

Automatic calculation of the local aortic compliance from thoracic cine-MRI

Emmanouil MARKODIMITRAKIS

Supervisor: Alain LALANDE

International Department of Computer Vision
University of Burgundy



A Thesis Submitted for the Degree of MSc in Vision and Robotics (VIBOT)

· 2022 ·

Abstract

Studying the elastic properties of the aorta is suitable for patients with dilated aorta. One measurement which quantifies the aortic elastic properties is aortic compliance. This property is defined as the relative change in the aortic cross-sectional area divided by the change in arterial pressure. This value decreases with age and some specific diseases and could predict upcoming acute events. Magnetic resonance imaging (MRI) is a valuable imaging modality for the noninvasive evaluation of aortic compliance. However, to obtain this measurement, the contour of the aorta is mandatory on a series of images covering the cardiac cycle. Manual tracing of these contours is subject to critical inter-observer variations. The composition of the thoracic aortic wall is heterogeneous, and it is well known that the elastic properties of the aorta vary according to the considered wall. The main objective of this work is to assess the elastic properties of four different quadrants of the aorta and compare the results with the elasticity measurement via Young's modulus. Young's modulus describes the stiffness properties of the aortic wall and is obtained ex-vivo on aortic samples harvested during surgery.

The database consists of 73 cine-MRI sequences of thoracic aorta from patients acquired in axial orientation at the level of the pulmonary trunk, of which 58 have a dilated aorta (with aortic aneurysm) and 15 without aortic dilation. The exams are acquired in the University Hospital of Dijon (France), each consisting of 30 slices on average across the cardiac cycle. Multiple deep learning architectures have been explored with different hyperparameters and settings to automatically segment the contour of the aorta on each image and automatically calculate the aortic compliance.

A semantic segmentation U-Net network outperforms the rest explored networks with a Dice score of 98.09% ($\pm 0.96\%$) and a Hausdorff distance of 4.88 mm (± 1.70). This network allows the automatic segmentation of the aorta on cine MRI sequence. A graphical user interface has been designed to assist the division in 4 quadrants, which is also implemented automatically and allows the calculation of local aortic compliance and strain.

Local aortic compliance, and aortic wall strain, are calculated from the segmented surfaces for each quadrant during the cardiac cycle. For each quadrant, the aforementioned elastic properties are compared with the values of Young's modulus. A significantly high correlation was observed between Young's modulus and in-vivo strain. Our results suggest that the lateral and posterior quadrants are the stiffest. In contrast, the medial and anterior quadrants have the lowest aortic stiffness. A contribution of this work is the GUI that enables medicals to calculate the local compliance and strain per quadrant. We conclude that our automatic segmentation method is robust and compatible with clinical practice.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 1.1 | Problem definition | 1 |
| 1.2 | Objectives | 2 |
| 1.3 | Thesis Structure | 3 |
| 2 | Clinical Background | 4 |
| 2.1 | Aorta Anatomy | 4 |
| 2.2 | Aortic Aneurysm & Dissection | 5 |
| 2.3 | Magnetic Resonance Imaging (MRI) | 7 |
| 2.4 | Aortic walls & Elastic properties of aorta | 9 |
| 3 | State of the art | 11 |
| 3.1 | Contour segmentation of aorta | 11 |
| 3.2 | Deep Learning | 11 |
| 3.3 | Fully Convolutional Networks & U-Net | 13 |
| 3.4 | Deep learning based segmentation | 15 |
| 3.4.1 | 2D methods | 15 |
| 3.4.2 | 2D plus time methods | 15 |
| 3.4.3 | Other methods (2.5D, 2D+3D, 3D) | 16 |

| | |
|---|-----------|
| 3.5 Comparison of two techniques (in-vivo and ex-vivo) for evaluating the elastic properties of the ascending aorta | 17 |
| 4 Methodology | 19 |
| 4.1 Dataset | 19 |
| 4.2 Preprocessing | 22 |
| 4.2.1 Minimun-Maximum Normalization | 22 |
| 4.2.2 Contrast stretching | 23 |
| 4.2.3 Central crop and pad | 23 |
| 4.3 Data Augmentation | 24 |
| 4.4 Model Training | 24 |
| 4.5 Post-processing | 27 |
| 4.6 Metrics | 28 |
| 4.6.1 Confusion Matrix | 28 |
| 4.6.2 Precision & Recall | 28 |
| 4.6.3 Intersection over Union | 29 |
| 4.6.4 Dice coefficient | 29 |
| 4.6.5 Hausdorff distance | 29 |
| 5 Results | 31 |
| 5.1 Deep Learning evaluation | 31 |
| 5.2 Aortic elasticity based evaluation | 32 |
| 5.2.1 Compliance | 32 |
| 5.2.2 Strain | 37 |
| 6 Conclusion | 40 |
| 6.1 Discussion | 40 |
| 6.2 Future Work | 42 |
| Bibliography | 51 |

List of Figures

| | | |
|-----|---|----|
| 2.1 | Anatomy of the aorta [5, 26] | 5 |
| 2.2 | Axial MRI at the level of pulmonary trunk. Modified image from our dataset. | 7 |
| 3.1 | Activation in details. Modified from [48] | 12 |
| 3.2 | Neural Networks diagram. Modified from [12] | 12 |
| 3.3 | U-Net architecture [53] Every blue box is a multichannel feature map, the number of the channel is denoted on the top of each box. The size information is provided at the lower left edge of each box. White boxes represent the feature maps that are copied thought the skip connections, and the arrows represent the different operations. | 14 |
| 4.1 | Workflow of the research strategy. Modified images from our dataset. | 20 |
| 4.2 | Examples of MR images with neither noise nor artifacts. [41] | 21 |
| 4.3 | Examples of MR images with artifact due to rapid or turbulent blood flow. [41] | 22 |
| 4.4 | Examples of MR images where (a) the aortic wall is not well-defined or (b) the aorta is highly dilated. [41] | 22 |
| 4.5 | Contrast stretching. Modified images from our dataset. | 23 |
| 4.6 | Crop and pad. Modified images from our dataset. | 24 |
| 4.7 | Hausdorff distance between the green line X and the blue line Y. [51] | 30 |

| | | |
|-----|--|----|
| 5.1 | Examples of automatic segmentations, a dilated aorta is presented on top images and a non-dilated in bottom ones. On the left, all the aortic surface is colored, in contrast to the right, where only the aortic wall is colored. | 33 |
| 5.2 | Comparison of non-rotated and rotated quadrants, with local compliance curves. | 34 |
| 5.3 | Example of automatic segmentation of dilated aorta divided by quadrants (c), accompanied by both the global (a) and local (b) compliance curves. Local compliance is presented with 4 different colors for each quadrant. | 34 |
| 5.4 | Example of automatic segmentation of non-dilated aorta divided by quadrants (c), accompanied by both the global (a) and local (b) compliance curves. Local compliance is presented with 4 different colors for each quadrant. | 35 |
| 5.5 | Bland Altman Plot of compliance for dilated (a) and non-dilated (b) aorta. . . . | 35 |
| 5.6 | Bland Altman Plot of Minimum surface area for dilated (a) and non-dilated (b) aorta. | 36 |
| 5.7 | Bland Altman Plot of Maximum surface area for dilated (a) and non-dilated (b) aorta. | 37 |
| 5.8 | Visualization of perimeter per quadrant with different colors. | 38 |
| 6.1 | Inaccurate prediction using noise augmented dataset. | 43 |
| 6.2 | A preview of the noise & wave augmented dataset. Noise is added gradually from top to bottom, while wave shaped artifacts are added in different opacity levels from left to right. | 43 |
| 6.3 | A visual comparison of U-Net predictions of two cases trained with non-augmented dataset on the left and augmented dataset on the right. | 44 |

List of Tables

| | | |
|-----|---|----|
| 4.1 | Confusion matrix | 28 |
| 5.1 | Hyperparameters & settings | 31 |
| 5.2 | Metric based evaluation of U-Net architectures | 32 |
| 5.3 | Correlation between ground truth and calculated values of min-max areas and compliance. The table consists of two different multi-cases datasets of dilated and non-dilated aortas. | 33 |
| 5.4 | Average compliance and standard deviation for each one of the four quadrants of dilated and non-dilated aorta datasets. | 36 |
| 5.5 | Average compliance of dilated aortas per quadrant, before and after discarding outliers images containing noise. A medical expert has implemented the discarding process. | 36 |
| 5.6 | Maximum Young's Modulus for different quadrants from the work of Siyu et al. [35], the highest the value, the highest the stiffness of aortic wall. | 37 |
| 5.7 | Comparison of perimeter's strain between non-dilated, and dilated aorta datasets. | 38 |
| 5.8 | Comparison of strain and Young's Modulus between our results and the results from the work of Siyu et al. [35]. We use only 22 cases from the ex-vivo dataset for this table, where the patients are common with our dilated dataset. | 39 |

| | | |
|-----|---|----|
| 6.1 | Comparison of elastic properties between our results and the results from the work of Siyu et al. [35]. We use only 22 cases from the ex-vivo dataset for this table, where the patients are common with our dilated dataset. | 41 |
| 6.2 | Metric based evaluation of noise augmented and non-augmented U-Net models. . | 42 |

Chapter 1

Introduction

There is an addressable amount of imaging data among the medical institutes, where researchers can leverage the effectiveness of medical treatment, prognosis, and decision-making. Most of these medical images are multi-modal, like Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and more.

Medical image analysis is a rapidly growing field of study and research due to the recent improvements in computer technology and the availability of open-source data. Heavily, expensive image processing tasks like segmentation and classification are easier and faster to implement with significant accurate results. As the hardware becomes more computational effective, the software becomes less time-consuming, allowing researchers and scientists to improve and evolve in every field. More specifically, advances in Graphics Processing Units (GPUs) have a specific important part as it concerns the field of Artificial Intelligence (AI) and Deep Learning (DL).

1.1 Problem definition

Many aortic wall pathologies (Marfan's syndrome, Elher-Danlos syndrome, Loeys-Detz syndrome) can be studied based on the evaluation of aortic elasticity, and distensibility [34, 46]. This study focuses on the pathology named Ascending Aortic Aneurysm (AAA), which leads the ascending aorta to dilate and lose its elastic properties. In most cases, dilation leads to dissection or rupture of the aortic artery and causes internal bleeding, a highly deadly condition that needs immediate surgery.

The most crucial stage of AAA is the prognosis. Using medical image analysis, doctors can determine if there is a need for medical treatment or even surgery to prevent the aortic wall

from tearing apart. This clinical decision is based on the size of the aneurysm [15] and the elastic properties of the ascending aorta. There are two potential ways to assess the parameters of the elasticity. Either by measuring Pulse Wave Velocity [34] which describes the overall (or global) stiffness of the aorta. Either by studying the aortic compliance, which describes the dynamic expansion of the aortic wall during systole [6].

This study concentrates on the second approach, assessing the aortic compliance, a measurement of the aortic stiffness, defined as the cross-sectional area variation during a cardiac cycle divided by the pulse pressure variation. To calculate the cross-sectional area from two-dimensional (2D) axial cine-MRI acquisitions, we need to localize and segment the ascending aorta at the level of the pulmonary trunk. The aortic wall certainly involves differences in the behavior of stiffness depending on the localization [35] meaning that taking into consideration four different quadrants, each one of them should result in different compliance. Local measurement of elasticity from ex-vivo experiments [35, 42, 43] named Maximum Young's Modulus (MYM) is compared with local compliance of each quadrant to analyze the possible correlation between them. Ex-vivo experiments referred to aortic samples retrieved during surgery and tested in a bi-axial stretching test to measure their elastic properties. The purpose of this comparison is to investigate if our proposed non-invasive method for the calculation of aortic elasticity is reliable.

Manual segmentation of the aorta introduces intra-observer variability and bias error, and it is a high time-consuming task. To deal with the problems mentioned above, automatization of the segmentation process is necessary.

One of the primary objectives of this study is to provide a robust automated technique for two-dimensional ascending aorta segmentation from cine-MRI at the level of the pulmonary trunk using deep learning image processing methods. Deep learning is state-of-the-art for medical image analysis, used for registration, segmentation, and classification. Regrettably, there is a lack of publications on 2D aorta segmentation employing DL to evaluate aortic elasticity. In this study, we discover recent publications on this topic and evaluate state-of-the-art DL architectures used to segment the thoracic axial aorta.

1.2 Objectives

This research aims to propose a robust and stable DL method for the segmentation of ascending aorta from cine-MRI in axial orientation at the level of the pulmonary trunk.

- Firstly, multiple architectures are trained on the clinical data to compare them in the

next steps and select the most appropriate one.

- In order to study the variability of elasticity and compliance between the different wall segments, we partition the segmented aorta from the previous step into four quadrants relative to medial (MED), posterior (POST), lateral (LAT), and anterior (ANT) according to the aortic wall.
- The evaluation of those architectures is divided into two parts.
 - The first one is the evaluation between aorta segmentations (predicted results of the DL model) and ground truth provided by physicians (with the semi-automated software QIR [41]), using metrics like Dice coefficient and Hausdorff distance, Precision, and Recall.
 - The next part of the evaluation is based on the elasticity properties of the aortic wall, and the comparison of our results with ex-vivo results, recommended by [35, 42, 43].
- All of the above objectives should be incorporated into a user-friendly graphical user interface (GUI) so that a clinician can perform real-world experiments.

1.3 Thesis Structure

The dissertation is organized as follows: An overview of the clinical and technical background is first presented in Section 2, followed by the state-of-the-art implementations in the field. The core of this thesis is Chapter 3, which explains the proposed methodology, including the training and evaluation process. Different experimental results are presented in Section 4, followed by conclusions about the different aspects of the project and proposed future work in Section 5.

Chapter 2

Clinical Background

2.1 Aorta Anatomy

The aorta is one of the most critical and largest arteries in the human body. It has the significant task of supplying oxygen-rich blood and nutrients to vital organs and major arteries. It begins at the left ventricle and terminates around the level of the starting of the iliac arteries [19, 27]. If disease or injury affects blood flow through this vessel, life-threatening complications can occur in minutes.

The aortic structure consists of three layers of tissue named inner, middle, and outer layers (or tunica intima, tunica media, and tunica adventitia, respectively). As displayed in figure 2.1 it can be divided into four sections across its length: the ascending aorta, the aortic arch, the thoracic (or descending) aorta, and the abdominal aorta [26]. Three major branches arise from the aortic arch: the brachiocephalic trunk, the left common carotid artery, and the left subclavian artery.

The first section is the ascending aorta, approximately two inches or five centimeters long. It arises from the aortic orifice and ascends to become the aortic arch. It travels with the pulmonary trunk within the pericardial sheath. The sinuses of Valsalva, which supply the myocardium with blood, are dilations in the ascending aorta located at the aortic valve level. These sinuses give rise to the left and right coronary arteries that supply the heart's muscle itself, the myocardium.

