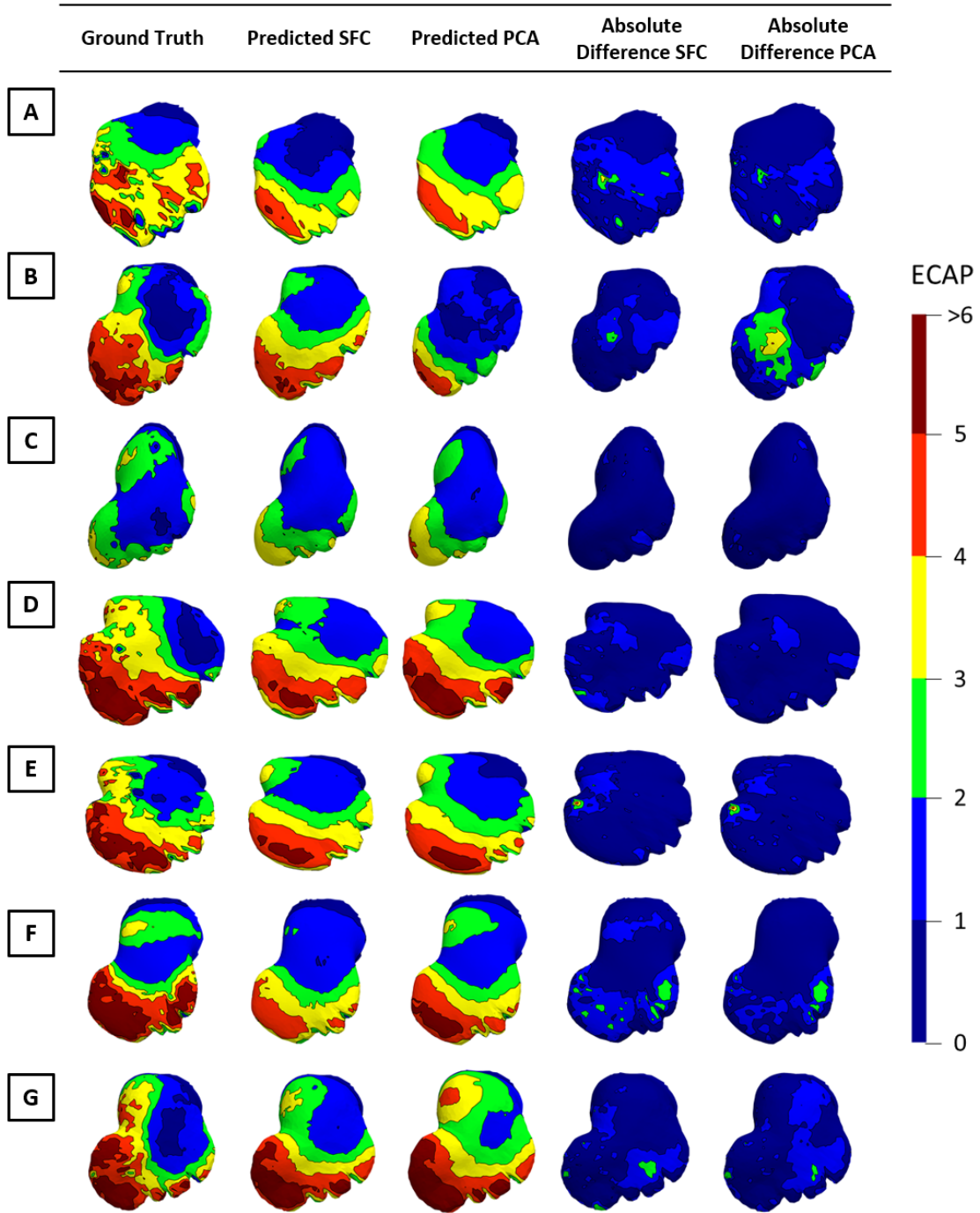


Figure 12: Set of ECAP values on a random test dataset of 21 shapes, 7 of which are shown in here. Each row corresponds to a given geometry while the columns from left to right represent: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA

Table 5: Mean average error (MAE) for the first 7 LAA geometries in Figure 12

Networks	A	B	C	D	E	F	G
PCA MAE	0.6498	1.0815	0.5509	0.4652	0.5138	0.5682	0.6774
SFC MAE	0.7950	0.6591	0.6861	0.5259	0.5459	0.7537	0.6254

ground truth (GT) can be seen along the predictions from both DL models in a discrete scale ranging from 0 to 6. Besides, the absolute error of the predictions can be observed in the rightmost columns at the same scale. The posterior face of the first 7 geometries shown in here and the results for the remaining 14 geometries can be seen in the Appendix A.1 and A.2. In addition, the MAE for the first 7 geometries are reported on Table 5, while the metrics for the rest of instances can be found on Appendix B.2.

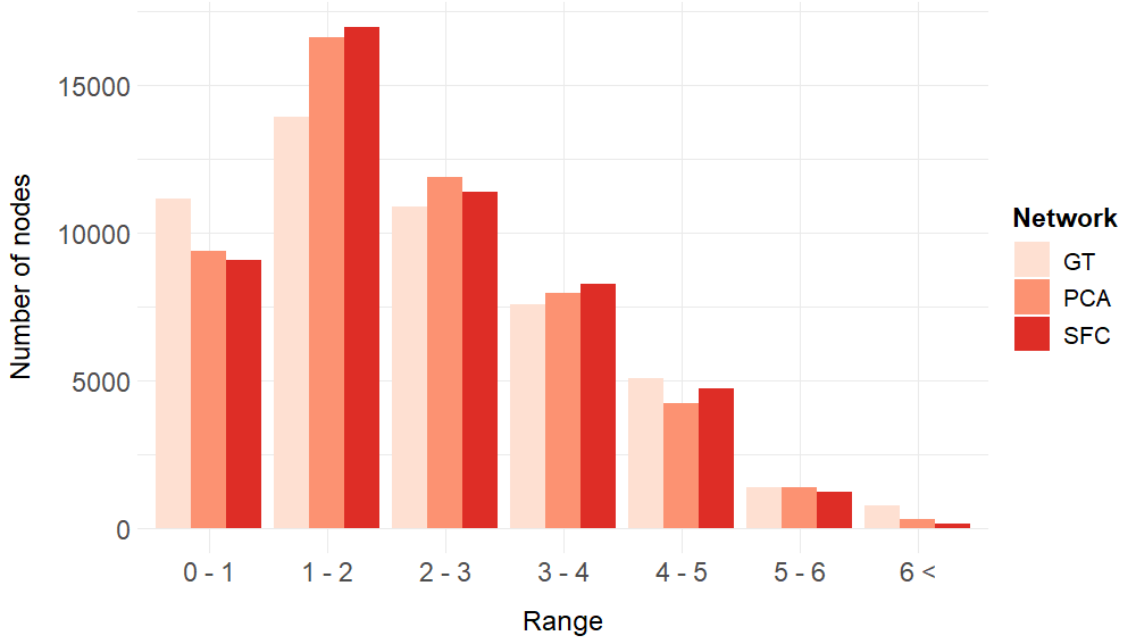


Figure 13: Total node count of the 21 test simulations in each range of ECAP values

Subsequently, the decision was made to divide the nodes in a set of ranges according to their ECAP value following the discrete scale on Figure 12. The results are reported on the histogram on Figure 13, making it fairly clear that the DL models tend to underestimate the number of nodes with the most extreme values. The specific values for each range and network are reported on Table 7 on Appendix B.

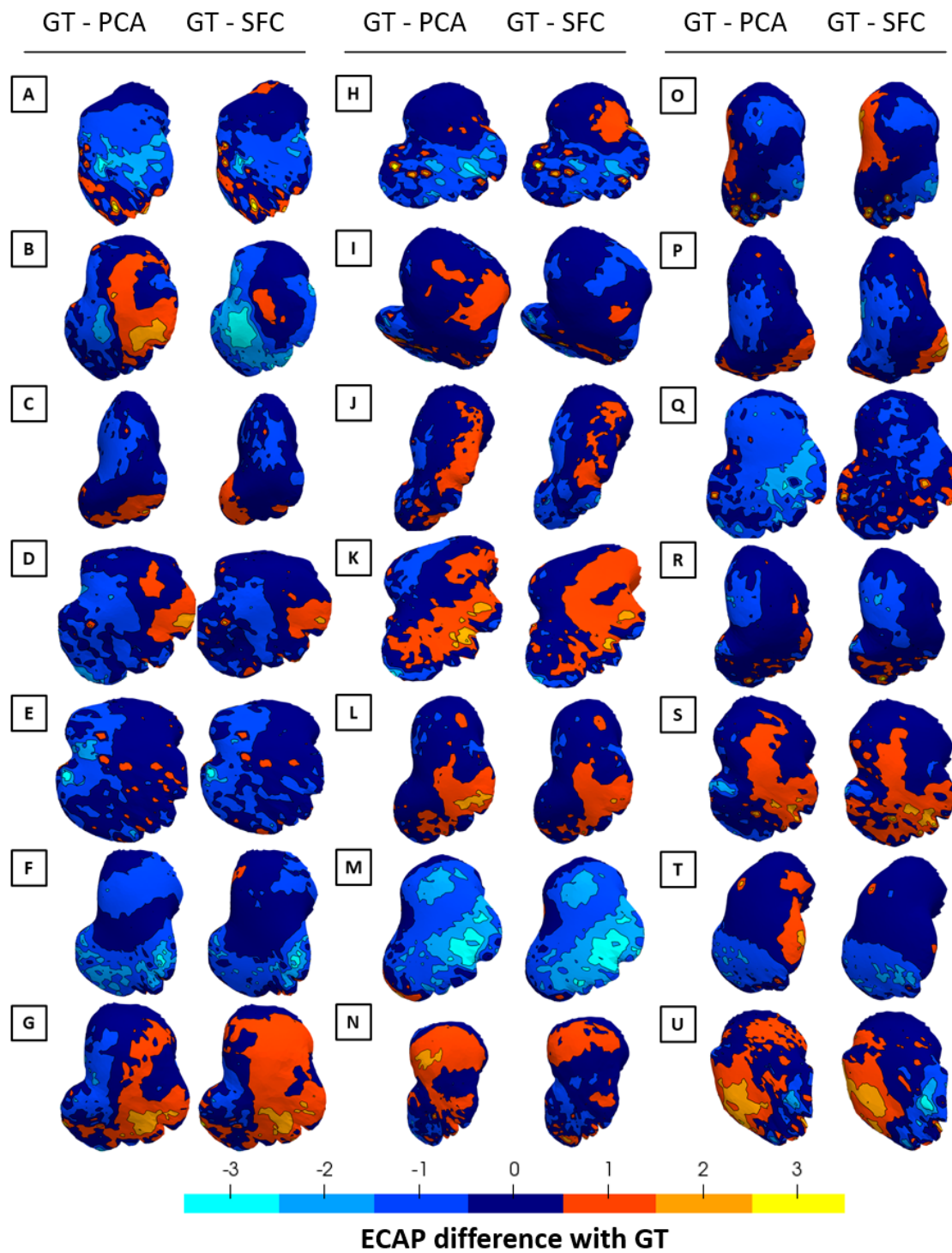


Figure 14: ECAP difference computed as the subtraction between the ground-truth (GT) and the PCA based DL model on the left columns, and the GT and the simple fully-connected (SFC) model on the right columns. The front-face is shown in here. The posterior faces can be found on Figure 22

To analyze the central tendency of the ECAP mappings the mean was computed, resulting in a slightly higher mean for the GT with $\mu_{GT} = 2.3446 \pm 0.2540$, while the mean for both predictive models turned out be almost identical with $\mu_{PCA} = 2.2398 \pm 0.2456$ and $\mu_{SFC} = 2.2428 \pm 0.2021$. Similarly for the standard deviation, the GT comes up on top $\delta_{GT} = 1.4982$, compared to its DL counterparts which are almost identical $\delta_{PCA} = 1.3019$, $\delta_{SFC} = 1.2967$.

