UNIVERSITÄT ZU LÜBECK

Aus dem Institut für Robotik und Kognitive Systeme
der Universität zu Lübeck
Direktor: Prof. Dr.-Ing. Achim Schweikard

Pseudohealthy Synthesis for Personalized Cardiovascular Prosthesis Shaping

Inauguraldissertation
zur
Erlangung der Doktorwürde
der Universität zu Lübeck

Aus der Sektion Informatik/Technik

vorgelegt von
Jannis Hagenah
aus Hildesheim

Lübeck, 2021

1. Berichterstatter: Prof. Dr. rer. nat. habil. Floris Ernst

2. Berichterstatter: Prof. Dr.-Ing. Erhardt Barth

Tag der mündlichen Prüfung: 17.12.2021

Zum Druck genehmigt. Lübeck, den 05.01.2022

Für Samuel

Mein klitzekleines Puzzleteil, damit die Welt,
in der du leben wirst, eine bessere sein wird.

Seht ihr den Mond dort stehen?
Er ist nur halb zu sehen
und ist doch rund und schön.
So sind wohl manche Sachen
die wir getrost belachen
weil unsre Augen sie nicht sehn.

Matthias Claudius

Kurzfassung

Die Anatomie der einzelnen Teile des menschlichen Herzens ist von Mensch zu Mensch sehr unterschiedlich. Gemeinsam bilden sie ein komplexes, dynamisches System, das seine Leistungsfähigkeit durch ein präzises biomechanisches Zusammenspiel erreicht. Muss eins dieser Teile durch eine Prothese ersetzt werden, sollte die Prothese daher das komplexe Zusammenspiel weitgehend erhalten, indem sie die individuelle Form so gut wie möglich nachahmt. Kardiovaskuläre Prothesen sind bisher jedoch nicht an die individuelle Anatomie des Patienten angepasst, da die gewünschte Prothesenform oft nicht direkt aus medizinischen Bildern extrahiert werden kann. Dieses kann an einer unzureichenden Bildauflösung liegen, wie es bei Aortenklappenprothese der Fall ist, oder an einer pathologischen Verformung des Organs, z.B. bei einer Aortenwurzelprothese. In diesen Fällen enthält das medizinische Bild des zu ersetzenden Organs nur Surrogatinformationen über den ursprünglichen, gesunden Zustand, d.h. die gewünschte Prothesenform. Das Ziel dieser Arbeit besteht darin, diese methodische Lücke zu schließen und die gewünschte gesunde Organform auf Basis der verfügbaren Surrogatinformationen zu schätzen. Dazu wurden neuartige Methoden und Konzepte zur Synthese eines pseudogesunden Zustands entwickelt. Außerdem wurden ex-vivo Datensätze von Schweinen gesammelt, die Proof-of-Concept-Evaluationen ermöglichen. In dieser Arbeit konnte nicht nur gezeigt werden, dass die Surrogatdaten ausreichend Informationen beinhalten, um Merkmale der gewünschten Prothesenform vorherzusagen, sondern es konnten auch neuartige Methoden zur Synthese eines pseudogesunden Zustands entwickelt werden. Darüber hinaus werden Ansätze zur Einbeziehung des Chirurgen in die Entscheidungsfindung vorgestellt. Um die Hürden bei der Anwendung im klinischen Betrieb abzubauen, wurde zusätzlich ein Ansatz zur diskreten Bildsynthese entwickelt. Dabei konnte gezeigt werden, dass die vollständig datengetriebene Identifikation von typischen anatomischen Formen sowie die Klassifikation des individuell optimalen Formtyps möglich sind. Diese Arbeit ist die erste zur pseudogesunden Formsynthese für die personalisierte Prothetik sowie die erste Arbeit zur Personalisierung von kardiovaskulären Prothesen. Da die vorgestellten Methoden und Konzepte nicht auf kardiovaskuläre Anwendungen beschränkt sind, bietet sie die Basis für personalisierte Prothetik durch maschinelles Lernen.

Abstract

The anatomy of the heart and its components highly varies from patient to patient. Combined, they form a complex, dynamic system that reaches its efficiency through a precise biomechanical interplay that is highly influenced by the individual shapes. Hence, if one of these components has to be replaced, the prosthesis should preserve this complex interplay by mimicking the individual shape as closely as possible. However, state-of-the-art cardiovascular prostheses are not tailored to the patient's individual anatomy. One challenge in personalizing the shape of cardiovascular prostheses is that, typically, the desired prostheses shape cannot be extracted from medical images of the organ at risk. This might be due to insufficient image resolution, as it is the case for aortic valve prostheses, or due to a pathological deformation of the organ, e.g. in the case of aortic root prostheses. In these cases, the medical image of the pathological organ only contains surrogate information about the original, healthy state, i.e. the desired prosthesis' shape. The aim of this work is to bridge this gap by estimating the desired healthy organ shape based on the available surrogate information. Thus, novel methods and concepts for pseudohealthy synthesis for personalized prosthesis shaping are developed. Additionally, ex-vivo porcine data sets were collected, allowing for proof-of-concept evaluations. In this thesis, it could not only be shown that the surrogate data carries a sufficient amount of information to predict features of the desired prosthesis shape, but also novel methods for pseudohealthy synthesis could be developed to synthesize an image of this desired shape. Furthermore, approaches to include the surgeon into the decision making pipeline are presented. Since personalized tailoring of prostheses for each and every patient comes at high financial, logistical and regulatory cost, an approach for discrete pseudohealthy synthesis was developed. Therefore, it could be shown that the fully data-driven identification of a set of typical anatomical shape types as well as the classification of the individually optimal prosthesis type is possible. This thesis presents the first work on pseudohealthy synthesis for personalized prostheses shaping and the first work on personalizing cardiovascular prostheses in general. As the presented methods and concepts are not limited to cardiovascular applications, they lay the basis for personalized prosthetics utilizing machine learning.