The next part of the aorta is the aortic arch, which is the continuation of the ascending

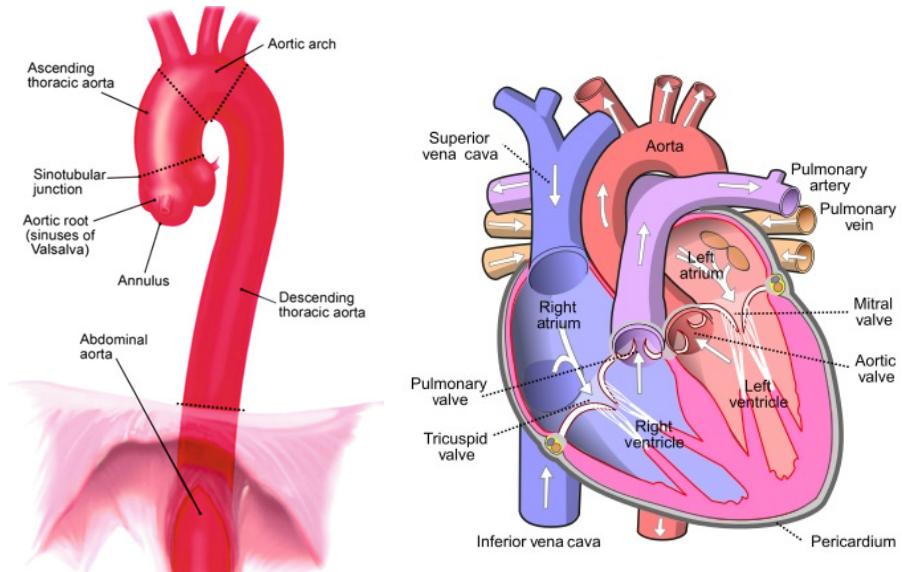


Figure 2.1: Anatomy of the aorta [5, 26]

aorta. It begins at the level of the second sternocostal joint (T2) and arches superiorly, then posteriorly, and to the left before moving inferior. The aortic arch is approximately five centimeters long, connected to the pulmonary trunk, and terminates at the level of about the T4 vertebra. Three major branches arise from the aortic arch, which supplies blood to the neck, head, and upper limbs.

The third part is the thoracic aorta (or descending) which continues from the aortic arch and spans from the level of T4 to T12. Initially, it begins to the left of the vertebral column, but approaches the midline as it descends. It exits the thorax via the diaphragm, where it becomes the abdominal aorta.

The abdominal aorta is the last part of the aorta and continues to the thoracic aorta. It begins at the level of T12 vertebrae. It terminates at the level of L4 vertebrae by bifurcating into the right and left common iliac arteries that supply blood to the lower limbs.

2.2 Aortic Aneurysm & Dissection

A serious condition where the aorta dilates even 50 percent more than its average diameter is defined as an ascending aortic aneurysm, leading the aortic wall to become weakened and stiff.

This condition can occur in any part of the aorta. However, it is more likely to occur in the abdomen (Abdominal Aortic Aneurysm) or thorax (Thoracic Aortic Aneurysm). Although it may occur in young people, this condition is more common in older adults.

When the diameter increases so drastically, it is possible the artery can rupture and cause internal bleeding. When there is an injury in the aorta wall because of dilation, it is possible to lead to Aortic Dissection (AD), which threatens to tear apart the aortic wall [26]. This condition can lead to blood leakage inside the wall so that the blood can flow on two channels instead of one. The first channel is the lumen, where the blood should normally flow. However, the second is a false channel where the blood remains stationary inside the wall—resulting in stenosis of the aortic lumen, leading to a reduction of blood flow to the rest of the body.

Treatments of AD depend on the type of dissection; type A develops in the ascending aorta, whereas type B involves a tear in the descending part of the aorta and may extend to the abdomen. Type A dissection involves surgery, where the surgeon removes as much of the dissected aorta as possible to stop blood from leaking into the aortic wall, and a graft is used to reconstruct the aorta. Medications are also used to reduce heart rate and lower blood pressure to prevent the worsening of the dissection. The treatment for type B dissection involves mainly lifelong medication to reduce blood pressure.

Aortic dissection requires urgent medical intervention, as it can result in death due to insufficient blood flow to the heart or complete rupture of the aorta. The most common occurrence of AD is the beginning of the ascending aorta, although it can occur anywhere along the length of the aorta.

Patients with AD will classically present with a tearing central chest pain radiating to the back. Other symptoms arise due to decreased blood supply to other organs, like stroke or ischemia. Other AD causes include chronic hypertension or weakened aortic wall, which can be due to Marfan Syndrome, or aortic aneurysm [26]. The first-line investigation for an AD is a CT angiogram. Likewise, MRI and Transesophageal echocardiography (TOE) can also be used. However, to obtain a complete representation of the heart function throughout the cardiac cycle, dynamic image sequences (cine) are necessary [39, 57]. CT scan outputs only one static image, so it is impossible to calculate compliance because it requires a series of images along the cardiac cycle. Furthermore, TOE is an invasive technique and does not provide information about the thoracic aorta.

2.3 Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging (MRI) has unquestionably contributed to a medical revolution. MRI is a well-established, and rapidly evolving field of cardiovascular medical imaging [49]. Diagnosis of cardiac disorders necessitates a precise assessment of the heart's function, and morphology [57]. Cardiac MRI (CMRI) meets these criteria. CMRI is a reference standard for cardiology practice due to several factors. Its benefits include, among other things, adaptability, excellent repeatability, and accuracy that no other individual imaging modality can achieve [14]. CMRI is non-invasive and does not employ ionizing radiation. It also has a high spatial resolution, a wide field of view, and outstanding soft-tissue contrast [20, 28], as shown in figure 2.2. CMRI can also provide a comprehensive cardiovascular assessment of a patient in a single setting. Moreover, it is particularly useful for evaluating global and regional left and right ventricular function by measuring stroke volume, ejection fraction, end-diastolic and end-systolic volumes, and masses.

Magnetic Resonance Imaging is a non-invasive method that creates comprehensive three-dimensional anatomical images based on the magnetic properties of the nuclei of proton atoms. Specifically, MRI measures and maps the portion of water in different body tissues, and through this process, a detailed image is generated. It relies on the magnetic properties of the proton in the nucleus of hydrogen atoms that, alongside Oxygen atoms, constitute the water molecule (H_2O) [29]. MRI is widely used for captiviting the human brain and body physiology and for heterogeneity (i.e., tumors) evaluation [62].

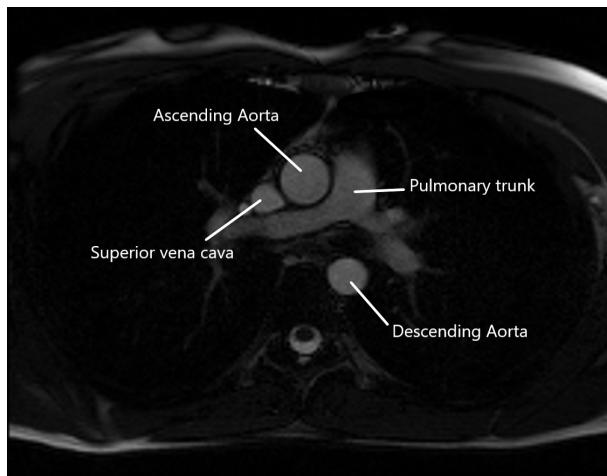


Figure 2.2: Axial MRI at the level of pulmonary trunk. Modified image from our dataset.

Despite the benefits highlighted above, CMRI is not a first-line research [49]. It is frequently obtained following other studies, such as echocardiography, radionuclide imaging, angiocardiology, or cardiothoracic CT [28]. This is due to the high cost of MRI technology, its limited availability, a lack of skilled personnel, clinician unfamiliarity, and patient compliance. It should be noted that CMRI is not always the best study for some patients. For example, claustrophobic, uncooperative, and pediatric patients impede the CMRI evaluation. In many circumstances, some form of sedation is required.

Dynamic CMRI is a challenging imaging modality to master. The difficulty is to speed up the fundamentally slow data acquisition while maintaining the high resolution and image quality requirements. Because of its speed, MRI has traditionally had significant limits in imaging moving organs [49]. Motion during the MRI scan process is the leading cause of image degradation. Any movement, even minor displacements, causes characteristic artifacts in the reconstructed images. Image blurring, ghosting, and misregistration are examples of undesirable consequences. This is especially challenging in cine CMRI, were dealing with motion caused by heartbeat and patient breathing remains one of the most difficult challenges.

In comparison to Computed Tomography (CT), MRI offers more information regarding image detail [44]. Zhao et al. in [66] argued that if one relies only on data from a single source, the understanding of diseases will be limited. They specifically pointed out that the data must be acquired from different medical institutes so that the data load will be larger. The reason behind this is the different equipment used in each medical institute. Images taken from different scanners have different parameters, while the volume of their histograms varies as well. More centers lead to more differences, even if they show the same part of the body. MR images, taken from different equipment or with different parameters, might have considerable intensity differences, significantly impacting image analysis results [62]. In preprocessing medical images, enhancement techniques are used to improve image quality. Salem et al. in [58] discussed the improvement of medical images that lead to the best possible perception of information in the fields of space and frequency. The improved image may give away more details compared to the original. Thus, MRI preprocessing requires special treatment to avoid data loss and achieve significant detail recovery.

2.4 Aortic walls & Elastic properties of aorta

Aortic walls quadrants

Aortic walls are the main structural component of the aorta; they must be elastic to expand and contract during the cardiac cycle (between systole and diastole), where the blood pressure changes rapidly. The thickness and composition of aortic walls vary depending on their location within the organ. The aortic wall can be partitioned relative to medial (MED), posterior (POST), lateral (LAT), and anterior (ANT) quadrants [35, 43].

Compliance

A measure to describe the elasticity of the aortic walls is the aortic compliance, defined as the variation of the cross-sectional area during a cardiac cycle divided by the pulse pressure variation. The mathematic formula is specified in the equation 2.1. In this study, the cross-sectional area is calculated based on 2D thoracic MRI slices on the level of the pulmonary trunk along the cardiac cycle. The pressure is provided by physicians acquired during the MRI acquisition. To put it briefly, we divide the difference between the maximal and minimal area of the aorta by the difference in systolic and diastolic pressures.

$$\text{Compliance} = \frac{\text{MaximumArea} - \text{MinimumArea}}{\text{SystolicPressure} - \text{DiastolicPressure}} (\text{mmHg}) \quad (2.1)$$

Young's modulus & Strain

Young's modulus is a metric for determining how difficult it is to produce elastic deformation in solid materials. Applying a load F to elastic materials can result in a change in length l . The stress σ is the external force perpendicular to the cross-sectional area A , and the engineering strain ε is the length variation in its vertical direction.

The force and the length determine the stresses and strains. Engineering strain is defined as the ratio of the resting specimen length (after preconditioning) l_0 divided by the deformed length Δl , where Δl is the difference in length between the resting and load-filled specimens.

$$\varepsilon = \frac{l - l_0}{l_0}$$

The amount of tensile load F recorded during the test per unit loaded cross-sectional area A of the specimen is referred to as stress σ [21]. Incompressible tissue is a term used to describe

aortic tissue. Area A is equal to the cross-sectional area of the resting specimen A_0 , where A_0 is calculated using the load-free specimen thickness and length l_0 .

Stress σ can be computed as:

$$\sigma = \frac{F}{A_0}$$

The elastic modulus, also known as Young's modulus, is a widely used term to describe the biomechanical properties of human tissue [1, 22, 31, 60]. The calculation of Young's modulus in aortic tissue can be done not only in uni-axial tensile tests [17], but also in biaxial tensile tests [11, 25, 64]. Young's modulus at various stress levels can be calculated as the first derivative of stress overstretch using the stress-strain curve.

$$E = \frac{\sigma}{\varepsilon}$$

The aforementioned equations describe the ex-vivo measurements. However, to calculate in-vivo strain, we take into account the perimeter of the aorta (aortic wall). Strain is defined as the difference between the maximum and minimum perimeter divided by the minimum perimeter:

$$\text{Strain} = \frac{\text{MaximumPerimeter} - \text{MinimumPerimeter}}{\text{MinimumPerimeter}} \quad (2.2)$$

Chapter 3

State of the art

3.1 Contour segmentation of aorta

Some published studies on aorta segmentation using classical image processing techniques are described briefly below. Rose et al. [54, 55] take into consideration two static images during systole and diastole. In this manner, we lose information on the aortic cross-sectional area during the cardiac cycle. Some authors take into account the whole cardiac cycle, using fully or semi-automated but not deep learning methods, allowing the creation of cross-sectional area vs. time curves [32, 33]. Likewise, Miteran et al. [41] detect the aortic wall using an adaptation of a curvilinear region detector, with significant low segmentation error and robust results. Following the segmentation, they get the maximum and minimum areas without any manual selection of the systole phase, allowing us to evaluate the elastic modulus of the aortic wall [30, 47]. Although some of those methods derive promising results, most are semi-automated or need user input, for example, in the center of the aorta. From the methods mentioned above, only the work of Miteran et al. [41] is fully automated; thus, we conclude there is a lack of fully automated or deep learning methods.

3.2 Deep Learning

Image segmentation is a common problem in computer vision research, which refers to dividing an image into sub-images by grouping specific features such as texture, color, etc. It can be divided into semantic and instance segmentation, where semantic refers to categorizing all the pixels of the same class into one, and an instance refers to identifying each instance of the same category as a separate object. Despite their advances, classic image segmentation algorithms

are outperformed by methods based on Deep Learning and Neural Networks, a subcategory of Machine Learning and Artificial Intelligence. Each network neuron can be described as a small information-processing unit with simple tasks like convolution, pooling, or activation. A combination of neurons connected forms a hidden layer, where each layer receives a set of weighted inputs, processes it through an activation function, and produces an output for the next layer, as shown in figure 3.1. Eventually, a collection of an input layer, an output layer, and multiple hidden layers in between form a neural network, displayed in figure 3.2.