On the other hand, the subtraction of the CFD obtained ECAP maps and its predictions can be seen on Figure 14, with the PCA and SFC results being side by side. This time all the 21 test geometries are shown together avoiding the computation of the absolute difference so that any tendency of the predicted ECAP maps to have lower or higher values than the GT could be better visualized. The scale was set to $[-3, 3]$ to better appreciate the subtle changes between the models. The figures for the posterior faces can be found on Appendix A.4.

Similarly, to learn if each of the different DL models was better suited to capture specific features of the geometry and to better spot the disparity in their predictive capabilities, their subtraction was also computed and it is shown on Figure 21.

3.1.2 Detection of potentially high risk thrombotic areas

While previously the whole ECAP value range has been considered to measure the accuracy of the predictions, throughout this section the focus is going to be set on the highest ECAP values, as they are the most relevant when assessing the potential risk of thrombogenesis. Following the study by Di Achille *et al.* a value of 4 was set as the lower bound for high thrombogenic risk [49]. Consequently, any node with an ECAP value above such threshold is considered to be on high risk for thrombi generation.

With this in mind, a binary classification of the nodes with ECAP value above 4 was performed, with the positive condition being $ECAP > 4$, that is, corresponding to a region with potential thrombi risk and the negative condition represented by $ECAP \leq 4$. The results are shown on Figure 15. Following the statements above, nodes

Table 6: Confusion matrices for both SFC and PCA predicted ECAP values

		SFC Prediction outcome		PCA Prediction outcome		
		DL Risk	DL Safe	DL Risk	DL Safe	Total
True Value	Risk	233.95 0.5967	158.14 0.4118	230.62 0.5882	161.48 0.4033	2073.9
	Safe	55.57 0.0268	2018.3 0.9732	50.19 0.0242	2023.7 0.9758	392.1
Total		289.5	2176.5	280.8	2185.2	

above the threshold of 4 in both the GT and DL predictions were considered as true positives (TP), those that stood above just on the GT were labeled as false negatives (FN), ECAP values that were mistakenly predicted as high risk were defined as false positives (FP) and finally, nodes below the threshold on both cases were classified as true negatives (TN). The above labels have been color-coded as follow on Figure 15: TN - Green, FN - Blue, FP - Red and TN - Grey.

Furthermore, to perform a qualitative assessment of the networks classifying capabilities, a couple of confusion matrices were constructed for each of the DL networks which are reported on Table 6. The mean number of nodes across the 21 sample shapes for each of the labels appears on the top row of the matrices with their corresponding rates shown below. On the right column the true amount of positive and negative condition values is shown, while the predicted numbers are shown below.

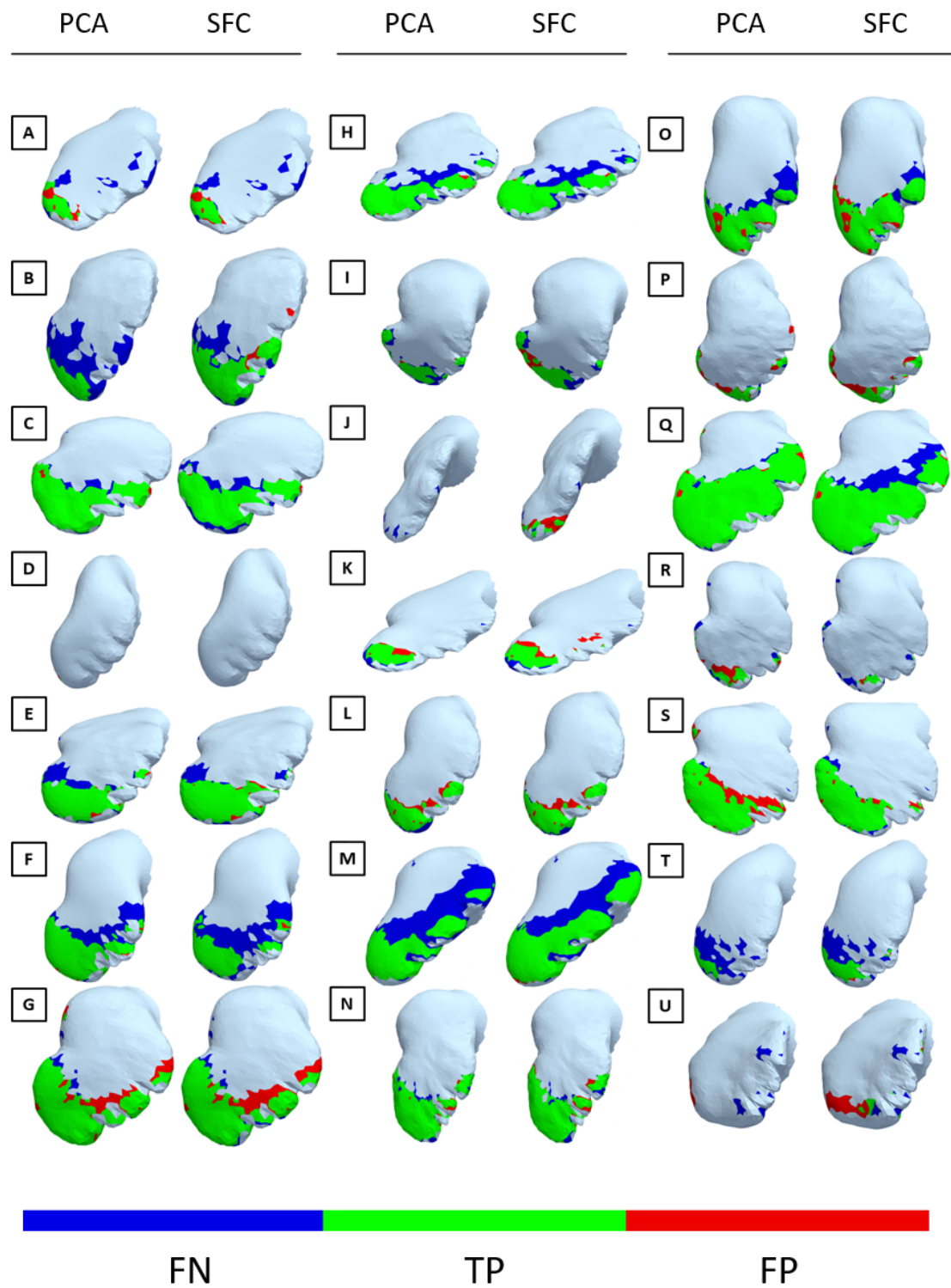


Figure 15: An ECAP value of 4 has been chosen as a threshold for potential thrombi formation risk. The performance of the DL model to detect such risk areas is evaluated by measuring the areas in which the $ECAP > 4$ on both the GT and the DL predictions (True Positive in green) intersect, areas that only appear to be risky on GT (False negative in blue) and areas that have been misdiagnosed by the DL model (False positive in red)

Chapter 4

Discussion and conclusions

In this thesis, a DL based surrogate of CFD simulations was proposed, with the aim of reducing computational time and cost of hemodynamic characterization, by directly estimating the ECAP values of the LAA from its geometrical shape alone. With this objective in mind, two different DNN models were developed to perform a non-linear regression between the variables. The first and simplest approach, consisted on a simple fully-connected feed-forward network. The latter, combined supervised and unsupervised learning by performing dimensionality reduction of the involved variables through truncated PCA, and reconstructing them back once the non-linear mapping was over. To our knowledge, the work developed in this thesis represents, only behind the study performed by Liang *et al.* [35], one of the first successful implementations of DL techniques for fast and accurate prediction of hemodynamic parameters conventionally attained through CFD.

First and foremost, the main objective of the study being the reduction of the processing time and computational requirements of finite element analysis, the time required to obtain the ECAP distributions has dropped down several orders of magnitude thanks to the implementation of the DNNs. As a matter of fact, while the CFD simulations performed to obtain the ground truth ECAP values took an hour and a half on average, once trained and given a set of input shapes, the DL networks required less than a second to obtain their corresponding predictions. Furthermore,

the training of the DNNs took less time than the completion of one single CFD simulation. Yet, it must be stated that the SFC approach took much longer to train, 25 minutes approximately, than the dimensionality reduction model, which completed the training of the multilayer neural network in under 30 seconds.

Secondly, the performance of both DNNs was benchmarked against the ground truth ECAP maps obtained from the CFD simulations through a 100-fold Monte Carlo cross-validation. This led to a MAE of ~ 0.64 for both DL models. Thus, it is clear that DNNs had remarkable success in providing highly consistent predictions of ECAP distributions with those obtained through traditional finite element analysis. The performance of the DL networks is going to be analyzed more in depth on the sections below.

4.1 Oval geometry and dynamic mesh

After thoroughly analyzing the ECAP results obtained on the 4 different configurations, it was decided that the more sensible approach was combining the ovoid representation of the LAA together with the implementation of the dynamic mesh.

First of all, the CFD simulations in which the LA walls were considered rigid, assumption that might be partially correct on patients with AF, was completely discarded due to the abnormal ECAP distribution on the main lobe and exceedingly high peak values unlike those seen in studies such as the one by Di Achille *et al.* [49]. The rigid wall assumption means that the volume of the LAA cavity does not change during the cardiac cycle. As a consequence, the continuous blood inlet through the PVs, as defined in Figure 8, coupled with the closed MV during ventricular systole breaks the mass conservation law, resulting in mass imbalance. As the ECAP incorporates the TAWSS parameter, which is time averaged, and the systole being the most relevant time frame influencing ECAP, motionless simulations showcase much higher overall ECAP values.