Contents

List of Abbreviations

Acronym	Full Form
ADAM	Adaptive Moment Estimation
AHD	Angle-compensated Homogeneous Dilation
ASCD	Average Symmetric Contour Distance
CNN	Convolutional Neural Network
CT	Computed Tomography
CVAE	Conditional Variational Autoencoder
CVG	Composite Aortic valve Graft
DE	Direct Estimation
DM	Deformation Modeling
DOF	Degree of Freedom
FD	Free Deformation
GAN	Generative Adversarial Network
GMM	Gaussian Mixture Model
HD	Homogeneous Dilation
LED	Light-Emitting Diode
ML	Machine Learning
MLP	Multi-Layer Perceptron
MRI	Magnetic Resonance Imaging
NN	Neural Network
PCA	Principal Component Analysis
QP	Quadratic Programming
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
RF	Random Forest
RGB	Red-Green-Blue color model
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network

SSM	Statistical Shape Model
SVM	Support Vector Machine
SVR	Support Vector Regression
TAVI	Transcatheter Aortic Valve Implantation
TEE	Transesophageal Echocardiography
tSNE	t-distributed Stochastic Neighborhood Embedding
US	Ultrasound
VAE	Variational Autoencoder

Dataset Nomenclature

Symbol	Description
$\tilde{D}_S$	• Dataset for structural pathology • Computed 3D images of aortic root and leaflet shape images • 10 hearts
D_S	• Dataset for structural pathology • 3D images of aortic root and leaflet shape images • 30 hearts
D_S^{aux}	• Dataset for structural pathology • Leaflet shape images • 56 hearts • Also referred to as auxiliary dataset
D_M	• Dataset for morphological pathology • Computed 3D images of aortic root in healthy and pathological state • 24 hearts
$\tilde{D}_M$	• Dataset for morphological pathology • 2D slice images of aortic root in healthy and pathological state • Each slice image shows commissure plane • 24 hearts

1 Introduction

The field of personalized medicine gained importance during the last years [128]. The key idea is to accept that every body is individual and differs from other bodies, leading to the concept of tailoring medical treatment according to these individual features [74]. Besides applications in disciplines like oncology, one prominent example of personalized medicine is the personalized shaping of prostheses. When an organ or a part of an organ has to be replaced, the prosthesis should ideally mimic the original, healthy state of this organ as closely as possible to achieve optimal outcome for the patient.

Due to disruptive and promising advancements in the field of rapid manufacturing, fast production of unique prosthesis in 3D printers becomes more and more realistic [6], [138]. It is even possible to produce personalized 3D knitted scaffolds that can be cultivated using tissue engineering. Other research groups focus on direct 3D printing of biological tissue [38]. For example, it is possible to use these techniques for creating artificial aortic valves [68] or aortic root graft prostheses [76]. Hence, manufacturing personalized prostheses for each patient in a clinical scenario seems to be possible in the near future.

However, state-of-the-art cardiovascular prostheses do not aim at remodeling the individual shape. While biological aortic valve prosthesis assume that the valve and its leaflets is symmetric, which is barely the case in human [85], the shape mechanical aortic valve prostheses does not resemble a native valve at all [39]. Aortic root prostheses typically ignore the root's substantial anatomical geometry and are shaped as tubes [39]. The only parameter of these prostheses that is chosen individually is its diameter. This choice is either done using image processing tools to measure an optimal diameter [5] or without any computational guidance. The latter case is specifically common for valve-sparing aortic root reconstruction surgery, where the surgeon has to rely on his or her experience to choose the individually optimal diameter of the aortic root prosthesis.

This limited personalization and lack of decision support tools is mainly due to the limited capabilities of assessing the desired prosthesis shape. Typically, the individual shape of the part of the organ that should be replaced is acquired utilizing medical imaging. From these images, the shape of the structure can be extracted using image segmenta-

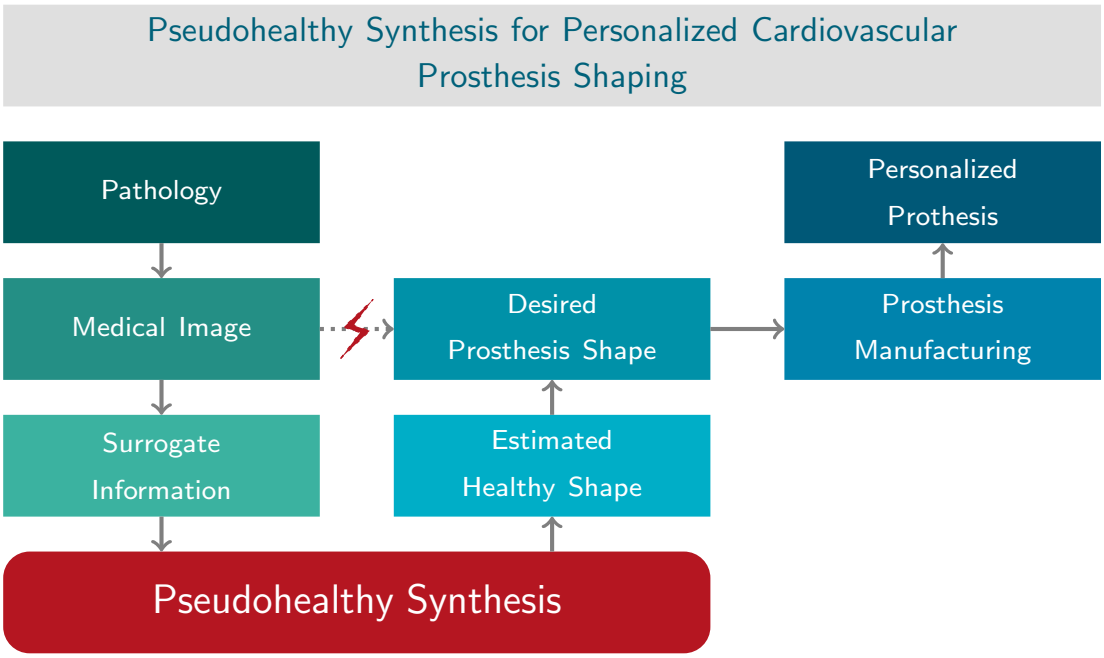

Figure 1.1: Workflow of personalized prosthesis shaping. Typically, the desired shape cannot be extracted from the medical image. Thus, it is proposed to only extract surrogate information and use this information to estimate a healthy state. This process is referred to as pseudohealthy synthesis. In this work, concepts and methods for pseudohealthy synthesis are developed, aiming at personalizing cardiovascular prostheses.

tion methods. The design of the prosthesis can then be adapted to these segmentations (see Fig. 1.1). However, this naive approach is limited as there are scenarios where the desired shape cannot be extracted from the medical images.