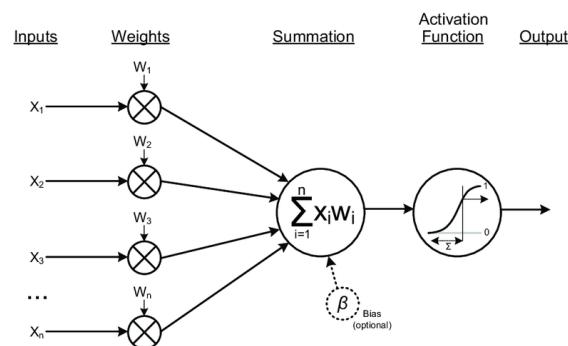


Figure 3.1: Activation in details. Modified from [48]

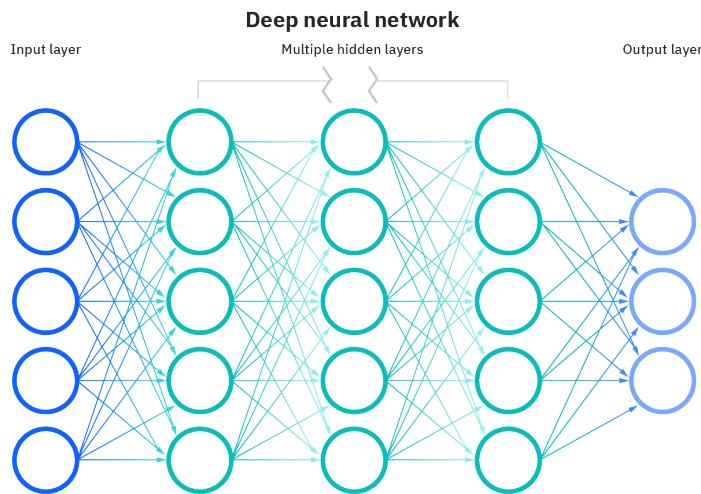


Figure 3.2: Neural Networks diagram. Modified from [12]

One of the most famous advances of deep learning is Convolutional Neural Networks (CNN)

with supervised learning, which has many breakthroughs in image analysis and processing. CNNs leverage the convolution operation to extract high-level features from an image while reducing its dimensions, to improve training performance. Convolution is the sequential multiplication of a small size kernel (ex. 3×3) across the spatial dimensions of the image. This operation downsamples the image while keeping some crucial features. The importance of those features depends on the kernel values, which are the trainable properties or weights of the network, and its size, which is predefined.

3.3 Fully Convolutional Networks & U-Net

In this section, we briefly define Fully Convolutional Networks (FCN) [36], and U-Net [53] architecture with the most famous modified versions of it that we have explored and experimented. Defining those architectures in depth would be out of the scope of this study. Subsequently, we present some state-of-the-art publications concerning the field of aorta segmentation, using a variety of techniques like U-Net or even semi-automatic image processing.

In general, CNN architectures have some fully connected layers at the end of the network to perform classification (the previous layer's output neurons are fully connected with the next layer's input). Those fully connected layers increase complexity and millions of trainable parameters, decreasing the model's overall performance. A straightforward solution is to remove the fully connected layers and replace them with convolutional layers, as suggested by VGG [59] and ResNet [23]. However, the results of shallow FCN are not satisfying enough for medical image analysis, although they look promising. Some encoder-decoder architectures, like SegNet and U-Net, have been proposed to improve FCN. The encoder-decoder has a symmetric structure consisting of two FCNs, one for extraction of low resolution but high-level feature maps from input images, connected to another inverted FCN for the decoder. The decoder's task is to map the low-resolution features learned by the encoder to high-resolution pixel space information, as the labels of each pixel in the image.

U-Net [53] has been invented for medical image segmentation, and it is one of the most cited architectures in this field. It consists of one encoder and one decoder, in the shape of U shown in figure 3.3 similarly with SegNet, but with one significant addition, the skip connections. The advantage of skip connections is that they provide additional higher-level spatial feature information to the decoder. The encoder consists of four submodules named convolutional blocks; each contains convolutional layers followed by a max-pooling layer for downsampling. The resolution decreases as the image passes through each encoder block, but the output is successively

a higher-level feature map. The decoder follows the opposite logic with four de-convolutional blocks; it up-samples the image increasing the dimensions and giving predictions for each pixel. Between the encoder and the decoder, there is an extra convolutional block named "bottleneck" where the dimensions of the feature map become flattened and spatial information is lost. This lack of dimensionality leads our decoder to weak predictions because there is no information about where each pixel belongs on the image. In this manner, the inventors of U-Net added skip connections to bring higher-level spatial information from the output of the encoder's block on the same level to the decoder block. This combination of high and low-level information improves the model's accuracy, giving the opportunity at the same time to extract complex features.

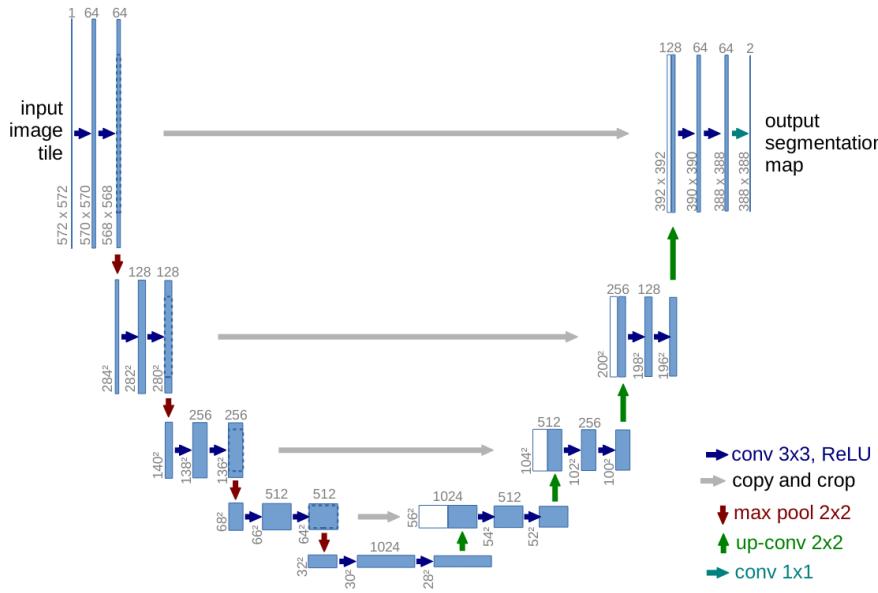


Figure 3.3: U-Net architecture [53] Every blue box is a multichannel feature map, the number of the channel is denoted on the top of each box. The size information is provided at the lower left edge of each box. White boxes represent the feature maps that are copied thought the skip connections, and the arrows represent the different operations.

3.4 Deep learning based segmentation

3.4.1 2D methods

Tianling et al. [38] introduce a DL-based algorithm for dissected aorta segmentation on computed tomography angiography (CTA) images. The algorithm consists of two steps. Primarily, a 3D CNN is implemented to divide the 3D volume into two anatomical portions. Secondly, two 2D CNNs based on a pyramid scene parsing network (PSPnet) proposed in [67] segment each specific piece separately. PSPnet performs a spatial pyramid pooling (SPP) to address the diameter variations of the aorta over the slices. Moreover, the accuracy is enhanced by joining a boundary extraction branch into the model. This multi-branch architecture is a modification of the model recommended in [63]. To gain higher accuracy, the authors divide the model into three branches after convolution operations, handling the three detection, segmentation, and boundary extraction tasks. In this approach, the 2D PSPnet model only requires two branches for segmentation and boundary extraction, whose outputs are later fused. As they mention, the performance of this study is measured with an average Dice index of 92 percent. The combination of 3D and 2D models improves the performance of accuracy and robustness compared to 3D or 2D only models. The edge extraction branch improves the Dice Coefficient near aorta boundaries from 73.41% to 81.39%.

Fadil et al. [18] trained multiple sequence-specific U-Net 2D [53] models to segment different areas of the heart like the left ventricle (LV), right ventricle (RV), and aorta in the cine short-axis slices. Metrics like Dice similarity, Hausdorff distance, mean surface distance, and Pearson correlation were used to validate and compare the results of the suggested pipeline with intra-observer variability caused by the experts who segmented the ground truth. All models were trained using Adam optimizer with a weighted cross-entropy loss. The results indicate precise segmentation with Dice scores (0.91 for LV, 0.92 for RV, 0.93 for aorta on average), comparable to intra-observer Dice scores of (0.86 for LV, 0.87 for RV, 0.95 for aorta flow in average). Despite the high values of Dice scores on the LV and RV, we can observe that the aorta obtains lower Dice scores compared to the experts.

3.4.2 2D plus time methods

As Wenjia et al. [4] mention, direct segmentation to each time frame of a sequence may ignore the temporal continuity inherent in the series. In this concept, 2D plus time (2D+t) studies consider an image sequence concerning acquisition MRI time. Instead of feeding the algorithm a single slice, we have the advantage of prior knowledge during training. For example, when the aorta dilates, we can track the cross-area variability for every case and train our model.

One possible implementation of 2D+t is the proposed architecture of Wenjia et al. [4]. They propose a network composed of two subnetworks. This network combine a Fully Convolutional Network (FCN) with a Recurrent Neural Network (RNN). This approach benefits the segmentation with both spatial and temporal information. One problem they try to eliminate is sparse data annotation, which was faced by performing a non-rigid image registration method [56] to propagate the labels from the annotated frames to the unlabeled neighboring ones in the cardiac cycle. In other words, pseudo-images were generated for the missing slices of the sequence. Furthermore, an exponentially weighted loss function for training was introduced. As the FCN which is the first part of the architecture, they use U-Net architecture to analyze spatial features and extract feature maps. The next step is to feed those feature maps into the RNN, exploring the temporal features. As RNN network, they selected convolutional LSTM (C-LSTM) [61] which can analyze multidimensional images across the temporal domain. This method achieved an average Dice coefficient of 0.96 and 0.95 for the ascending aorta and the descending aorta, respectively, over a test set of 100 subjects. The proposed method suggests that a sequence of images instead of a single image may significantly improve segmentation accuracy and temporal smoothness compared to other methods based only on spatial information.

3.4.3 Other methods (2.5D, 2D+3D, 3D)

An in-between method of 2D and 3D is called 2.5D, which is configured 2D strategy that takes into consideration also information from the adjacent slices [40]. The lack of spatial data in a 2D network is a severe issue. Although 3D networks have the spatial information of the whole volume, the computational complexity and the training time are also an impediment. The implementation of 2.5D networks is based on 2D architectures, but instead of 1 image as an input, the model gets the reference slice stacked with the adjacent slices. This method is based on RGB input of the neural network, where the channels of the image are 3. Similarly, in medical imaging modalities (MR, CT), the slices are commonly grayscale; thus, the color channels are replaced by the adjacent slices of reference slice. Dual et al. [16] proposed a 2.5D fully convolutional network for cardiac MRI using atlas propagation for semantic segmentation. The key feature of this approach is the prior shape knowledge through the refinement step, improving the workflow quality. Following the 2.5D strategy, Zhou et al. [68] proposed FCN for three-dimensional CT images for 19 anatomical structures. The images are stacked by employing majority voting over 2D images; each image in the stacked image set corresponds to a different view (axial, sagittal, coronal).

The authors of ComboNet [2] propose the combination of 2D and 3D images to limit the

sparse outliers and prevent the leakage of the surrounding soft tissues caused by 2D images. Another research [7], based on 3D MRI segmentation of the aorta, suggests a fully automated method to generate 4D-flow for reproducible quantification of aortic flow, peak velocity, and dimensions.

3D MRIs consists of a sequence of slices, where each slice encapsulates three different dimensions of the acquisition (axial, sagittal, coronal). Summing up, 3D architectures are out of the scope of this study. We cannot reproduce or implement those approaches because of our 2D plus time dataset, which consists only of axial sequences of the thoracic aorta.

Similarly, studies like [10,13,24] where clinical data consists of CT images are not explored. It is not possible to assess the elastic properties of the aorta using CT because it does not provide temporal information.

3.5 Comparison of two techniques (in-vivo and ex-vivo) for evaluating the elastic properties of the ascending aorta.

The current work is part of the studies [35, 42, 43] which intends to assess the correlation between ex-vivo biaxial stretching and in-vivo evaluation of the elastic properties of the aorta. Ex-vivo biaxial stretching allows the evaluation of stiffness of the aorta on tissue collected during aortic surgery. While in-vivo, in this work, refers to the experiments based on medical imaging. Aneurysms of ascending aorta (AA) impact the aorta over the years, progressively dilating, and the elasticity decreases while the stiffness increases. The main complication of AA is type A aortic dissection, which is associated with high levels of mortality. Study population consists of 70 in studies [42,43] and 100 patients in study [35]. All patients having an aneurysm of ascending aorta require surgery. MRI acquisition is performed pre-operatively for ascending and descending aorta. The MRI acquisition protocol performed in this work include a sequence of images specific for the evaluation of aortic compliance.

For all the studies, the aorta samples are collected during the aorta surgery for the biomechanical laboratory stretching tests (post-operative). Physicians can decide if the aneurysm is crucial (if there is a need for surgery), based on the diameter of the ascending aorta. Current guidelines recommend prophylactic replacement if the diameter is greater than 55 mm. There are some tissue disorders where the threshold should be lower than the average patient. Al-

though, some studies [30, 47] emerge that in many cases of acute aortic dissection, the diameter is less than 55 mm. This study declares that there is a need for new, more effective criteria for the prevention of aneurysm compilations based on the biomechanical properties of the aorta.

For each image, the ascending and descending aorta contours are automatically extracted in planes perpendicular to the sino-tubular junction and the diaphragm, respectively, and the maximal and minimal areas of them are recorded. In advance, the blood pressure of the patient is used to compute the compliance.

The collected aortic wall samples from the replacement procedure of studies [35, 43] are partitioned relative to medial, posterior, lateral, and anterior quadrants. The average thickness is measured for each quadrant. Then they stretch it biaxial until it ruptures. To quantify those experiments, the authors suggest that "Young's modulus, which measures the ability of a material to withstand changes in length (strain) when under lengthwise tension, or compression (tensile stress)." The correlation between MRI and bi-axial testing measurements has been studied.

Results from the work of Morgan et al. [43] implies there is a high correlation between the aortic elasticity measured by Maximum's Young Modulus (MYM) and the aortic compliance measured from MRI imaging. Although those results are inefficient in proving that aortic compliance is a criterion for operative indication, they proved the robustness of cardiac MRI imaging to assess the elastic properties of the aorta.

The work of Siyu et al. [35] concludes there is variability across the four quadrants of the aortic wall. The lateral quadrant was the stiffest among all, while the medial quadrant had the lowest aortic stiffness across the walls. The main objectives of our work, is to assess the elastic properties of the aorta (compliance and strain) and compare them with the above results.

Chapter 4

Methodology

In summary, this chapter describes in detail all the necessary information and action implemented during the research of this work. The workflow, presented briefly in figure 4.1, starts from data acquisition, followed by semi-automated segmentation, which has been implemented by experts providing the ground truth for model training. Multiple architectures and hyperparameters have been explored, and the model with higher performance has been selected, leading to the next step, the aorta segmentation and evaluation. The last step of this project is the division of the aortic surface into different quadrants to explore if the local compliance has a significant correlation with the results ex-vivo provided by the research of Siyu et al. [35]. Another purpose of this division in local compliance computed per quadrant is to explore if elasticity is reduced or increased on a different side of the aortic wall.