In fact, this is what can be observed in Figure 6, ECAP peaks being concentrated at the end of the main lobe, with big patches showcasing an ECAP value well above 10,

exceeding by far the 99th upper higher ECAP percentiles presented by Di Achille *et al.* [49]. As a result, motionless simulations were deemed unsuitable for the study. In contrast, the upper percentiles observed in the dynamic simulations are much closer to those observed in the study, rarely exceeding values above 5.

Secondly, a decision had to be made between the artificial or anatomical representation of the LAA. The ECAP distribution of both morphologies share some resemblance, as well as more “realistical” peak values, but making use of real geometries would be always preferable. For that purpose, a code was provided by the research group to generate meshes from patient geometries that would share the same connectivity and number of nodes. Nevertheless, since the performed CFD simulations would not have had any clinical relevance regardless of the chosen LA, due to artificial LAA populations being used for the geometry assemble, the oval was chosen over the real geometries due to time constraints and in order to simplify the workflow.

4.2 DL Performance

As has been noted, both DNN networks developed in this thesis show more than adequate predictive capability given the difficulty of the task in hand, as it is clearly shown on Figures 12, 16 and 17. It seems that the multilayer networks are able to capture the non-linear relationship between the LAA geometry and the shape, with great overall resemblance between the ECAP distribution patterns of the GT and the DL predictive models.

On the other hand, it is well known that DL methods are considered as “black boxes” due to the difficulty in interpreting the results, that is, figuring out which of the input features are the most relevant or how they relate to the observed output. The DNNs on this thesis are no exception, and while we can measure how accurate the obtained prediction is, it is hard to discern any clear pattern or geometrical feature with which the neural network is struggling to capture. Giving a look at the absolute differences between GT and predictions on the aforementioned figures, it

can be observed that the differences stay mostly below one, with the exception of a couple of geometries on which the DL model struggles more for reasons beyond understanding.

Machine learning techniques, and more precisely DNNs, have not been widely implemented yet as a substitute to CFD simulations, and the studies looking to predict clinical hemodynamic parameters are fewer still. In fact, to our knowledge Liang *et al.* [35] have been the first in successfully implementing these networks with such objective in mind. Therefore, their work is the only reference available against which to benchmark our results. Overlooking obvious differences between both studies, since they are measuring several components of the stress tensor instead of a single parameter such as the ECAP, Liang *et al.* managed to obtain an average error of 0.492%, an order of magnitude lower than the NMAE of 4.72% and 5.756% obtained by the PCA and SFC DL networks employed on this study. Although at first glance it may look like a remarkable drop on the DL prediction accuracy, there are a number of reasons that explain such a disparity.

To begin with, perhaps the most influential factor is the complexity of the studied geometry. Compared to the LAA, the aorta is a very simple geometry with mainly laminar flow, making DNN-based surrogate estimation easier. Furthermore, a simplified version of the aorta was employed by Liang *et al.*, which had the branches of the aortic arch removed. In contrast, the LAA morphology from the SSM employed on this study, presents much higher heterogeneity with high size and cavity volume variability, different number of lobes or distinct direction in the folding of the main lobe among others. That being said, while it is true that the employed SSM of the LAA manages to capture major variations of a real anatomical appendage, by no means it encompasses the astonishing complexity and heterogeneity found on real anatomical LAA populations.

Secondly, this very complexity of the LAA morphology, severely restricted the possibility of applying structured grids when meshing the different shapes. This completely negated the employment of convolutional neural networks on regular patches of the geometry, similarly to what Liang *et al.* implemented, in order to successfully

extract features from the ECAP maps, which we hypothesize could vastly improve the prediction accuracy obtained in the study. Lastly, the size of the final dataset used in the training of the neural network was 3 times smaller than the one employed by Liang *et al.*. Nevertheless, this last point seems to be quite irrelevant, since the DNNs were trained with sequentially increasing dataset sizes and the performance of both networks hardly improved from 160 training geometries onward.

Alternatively, several attempts were made to improve the overall accuracy of both DL networks by including learning rate schedulers or increasing the depth of the network with little or no success. All things considered, it is believed that DL models employed throughout this thesis manage to approximate the non-linear function relating the LAA geometry and its ECAP maps quite successfully.

4.3 DL network model comparison

Despite the fact, that the input and output tensors for both models are drastically different, both in value and in size, due to the dimensionality reduction implemented on the PCA network, identical accuracy values are obtained for all relevant metrics. Conceptually, the dimensionality reduction network should have delivered superior performance since the non-linear function to be fitted is far more complex for the SFC model, *a priori*. Nonetheless, the set of input-output variables are so distinct that is very difficult to perform a proper assessment of the causes. This surprising result is ratified by the comparison made on Figure 14, where the subtraction between the GT and both predictions are plotted side by side. The overall patterns of difference between the ECAP mappings of both models seem to be identical. Interestingly, looking at the MAE values reported on Table 5 there seems to be no visible trend by which one network obtains a better precision than the other in a specific geometry.

However, Table 7 seems to showcase the most apparent difference between the two models. First of all, the central tendency for both ECAP predictions (more of the predicted nodes fall in the 1-4 ECAP value range than in the ground truth) seems to be a little bit lower than the ground truth. This is corroborated by the slightly

higher mean of the CFD attained ECAPs. As for the difference between the two networks in each of the 7 ranges, here seem to be no substantial differences except for the highest ECAP values scoring above 6, in which the PCA approach appears to obtain substantially higher number of nodes. In fact, looking on Figures 12,16 and 17 at the geometries in which the ECAP value seems to be the higher, such as the shapes C,F,M and Q, it can be observed that the SFC slightly underestimates peak ECAP values, which is translated in worse MAE score as seen in Table 5. Furthermore, Figure 21, which is the result of the subtraction of both predictions, showcases that in the aforementioned C,F,M,Q geometries the SFC approach has lower values on the final end of the main lobe. The histogram of Figure 13 further confirms this hypothesis as it clearly showcases that the SFC has the lower count on the total amount of nodes above 5 in comparison with the ground truth.

The absolute error on the peak ECAP values and the upper percentiles are shown in Tables 3 and 4. Similarly to the study by Di Achille *et al.* [49], the upper percentiles were also taken in account as it might be more convenient and robust than only considering peak nodal values so that areas with high values over a non-negligible region are also included instead of just isolated outliers nodes. It can be seen that while the AE is quite bad for the peak values as we go down on the percentiles the value improves. Therefore, it seems that the DL networks are able to successfully predict the overall areas in which ECAP values are high, but not so much the values of the outlier isolated nodes. In fact, if looked closely at the histogram on Figure 13, it is obvious at first glance that both DL predicted ECAP values have a stronger central tendency than the GT, which has its values much more evenly spreaded. This is supported by the slightly higher standard deviation $\delta_{GT} = 1.4982$, compared to the predictions with $\delta_{PCA} = 1.3019$ and $\delta_{SFC} = 1.2967$. Given this points, it seems that the developed DL networks are not capable of properly adjusting to extreme values, probably due to the lower frequency of occurrence on the training dataset.

Even though the accuracy results for both approaches may seem equal, the same cannot be said about the time performance of the neural network training. This

represents the most remarkable advantage of the PCA approach over its counterpart. Due to the implemented truncated PCA, the dimensions of the input and output space of the multilayer neural network performing the non-linear regression are much lower, going from the 7500 input units of the SFC model, to the comparatively insignificant 32 units of the PCA approach. This also extends to the hidden layers, while the difference is smaller, going from 5000 to 512 ReLU units respectively. Since the unsupervised learning method implemented on the latter is completed almost instantaneously, it represents enormous savings not only in time, but on CPU and GPU resources employed. This resulted in the PCA network being trained astonishingly fast, just over 30 seconds, while the SFC based model required near 25 minutes to complete its training.

Overall, both approaches seem to generate acceptable predictions of the ECAP mappings given the input LAA shapes. Nonetheless, since the same result is achieved with fewer computer resources and orders of magnitude faster, the dimensionality reduction network is thought to have an edge over the simple fully-connected network.

4.4 High ECAP values

Despite the fact this thesis lacks any clinical relevance and focuses on the feasibility of the concept rather than its clinical implications, studying the prediction capabilities of the networks in detecting the highest ECAP values was thought to be of interest. Constrained by the lack of literature and the absence of validated ECAP values on the LA, the study performed by Di Achille *et al.* [49] on thrombotic abdominal aortic aneurysms was taken as the reference for high ECAP values, with a value of the top 99th percentile of 4. This value was set as the threshold above which we consider a node, to have a higher risk of thrombi formation. Figure 15 makes it easier to visualize the accuracy of the DL models on predicting the highest ECAP values. In addition the posterior face can be seen on Appendix A.5. Green areas correspond to true positive nodes (TP), that is, nodes that have been correctly classified above an ECAP value of 4 by the DL networks, while blue and red areas represent false

negatives (FN) and false positives (FP) respectively.

The quantitative analysis performed through the confusion matrices seems to reinforce the idea that both predictive DL networks tend to underestimate the most extreme ECAP values, that is, the number of nodes that are the farthest from the mean value. As a matter of fact, it can be observed at first glance in Figures 15 and 23 that the number of FN nodes far exceeds that of FP nodes. As reported on Table 6 the FN nodes are 3 fold more numerous, which means that the number of nodes that the DL networks have wrongly predicted to be below the threshold is 3 times higher than the nodes that have been incorrectly assigned to be in thrombosis risk.

On the other hand, although the true positive rate is rather low at ~ 0.6 , it must be stated that almost all significantly large FP and FN areas are found surrounding the edges of TP patches colored in green. This means that while the number of wrongly classified nodes is not negligible, the networks manage to capture the approximated areas of high ECAP values and that no missclassifications occur outside of this regions. Therefore, the developed DL architectures can give insight into the general layout of the areas on risk of thrombus formation, but are unable to capture the full extent of the areas at risk, with a tendency to underestimate their size compared to the GT distributions.

4.5 Study limitations

As the goal of this study was to assess the feasibility of DNNs as a possible surrogate of CFD, several assumptions and simplifications were made to alleviate the workload.

First and foremost, employing an artificial ovoid representation of the LA completely neglects any possible clinical relevance that the predicted ECAP maps might have. Moreover, the same ovoid geometry was employed for all artificial patients without any inter-patient variation. However, these limitations can be easily overcome if a sufficiently big dataset of real LA patients is obtained. Similarly to the LAA, some kind of registration would be needed in order to establish the same nodal connectivity and number of nodes.