In general, there are two different kinds of indications for replacing an organ or part of an organ by a prosthesis: a structural pathology and a morphological pathology. In the case of a structural pathology, the organ that should be replaced suffers from a change in its structure, leading to material weakness or malfunctioning due to changes in the biomechanical properties. One example is the calcification of aortic valve leaflets that leads to aortic valve stenosis as the valve cannot open properly due to increased leaflet stiffness [39]. In contrast, the shape of the organ is deformed in the case of a morphological pathology. An example for this is the pathological dilation of the aortic root due to an aneurysm [39].

For the first case, the optimal prosthesis shape is generally assessible using medical imaging as the prosthesis can be shaped exactly as the organ as it is observed in the image. As long as the desired biomechanical properties are sufficient, the approach described above is feasible. However, this only works if it is possible to acquire images of the desired organ shape. This is a big challenge for tiny structures like the aortic valve that consists of extremely thin material that moves very fast in the blood stream. Imaging with classical tomographic imaging modalities like computed tomography (CT) or magnetic resonance imaging (MRI) is not possible, and even in ultrasound imaging, a full segmentation is only possible in a model-based way [43]. Furthermore, even if the macro-structure of the leaflets can be assessed, the important micro-structure, i.e. the distribution of prominent collagen fibers with a thickness in the micrometer range remains unknown. In the second case, the desired shape is not assessible at all as the acquisition of images showing the organ will only lead to the deformed, pathological shape. Even with optimal imaging modalities, the natural, healthy state remains unknown.

Hence, both cases can lead to situations where the desired shape of the prosthesis is not known. To achieve personalization, the desired healthy shape has to be estimated based on surrogate information. This process is known as pseudohealthy synthesis, i.e. the synthesis of a shape, typically by synthesizing an image, that is assumed to be healthy. As one can not be sure that the shape really represents a healthy organ, it is referred to as pseudohealthy. The term pseudohealthy synthesis was introduced by [19] in 2016 and lead to an expansion of the field in recent years. Typically, this method is applied for segmentation, for example of lesions in the brain. Thus, a pseudohealthy image is synthesized based on the pathological image and the difference of both images is computed. If the pseudohealthy synthesis works optimally, this difference image represents the areas affected by the pathology. However, state-of-the-art methods for pseudohealthy synthesis are not suitable for the given problem as personalizing prosthesis shapes only based on surrogate information poses novel challenges on sparing the patient's individual anatomical traits while still allowing for strong domain differences between the surrogate information and the desired healthy shape.

The goal of this work is to examine whether personalized prosthesis shaping is possible using pseudohealthy synthesis. Thus, both cases, the structural pathology as well as the morphological pathology, are investigated exemplarily for cardiovascular prosthesis shaping. For structural pathology, the personalization of aortic valve prostheses is assessed. Thus, the desired shape of the three individual valve leaflets should be estimated based on surrogate tissue, i.e. the shape of the aortic root. In the morphological pathology case, the personalization of aortic root prostheses is investigated where the

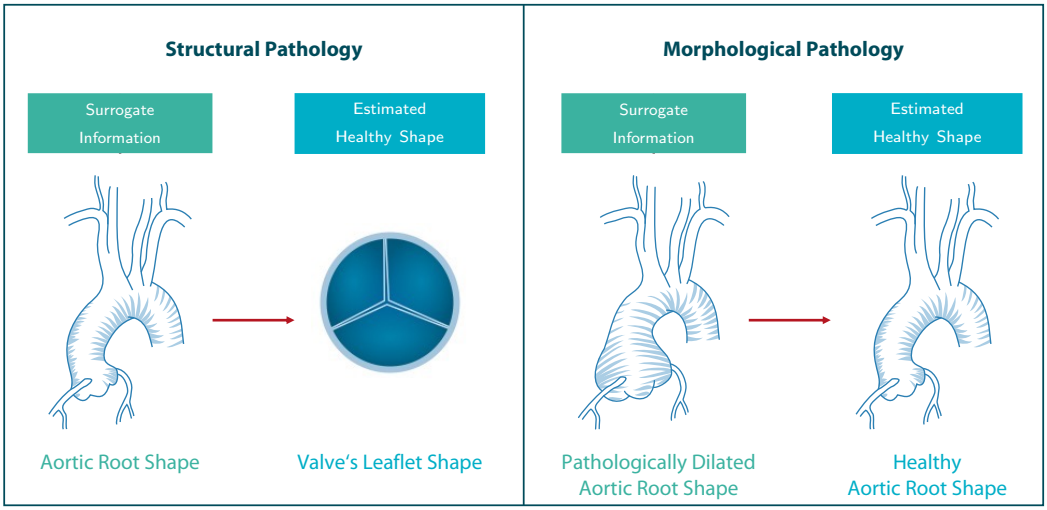

Figure 1.2: Visualization of the general problem formulation for the cases of a structural and a morphological pathology.

healthy shape is estimated based on the pathologically dilated shape of the aortic root. Fig. 1.2 visualizes these problem formulations.

Several approaches to solve these problems focused on biomechanical modeling and simulation of the prosthesis' behaviour to identify the individually optimal one. However, all these approaches suffered from very rough approximations of the complex biomechanical properties and were limited to only predicting specific features of the prosthesis like the diameter. The identification of the individual biomechanics and the resulting blood flow, including a qualitative evaluation of different behaviours, remains an open research questions. Hence, modeling-based approaches are not capable of solving the given pseudohealthy synthesis problem.

In recent years, data-driven approaches from the field of machine learning have developed rapidly and provide impressive results in complex pattern recognition. Deep neural networks, generally also referred to as deep learning, revolutionized the field of computer vision and showed unprecedented performance in nearly all areas of medical image analysis. Pseudohealthy synthesis is no exception and it could be shown that deep learning based approaches significantly increased the quality of the results. This is mainly due to the big advances of generative modeling using deep neural networks, e.g. using au-

toencoders or Generative Adversarial Networks (GANs). Hence, in this work, machine learning based approaches are examined to solve the given problem of pseudohealthy synthesis for cardiovascular prosthesis shaping.

This thesis presents the first work on personalized cardiovascular prosthesis shaping utilizing pseudohealthy synthesis. The goal is to develop methods and concepts to estimate the desired healthy organ shape only based on the available surrogate information (see Fig. 1.1). As both pathological cases, i.e. the structural and the morphological one, differ in their domain gap between the available surrogate information and the desired healthy shape, both cases are investigated and case-specific methods are proposed. As manufacturing prostheses in a highly individual way is already possible, the aim of this work is to close the gap between a medical image of the patient's pathological state and the desired prosthesis shape. Thus, manufacturing personalized prostheses, benchmarking them against state-of-the-art ones or performing clinical studies on human probands or data is out of scope. The thesis rather focuses on methodological development and proof-of-concept studies on ex-vivo porcine data sets.