4.1 Dataset

A cardiac MRI is performed before the cardiac surgery to determine in particular aortic compliance. Images are collected using a 3T MRI (Skyra, Siemens Healthineers, Erlangen, Germany), with only four separate acquisitions in addition to the standard protocol. The standard approach includes a series of T1-weighted Turbo Spin Echo images in oblique-sagittal orientation that encompass the aorta, cine-MRI in oblique sagittal and oblique coronal LVOT (Left ventricular outflow tract) planes, as in sinus plane, and angio-MRI without contrast agent (for a total duration of around 20 minutes). A FLASH-type sequence is undertaken during a short breath-hold in the transverse plane at the level of the pulmonary artery bifurcation. Because it is less sensitive to noise caused by rapid or turbulent flow at 3T, this sequence is chosen over a Steady State Free Precession (SSFP)-type sequence. This plane allows operators to examine both the ascending and descending aortas. Images at all phases of the cardiac cycle are obtained

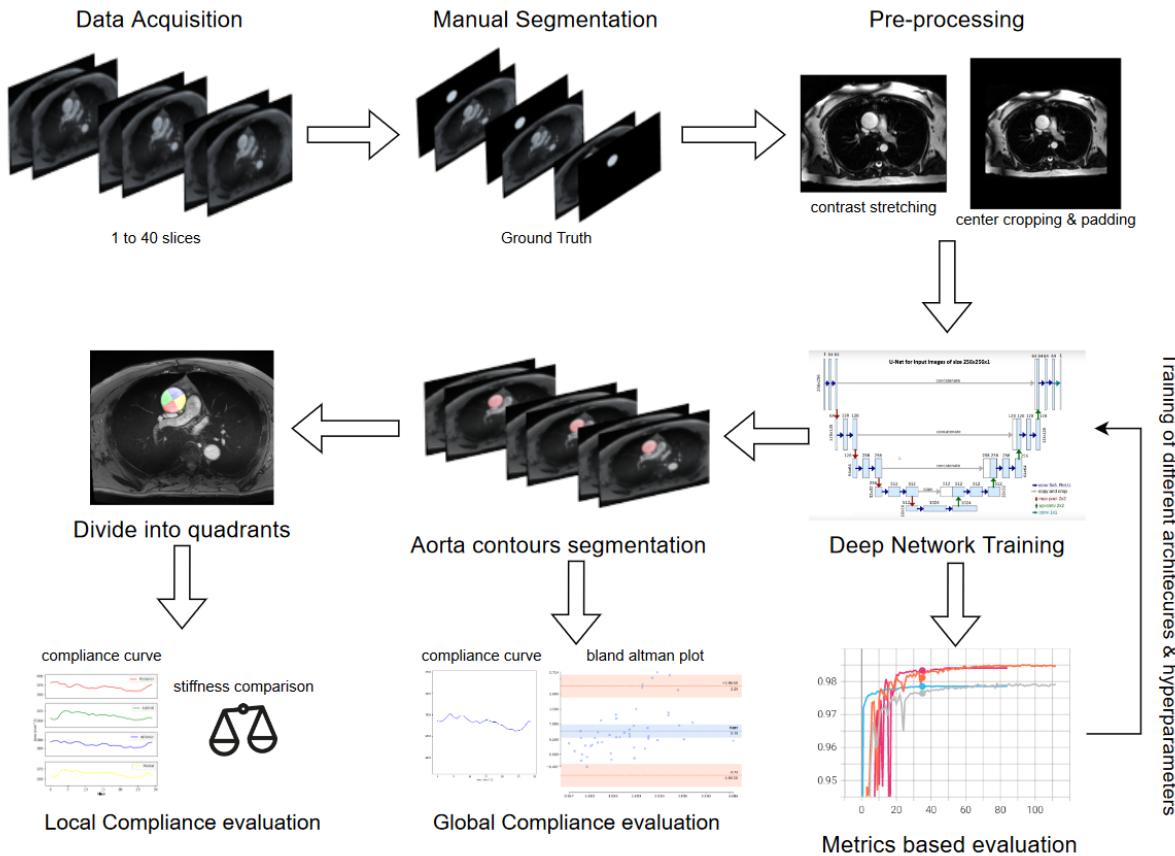


Figure 4.1: Workflow of the research strategy. Modified images from our dataset.

with a temporal resolution of 20 msec to 34 msec, with the following sequence settings: echo time of 3.42 msec, repetition time of 7.21 msec, flip angle of 12°, the spatial resolution of 1.09 mm²/pixel, and 1.25 mm²/pixel (corresponding to a field of view ranging from 350 mm to 400 mm), and slice thickness of 5 mm. A generalized auto-calibrating partially parallel acquisitions (GRAPPA) is performed, with an acceleration factor of 2. A pre-scan normalized filter and distortion correction were also applied.

The data has been acquired from the University Hospital (CHU) of Dijon, consisting of 73 exams, with an average of 30 to 40 slices per case across the cardiac cycle. 58 cases include aortic dilation, and 15 do not contain dilation. During the cardiac cycle, each patient's systolic and diastolic pressure has been recorded and provided for our study. The dataset consists of a total of 2,189 slices of cardiac axial MRIs on the height of the pulmonary trunk, with the same

number of ground truth masks segmented from physicist experts, using the semi-automated software QIR [55]. The process of ground truth segmentation is not fully automated. Instead, the expert must indicate the center of the ascending aorta. The software proposes a potential segmentation, the area of interest (ROI), which is not always precise, mainly when the image contains a dilated aorta. Then the expert must correct the exported segmentation to fit precisely the desired ROI. Instead of this laborious and time-consuming task, we propose a fully automated deep learning segmentation method.

Some images of the dataset are at a higher resolution, or the aorta walls are well-defined; thus, they are more accessible to segment, as depicted in the examples of figure 4.2. However, as presented in figure 4.3, there is another case where the noise or artifacts are often present in the critical part of the cardiac cycle, especially for the ascending aorta, as the 3 Tesla MRI sequence is sensitive to rapid and turbulent flow. Finally, other common difficulties can be encountered, such as highly dilated ascending aorta presented in figure 4.4(a). When the aortic border is poorly visible due to the proximity of some other structures, as shown in figure 4.4(b).

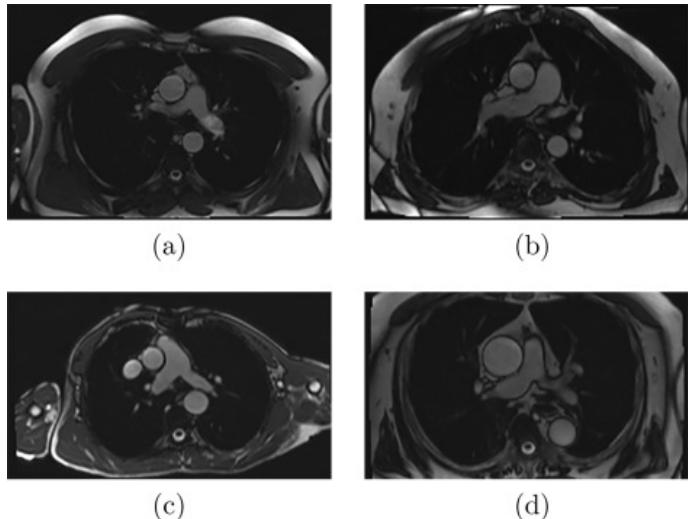


Figure 4.2: Examples of MR images with neither noise nor artifacts. [41]

The dataset has been split into training and test sets consisting of 1,784 and 405 images to train and evaluate different deep learning architectures. For this work, a validation set is not necessary because all the training experiments have been implemented with K-field cross-validation [50].

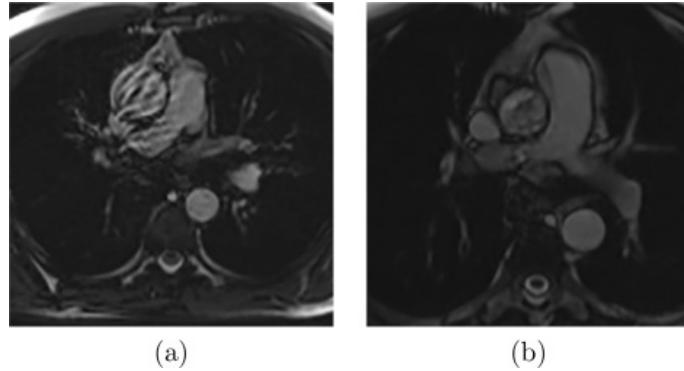


Figure 4.3: Examples of MR images with artifact due to rapid or turbulent blood flow. [41]

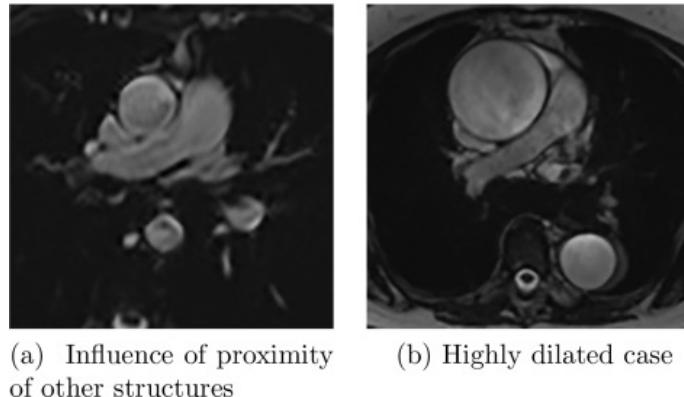


Figure 4.4: Examples of MR images where (a) the aortic wall is not well-defined or (b) the aorta is highly dilated. [41]

4.2 Preprocessing

Every image is represented to the software as an array of pixels; grayscale images are two-dimensional arrays with dimensions equal to image dimensions. In the subsections below, we refer to images as arrays with the capital letter "X" and each image pixel as the lower letter "x." During the pre-processing phase, each one of the following techniques has been applied to every image in the same order as they appear, normalization, contrast stretching, central crop and pad.

4.2.1 Minimun-Maximum Normalization

Usually, deep learning architectures require images to be scaled to a specific range; more specifically, the pixel values must be scaled between zero and one. In this case, Unet architectures

do not require normalization, but when normalization is implemented, there are improvements concerning computation time and memory performance. On this ground, we chose to perform min-max normalization, to scale pixels at the range of 0 to 1 according to:

$$\frac{X - X_{min}}{X_{max} - X_{min}}$$

4.2.2 Contrast stretching

Improving the image's quality and visibility in some image processing tasks is mandatory. Many techniques are invented for this purpose, like histogram matching, histogram equalization, etc. One of them is histogram stretching, shown in figure 4.5, which stretches pixel values of an image among the range of histograms and increases the quality of the MRIs by increasing or decreasing the contrast between adjacent pixels. This can be useful for enhancing images that are difficult to see because of low contrast or poor lighting.

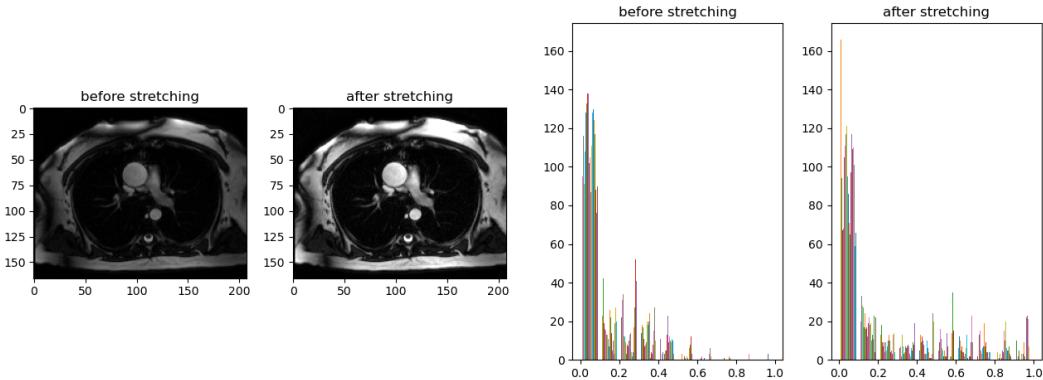


Figure 4.5: Contrast stretching. Modified images from our dataset.

4.2.3 Central crop and pad

We use a custom method implemented in Python in order to crop the big images to the desired 256×256 dimensions, shown in figure 4.6. On the other hand, to handle small images, we add zero padding around the 4 sides of the image. The cropping has been implemented in respect to the center of the image.

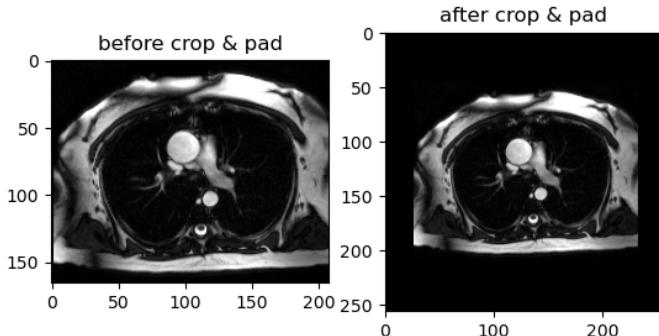


Figure 4.6: Crop and pad. Modified images from our dataset.

4.3 Data Augmentation

Limited datasets can introduce overfitting and low accuracy. Data augmentation is necessary for some complex deep learning tasks to improve accuracy and tackle overfitting. Nevertheless, it can be challenging, primarily when we process medical images. We perform random transformations on images using an automated software [8]. Those transformations consist of random rotation, random horizontal and vertical flip, random affine transformation, and random brightness and transformation.

4.4 Model Training

A neural network learns or approximates a function using samples in the training dataset to map inputs to outputs optimally. The stochastic gradient descent algorithm is used to train deep learning neural networks. Stochastic gradient descent is an optimization algorithm that estimates the error gradient for the model’s current state using examples from the training dataset, then updates the model weights using the backpropagation of errors algorithm.

Model Architecture

In this work, we experimented with U-Net and some other modified architectures of U-Net, like Residual-U-Net [65] where each submodule of U-Net is replaced with a residual connection and a dense layer. Attention-U-Net introduced by [45], where they added an attention mechanism to readjust the encoder’s output characteristics and provide the decoder knowledge about where is the most important spatial information through attention gates. Attention-Residual-U-Net [9] inspired by Residual Network [23] implement both residual blocks and attention gates.

Similarly, inspired by Recurrent Neural Networks (RNN) proposed Recurrent-Residual-U-Net (R2-U-Net) [3], which is a combination of RNNs and residual blocks. In advance, U-Net++ is a modification of the U-Net encoder-decoder network where the encoder and decoder subnetworks are connected through a series of nested, dense skip pathways [69].

There are no fully connected layers in U-Net architectures, allowing the user to select the input dimensions. We chose input dimensions 256×256 because our images are low resolution. Our implementation of U-Net has some slightly minor modifications from the original one; on the convolutional block, we added Instance Normalization layers after each convolutional layer and before each Rectified Linear Unit (ReLU) Activation.