Secondly, while the SSM is based on a dataset of 103 patient geometries provided by the Technical University of Denmark (DTU), it still does not capture the full variability and the complexity observed on real LAA geometries. Therefore, at the moment, the neural networks would probably fail at predicting the ECAP maps of real LAA geometries. Nevertheless, this can be easily fixed in the future by constructing the SSM with a much bigger dataset of real patient geometries, encompassing the whole range of variability showcased by anatomical LAAs.

Another important drawback of the study, is the complete absence in the literature of validated ECAP maps on the LAA. Without it, no matter how the DL model is improved, the obtained predictions will never have any clinical relevance, because there is no way to validate neither the ground truth ECAP distributions from the CFD simulation, nor the predicted ECAP maps from the DNNs. In order to overcome this problem, technology such as 4D-Flow MRI imaging would need to further develop so that intracardiac hemodynamics could be characterized without the need for computational simulations.

It is worth noting that at first the thesis aimed to predict the flow velocity vector, but due to obvious time constraints imposed on the completion of a master thesis this idea was discarded in favour of the ECAP prediction. Since, the ECAP is calculated with the TAWSS, which is time averaged, it avoided getting involved in much more complex neural network architectures than the simple multilayer fully-connected networks presented in here.

4.6 Conclusions

In this study a successful DL surrogate of CFD was developed for the estimation of the ECAP parameter in the LAA, aiming to evaluate the potential thrombogenic risk. We believed this study to be one of the earliest implementations of DL as a substitute of finite element analysis, and the first to be successfully employed on analyzing the LAA. The developed DNNs managed to capture the non-linear function relating the geometry of the LAA and its ECAP distributions which resulted

in accurate predictions consistent with the ground truth values obtained through CFD simulations. Moreover, the task was completed order of magnitudes faster than any conventional CFD approach. Hence, with this study, the feasibility of DL approaches for the analysis of bodily fluids has been established, that could potentially enable quasi-instantaneous hemodynamic analysis in the future, opening up a wide range of innovative clinical applications.

Despite all of the work presented on the previous sections, there is still much to be done to improve the model. Furthermore, once the feasibility of the DL models as CFD surrogate has been proven, it could be used to characterize many other hemodynamic parameters.

First of all, it would be interesting to test the robustness of the DL networks to real-patient geometries. For this purpose, the non-volumetric registration performed by Slipsager *et al.* [39] would be needed to generate correspondence on the new geometries. Afterwards, it would be as simple feeding the space coordinates of the new geometry dataset to the DL networks to evaluate its performance.

On the other hand, data augmentation could be included on the training dataset of LAA shapes, by rotating the geometries employed on this thesis. Therefore, the atrial hemodynamics would not be only influenced by the LAA shape, but also by its orientation respect to the LA. A further step would be the variation of the morphological parameters of the employed LA ovoid, similarly to the study performed by Garcia-Isla *et al.* [28]. Different cavity volumes and PV and MV sizes could be included, and later, evaluate the performance of the DL model over this changes. Lastly, the final challenge would suppose the inclusion of realistic LA geometries and imposing correspondence in the whole LA and LAA assemble. The great complexity of the LA geometries, which present very heterogeneous morphology such as different numbers of PVs, could require the implementation of more elaborated DL architectures.

As a final objective for future studies, the prediction of the velocity flow field could be also considered. This would require a completely different pipeline and alternative

DL architectures such as physics-informed deep generative models [62], in order to predict the components of the velocity vector on the volumetric voxels. In addition, many different neural networks have been successfully employed on graphic simulations of complex turbulent flow, such as on smoke particles, based on convolutional neural networks that could be soon implemented in realistic physical simulations [63].

List of Figures

1	Main classification of the LAA morphology (from Di Biase <i>et al.</i> [12])	2
2	LAA occlusion devices (a) Watchman device (b) Amplatzer device [19]	4
3	General pipeline of the project. Following the enumeration, the geometries are generated from the SSM [1], assembled [2] and aligned [3]. Afterwards, the 3D volumetric mesh is generated [4] and CFD simulations are carried out [5] before calculating the ECAP maps [6]. Once the training data has been generated, the deep neural networks are trained (in red). As the final step the accuracy of the networks is tested with a set of new geometries (in green). LAA: Left Atrial Appendage, SSM: Statistical Shape Model, CFD: Computational Fluid Dynamics, ECAP: Endothelial Cell Activation Potential.	10
4	SSM generated with varying number of retained principal components (PC) A) $PC = 5$, B) $PC = 10$, C) $PC = 15$. In red a set of self-intersecting faces responsible for the divergence of the CFD simulations	14
5	Oval representation of the LA cavity. sRPV: Superior right pulmonary vein, iRPV: Inferior right pulmonary vein, sLPV: Superior left pulmonary vein, iLPV: Inferior left pulmonary vein, MV: Mitral valve.	15
6	ECAP maps for both real patient and SSM generated LA surfaces. On the left the LA cavity is supposed to be motionless, while on the right a dynamic mesh is implemented on the MV annulus.	16

7	Assembly procedure of the geometry. 1. LAA shape obtained from the SSM 2. Seal the ostium 3. Extrude twice the ostium cover 4. Assemble with the Oval 5. Separate the second extrude and perform a boolean union with the Oval 6. Combine back the oval and second extrude to the LAA and the first extrude, eliminate LAA 7. Load the LAA again 8. Seal and smooth the elements 9. Align the MV on the z-axis.	17
8	PVs inlet velocity profile throughout the whole cardiac cycle from systole 0 – 0.40s to diastole 0.41 – 1.05s. The red line indicates the transition between the two cardiac cycles.	20
9	70 nodes (in red) where eliminated on the outer ring of the LAA. The rest of the 2466 nodes (in blue) where retained for the deep learning analysis.	23
10	Network structure for the simple fully-connected feedforward network. no = Number of mesh nodes, U = Number of activation units.	25
11	Network structure for the dimensionality reduction DL model. no = Number of mesh nodes, U = Number of activation units. Figure adapted from Liang <i>et al.</i> [35].	26
12	Set of ECAP values on a random test dataset of 21 shapes, 7 of which are shown in here. Each row corresponds to a given geometry while the columns from left to right represent: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA	32
13	Total node count of the 21 test simulations in each range of ECAP values	33

- 14 ECAP difference computed as the subtraction between the ground-truth (GT) and the PCA based DL model on the left columns, and the GT and the simple fully-connected (SFC) model on the right columns. The front-face is shown in here. The posterior faces can be found on Figure 22 34

- 15 An ECAP value of 4 has been chosen as a threshold for potential thrombi formation risk. The performance of the DL model to detect such risk areas is evaluated by measuring the areas in which the $ECAP > 4$ on both the GT and the DL predictions (True Positive in green) intersect, areas that only appear to be risky on GT (False negative in blue) and areas that have been misdiagnosed by the DL model (False positive in red) 37

- 16 Set of ECAP values on a random test dataset of 21 shapes, the front of geometries H to N is shown in here. From left to right: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA 62

- 17 Set of ECAP values on a random test dataset of 21 shapes, the front of geometries O to U is shown in here. From left to right: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA 63

- 18 Set of ECAP values on a random test dataset of 21 shapes, the back of geometries A to G is shown in here. From left to right: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA 64

- 19 Set of ECAP values on a random test dataset of 21 shapes, the back of geometries H to N is shown in here. From left to right: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA 65
- 20 Set of ECAP values on a random test dataset of 21 shapes, the back of geometries O to U is shown in here. From left to right: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA 66
- 21 ECAP difference computed as the subtraction between the PCA based prediction and the SFC generated ECAP maps. 67
- 22 ECAP difference computed as the subtraction between the ground-truth (GT) and the PCA based DL model on the left columns, and the GT and the simple fully connected model on the right columns. The back-face is shown in here. 68
- 23 Back-faces corresponding to Figure 15. The performance of the DL model to detect areas in which the $ECAP > 4$ (thrombogenesis risk) is evaluated. True positive (TP) in green, where both the GT and the DL predictions intersect, False negatives (FN) in blue, False positives (FP) in red. 69

List of Tables

1	Ovoid LAA dimensions, VLA: Volume Left Atrium, VLAA: Volume Left Atrial Appendage, ALPV: Area Left Pulmonary Veins, ARPV: Area Right Pulmonary Veins, AMV: Area Mitral Valve	15
2	Performance of DL models on a 100 fold Monte Carlo cross-validation. MAE: Mean Absolute Error, RMSE: Root Mean Square Error, NMAE: Normalized Mean Absolute Error.	31
3	Absolute error (AE) error of peak ECAP values. Max: Maximum peak values, 99th: Upper 99th percentile, 90th: Upper 90th percentile.	31
4	Normalized absolute error (NAE) of peak ECAP values. Max: Maximum peakvalues, 99th: Upper 99th percentile, 90th: Upper 90th percentile.	31
5	Mean average error (MAE) for the first 7 LAA geometries in Figure 12	33
6	Confusion matrices for both SFC and PCA predicted ECAP values	36
7	Number of nodes in each range of ECAP values. The rows 1-3 represent the total number of nodes of the 21 test simulations in each range. The last two rows represent the ratio between the amount of nodes detected for the predictions and the ground truth.	71
8	Mean average error (MAE) for geometries H to N in Figure 16	71
9	Mean average error (MAE) for geometries H to N in Figure 17	71

Bibliography

- [1] Andrade, J., Khairy, P., Dobrev, D. & Nattel, S. The Clinical Profile and Pathophysiology of Atrial Fibrillation. *Circulation Research* **114**, 1453–1468 (2014).
- [2] Beigel, R., Wunderlich, N. C., Ho, S. Y., Arsanjani, R. & Siegel, R. J. The Left Atrial Appendage: Anatomy, Function, and Noninvasive Evaluation. *JACC: Cardiovascular Imaging* **7**, 1251–1265 (2014).
- [3] Go, A. S. *et al.* Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* **285**, 2370–5 (2001).
- [4] What is Atrial Fibrillation (AFib or AF)? | American Heart Association. URL <https://www.heart.org/en/health-topics/atrial-fibrillation/what-is-atrial-fibrillation-afib-or-af>.
- [5] Chugh, S. S. *et al.* Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation* **129**, 837–47 (2014).
- [6] Samol, A. *et al.* Prevalence of unknown atrial fibrillation in patients with risk factors. *EP Europace* **15**, 657–662 (2013).
- [7] Healey, J. S. *et al.* Subclinical Atrial Fibrillation and the Risk of Stroke. *New England Journal of Medicine* **366**, 120–129 (2012).