1.1 Research Questions

Pseudohealthy synthesis and machine learning in general has never been investigated in the scope of cardiovascular prosthesis shaping before. Therefore, the overall goal of this thesis is divided into answering three research questions:

Q1: Is it possible to predict individual prosthesis features?

Q2: Is it possible to synthesize personalized healthy shapes?

Q3: Is it possible to identify optimal sets of general prostheses?

The research questions are consecutive and are answered for both cases, the structural and the morphological pathology, respectively. The key idea as well as the contribution of each research question is described in detail in the following paragraphs:

Q1: Is it possible to predict individual prosthesis features?

In the scope of this question, it is investigated whether the surrogate actually carries enough information to infer information about the desired, but unknown, shape. Thus, classical machine learning methods are trained to predict specific geometric features of the prosthesis based on the surrogate, i.e. the surrounding tissue or the pathological

shape, depending on the case of pathology. This question serves as a proof-of-concept study to ensure that the general concept of solving the prosthesis shaping problem using pseudohealthy synthesis.

Q2: Is it possible to synthesize personalized healthy shapes?

Based on the results of Research Question *Q1*, it is examined whether it is possible to not only predict specific geometric features of the prosthesis but synthesizing the full shape. Thus, generative models are combined with models mapping from the surrogate to the pseudohealthy state. These newly developed models are evaluated to assess whether pseudohealthy synthesis for cardiovascular prosthesis shaping is possible.

Q3: Is it possible to identify optimal sets of general prostheses?

Even though fully individual prosthesis shaping as examined in *Q2* follows the principle of personalized medicine, it comes with huge logistical, economical and regulatory challenges and issues. Hence, it is investigated whether a solution in between the two extremes, i.e. full personalization and a one-size-fits-all approach, could lead to an improved patient outcome combined with an easier translation into clinical application. Thus, it is investigated whether it is possible to synthesize a set of optimal prosthesis shapes in a completely data-driven way so that each patient can get one out of this set and gets a well-fitting prosthesis.

1.2 Structure of this Work

The structure of this work follows the three research questions. At first, a rough overview over the medical background is given in Chap. 2, followed by a presentation of the relevant machine learning concepts in Chap. 3. Then, the data sets that were collected and used within this thesis are described in Chap. 4. Based on this, three chapters present the methodological development and proof-of-concept studies performed to address the three research questions. Thus, *Q1* is tackled in Chap. 5, *Q2* in Chap. 6 and *Q3* in Chap. 7. Each of these chapters follows the same structure: After a short introduction, the contribution of the chapter is described, followed by an overview of related work. Then, the research question is addressed for both kinds of pathologies, i.e. structural and morphological, in two sections. Finally, the results are wrapped up and discussed in a conclusion, giving an answer to the research question. This structure is present throughout the three

Chapters 5, 6 and 7. Finally, the final result of the thesis is formulated and discussed in the Chapters 8 and 9.

During his time at the University of Lübeck, the author published 4 journal papers (3 as first author, 3 under review or in revision) and 14 conference papers (12 as first author). Therefore, parts of this work have been published previously. These publications are referenced in a footnote in the beginning of each chapter.

1.3 Notes on Terminology

Throughout this thesis, several terms are used in a way that might appear differently in related literature. Hence, this section aims at providing a clear framing of this potentially ambiguous terminology.

In the following, the term *agent* refers to any kind of model that is capable of learning from observation. The underlying learning principle is not restricted to reinforcement learning but could also span over e.g. supervised or unsupervised learning tasks, following the broad definition of an agent from [116]. Please refer to Sec. 3.1.2 for details.

The aim of this thesis is to estimate anatomical shapes. A shape is represented by an object that is visible within an image. Thus, a structure of connected pixels within a 2D image is denoted as a shape in the scope of this thesis. Please note that this does not restrict a shape to be represented by pixels in a 2D image, it could also consist of a point cloud or a mesh in general.

Throughout the thesis, the performance of different models is assessed using cross-validations, e.g. 10-fold ones. In these cases, the evaluated metric is typically given as the average metric over all folds, together with its standard deviation. The author is aware that a standard deviation typically requires more than ten observations to be reliable. However, an assessment of the variance of a model's performance regarding different data sets is still beneficial for a profound comparison of different models. Thus, the standard deviation is still computed and given even for small numbers of observations. The given value can be interpreted as a rough approximation of the standard deviation that does not aim at statistical exactness but should support the reader in analysing the results.

2 Medical Background

In this chapter, the medical background of this thesis is presented. After describing the aortic valve apparatus' anatomy and physiology, an overview of relevant cardiovascular implants, divided into aortic valve prosthetics and aortic root prosthetics. Thus, the medical indications as well as surgical intervention methods are presented and examples of different prosthesis types are given.

2.1 Aortic Root and Valve Anatomy and Physiology

The explanations in this chapter mainly follow [96] and [39]. The aortic root presents the junction between the left ventricle of the heart and the ascending aorta. Its shape is defined by three bulbs emerging from the circular root shape, the sinuses of valsalva. From two of these sinuses arise the coronary arteries. The nomenclature of the sinuses follows the coronary arteries as seen from a caudal view ("from above"), the sinuses are called right-coronary, left-coronary and non-coronary sinus. The connection between the aortic root and the left ventricle of the heart is referred to as ventriculoarterial junction. Fig. 2.1 illustrates the aortic valve apparatus.