Learning Rate

The step size or "learning rate" is the amount by which the weights are adjusted during training. The learning rate is a customizable hyperparameter used in neural network training with a modest positive value, between 0.0 and 1.0. During training, the backpropagation of error calculates the amount of error caused by a node's weights in the network. Instead of updating the weight with the total amount, the learning rate is used to scale it. This means that a learning rate of 0.1, a commonly used default value, means that each time the weights in the network are updated, they are modified by $0.1 * (\text{estimated weight error})$, or 10% of the estimated weight error.

The learning rate hyperparameter controls the rate or speed at which the model learns. Assuming a well-configured learning rate, the model will learn to best approximate the function. In general, a high learning rate permits the model to learn faster but at the expense of producing a suboptimal final set of weights. A slower learning rate may allow the model to acquire more optimal, or even globally ideal, weights, but training time will be significantly longer. A learning rate that is too high will result in weight updates that are too large, and vice versa. During Overtraining epochs, the model's performance (such as its loss on the training dataset) will oscillate. A slow learning rate may never converge or become trapped on an inferior solution. In the worst-case scenario, excessive weight updates may lead the weights to explode (i.e., result in a numerical overflow). As a result, we should not use a learning rate that is either too high or too low. Nonetheless, we must construct the model so that a "sufficient" combination of weights is obtained on average to approximate the mapping issue as represented by the training dataset.

Unfortunately, we cannot analytically calculate the optimal learning rate for a given model on a particular dataset. Instead, a reasonable learning rate must be established through trial

and error. The range of values to consider for the learning rate is less than 1.0 and more than 10-6.

Learning rate scheduler

Dynamic learning rates are an alternative to employing a fixed learning rate during the training process. The learning rate schedule or decay refers to how the learning rate changes over time (training epochs). One of the most basic learning rate schedules is to decrease the learning rate linearly from a high initial value to a low one. This allows for significant weight adjustments at the start of the learning phase and minor modifications or fine-tuning near the end.

While training neural networks, utilizing a learning rate schedule may be a best practice. The setup problem involves selecting an initial learning rate and a learning rate schedule rather than a fixed learning rate hyperparameter. Given the improved performance that a learning rate schedule may allow, it is possible that selecting an initial learning rate is less sensitive than selecting a fixed learning rate. The learning rate can be reduced to near-zero levels. Alternatively, the learning rate can be degraded over a set number of training epochs and then held constant at a low value for the remaining training epochs to allow for more time for fine-tuning.

ADAM Optimizer

Adaptive Moment Estimation (ADAM) is a technique for optimizing gradient descent algorithms. When dealing with complex data involving many information or parameters, the method is extremely efficient. It uses less memory. It appears to be a hybrid of the gradient descent with momentum and the RMSprop algorithms:

- Gradient descent with momentum: Adding history to the weight update makes training a neural network easier. An exponentially weighted average of the previous weight updates can be included when the weights are changed. This transition to stochastic gradient descent is referred to as "momentum." It is used to smooth the optimization process and accelerate the gradient descent algorithm by taking the 'exponentially weighted average' of the gradients into account. Using averages causes the algorithm to converge to the minima more quickly. Momentum is set to be larger than 0.0 and less than one, with typical values like 0.9 and 0.99 being utilized in practice.
- Root Mean Square Propagation (RMSprop) [70]: It is an adaptive learning technique that attempts to enhance AdaGrad [37]. It uses the 'exponential moving average' rather than

the cumulative sum of squared gradients as AdaGrad does.

Adam Optimizer leverages the strengths or good characteristics of the previous two methods and improves them to provide a more efficient gradient descent. We adjust the gradient descent rate to minimal oscillation when it reaches the global minimum while taking large enough steps (step-size) to avoid the local minima barriers along the way. As a result, we are efficiently integrating the preceding strategies' features to obtain the global minimum.

Optimal hyperparameters & settings

After exploring multiple hyperparameters and settings, the optimal ones have been chosen to train the final U-Net model, as described below. The initial learning rate has been set to 0.001 with a minimum of 0.0000001. The learning rate has been reduced by 0.1 every five epochs, and the validation loss did not improve. As the loss function, Dice loss has been preferred, defined as:

$$\text{DiceLoss} = 1 - \text{DiceCoefficient} = 1 - \frac{2TP}{2TP + FP + FN}$$

Furthermore, the ADAM gradient descent optimizer has been selected. The initial epoch size is 200, but the training process is terminated earlier due to early stopping. As a normalization technique, instance layer normalization has been performed after each convolutional block. As mentioned earlier, the training process has been implemented using K-fold cross-validation with several K equals to 5. In this manner, we do not need a validation set, leading to more training data availability.

Metric-based evaluation is performed on multiple architectures and hyperparameter settings experiments to select the one with the highest ranking. Those experiments are described in more detail in the next chapter.

4.5 Post-processing

The pipeline's output images are not the same size as the input because our input was pre-processed to fit the architecture's specifications. In other words, the output image must go through the opposite process to be restored to the original image's dimensions. To accomplish this, we must undo the zero padding and the center cropping to restore the output mask. Furthermore, in some slices of the image series, we use the Convex Hull method to fix the shape of the predicted mask when there are missing pixels or holes. This is a geometric technique used to smooth the output masks. It can be defined as the set of all convex combinations of

points in the subset, where a convex combination is a linear combination of points where all coefficients are non-negative, and sum to 1 [52].

4.6 Metrics

This section provides a brief overview of the confusion matrix and the performance measures used in this work.

4.6.1 Confusion Matrix

The confusion matrix as shown in table 4.1 must be defined, prior to the description of the metrics. A confusion matrix is a table that describes performance of a classification model, in other words it answers the following question: how many of the predicted values are correctly predicted, or not. It is mandatory though to have a set of test data, named as ground truth, for which the true values are known. If a prediction is classified as True and the ground truth is True, we have a True Positive (TP). If the ground truth is False, we have a False Positive (FP). On the other hand, if the prediction is classified as False and the ground truth is False, we have a True Negative (TN). At last, if the prediction is True, then we have a False Negative (FN).

| | | Predicted | |
|--------|-------|-----------|-------|
| | | True | False |
| Actual | True | TP | FP |
| | False | FN | TN |

Table 4.1: Confusion matrix

It is necessary to mention that the more True Positives and True Negatives we have, the better our results are. In other words, as we try to improve our model, we try to eliminate the False values.

4.6.2 Precision & Recall

The Precision is defined as the TP divided by the TP plus the FP predictions, and it can answer the question of how many of the retrieved values are relevant.

$$Precision = \frac{TP}{TP + FP}$$

While Recall is defined as the TP divided by the TP plus the FN predictions and answers the question, how many relevant elements are retrieved.

$$\text{Recall} = \frac{TP}{TP + FN}$$

4.6.3 Intersection over Union

Intersection over Union is a state-of-the-art metric method for semantic segmentation tasks. It can be computed by dividing the overlapping area divided by the union area between the ground truth and the prediction images, respectively. The formula is defined below:

$$IoU = \frac{TP}{TP + FP + FN}$$

4.6.4 Dice coefficient

The Dice coefficient is also a state-of-the-art metric in the semantic segmentation field. It can be described as two times the area of overlap divided by the total area between ground truth and prediction images.

$$\text{DiceCoefficient} = \frac{2TP}{2TP + FP + FN}$$

4.6.5 Hausdorff distance

The Hausdorff distance, which measures how far two subsets of a metric space are from each other, is used to calculate the distance between two images. The primary objective of this metric is to detect and highlight outliers. The narrower the distance, the more accurate the predicted segmentation. To define the Hausdorff distance of two subsets, we assume that X , Y are two non-empty subsets of a metric space as shown in 4.7.

(M, d) .

$$d_H(X, Y) = \max \left\{ \sup_{x \in X} d(x, Y), \sup_{y \in Y} d(Y, x) \right\}$$

Where \sup represents the supremum, similar to \inf the infimum, and where

$$d(a, B) = \inf_{b \in B} d(a, b) \quad d(a, B) = \inf_{b \in B} d(a, b)$$

quantifies the distance from a point $a \in X$ to the subset $B \subseteq X$.

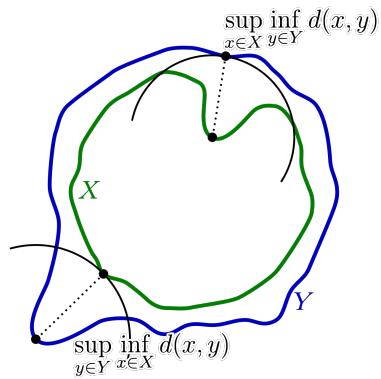


Figure 4.7: Hausdorff distance between the green line X and the blue line Y. [51]

Chapter 5

Results

In this chapter, we summarize the results in two sections, where the first one evolves the evaluation of deep learning architectures, and the second is based on aortic elasticity parameters.

5.1 Deep Learning evaluation

Since manual segmentation is a time-consuming task, initially, our dataset size was small, and progressively as the physicians provided more MRIs with ground truth segmentation, our dataset increased. As the dataset increased, the performance also tended to increase. Dice loss was selected as it outperforms the Focal and Tversky losses. Different architectures, exhibited in table 5.2, have been trained with different hyperparameter settings as shown in table 5.1. All hyperparameter settings have been tuned in the same manner. The initial learning rate is set to 0.001 with a decay of 0.1 over the epochs when the validation loss is not improved for five epochs. Likewise, instance layer normalization has been chosen instead of batch normalization. On average, total training time is around 30 to 40 epochs with a maximum training process of two hours, as we implement early stopping with the patience of 16.

| | | | | | |
|-------------------------|------------------|-----------------|--------------|--------|-----------|
| Early stopping patience | 1 | 2 | 4 | 8 | 16 |
| Starting learning rate | 0.1 | 0.01 | 0.001 | 0.0001 | |
| Loss function | Dice Loss | Focal Loss | Tversky Loss | | |
| Layer Normalization | Batch | Instance | Group | | |

Table 5.1: Hyperparameters & settings

Compared to our final proposal (Dice=98.09%, IoU=96.27%, HD=4.88mm), the initial ar-

chitecture of U-Net with less training data, no hyperparameter tuning, and no adjustments performed worse (Dice=97.07%, IoU=94.70%, HD=7.02mm). Our model's outstanding performance is indicated by a visual representation of the segmentation in figure 5.1. This difference in performance has been achieved gradually. One of the first actions is normalizing the pixel values and performing center crop and padding instead of resizing the image. The next stage is augmenting the dataset, random horizontal or vertical flipping, random rotations, random brightness contrast, and random affine transformations. Finally, hyperparameter tuning introduces a significant difference in performance, especially between different architectures. Concluding, U-Net with hyperparameter tuning and settings described above performs excellent compared to the rest, as displayed in table 5.2. Although Residual-Unet has the same Dice coefficient as Unet, similarly to Attention-Residual-Unet, we chose as a final model to use Unet architecture for the following experiments. Having the minimum Hausdorff distance makes this model less sensitive to outliers and misclassifications, as well as the architecture is not so complex as the rest.

| Model | Dice coeff | IoU | Hausdorff | Precision | Recall |
|--------------|------------------|------------------|-----------------------|------------------|------------------|
| | Mean ± Std. Dev. | Mean ± Std. Dev. | Mean ± Std. Dev. (mm) | Mean ± Std. Dev. | Mean ± Std. Dev. |
| Unet | 98.09% (± 0.96%) | 96.27% (± 1.83%) | 4.88 (± 1.70) | 98.47% (± 1.45%) | 97.80% (± 1.74%) |
| Res-Unet | 98.09% (± 0.88%) | 96.26% (± 1.68%) | 4.91 (± 1.74) | 98.28% (± 1.53%) | 97.99% (± 1.54%) |
| Att-Unet | 97.64% (± 1.61%) | 95.43% (± 2.85%) | 5.66 (± 2.01) | 98.62% (± 1.15%) | 96.83% (± 3.01%) |
| Att-Res-Unet | 98.03% (± 0.78%) | 96.14% (± 1.47%) | 5.24 (± 1.72) | 98.14% (± 1.27%) | 97.99% (± 1.64%) |

Table 5.2: Metric based evaluation of U-Net architectures

5.2 Aortic elasticity based evaluation

5.2.1 Compliance

Initially, the division of the quadrants was based on 90 degrees of horizontal and vertical splitting, dividing the aorta into four surfaces. This approach results in highly oscillated compliance curves per quadrant, as shown in figure 5.2(a). With a further examination of the work of Siyu et al. [35], we decide to rotate the quadrants 45 degrees clockwise, as displayed in figure 5.2(b). This second approach derives smooth compliance curves without oscillation.

Figures 5.3 and 5.4 demonstrate two examples of automatic segmented contours divided into four quadrants. Figure 5.3 presents a case of the dilated aorta where both global and local compliance curves are flat with low variation, suggesting a low elasticity value. On the other hand, figure 5.4 demonstrates a case of a non-dilated aorta, where the area values are initially low, expanding rapidly in the first five slices and decrease afterwards, as displayed in figure 5.4

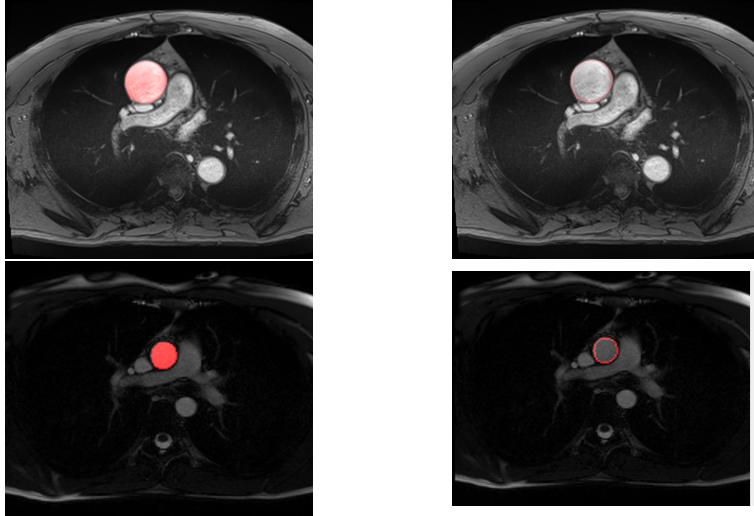


Figure 5.1: Examples of automatic segmentations, a dilated aorta is presented on top images and a non-dilated in bottom ones. On the left, all the aortic surface is colored, in contrast to the right, where only the aortic wall is colored.

(a). Similar behavior occurs in each quadrant for the local compliance curves in figure 5.4 (b).