- [8] Lin, H.-J. *et al.* Stroke Severity in Atrial Fibrillation. *Stroke* **27**, 1760–1764 (1996).
- [9] Marini, C. *et al.* Contribution of Atrial Fibrillation to Incidence and Outcome of Ischemic Stroke. *Stroke* **36**, 1115–1119 (2005).
- [10] Chugh, S. S., Blackshear, J. L., Shen, W.-K., Hammill, S. C. & Gersh, B. J. Epidemiology and natural history of atrial fibrillation: clinical implications. *Journal of the American College of Cardiology* **37**, 371–378 (2001).
- [11] Al-Issa, A. *et al.* Regional function analysis of left atrial appendage using motion estimation CT and risk of stroke in patients with atrial fibrillation. *European heart journal cardiovascular Imaging* **17**, 788–96 (2016).
- [12] Di Biase, L. *et al.* Does the Left Atrial Appendage Morphology Correlate With the Risk of Stroke in Patients With Atrial Fibrillation? *Journal of the American College of Cardiology* **60**, 531–538 (2012).
- [13] Tabata, T. *et al.* Role of left atrial appendage in left atrial reservoir function as evaluated by left atrial appendage clamping during cardiac surgery. *The American journal of cardiology* **81**, 327–32 (1998).
- [14] Chanda, A. & Reilly, J. P. Left Atrial Appendage Occlusion for Stroke Prevention. *Progress in Cardiovascular Diseases* **59**, 626–635 (2017).
- [15] Al-Saady, N. M., Obel, O. A. & Camm, A. J. Left atrial appendage: structure, function, and role in thromboembolism. *Heart (British Cardiac Society)* **82**, 547–54 (1999).
- [16] Lupercio, F. *et al.* Left atrial appendage morphology assessment for risk stratification of embolic stroke in patients with atrial fibrillation: A meta-analysis. *Heart Rhythm* **13**, 1402–1409 (2016).
- [17] Wu, L. *et al.* Relation of Left Atrial Appendage Morphology Determined by Computed Tomography to Prior Stroke or to Increased Risk of Stroke in Patients With Atrial Fibrillation. *The American Journal of Cardiology* (2019).

- [18] Watson, T., Shantsila, E. & Lip, G. Y. Mechanisms of thrombogenesis in atrial fibrillation: Virchow's triad revisited. *The Lancet* **373**, 155–166 (2009).
- [19] Cruz-Gonzalez, I. *et al.* Left atrial appendage closure devices for cardiovascular risk reduction in atrial fibrillation patients. *Research Reports in Clinical Cardiology* **6**, 47 (2015).
- [20] Alli, O. & Holmes, D. R. Left Atrial Appendage Occlusion for Stroke Prevention. *Current Problems in Cardiology* **37**, 405–441 (2012).
- [21] Fender, E. A., Kiani, J. G. & Holmes, D. R. Left Atrial Appendage Closure for Stroke Prevention in Atrial Fibrillation. *Current Atherosclerosis Reports* **18**, 65 (2016).
- [22] Singh, I. M. & Holmes, D. R. Left Atrial Appendage Closure. *Current Cardiology Reports* **12**, 413–421 (2010).
- [23] Fauchier, L. *et al.* Device-related thrombosis after percutaneous left atrial appendage occlusion for atrial fibrillation. *Journal of the American College of Cardiology* **71**, 1528 – 1536 (2018).
- [24] Lempereur, M. *et al.* Device-associated thrombus formation after left atrial appendage occlusion: A systematic review of events reported with the Watchman, the Amplatzer Cardiac Plug and the Amulet. *Catheterization and Cardiovascular Interventions* **90**, E111–E121 (2017).
- [25] Markl, M., Frydrychowicz, A., Kozerke, S., Hope, M. & Wieben, O. 4D flow MRI. *Journal of Magnetic Resonance Imaging* **36**, 1015–1036 (2012).
- [26] Fraser, K. H., Taskin, M. E., Griffith, B. P. & Wu, Z. J. The use of computational fluid dynamics in the development of ventricular assist devices. *Medical Engineering & Physics* **33**, 263–280 (2011).
- [27] Otani, T. *et al.* A Computational Framework for Personalized Blood Flow Analysis in the Human Left Atrium. *Annals of Biomedical Engineering* **44**, 3284–3294 (2016).

- [28] García-Isla, G. *et al.* Sensitivity analysis of geometrical parameters to study haemodynamics and thrombus formation in the left atrial appendage. *International Journal for Numerical Methods in Biomedical Engineering* **34**, e3100 (2018).
- [29] Bosi, G. M. *et al.* Computational Fluid Dynamic Analysis of the Left Atrial Appendage to Predict Thrombosis Risk. *Frontiers in cardiovascular medicine* **5**, 34 (2018).
- [30] Masci, A. *et al.* The Impact of Left Atrium Appendage Morphology on Stroke Risk Assessment in Atrial Fibrillation: A Computational Fluid Dynamics Study. *Frontiers in Physiology* **9**, 1938 (2019).
- [31] Aguado, A. M. *et al.* In silico Optimization of Left Atrial Appendage Occluder Implantation Using Interactive and Modeling Tools. *Frontiers in Physiology* **10**, 237 (2019).
- [32] Guo, X., Li, W. & Iorio, F. Convolutional Neural Networks for Steady Flow Approximation. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, 481–490 (ACM Press, New York, New York, USA, 2016).
- [33] Hennigh, O. Lat-Net: Compressing Lattice Boltzmann Flow Simulations using Deep Neural Networks (2017). 1705.09036.
- [34] Liang, L., Liu, M., Martin, C., Elefteriades, J. A. & Sun, W. A machine learning approach to investigate the relationship between shape features and numerically predicted risk of ascending aortic aneurysm. *Biomechanics and Modeling in Mechanobiology* **16**, 1519–1533 (2017).
- [35] Liang, L., Liu, M., Martin, C. & Sun, W. A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis. *Journal of The Royal Society Interface* **15**, 20170844 (2018).

- [36] Brinker, T. J. *et al.* Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *European Journal of Cancer* **113**, 47–54 (2019).
- [37] Buetti-Dinh, A. *et al.* Deep neural networks outperform human expert’s capacity in characterizing bioleaching bacterial biofilm composition. *Biotechnology Reports* **22**, e00321 (2019).
- [38] Kutz, J. N. Deep learning in fluid dynamics. *Journal of Fluid Mechanics* **814**, 1–4 (2017).
- [39] Slipsager, J. M. *et al.* Statistical Shape Clustering of Left Atrial Appendages. In *Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges*, 32–39 (Springer International Publishing, 2019). URL http://link.springer.com/10.1007/978-3-030-12029-0_{_}4.
- [40] Meshmixer. URL <http://www.meshmixer.com/>.
- [41] Meshlab. URL <http://www.meshlab.net/>.
- [42] Free Student Software | ANSYS Student. URL <https://www.ansys.com/academic/free-student-products>.
- [43] Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. URL <http://gmsh.info/>.
- [44] MATLAB. URL <https://es.mathworks.com/products/matlab.html>.
- [45] ANSYS Fluent Software | CFD Simulation. URL <https://www.ansys.com/products/fluids/ansys-fluent>.
- [46] ParaView. URL <https://www.paraview.org/>.
- [47] 10.6.2.Dynamic Mesh Update Methods. URL https://www.sharcnet.ca/Software/Ansys/16.2.3/en-us/help/flu_ug/flu_ug_dynam_mesh_update.html.

- [48] Mill, J. Optimal boundary conditions in fluid simulations for predicting occlude related thrombus formation in the left atria. In *Computational and Mathematical Biomedical Engineering, Sixth International Conference* (2019).
- [49] Di Achille, P., Tellides, G., Figueroa, C. A. & Humphrey, J. D. A haemodynamic predictor of intraluminal thrombus formation in abdominal aortic aneurysms. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **470**, 20140163–20140163 (2014).
- [50] Fernández-Pérez, G., Duarte, R., Corral de la Calle, M., Calatayud, J. & Sánchez González, J. Analysis of left ventricular diastolic function using magnetic resonance imaging. *Radiología (English Edition)* **54**, 295–305 (2012).
- [51] Veronesi, F. *et al.* Quantification of Mitral Apparatus Dynamics in Functional and Ischemic Mitral Regurgitation Using Real-time 3-Dimensional Echocardiography. *Journal of the American Society of Echocardiography* **21**, 347–354 (2008).
- [52] Keras. URL <https://keras.io/>.
- [53] TensorFlow. URL <https://www.tensorflow.org/>.
- [54] Anaconda Python. URL <https://www.anaconda.com/>.
- [55] CUDA Toolkit | NVIDIA Developer. URL <https://developer.nvidia.com/cuda-toolkit>.
- [56] cuDNN | NVIDIA Developer. URL <https://developer.nvidia.com/rdp/cudnn-download>.
- [57] Heaton, J. *Introduction to Neural Networks for Java, 2Nd Edition* (Heaton Research, Inc., 2008), 2nd edn.
- [58] Muknahallipatna, S. & Chowdhury, B. Input dimension reduction in neural network training-case study in transient stability assessment of large systems. In *Proceedings of International Conference on Intelligent System Application to Power Systems*, 50–54 (IEEE, 2002).

-
- [59] Wiesler, S. & Ney, H. A Convergence Analysis of Log-Linear Training. *Advances in Neural Information Processing Systems* 1–9 (2011).
- [60] Jolliffe, I. T. & Cadima, J. Principal component analysis: a review and recent developments. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* **374**, 20150202 (2016).
- [61] Ruder, S. An overview of gradient descent optimization algorithms. *CoRR abs/1609.04747* (2016). 1609.04747.
- [62] Yang, Y. & Perdikaris, P. Physics-informed deep generative models (2018). 1812.03511.
- [63] Kim, B. *et al.* Deep Fluids: A Generative Network for Parameterized Fluid Simulations. *Computer Graphics Forum* **38**, 59–70 (2019).