The most prominent structure of the aortic root is the aortic valve. It is one of the four heart valves and prevents the blood in the aorta from flowing back into the ventricle during the diastole, i.e. the relaxation phase of the heart after contraction. The aortic valve consists of three semilunar leaflets that are attached to the root wall within one sinus, respectively. The leaflets act like a passive valve as they let the blood flow through the root during systole and, due to a higher pressure in the aorta than in the relaxing ventricle, lie together during diastole closing the valve. Similar to the sinus nomenclature, the leaflets are called right-coronary, left-coronary and non-coronary leaflet. The line where the leaflet is attached to the root wall is called commissure line, while the highest points of this line are referred to as commissure points. As these are also the points where the commissure lines of two neighbored leaflets meet, there are three commissure points in total. The imaginary circular line that cuts through the three commissure points

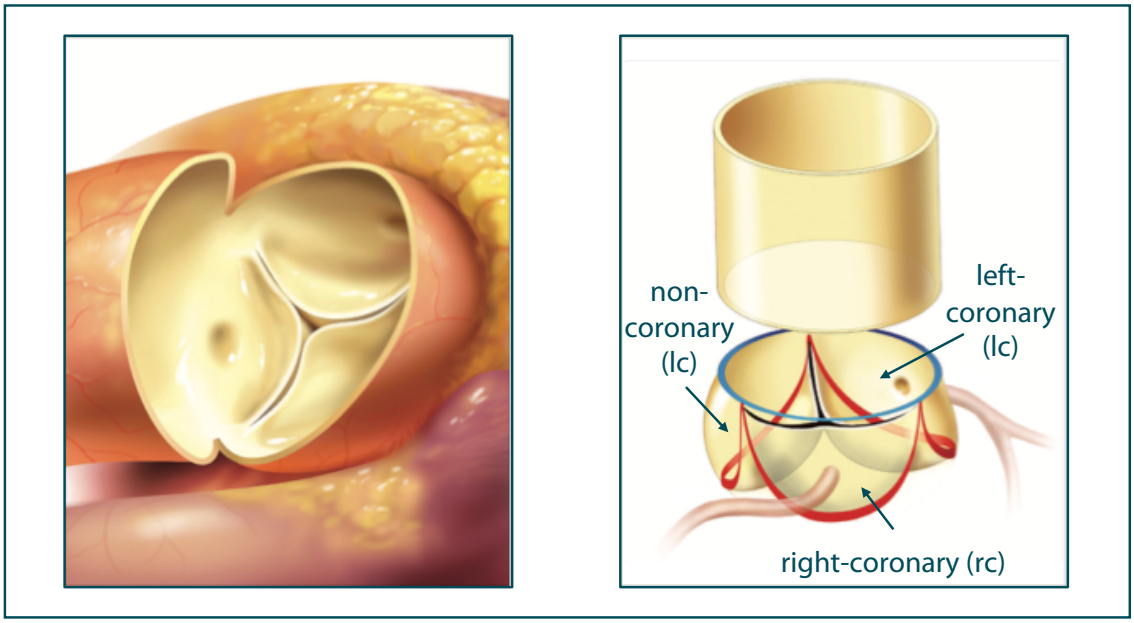

Figure 2.1: Illustration of the aortic valve apparatus. Left: The most prominent structure is the aortic valve that sits within the aortic root. It consists of three leaflets. Right: Sketch with right-coronary, left-coronary and non-coronary sinus, commissure lines (red) and sinotubular junction (blue). Adapted with permission from [39].

is referred to as sinotubular junction and marks the transition between the aortic root and the ascending aorta (see Fig. 2.1).

The leaflets consist of three tissue layers, the ventricularis at the ventricular side, the fibrosa on the arterial side and the spongiosa as an intermediate layer in between. While the ventricularis and spongiosa layers are comparably homogeneous, the fibrosa features a network of prominent collagen fibers that are thought to be the main stress bearers in the closed state [118]. The part of the leaflets that is not attached to the root wall is called the leaflet's free edge. Along the free edge, there is a thick node, the nodulus arantius. This node is the thickest point of the leaflet with a thickness of about $0.7 - 1.5\,mm$, depending on the age [121]. In the closed state of a healthy aortic valve, the noduli arantii of all three leaflets lie together, forming the coaptation point. The area where the free edges of two leaflets touch each other in the closed state is called coaptation.

Even though the aortic valve had been described as a passive structure for a long time, modern studies have shown the complex dynamical processes within the aortic valve apparatus [148], [10]. Over the phases of a heart cycle, the aortic root behaves highly dynamically and increases its volume by up to 40%. This expansion first appears during

systole at the ventriculoarterial junction and then propagates in cranial direction, passes the sinotubular junction and passes on towards the ascending aorta. This supports the opening of the valve. During diastole, the root wall shrinks back to its original size while the pressure gradient changes due to ventricle relaxation. This causes non-laminar turbulences within the three sinus' of valsalva that press the leaflets towards each other and hence accelerate the valve's closing process. The large expansion of the aortic root also alters its height, i.e. the height difference between the ventriculoarterial junction and the sinotubular junction, as well as its relative orientation. This can be assessed by the tilt angle between the commissure plane, i.e. the plane of the sinotubular junction, and the plane through the ventriculoarterial junction. This tilt angle is minimal during systole to provide a minimal flow resistance and maximal during diastole to achieve a high counterpressure for the backflowing blood [148]. Hence, the aortic root and valve apparatus form a complex dynamical system that reaches its efficiency through the precise interplay of many biomechanical components and effects [147].

2.2 Cardiovascular Implants

Despite remarkable progress in cardiology and cardiovascular surgery, five out of the ten most common causes of death are cardiac diseases [11]. Therefore, intense research and development in the field of cardiovascular surgery started in the early 20-th century and has since been intensively continued. The history of cardiovascular implants dates back to 1952 when Charles Hufnagel implanted the first valvular device in the descending aorta to prevent regurgitation [117]. The valve replacements by artificial prostheses were performed in 1960 by Nina Braunwald, Andrew Morrow and Albert Starr [39]. Since then, the amount of cardiovascular implants grew steadily [11]. In the area of aortic valve surgery, there are two surgical paradigms: Valve replacement and valve-sparing reconstruction. As the name suggests, the native aortic valve is replaced by a valve prosthesis in the first case. In contrast, the patient's native valve is preserved in valve-sparing reconstruction by remodeling the aortic root using a tubular prosthesis. Both kinds of implants, i.e. aortic valve prosthetics and aortic root prosthestics, are described in the following.