To quantify the relationship between our model's results and the expected results, in table 5.3 we describe the statistical correlation between the ground truth and predicted values of min-max surface aortic areas, where we observe a significantly high correlation. Instead, a low correlation can be observed for compliance between ground truth and our output for dilated aorta, calculated by the formula described in equation 2.1.

| | Non-dilated aorta | | Dilated aorta | |
|-------------------|--------------------------|-------------|----------------------|-------------|
| | Correlation | P-Value | Correlation | P-Value |
| Min Area | 99% | $< 10^{-5}$ | 99% | $< 10^{-5}$ |
| Max Area | 99% | $< 10^{-5}$ | 99% | $< 10^{-5}$ |
| Compliance | 93% | $< 10^{-5}$ | 66% | $< 10^{-5}$ |

Table 5.3: Correlation between ground truth and calculated values of min-max areas and compliance. The table consists of two different multi-cases datasets of dilated and non-dilated aortas.

The evaluation of compliance, minimal, and maximal surfaces are between the prediction

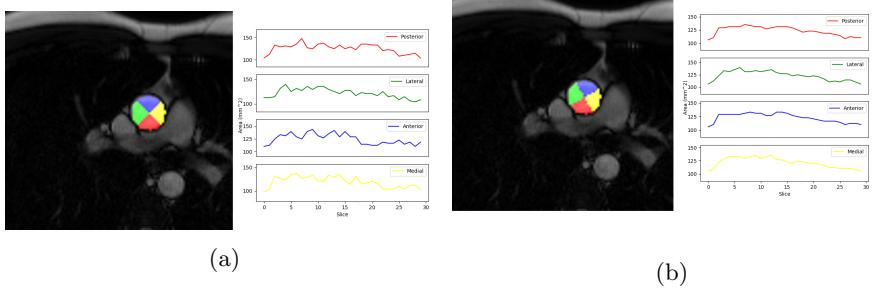


Figure 5.2: Comparison of non-rotated and rotated quadrants, with local compliance curves.

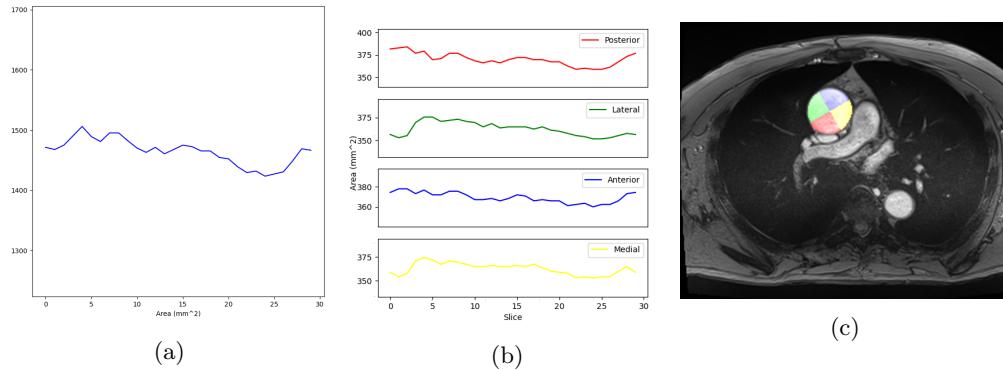


Figure 5.3: Example of automatic segmentation of dilated aorta divided by quadrants (c), accompanied by both the global (a) and local (b) compliance curves. Local compliance is presented with 4 different colors for each quadrant.

and ground truth is summarized below in three Bland Altman plots 5.5,5.6,5.7. For the compliance of the dilated aorta dataset in figure 5.5 (a), the average error rate is high (0.78 mmHg) concerning the range of values from 0.55 to 4 mmHg . However, for the non-dilated aorta dataset in figure 5.5 (b), the average error is poor (0.17 mmHg). In addition, the range of compliance specified in the X-axis is from 1.3 to 4 mmHg , so an error of size 0.17 mmHg is insignificant, resulting in higher accuracy from the non-dilated aorta dataset.

Concerning the maximum area in figure 5.7 the plots are similar for both datasets with low average error compared to the range of the maximum area. Concerning the minimum area, the average error is low (61.8 mm^2) on the non-dilated aorta dataset in figure 5.6 (b). But significant low value (19.33 mm^2) can be observed for the dilated aorta dataset in figure 5.6 (a).

The above results indicate a high compliance correlation with the ground truth from non-

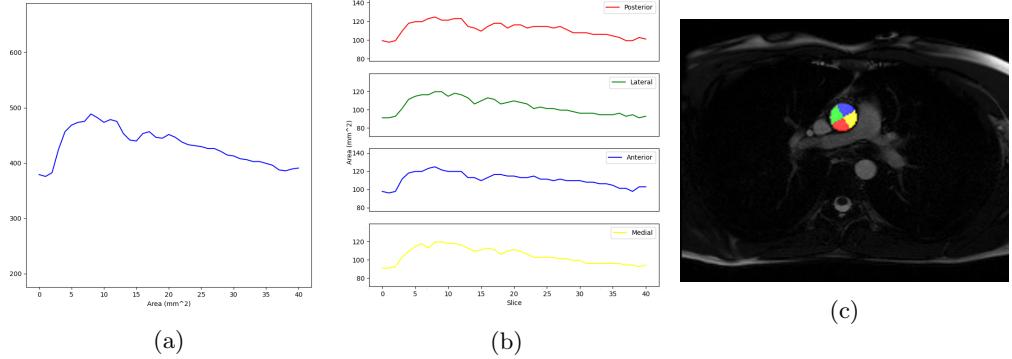


Figure 5.4: Example of automatic segmentation of non-dilated aorta divided by quadrants (c), accompanied by both the global (a) and local (b) compliance curves. Local compliance is presented with 4 different colors for each quadrant.

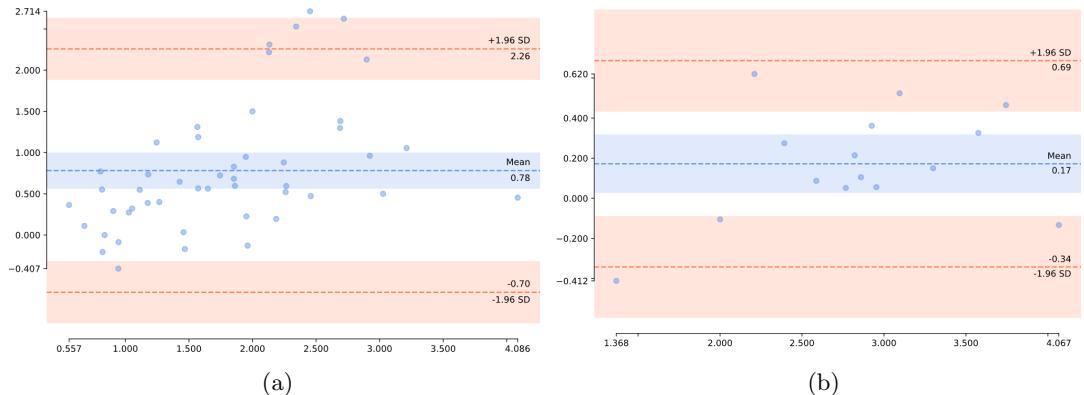


Figure 5.5: Bland Altman Plot of compliance for dilated (a) and non-dilated (b) aorta.

dilated aorta images compared to non-dilated ones. Likewise, as it concerns the surface areas, the results are credible, so our method seems to be robust.

As it concerns the compliance of different quadrants of the aortic wall, table 5.4 suggests that compliance is equal on average for each quadrant. Although, correlation of compliance is significantly higher for non-dilated aortas.

An explanation of low performance in compliance calculation is based on the method that medical experts use to extract the ground truth. When outlier images occur in the sequence, they discard them according to the compliance curve. A contribution of this work is the GUI that provides medicals a tool to calculate the local compliance per quadrant. Using this GUI,

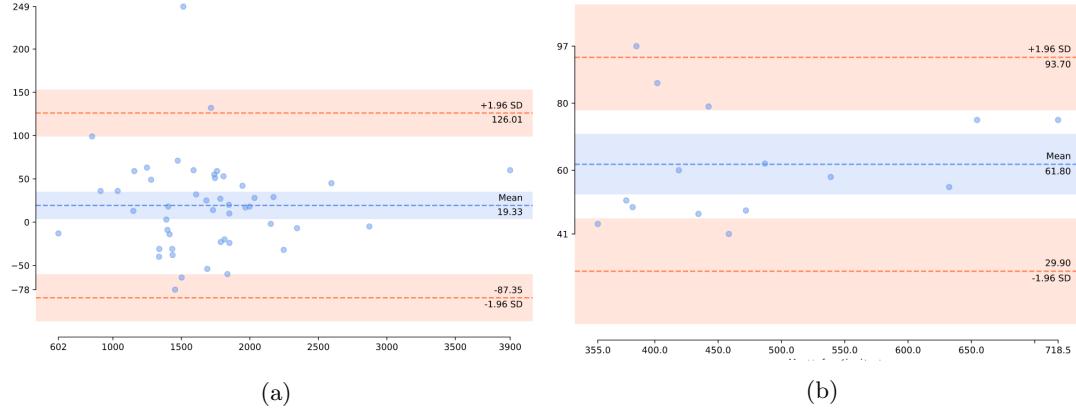


Figure 5.6: Bland Altman Plot of Minimum surface area for dilated (a) and non-dilated (b) aorta.

| | Compliance Mean \pm Std. Dev. (mmHg) | | | |
|-------------------|---|---------------------------|---------------------------|--------------------|
| | Posterior | Lateral | Anterior | Medial |
| Non-dilated aorta | 0.73 (\pm 0.13) | 0.74 (\pm 0.12) | 0.74 (\pm 0.12) | 0.73 (\pm 0.12) |
| Dilated aorta | 0.55 (\pm 0.23) | 0.57 (\pm 0.25) | 0.54 (\pm 0.24) | 0.55 (\pm 0.26) |

Table 5.4: Average compliance and standard deviation for each one of the four quadrants of dilated and non-dilated aorta datasets.

a medical expert performs segmentations for 47 dilated cases and discards the outliers, as described above, to compute the local compliance with more specificity. Table 5.5 summarize the results of this experiment, compared with the results provided by a non-expert (without discarding the outliers).

| | Compliance Mean \pm Std. Dev. (mmHg) | | | |
|------------------|---|---------------------------|--------------------|--------------------|
| | Posterior | Lateral | Anterior | Medial |
| Dilated aorta | | | | |
| With outliers | 0.55 (\pm 0.23) | 0.57 (\pm 0.25) | 0.54 (\pm 0.24) | 0.55 (\pm 0.26) |
| Without outliers | 0.51 (\pm 0.22) | 0.56 (\pm 0.25) | 0.51 (\pm 0.22) | 0.54 (\pm 0.27) |

Table 5.5: Average compliance of dilated aortas per quadrant, before and after discarding outliers images containing noise. A medical expert has implemented the discarding process.

In addition, without the outliers, the compliance correlation between ground truth and predictions have increased from 0.66% to 0.80% (p-value 0.004). Even though the results from the

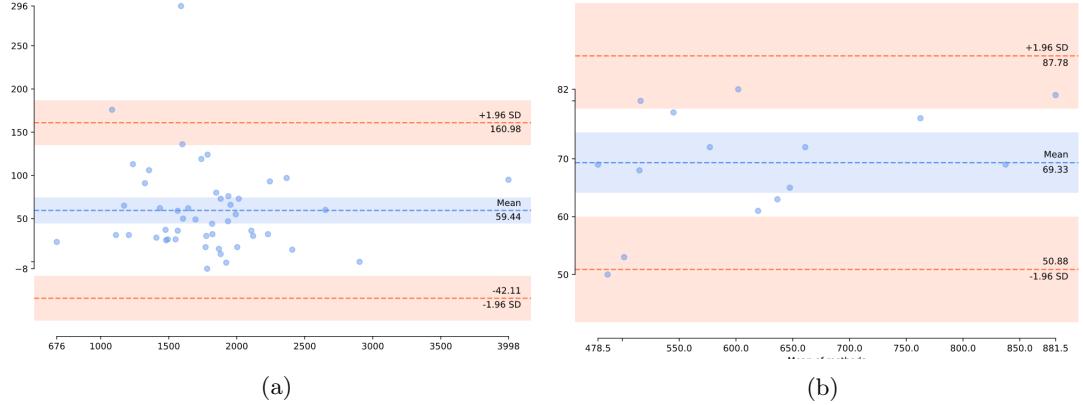


Figure 5.7: Bland Altman Plot of Maximum surface area for dilated (a) and non-dilated (b) aorta.

expert suggest a difference in the average local compliance, the standard deviation is still significantly high, so there is no robust outcome. Compared to Maximum Young's Modulus data from [35], based on 23 common cases along with the studies, there is no correlation between our results. According to Siyu et al. [35] lateral quadrant is the stiffest, while the medial quadrant has the lowest stiffness value as shown in table 5.6. This conclusion does not correspond to our results, where the lateral quadrant has the highest compliance, implying the lowest stiffness.

| | Posterior | Lateral | Anterior | Medial |
|--|-------------------|-------------------------------------|--------------------|--------------------|
| Young's Modulus Mean \pm Std. Dev. (MPa) | 0.99 (\pm 0.3) | 1.18 (\pm 0.46) | 0.77 (\pm 0.23) | 0.62 (\pm 0.18) |

Table 5.6: Maximum Young's Modulus for different quadrants from the work of Siyu et al. [35], the highest the value, the highest the stiffness of aortic wall.

5.2.2 Strain

The local compliance is a questionable measurement. This may occur because the division implementation uses the center of gravity to divide the aorta, dividing the area into nearly symmetrical and equal surfaces. The aorta is also moving, so the center of gravity can change one image to another. Considering the previous assumptions, the measurement of local compliance may not be credible. As far as we acknowledge, there is no other process to divide the aorta into quadrants except considering the center of gravity in a dynamic series of images. A possible solution for this problem is to calculate the local strain of each quadrant's wall instead

of local compliance. To implement that, the local outer perimeter of each quadrant is required to quantify the difference between quadrants.

The aforementioned local outer perimeters for each quadrant are visualized in figure 5.8, with different colors. The results of the local strain are reported in table 5.7.

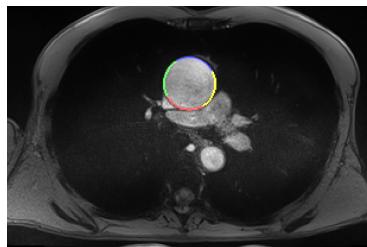


Figure 5.8: Visualization of perimeter per quadrant with different colors.

| | Strain Mean \pm Std. Dev. | | | |
|-------------------|--------------------------------|---------------------|---------------------|---------------------|
| | Posterior | Lateral | Anterior | Medial |
| Non-dilated aorta | 0.266 (\pm 0.07) | 0.25 (\pm 0.07) | 0.246 (\pm 0.05) | 0.249 (\pm 0.09) |
| Dilated aorta | 0.112 (\pm 0.08) | 0.093 (\pm 0.05) | 0.084 (\pm 0.05) | 0.086 (\pm 0.04) |

Table 5.7: Comparison of perimeter's strain between non-dilated, and dilated aorta datasets.