Appendix A

First Appendix

A.1 Prediction figures - Front

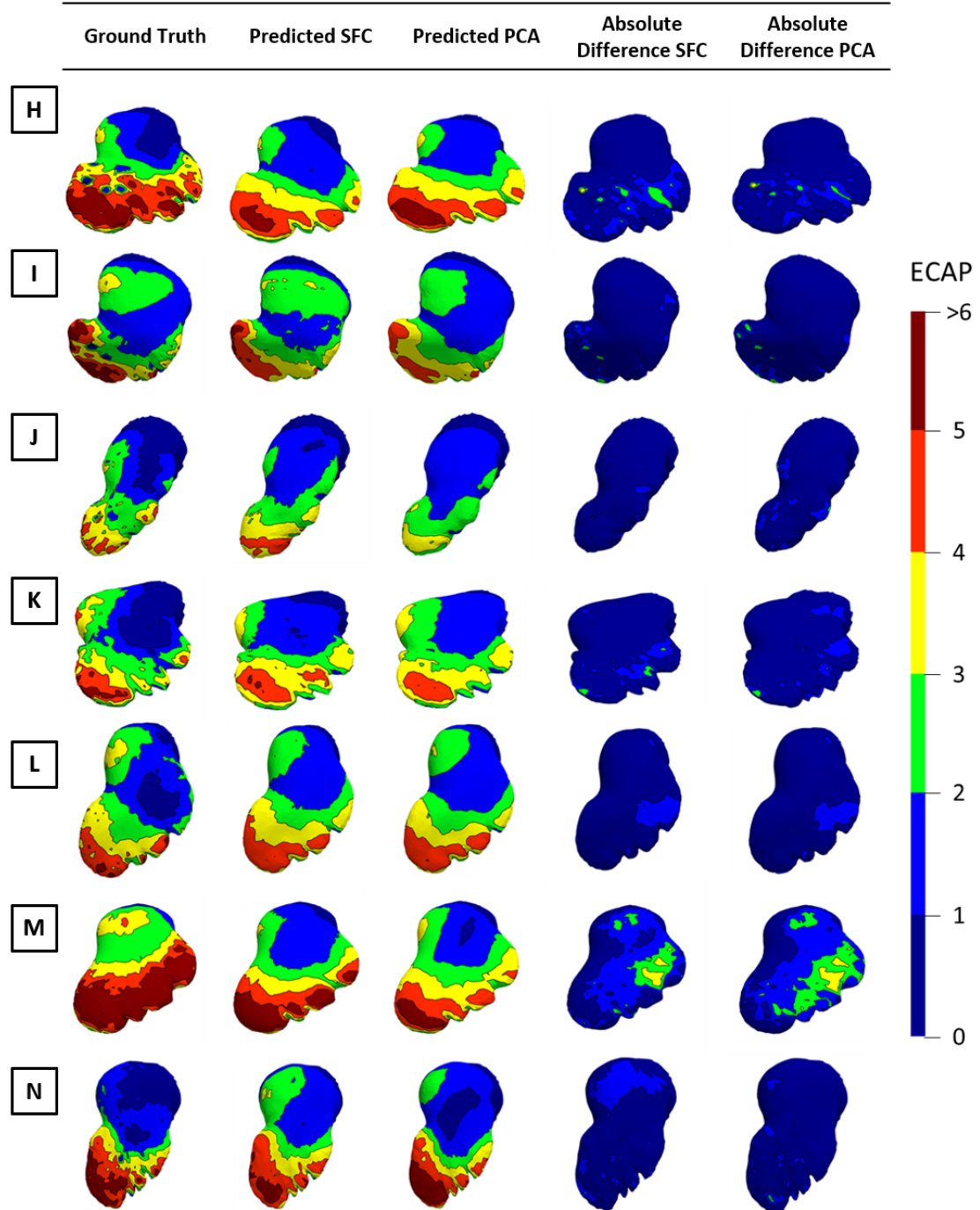


Figure 16: Set of ECAP values on a random test dataset of 21 shapes, the front of geometries H to N is shown in here. From left to right: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA

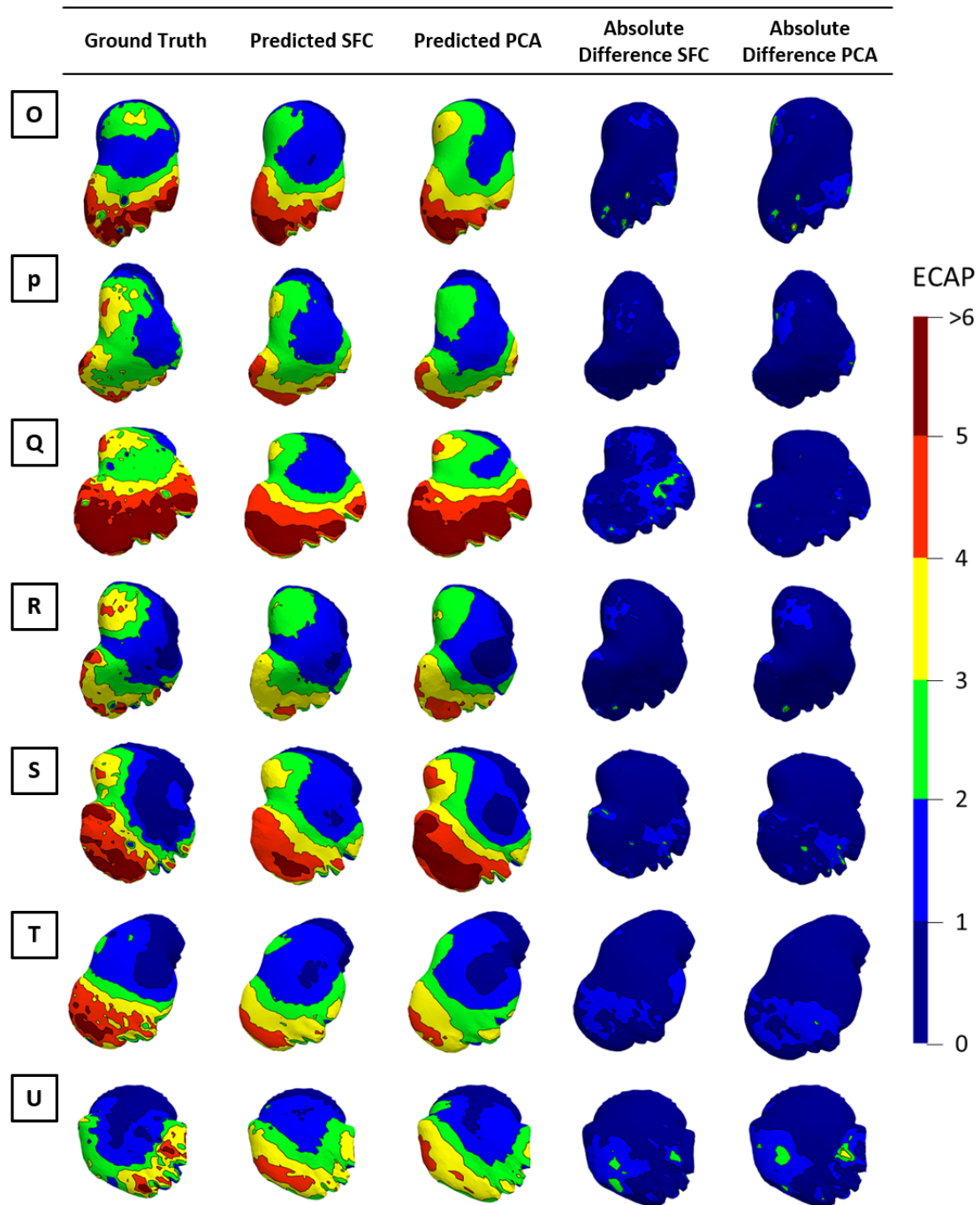


Figure 17: Set of ECAP values on a random test dataset of 21 shapes, the front of geometries O to U is shown in here. From left to right: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA

A.2 Prediction figures - Back

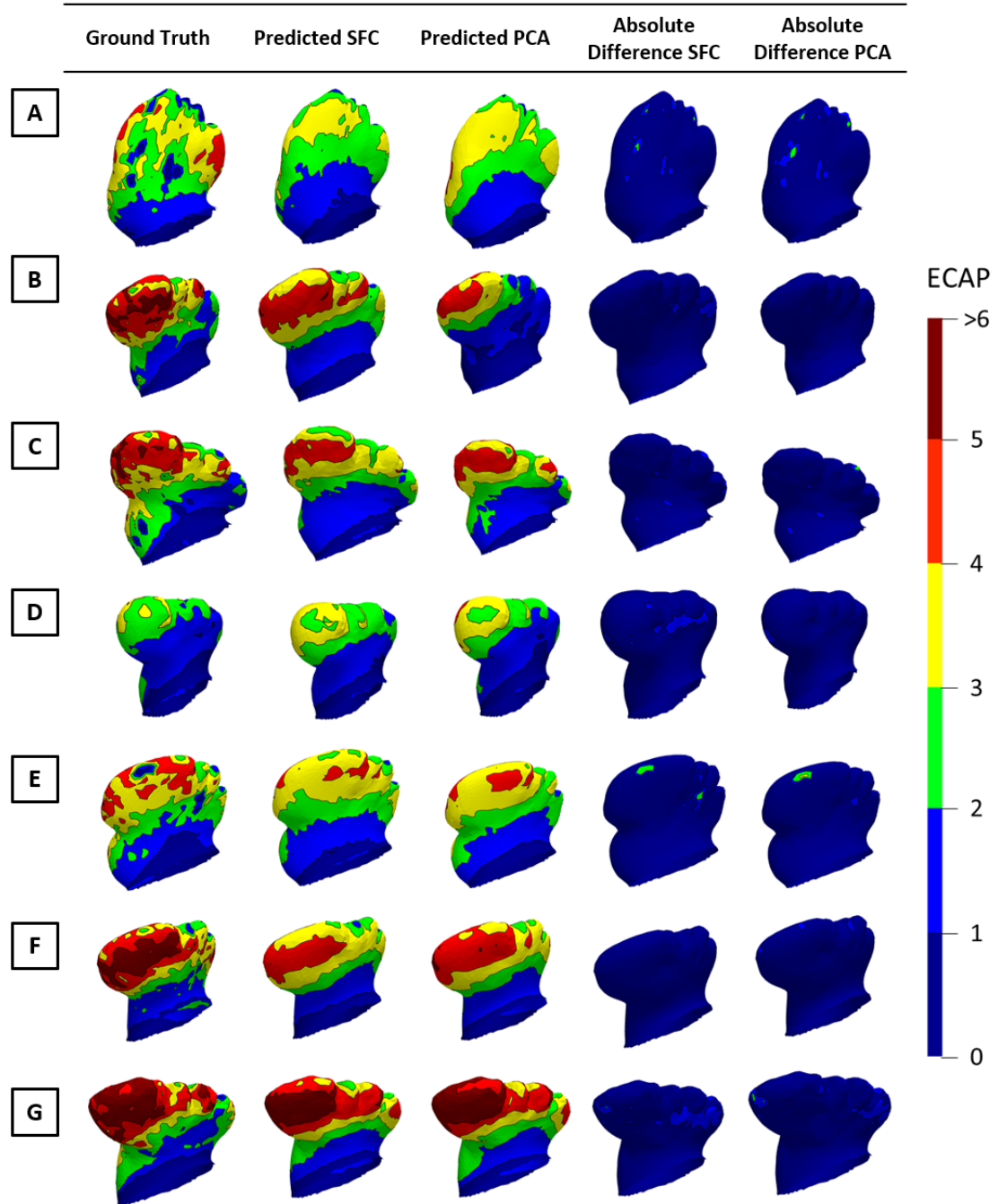


Figure 18: Set of ECAP values on a random test dataset of 21 shapes, the back of geometries A to G is shown in here. From left to right: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA

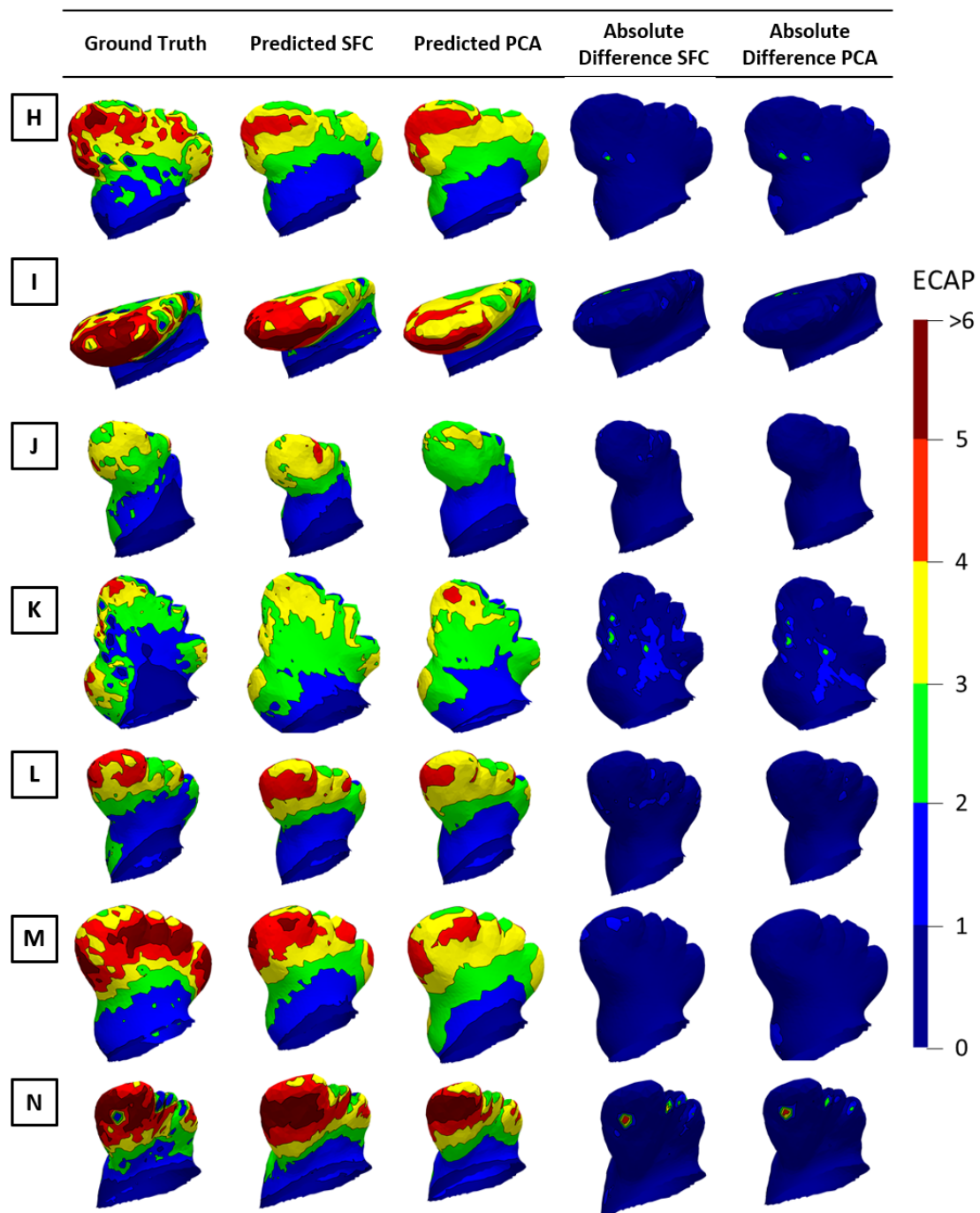


Figure 19: Set of ECAP values on a random test dataset of 21 shapes, the back of geometries H to N is shown in here. From left to right: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA

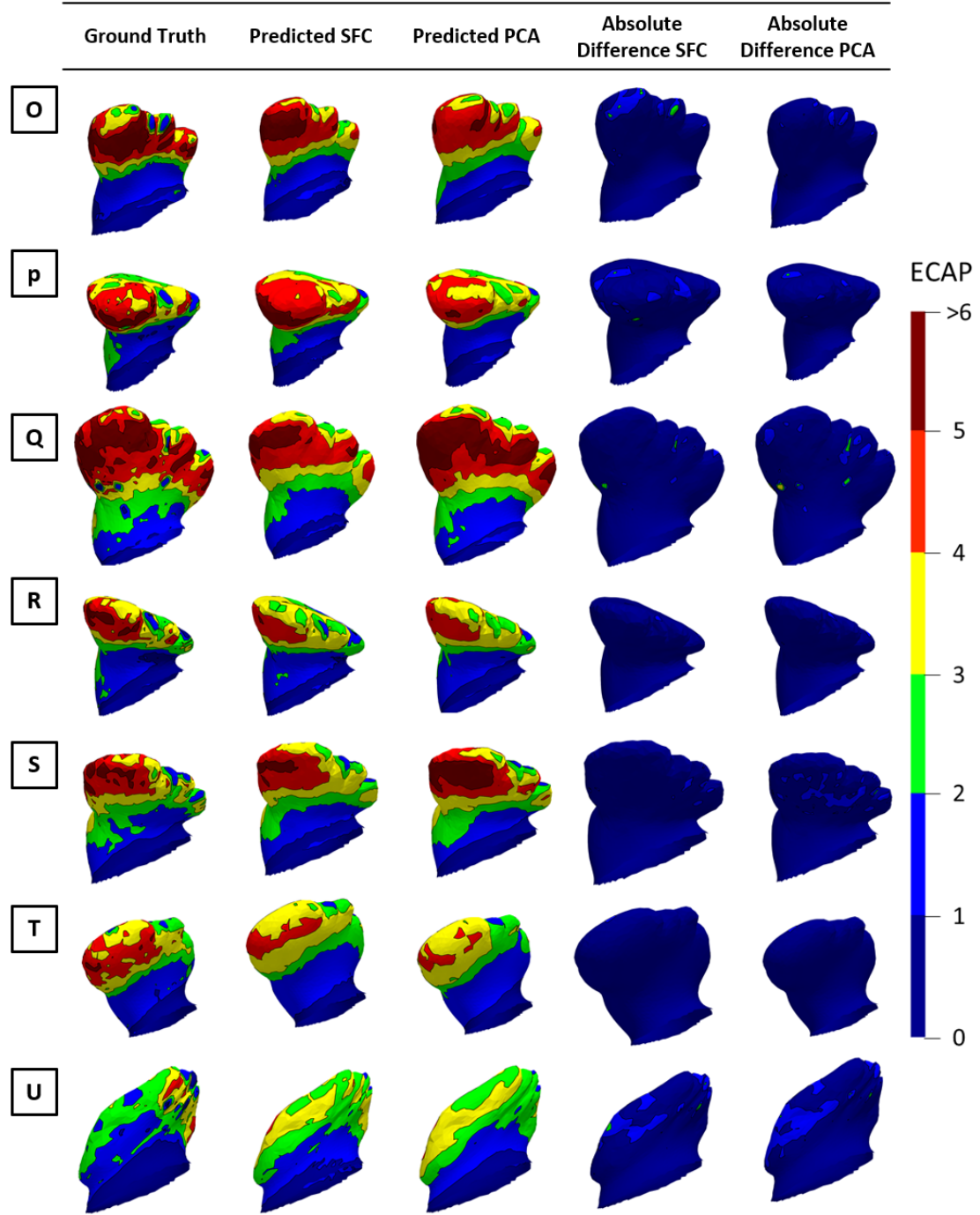


Figure 20: Set of ECAP values on a random test dataset of 21 shapes, the back of geometries O to U is shown in here. From left to right: Ground truth (GT) obtained from CFD simulations, predicted ECAP from the Simple Fully connected (SFC) network, prediction of the PCA network, difference between GT - SFC, difference between GT - PCA

A.3 SFC - PCA difference

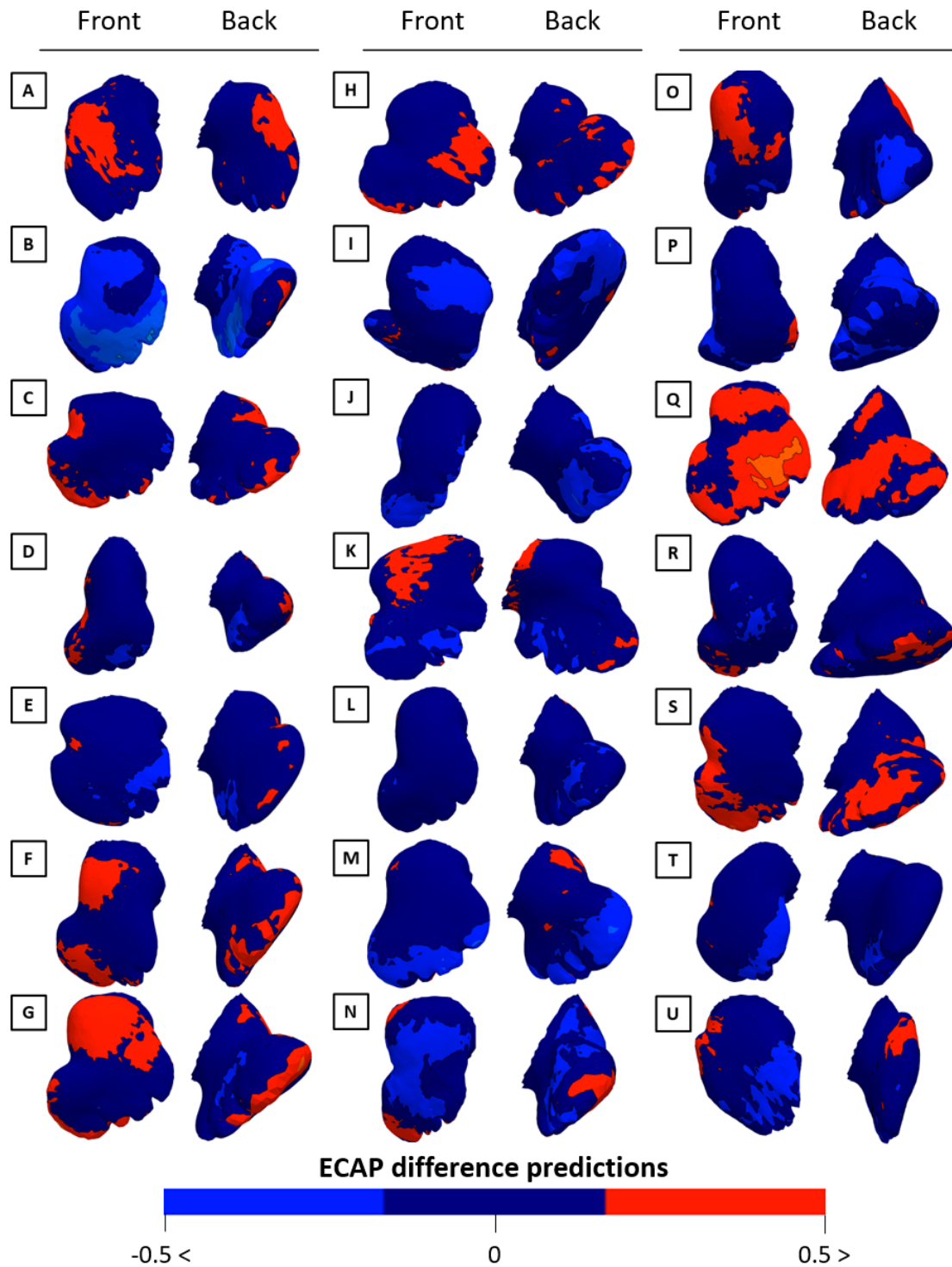


Figure 21: ECAP difference computed as the subtraction between the PCA based prediction and the SFC generated ECAP maps.