2.2.1 Aortic Valve Prosthetics

There are two main indications for a replacement of the aortic valve: aortic valve stenosis and aortic valve insufficiency. If the valve's orifice area is pathologically narrowed and

not enough blood can flow through into the aorta, the pathology is called aortic valve stenosis [69]. One common cause of a stenosis is the calcification of the valve leaflets, also called senile degenerative calcific aortic valve stenosis. In this pathology, calcific nodules settle around the leaflet tissue and increase its stiffness drastically [89]. Hence, the leaflets cannot open properly anymore, preventing the bloodstream to flow freely through the valve. The degenerative calcific stenosis is the most frequent cause for aortic valve replacement [39]. Another common type of stenosis is the congenital aortic valve stenosis. In this case, the valve only consist of two (bicuspid valvular stenosis) or, in very rare cases, of one leaflet (unicuspid valvular stenosis) [39]. Even though the leaflets are not suffering from increased stiffness, the pathological geometry provides a smaller orifice area compared to a native trileaflet valve. Besides these two main pathologies, there are several other very rare causes of aortic valve stenosis. Further details can be found in [39].

If the valve cannot close properly anymore, the blood can flow back from the aorta into the left ventricle. This is referred to as aortic valve insufficiency [39]. The typical reason is a pathological dilation of the aortic root. The state-of-the-art treatment method is valve-sparing aortic root reconstruction (see Sec. 2.2.2). However, in severe cases, the valve cannot be spared and must be replaced by a prosthesis.

In general, there are three different types of aortic valve prosthetics: mechanical prostheses, biological prostheses and auto-, allo- or xenografts [39]. Mechanical prostheses are made of synthetic materials, for example silastics or titanium. The are mainly aiming at providing fast and reliable opening and closing characteristics with a long durability. hence, the closing mechanism does not necessarily resemble the trileaflet structure of the aortic valve. Thus, mechanical valves utilize tilting or non-tilting disc valves, caged-ball valves or bileaflet valves [39]. While mechanical valves provide an impressive long-term durability, the patient has to be treated life-long with anticoagulants to avoid thrombosis and embolism.

In contrast, biological valves are aiming at mimicking the valve's native geometry more closely. They consist of three leaflets attached to a stent [39]. The leaflets are made of biological tissue, typically processed bovine pericardium. A long-term anticoagulation therapy is not necessary due to the biocompatible materials used, but the durability is lower and the risk of follow-up surgeries is higher than for mechanical valves.

In allo-, auto- or xenografts, the valve is replaced by a purified native valve. This can either come from an organ donor (allograft) or an animal, typically a pig (xenograft) [97]. One special case is the autograft where the implanted native valve is the patient's own tricuspid valve. This method is called Ross procedure [31].

The surgical techniques for aortic valve replacement can be divided into conventional surgery and transcatheter aortic valve implantation (TAVI) [39]. Conventionally, the valve is implanted in an open-thorax surgery. Alternatives are minimally invasive approaches like the hermisternotomy. Due to the high degree of freedom for the surgeon's movement and actions, all kind of prosthetics can be implanted following the conventional surgery. However, an open-thorax intervention presents a high stress for the body which is specifically a risk for elderly patients. Therefore, TAVI presents a minimally invasive alternative to the conventional method. In this approach, special kinds of biological prostheses are utilized that feature a foldable stent [3]. Initially folded, the prosthesis is brought into the body through a vessel, typically the femoral artery, using a catheter. An alternative is the transapical implantation where the catheter is inserted through the left ventricle [39]. The current position of the catheter is monitored, typically using X-ray, and the stent is deployed within the aortic. As the pathological leaflets remain in the root but are pressed to the root wall by the stent, this procedure is also called valve-in-valve implantation. Specifically due to the higher geometric requirements, i.e the necessity to be foldable and open up fast and reliably, the leaflets of TAVI prosthetics are designed simple, symmetric and only roughly follow the native aortic valve geometry. Fig. 2.2 shows examples of the different aortic valve prosthesis types.

In recent years, important advances have been made in the field of developing tissue engineered aortic valve prostheses [38]. For example, [68] implanted recellularized aortic valves into pigs showing good in-vivo performance. Such approaches indicate the possibility of fabricating fully personalized aortic valve prostheses in the near future.

2.2.2 Aortic Root Prosthetics

The replacement of the aortic root is necessary in the case of aortic regurgitation, i.e. when blood flows back into the left ventricle during diastole [39]. Even though there might be pathologies of the leaflets, e.g. perforations, that lead to regurgitation, the more common reason is a pathological dilation of the aortic root that causes a change of the aortic valve apparatus' geometry. Hence, the leaflets cannot seal the valve properly anymore. Especially patients suffering from Marfan syndrome tend to develop aortic root aneurysms [120].

While valve replacement formerly was a typical intervention for this indication, valve-sparing aortic root reconstruction surgery presents a promising alternative to avoid the downsides and risks of prosthetic heart valves [39]. This kind of technique spares the patient's own valve leaflets and only the dilated aortic root is replaced by a prosthesis.