Regarding the non-dilated aorta in table 5.7, we observe similar strain along with the quadrants, with minor superiority of posterior and lateral over the rest of the quadrants. Contrariwise, for the non-dilated aorta, the posterior and lateral quadrants have undoubtedly higher strain values. According to the results of Siyu et al. [35] in table 5.8, we observe maximum stiffness in the lateral quadrant, in contrast with our results suggest that the posterior has maximum stiffness. Although comparing the quadrants in couples, we observe that posterior and lateral quadrants are the stiffest between both works. Similarly, anterior and medial quadrants have the minimum stiffness.

| Dilated aorta | Posterior | Lateral | Anterior | Medial |
|---|--------------------------------------|--------------------------------------|---------------------|---------------------|
| Young's Modulus Mean \pm Std. Dev. (MPa) [35] | 0.99 (\pm 0.3) | 1.18 (\pm 0.46) | 0.77 (\pm 0.23) | 0.62 (\pm 0.18) |
| Strain Mean \pm Std. Dev. [35] | 0.339 (\pm 0.06) | 0.380 (\pm 0.05) | 0.336 (\pm 0.06) | 0.319 (\pm 0.05) |
| Strain (ours) Mean \pm Std. Dev. | 0.112 (\pm 0.08) | 0.093 (\pm 0.05) | 0.084 (\pm 0.05) | 0.086 (\pm 0.04) |

Table 5.8: Comparison of strain and Young's Modulus between our results and the results from the work of Siyu et al. [35]. We use only 22 cases from the ex-vivo dataset for this table, where the patients are common with our dilated dataset.

Chapter 6

Conclusion

6.1 Discussion

This dissertation aims to investigate the elastic properties of the aorta. To do so, we locate and segment the ascending aorta at the level of the pulmonary trunk from 2D axial MRIs. The aorta wall undoubtedly exhibits changes in stiffness behavior based on localization [35]; this means that taking into account four separate quadrants should result in varying compliance. Maximum Young's Modulus, a local assessment of elasticity from ex-vivo tests, is compared with the local compliance of each quadrant to examine the possible association between them. Ex-vivo tests involve obtaining aorta samples during surgery and testing them in a bi-axial stretching test to determine their elastic characteristics. This comparison aims to evaluate if our proposed non-invasive approach for calculating aortic elasticity is accurate. Manual aorta segmentation adds intra-observer variability and bias error, and it is a time-consuming task. To address the challenges outlined above, the segmentation process must be automated.

One of the primary objectives of this research is to develop a robust automated system for two-dimensional ascending aorta segmentation at the pulmonary trunk level utilizing deep learning image processing methods. We assess the local compliance concerning four different quadrants of the aorta to determine if different elasticity behavior occurs along with them. Firstly, we automatically segment the aorta with a high accuracy rate, and then a division into four quadrants is performed. The center of gravity is specified for each slice to implement the division. Afterward, local compliance for each quadrant is calculated based on minimal and maximal surface areas from all the slices of the exam. As displayed in table 5.4, the average surface area of each quadrant appears to be equal, implying no significant difference in elastic-

ity between quadrants. In contrast with the work of Siyu et al. [35], where the results imply different stiffness between the quadrants.

According to the metric-based results of the previous chapter, the proposed automated method is robust; nevertheless, when the experiments include images with dilated aortas, the compliance measurement is susceptible to outliers. Even if the predicted surfaces of the minimal and maximal areas are precise, and the average error is poor, the predicted compliance is not accurate, stating that minor pixel-level differences can lead to a high difference in compliance level.

Another limitation introduced by the acquisition modality, where 3 Tesla MRIs have a high signal-to-noise ratio. As a result of this noise, some slices of the MRI sequence introduce artifacts, as shown in figure 4.3 due to rapid or turbulent blood flow. Segmentation is becoming a complex task because noise makes the aortic wall less visible. Besides, it is possible in the series of images to experience different noise ratios on different slices. It is observed that, on average, in the middle of the slice sequence, the noise is higher due to the fast blood flow during the systole phase. Usually, the aorta dilation reaches the maximal area at this period of the cardiac cycle, as a result those images are the most important ones for the computation of compliance, defined in equation 2.1. This formula is based on the difference of minimal-maximal surface area, so the compliance becomes very sensitive to image noise. When an image contains high noise, the confidence of the prediction is lower (the aortic wall contours are not well-defined), so global compliance introduces high error.

| Dilated aorta | Posterior | Lateral | Anterior | Medial |
|---|--------------------------------------|--------------------------------------|---------------------|-------------------------------------|
| Young's Modulus Mean \pm Std. Dev. (MPa) [35] | 0.99 (\pm 0.3) | 1.18 (\pm 0.46) | 0.77 (\pm 0.23) | 0.62 (\pm 0.18) |
| Strain Mean \pm Std. Dev. [35] | 0.339 (\pm 0.06) | 0.380 (\pm 0.05) | 0.336 (\pm 0.06) | 0.319 (\pm 0.05) |
| Strain (ours) Mean \pm Std. Dev. | 0.112 (\pm 0.08) | 0.093 (\pm 0.05) | 0.084 (\pm 0.05) | 0.086 (\pm 0.04) |
| Compliance (ours) Mean \pm Std. Dev. (mmHg) | 0.51 (\pm 0.22) | 0.56 (\pm 0.25) | 0.51 (\pm 0.22) | 0.54 (\pm 0.27) |

Table 6.1: Comparison of elastic properties between our results and the results from the work of Siyu et al. [35]. We use only 22 cases from the ex-vivo dataset for this table, where the patients are common with our dilated dataset.

Local compliance is a questionable metric. Dividing the aorta, using the center of gravity, leads into nearly symmetrical and equal surfaces for each quadrant. In advance, the center of gravity can shift from one image to another, because the aorta is moving. Given the preceding assumptions, measuring local compliance may not be credible. As far as we know, the only

way to divide the aorta into quadrants is to consider the center of gravity in a series of images. Instead of calculating local compliance, one possible alternative is to calculate the local strain of each quadrant.

In table 6.1, we can observe a high correlation between our results and the results of Siyu et al. [35], although it is evident that the two different methods of image processing and ex-vivo biaxial testing, yield different range values for aortic wall strain. Considering that there is no information available for the wall thickness in our study, we can not handle the abovementioned problem. Nevertheless, some critical observations are possibly extracted from this comparison. Both studies imply that posterior and lateral aortic quadrants have higher strain than anterior and medial quadrants. Strain can be defined as described in section 2.4. Assuming that strain is positively correlated with Maximum Young's Modulus, we can anticipate that high strain values correspond to high stiffness values. The lateral and posterior quadrants are the stiffest, while the medial and anterior quadrants have the lowest aortic stiffness. According to strain and MYM results of Siyu et al. [35], it can be observed that the lateral quadrant has the lowest aortic stiffness, in correspondence with our results. Nevertheless, our results suggest the posterior quadrant as the stiffest, in contrast with the ex-vivo results, which suggest the lateral as the stiffest quadrant. Eventually, it can be observed that the highest compliance corresponds to the highest young modulus on the lateral quadrant.

6.2 Future Work

One suggestion to improve the performance of segmentation is to create synthetic images, extending the dataset by adding Gaussian noise and artifacts, as suggested by [41]. Gaussian noise is added gradually, as visualized in image 6.2 from top to bottom. Similarly, wave-shaped artifacts were added gradually at different opacity levels from left to right. This approach multiplies the size of the dataset by sixteen, resulting in 35024 images instead of 2189 in the initial dataset. Of course, this method requires a longer training time, but it may improve segmentation results, as shown in figure 6.3. The first metrics-based evaluation of this suggestion results in poor performance when compared with the non-augmented dataset, as shown in table 6.2.

| U-Net | Dice coeff | IoU | Hausdorff | Precision | Recall |
|----------------------|-----------------------|-----------------------|---------------------------|-----------------------|-----------------------|
| | Mean \pm Std. Dev. | Mean \pm Std. Dev. | Mean \pm Std. Dev. (mm) | Mean \pm Std. Dev. | Mean \pm Std. Dev. |
| No-augmented Dataset | 98.09% (\pm 0.96%) | 96.27% (\pm 1.83%) | 4.88 (\pm 1.70) | 98.47% (\pm 1.45%) | 97.80% (\pm 1.74%) |
| Augmented Dataset | 97.82% (\pm 2.13%) | 95.81% (\pm 3.53%) | 5.56 (\pm 2.76) | 98.12% (\pm 3.42%) | 97.62% (\pm 1.59%) |

Table 6.2: Metric based evaluation of noise augmented and non-augmented U-Net models.

Despite the poor evaluation metrics, a visual comparison of those models in figure 6.3 implies

that when we augment the dataset, the predictions are less noise-sensitive and more robust. Nevertheless, in some other cases, the predictions are inaccurate, as displayed in figure 6.1. This hypothesis must be studied and proven in more detail. These results are preliminary, and there is a need for further investigation of this approach. To compare the new model results with clinical data, a comparison with ex-vivo results should be implemented (as presented in this work).

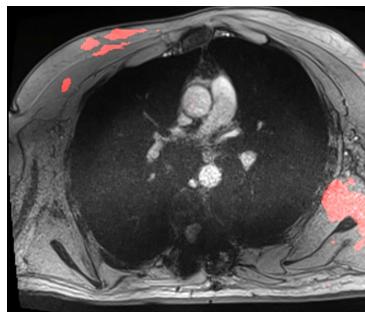


Figure 6.1: Inaccurate prediction using noise augmented dataset.

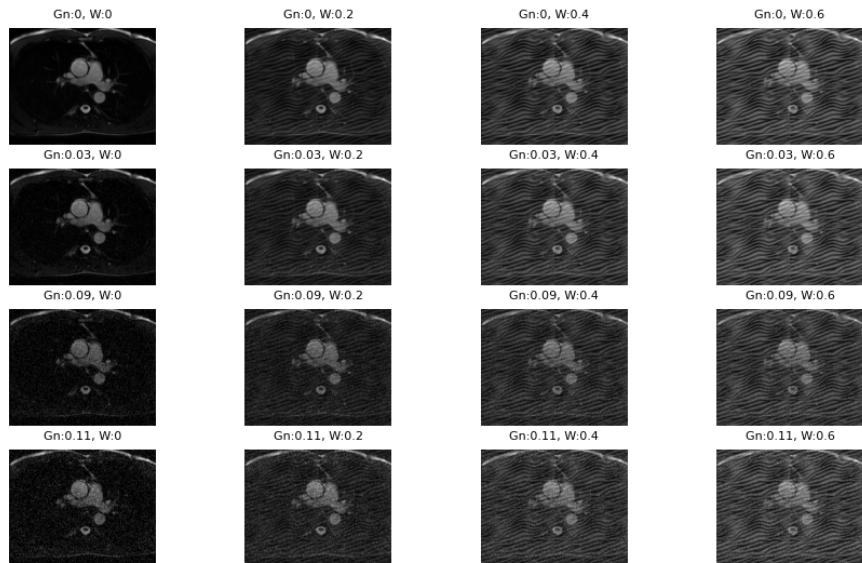


Figure 6.2: A preview of the noise & wave augmented dataset. Noise is added gradually from top to bottom, while wave shaped artifacts are added in different opacity levels from left to right.

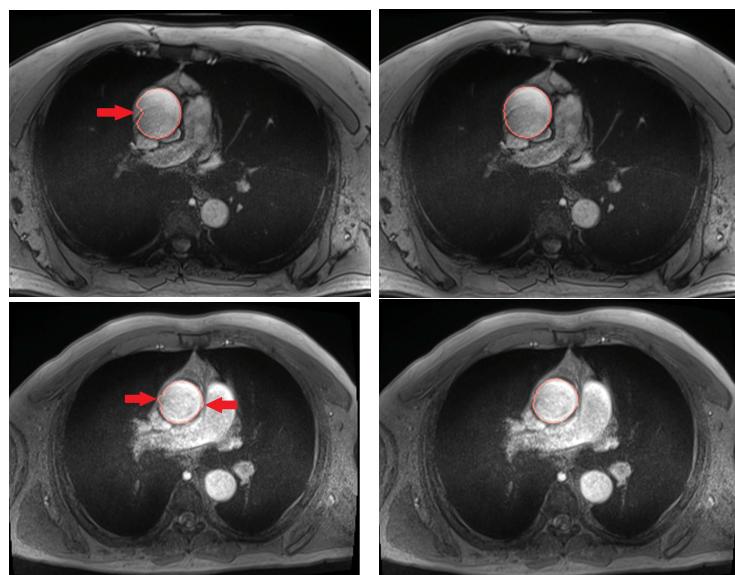


Figure 6.3: A visual comparison of U-Net predictions of two cases trained with non-augmented dataset on the left and augmented dataset on the right.

Bibliography

- [1] P. G. Agache, C. Monneur, J. L. Leveque, and J. De Rigal. Mechanical properties and young's modulus of human skin in vivo. *Archives of Dermatological Research*, 269(3):221–232, December 1980.
- [2] Orhan Akal, Zhigang Peng, and Gerardo Hermosillo Valadez. Combonet: Combined 2d amp; 3d architecture for aorta segmentation, 2020.
- [3] Md. Zahangir Alom, Mahmudul Hasan, Chris Yakopcic, Tarek M. Taha, and Vijayan K. Asari. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *CoRR*, abs/1802.06955, 2018.
- [4] Wenjia Bai, Hideaki Suzuki, Chen Qin, Giacomo Tarroni, Ozan Oktay, Paul M. Matthews, and Daniel Rueckert. Recurrent neural networks for aortic image sequence segmentation with sparse annotations. *CoRR*, abs/1808.00273, 2018.
- [5] Soos MP Bamalan OA. Anatomy, thorax, heart great vessels.
- [6] Gustav G Belz. Elastic properties and windkessel function of the human aorta. *Cardiovascular Drugs and Therapy*, 9(1):73–83, 1995.
- [7] Haben Berhane, Michael Scott, Mohammed Elbaz, Kelly Jarvis, Patrick McCarthy, James Carr, Chris Malaisrie, Ryan Avery, Alex J. Barker, Joshua D. Robinson, Cynthia K. Rigsby, and Michael Markl. Fully automated 3d aortic segmentation of 4d flow mri for hemodynamic analysis using deep learning. *Magnetic Resonance in Medicine*, 84(4):2204–2218, 2020.
- [8] Alexander Buslaev, Vladimir I. Iglovikov, Eugene Khvedchenya, Alex Parinov, Mikhail Druzhinin, and Alexandr A. Kalinin. Albumentations: Fast and flexible image augmentations. *Information*, 11(2), 2020.
- [9] Xiaocong Chen, Lina Yao, and Yu Zhang. Residual attention u-net for automated multi-class segmentation of covid-19 chest ct images, 2020.