A.4 GT - Prediction difference back-face

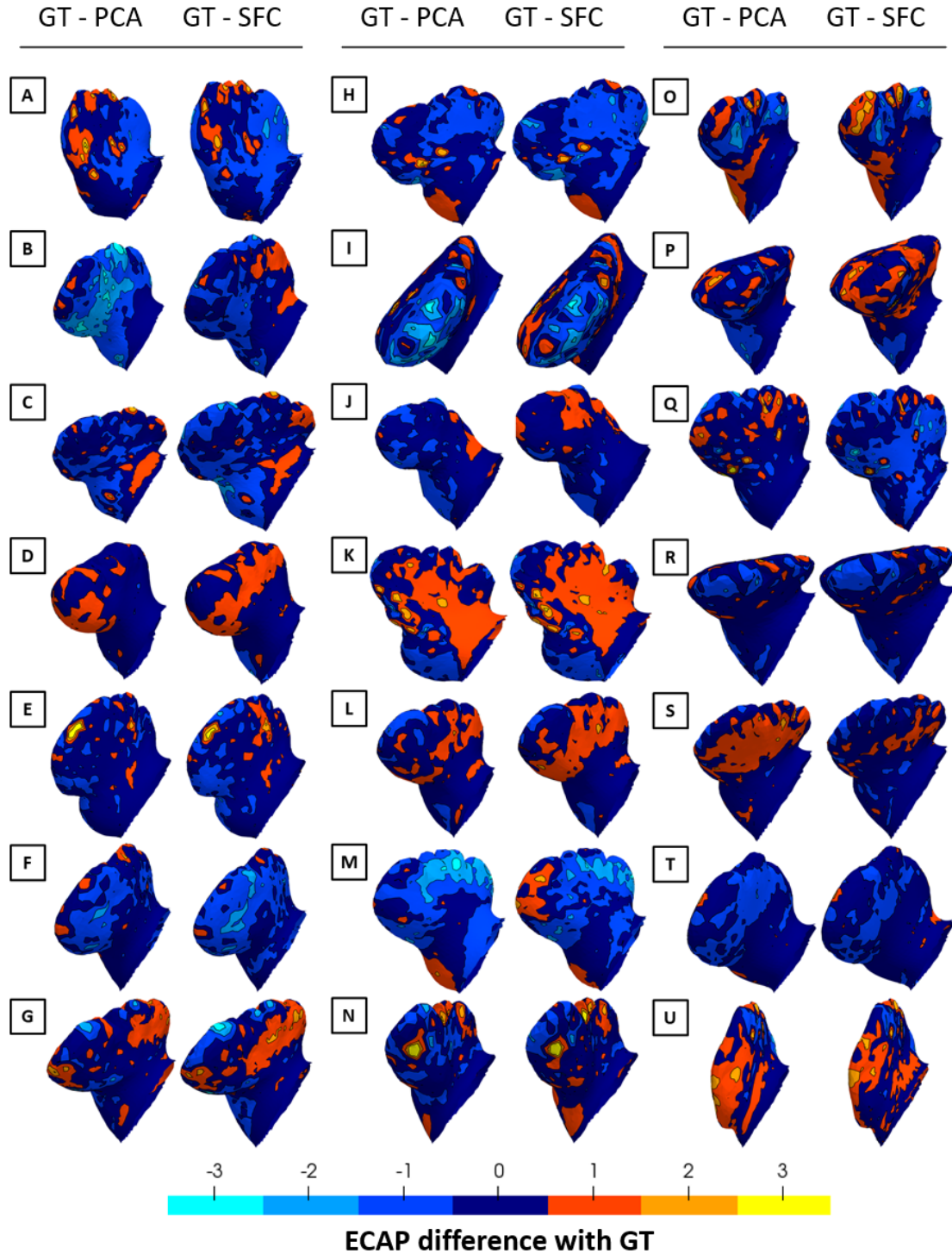


Figure 22: ECAP difference computed as the subtraction between the ground-truth (GT) and the PCA based DL model on the left columns, and the GT and the simple fully connected model on the right columns. The back-face is shown in here.

A.5 FP back-face

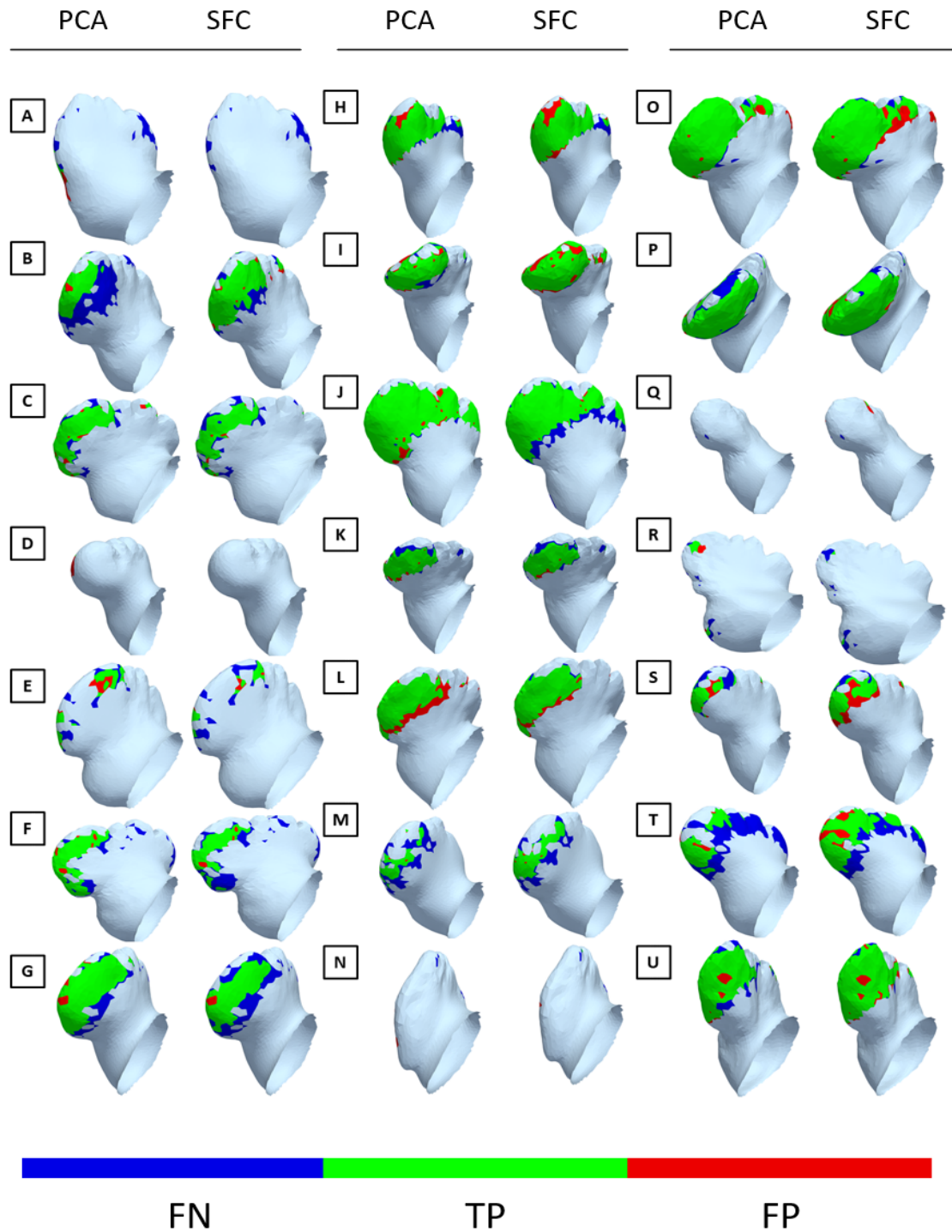


Figure 23: Back-faces corresponding to Figure 15. The performance of the DL model to detect areas in which the $ECAP > 4$ (thrombogenesis risk) is evaluated. True positive (TP) in green, where both the GT and the DL predictions intersect, False negatives (FN) in blue, False positives (FP) in red.

Appendix B

Second Appendix

B.1 Total amount of nodes in each ECAP range

Table 7: Number of nodes in each range of ECAP values. The rows 1-3 represent the total number of nodes of the 21 test simulations in each range. The last two rows represent the ratio between the amount of nodes detected for the predictions and the ground truth.

Networks	0-1	1-2	2-3	3-4	4-5	5-6	> 6
GT	11153	13943	10873	7583	5083	1380	762
SFC	9062	16978	11385	8281	4721	1237	122
PCA	9380	16642	11902	7965	4226	1380	291
SFC/GT	0.8125	1.2177	1.0471	1.0920	0.9288	0.5178	0.1601
PCA/GT	0.8410	1.1936	1.0946	1.0504	0.8314	0.5776	0.3819

B.2 Mean Absolute Error of simulations H-U

Table 8: Mean average error (MAE) for geometries H to N in Figure 16

Networks	H	I	J	K	L	M	N
PCA MAE	0.6229	0.5455	0.5656	0.6790	0.4823	1.0496	0.5453
SFC MAE	0.6965	0.5513	0.5161	0.7204	0.5314	0.9299	0.6262

Table 9: Mean average error (MAE) for geometries H to N in Figure 17

Networks	O	P	Q	R	S	T	U
PCA MAE	0.6814	0.5793	0.4868	0.4503	0.5833	0.5386	0.6877
SFC MAE	0.6141	0.5107	0.8221	0.4644	0.5169	0.5338	0.6386