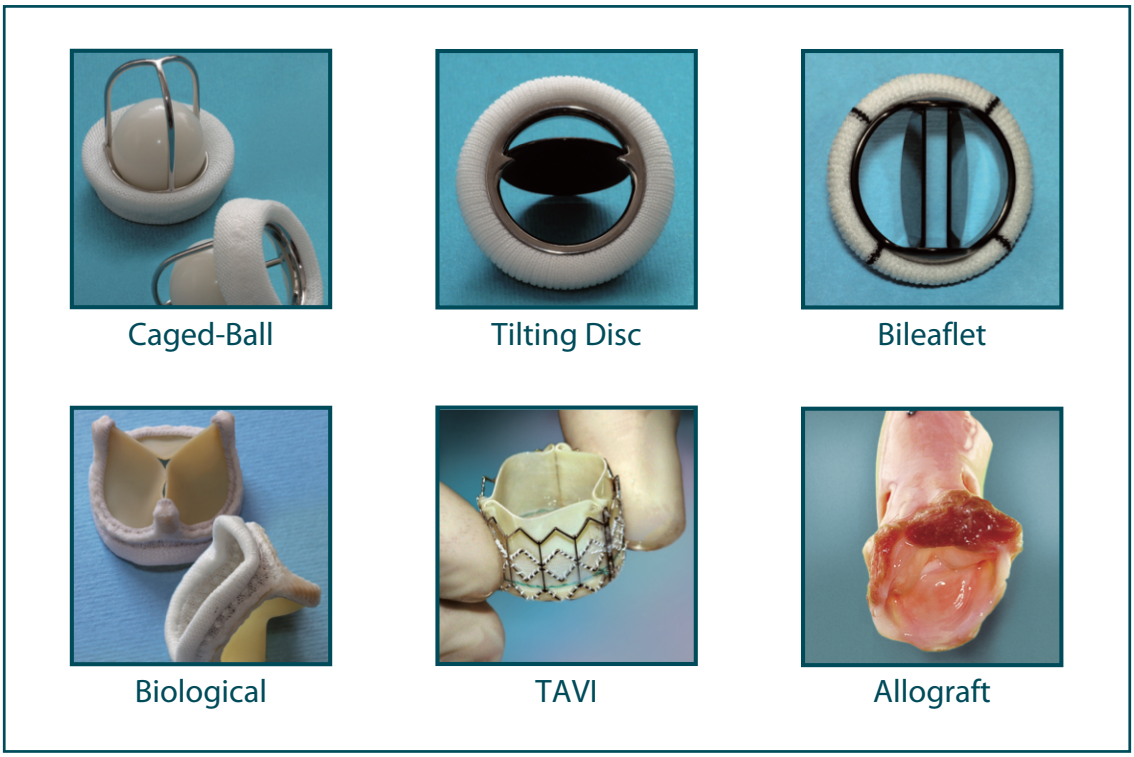

Figure 2.2: Different types of aortic valve prostheses. Examples of mechanical valves are a caged-ball valve (Starr-Edwards), a tilting disc valve (Omniscience) and a bileaflet valve (St. Jude Medical). Additionally, a biological valve (Edwards), a TAVI valve (Edwards) and an allograft are shown. Adapted with permission from [39].

Besides the important advantage of sparing as much of the patient's own tissue as possible to minimize the intervention in the complex system of the aortic valve apparatus, there is no need for life-long anticoagulation treatment as it is the case for mechanical aortic valve replacement [47].

There are two main approaches for valve-sparing aortic root replacement surgery: The reimplantation technique as proposed in [32] and the remodelation technique introduced in [124]. In reimplantation, the aortic root is cut away, the prosthesis is put over the remaining aortic valve and all remaining structures are sewed in. In the remodelation approach, the aortic root is cut away along the commissure lines and artificial sinus' of valsalva are formed by the root prosthesis. Even though this allows for a more realistic dynamical behaviour, it was shown that the patients have a higher risk of reoperation due to dilation of the remaining tissue below the prosthesis. Hence, the reimplantation

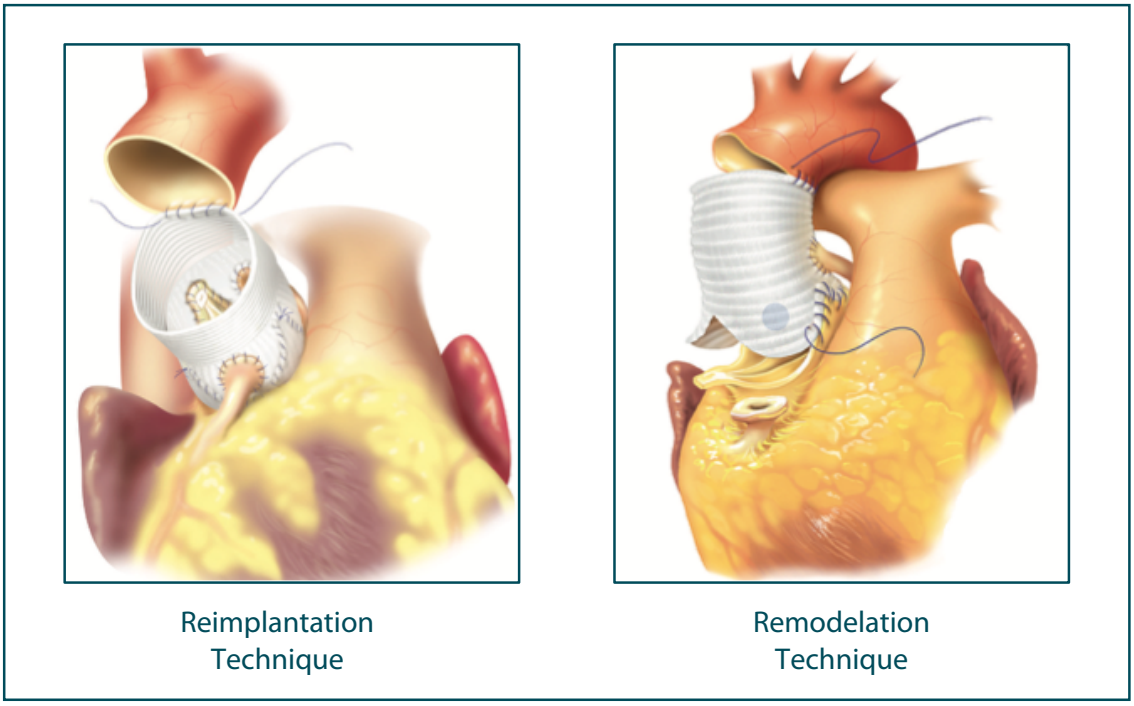

Figure 2.3: Illustration of valve-sparing aortic root reconstruction surgery using an aortic root prosthesis for the reimplantation (left) as well as the remodelation technique (right). Adapted with permission from [39].

technique is nowadays the more common approach [39]. Fig. 2.3 illustrates the two techniques for valve-sparing aortic root reconstruction surgery.

In clinical practice, different types of aortic root prosthesis, also called composite aortic valve graft (CVG) [7], are in use. In general, the CVGs are separated into mechanical composite grafts, biological composite grafts and xenografts [42], [23], [33]. While biological composite grafts make use of biologically harvested materials, typically bovine pericardium [44], mechanical composites are made of synthetic material, for example a combination of knitted polylactic acid (PLA) and a polycaprolactone matrix (PCL) [90]. In contrast to these crafted grafts, xenografts are purified transplants from animals, for example porcine aortic roots [88].