- [10] Junlong Cheng, Shengwei Tian, Long Yu, Xiang Ma, and Yan Xing. A deep learning algorithm using contrast-enhanced computed tomography (CT) images for segmentation and rapid automatic detection of aortic dissection. *Biomedical Signal Processing and Control*, 62:102145, September 2020.
- [11] Jennifer C.-Y. Chung, Edwin Wong, Mingyi Tang, Daniella Eliathamby, Thomas L. Forbes, Jagdish Butany, Craig A. Simmons, and Maral Ouzounian. Biomechanics of aortic dissection: A comparison of aortas associated with bicuspid and tricuspid aortic valves. *Journal of the American Heart Association*, 9(15), August 2020.
- [12] IBM Cloud Education. What are neural networks? accessed on 2022-05-16.
- [13] Albert Comelli, Navdeep Dahiya, Alessandro Stefano, Viviana Benfante, Giovanni Gentile, Valentina Agnese, Giuseppe M. Raffa, Michele Pilato, Anthony Yezzi, Giovanni Petrucci, and Salvatore Pasta. Deep learning approach for the segmentation of aneurysmal ascending aorta. *Biomedical Engineering Letters*, 11(1):15–24, November 2020.
- [14] Godwin Constantine, Kesavan Shan, Scott D Flamm, and Mohan U Sivananthan. Role of MRI in clinical cardiology. *The Lancet*, 363(9427):2162–2171, June 2004.
- [15] Ryan R. Davies, Amy Gallo, Michael A. Coady, George Tellides, Donald M. Botta, Brendan Burke, Marcus P. Coe, Gary S. Kopf, and John A. Elefteriades. Novel measurement of relative aortic size predicts rupture of thoracic aortic aneurysms. *The Annals of Thoracic Surgery*, 81(1):169–177, 2006.
- [16] Jinming Duan, Ghalib Bello, Jo Schlemper, Wenjia Bai, Timothy J. W. Dawes, Carlo Biffi, Antonio de Marvao, Georgia Doumoud, Declan P. O'Regan, and Daniel Rueckert. Automatic 3d bi-ventricular segmentation of cardiac images by a shape-refined multi- task deep learning approach. *IEEE Transactions on Medical Imaging*, 38(9):2151–2164, September 2019.
- [17] A. Duprey, K. Khanafer, M. Schlicht, S. Avril, D. Williams, and R. Berguer. In vitro characterisation of physiological and maximum elastic modulus of ascending thoracic aortic aneurysms using uniaxial tensile testing. *European Journal of Vascular and Endovascular Surgery*, 39(6):700–707, June 2010.
- [18] Hakim Fadil, John J. Totman, Derek J. Hausenloy, Hee-Hwa Ho, Prabath Joseph, Adrian Fatt-Hoe Low, A. Mark Richards, Mark Y. Chan, and Stephanie Marchesseau. A deep learning pipeline for automatic analysis of multi-scan cardiovascular magnetic resonance. *Journal of Cardiovascular Magnetic Resonance*, 23(1), April 2021.

- [19] Pierre-Loïc Felter. Creating hemodynamic atlas of aorta. 2017.
- [20] A.F. Frangi, W.J. Niessen, and M.A. Viergever. Three-dimensional modeling for functional analysis of cardiac images, a review. *IEEE Transactions on Medical Imaging*, 20(1):2–5, 2001.
- [21] Y. C. Fung. What are the residual stresses doing in our blood vessels? *Annals of Biomedical Engineering*, 19(3):237–249, May 1991.
- [22] Edward H. Harris, Leon B. Walker, and Bennett R. Bass. Stress-strain studies in cadaveric human tendon and an anomaly in the young's modulus thereof. *Medical & Biological Engineering*, 4(3):253–259, May 1966.
- [23] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.
- [24] Ho Aik Hong and U. U. Sheikh. Automatic detection, segmentation and classification of abdominal aortic aneurysm using deep learning. In *2016 IEEE 12th International Colloquium on Signal Processing & Its Applications (CSPA)*. IEEE, March 2016.
- [25] Dimitrios C. Iliopoulos, Rejar P. Deveja, Eleftherios P. Kritharis, Despina Perrea, George D. Sionis, Konstantinos Toutouzas, Christodoulos Stefanidis, and Dimitrios P. Sokolis. Regional and directional variations in the mechanical properties of ascending thoracic aortic aneurysms. *Medical Engineering & Physics*, 31(1):1–9, January 2009.
- [26] Eric M. Isselbacher. Thoracic and abdominal aortic aneurysms. *Circulation*, 111(6):816–828, 2005.
- [27] Eric M. Isselbacher. *Diseases of the Aorta*, pages 681–690. Humana Press, Totowa, NJ, 2006.
- [28] Jean Jeudy and Charles S. White. Cardiac magnetic resonance imaging: Techniques and principles. *Seminars in Roentgenology*, 43(3):173–182, July 2008.
- [29] Girish Katti, Syeda Ara, and Dr Shireen. Magnetic resonance imaging (mri) - a review. *Intl J Dental Clin*, 3, 03 2011.
- [30] Eun Kyoung Kim, Seung Hyuk Choi, Kiick Sung, Wook Sung Kim, Yeon Hyeon Choe, Jae K Oh, and Duk-Kyung Kim. Aortic diameter predicts acute type a aortic dissection in patients with marfan syndrome but not in patients without marfan syndrome. *The Journal of thoracic and cardiovascular surgery*, 147(5):1505–1510, 2014.

- [31] J.H. Kinney, M. Balooch, S.J. Marshall, G.W. Marshall, and T.P. Weihs. Hardness and young's modulus of human peritubular and intertubular dentine. *Archives of Oral Biology*, 41(1):9–13, January 1996.
- [32] Roland Krug, Jan M Boese, and Lothar R Schad. Determination of aortic compliance from magnetic resonance images using an automatic active contour model. *Physics in Medicine & Biology*, 48(15):2391, 2003.
- [33] Alain Lalande, Philippe Khau van Kien, Nicolas Salve, DOURIËD BEN SALEM, Louis Legrand, Paul Michael Walker, Jean-Eric Wolf, and François Brunotte. Automatic determination of aortic compliance with cine-magnetic resonance imaging: an application of fuzzy logic theory. *Investigative radiology*, 37(12):685–691, 2002.
- [34] Eldon D Lehmann. Clinical value of aortic pulse-wave velocity measurement. *The Lancet*, 354(9178):528–529, 1999.
- [35] Siyu Lin, Marie Catherine Morgant, Diana M. Marín-Castrillón, Paul M. Walker, Ludwig Serge Aho Glélé, Arnaud Boucher, Benoit Presles, Olivier Bouchot, and Alain Lalande. Aortic local biomechanical properties in the ascending aortic aneurysms (under review). *ACTA Biomaterialia*, 2022.
- [36] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. *CoRR*, abs/1411.4038, 2014.
- [37] Agnes Lydia and Sagayaraj Francis. Adagrad—an optimizer for stochastic gradient descent. *Int. J. Inf. Comput. Sci.*, 6(5):566–568, 2019.
- [38] Tianling Lyu, Guanyu Yang, Xingran Zhao, Huazhong Shu, Limin Luo, Duanduan Chen, Jiang Xiong, Jian Yang, Shuo Li, Jean-Louis Coatrieux, and Yang Chen. Dissected aorta segmentation using convolutional neural networks. *Computer Methods and Programs in Biomedicine*, 211:106417, 2021.
- [39] Rosa-María Menchón-Lara, Federico Simmross-Wattenberg, Pablo Casaseca de-la Higuera, Marcos Martín-Fernández, and Carlos Alberola-López. Reconstruction techniques for cardiac cine MRI. *Insights into Imaging*, 10(1), September 2019.
- [40] Jordi Minnema, Jan Wolff, Juha Koivisto, Felix Lucka, Kees Joost Batenburg, Tymour Forouzanfar, and Maureen van Eijnatten. Comparison of convolutional neural network training strategies for cone-beam CT image segmentation. *Computer Methods and Programs in Biomedicine*, 207:106192, August 2021.

- [41] J. Mitéran, O. Bouchot, A. Cochet, and A. Lalande. Automatic determination of aortic compliance based on mri and adapted curvilinear detector. *Biomedical Signal Processing and Control*, 40:295–311, 2018.
- [42] Marie Catherine Morgant, Siyu Lin, Chloe Bernard, Diana M. Marín-Castrillón, Aline Laubriet, Arnaud Boucher, Benoit Presles, Alain Lalande, and Olivier Bouchot. Elastic properties of ascending aorta evaluated by imaging and biaxial testing (under review).
- [43] Marie-Catherine Morgant, Siyu Lin, Diana Marin-Castrillon, Chloé Bernard, Aline Laubriet, Alexandre Cochet, Alain Lalande, and Olivier Bouchot. Comparison of two techniques (in vivo and ex-vivo) for evaluating the elastic properties of the ascending aorta: Prospective cohort study. *PLOS ONE*, 16(9):e0256278, September 2021.
- [44] Giorgio De Nunzio, Rosella Cataldo, and Alessandra Carlà. Robust intensity standardization in brain magnetic resonance images. *Journal of Digital Imaging*, 28:727–737, 2015.
- [45] Ozan Oktay, Jo Schlemper, Loïc Le Folgoc, Matthew C. H. Lee, Matthias P. Heinrich, Kazunari Misawa, Kensaku Mori, Steven G. McDonagh, Nils Y. Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention u-net: Learning where to look for the pancreas. *CoRR*, abs/1804.03999, 2018.
- [46] Michael F O’Rourke, Jan A Staessen, Charalambos Vlachopoulos, Daniel Duprez, and Gérard E Plante. Clinical applications of arterial stiffness; definitions and reference values. *American Journal of Hypertension*, 15(5):426–444, 2002.
- [47] Linda A Pape, Thomas T Tsai, Eric M Isselbacher, Jae K Oh, Patrick T O’Gara, Arturo Evangelista, Rossella Fattori, Gabriel Meinhardt, Santi Trimarchi, Eduardo Bossone, et al. Aortic diameter 5.5 cm is not a good predictor of type a aortic dissection: observations from the international registry of acute aortic dissection (irad). *Circulation*, 116(10):1120–1127, 2007.
- [48] Cameron Patterson. *Managing a real-time massively-parallel neural architecture*. PhD thesis, 01 2012.
- [49] Michael P. Pfeiffer and Robert W.W. Biederman. Cardiac MRI. *Medical Clinics of North America*, 99(4):849–861, July 2015.
- [50] Payam Refaeilzadeh, Lei Tang, and Huan Liu. *Cross-Validation*, pages 1–7. Springer New York, New York, NY, 2016.
- [51] Rocchini. Hausdorff distance sample, accessed 30 may 2022. 2007.

- [52] R. Tyrrell Rockafellar. *pp. 11–12.* Princeton University Press, 1970.
- [53] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.
- [54] Jean-Loïc Rose, Alain Lalande, El-Bey Bourennane, Paul Michael Walker, Olivier Bouchot, Eric Steinmetz, Louis Legrand, Yvon Voisin, Y.E. Wolf, and F. Brunotte. Automatic detection of vessel wall contours from cine-mri for aortic compliance determination. *Computers in Cardiology, 2005*, pages 411–414, 2005.
- [55] Jean-Loïc Rose, Alain Lalande, Olivier Bouchot, El-Bey Bourennane, Paul M. Walker, Patricia Ugolini, Chantal Revol-Muller, Raymond Cartier, and François Brunotte. Influence of age and sex on aortic distensibility assessed by mri in healthy subjects. *Magnetic Resonance Imaging*, 28(2):255–263, 2010.
- [56] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes. Non-rigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
- [57] Hajime Sakuma, Kan Takeda, and Charles B. Higgins. Fast magnetic resonance imaging of the heart. *European Journal of Radiology*, 29(2):101–113, February 1999.
- [58] Nema Salem, Hebatullah Malik, and Asmaa Shams. Medical image enhancement based on histogram algorithms. *Procedia Computer Science*, 163:300–311, 2019. 16th Learning and Technology Conference 2019Artificial Intelligence and Machine Learning: Embedding the Intelligence.
- [59] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv 1409.1556*, 09 2014.
- [60] E SJONTOFT and C EDMUND. determination of young's modulus for the human cornea. *Bulletin of Mathematical Biology*, 49(2):217–232, 1987.
- [61] Marijn F. Stollenga, Wonmin Byeon, Marcus Liwicki, and Jürgen Schmidhuber. Parallel multi-dimensional lstm, with application to fast biomedical volumetric image segmentation. *CoRR*, abs/1506.07452, 2015.
- [62] Yaping Wu, Guohua Zhao, Meiyun Wang, Wenyi Gao, Qian Zhang, and Yusong Lin. Automatic histogram specification for glioma grading using multicenter data. *Journal of Healthcare Engineering*, 2019:1–12, 12 2019.

- [63] Yan Xu, Yang Li, Yipei Wang, Mingyuan Liu, Yubo Fan, Maode Lai, I Eric, and Chao Chang. Gland instance segmentation using deep multichannel neural networks. *IEEE Transactions on Biomedical Engineering*, 64(12):2901–2912, 2017.
- [64] Yue Xuan, Andrew D. Wisneski, Zhongjie Wang, Matthew Lum, Shalni Kumar, Julia Pallone, Nick Flores, Justin Inman, Lilian Lai, Joanna Lin, Julius M. Guccione, Elaine E. Tseng, and Liang Ge. Regional biomechanical and failure properties of healthy human ascending aorta and root. *Journal of the Mechanical Behavior of Biomedical Materials*, 123:104705, November 2021.
- [65] Zhengxin Zhang, Qingjie Liu, and Yunhong Wang. Road extraction by deep residual u-net. *CoRR*, abs/1711.10684, 2017.
- [66] Guohua Zhao, Jie Bai, Peipei Wang, Guan Yang, Lei Shi, Yongcai Tao, Yusong Lin, and Jingliang Cheng. Hs-gs: A method for multicenter mr image standardization. *IEEE Access*, PP:1–1, 08 2020.
- [67] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2881–2890, 2017.
- [68] Xiangrong Zhou, Ryosuke Takayama, Song Wang, Takeshi Hara, and Hiroshi Fujita. Deep learning of the sectional appearances of 3d CT images for anatomical structure segmentation based on an FCN voting method. *Medical Physics*, 44(10):5221–5233, August 2017.
- [69] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. *CoRR*, abs/1807.10165, 2018.
- [70] Fangyu Zou, Li Shen, Zequn Jie, Weizhong Zhang, and Wei Liu. A sufficient condition for convergences of adam and rmsprop. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11127–11135, 2019.