While it is obviously not possible to personalize xenograft prostheses, several approaches were presented to achieve a better match between the patient's individual aortic root shape and the prosthesis in the scope of composite grafts. In [40] and [112], the authors present and evaluate mechanical composite grafts with leaflet cusps, mimicking the sinus' of valsalva. Even though this prosthesis matches the native root geometry more

closely and has positive effects on the valve functionality, the cusps are not tailored to the individual geometry but are shaped symmetrically. In [130], the authors present a prosthesis featuring personalized aortic cusp shapes by adjusting control points of Bezier-curves manually to achieve a sufficient similarity between the 3D model and the original native root. However, a lot of manual interaction is required in this approach and the amount of introduced human bias remains unclear. Modern techniques of 3D electrospinning allow for the individual fabrication of PLA-composed grafts with a high degree of freedom [15], [84]. Different kinds of these grafts have already been tested in vivo [113] and hence have a huge potential regarding personalized aortic root prosthesis fabrication. Furthermore, it is possible to let cell cultures grow on the knitted scaffold, leading to tissue engineered vascular grafts (TEVG) [109]. As the cell cultures are harvested from the patient, these grafts are highly biocompatible and drastically reduce the reaction of the immune system to the implant. A recent, concurring approach to electrospun scaffolds is 3D printing of aortic roots. In [76], the authors present first biomechanical evaluations of a printed aortic root, where the shape was extracted from micro-computed-tomography data of one patient. However, one remaining challenge is 3D printing the root using biocompatible materials with realistic biomechanical properties.

3 Basics of Machine Learning

In this chapter, the theoretical basis of machine learning is provided. First, a comprehensive overview of machine learning concepts is given in Sec. 3.1. Afterwards, the machine learning methods utilized in this thesis are presented and explained in detail, including Support Vector Machines (Sec. 3.3), artificial Neural Networks and Autoencoders (Sec. 3.4), Random Forests (Sec. 3.5) as well as unsupervised clustering (Sec. 3.6).

3.1 Introduction to Machine Learning

As a subfield of Artificial Intelligence, Machine Learning refers to adapting a model according to observations, also called data, so that predictions can be made for unobserved and hence unknown data points [116]. The adaption of the model is called learning or training. This section presents an attempt to integrate all kinds of machine learning into one unified model for an introductory purpose.

In general, machine learning algorithms can be divided using two different perspectives: the task perspective and the agent perspective. Both perspectives will be presented in detail in the following paragraphs.

3.1.1 Task Perspective

Two different coordinate spaces are relevant in machine learning: the *data space* D and the *feature space* X. The *data space* contains the observations of raw data, which could be for example images, short sequences of speech, written texts or sensor data. This data can be given in any format and dimensionality. In contrast, the *feature space* contains abstract features describing the data. Each dimension of the feature space corresponds to one feature of the data that, in the ideal case, carries a lot of information about the data distribution. The dimensionality of the feature space might be much lower than the one of the data space. Hence, each data sample from the data space is represented by a point in the feature space, and each point in feature space describes a potential

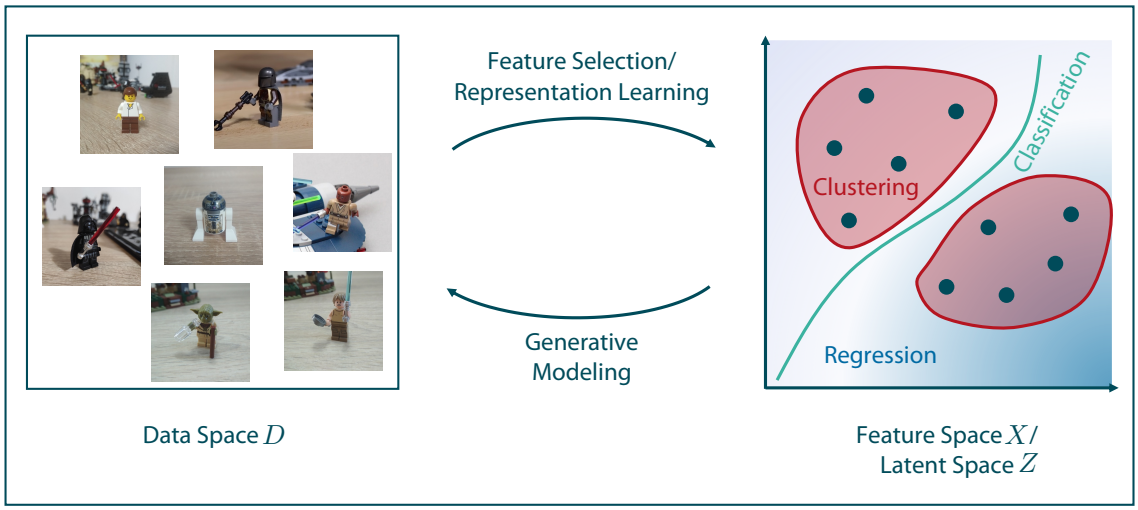

Figure 3.1: Nomenclature of machine learning algorithms: the task perspective.

data sample. As an example, one might think of the *data space* as a collection of landscape photographs, where each of the photographs is one data sample. One feature to describe this data set is the number of trees in an image. Hence, the *feature space* is one-dimensional and each photograph is represented by one point in this space, where the position of this point corresponds to the number of trees in this individual image. It is important to note that neither the mapping from data to feature space nor the mapping from feature to data space is unique. There might be a lot of very different images that show two trees.

Fig. 3.1 visualizes the nomenclature of different machine learning methods for the task perspective. The problem of defining a good feature space is called *representation learning*. The goal is to identify a mapping that converts the data into a space where each dimension is a feature describing the data, i.e. that carries information on the data distribution. In many approaches, this step was carried out by a human expert who defined sufficient features for the specific data set based on previous knowledge. This procedure is called *feature selection* and brings the advantage of fusing existing knowledge with data-driven learning. However, with the advances in the field of deep learning, *feature selection* has been more and more replaced by *representation learning* using deep neural networks. Details can be found in 3.4. Note that two different notations of the feature space are used within this thesis: A feature space point is denoted as $\mathbf{x}$ if the features were identified manually and it is called $\mathbf{z}$ if the feature description was derived using representation learning.

Finding a mapping from the feature to the data space is known as *generative modeling*. Here, the goal is to synthesize data only based on a given feature vector, i.e. one point in feature space. Typically, previous knowledge of the data distribution plays a key role in generative modeling. Following the example from above, a generative model could be capable of synthesizing a realistic image of two trees only based on the information that there should be two trees. Once again, generative modeling is not unique but usually seen as a probabilistic model. Hence, it is assumed that each feature vector leads to a distribution of data in the data space.

A sufficient feature space lays the basis for multiple tasks. If the data samples belong to different categorical classes, e.g. $\{cat, dog\}$, a *classification* aims at finding a separating line between these classes in feature space. The goal is that this line partitions the feature space in two parts were each point within one part belongs to the class cat and all points in the other part to the class dog. In contrast, it might also be that each data point is not assigned to a categorical class but a real number, e.g. a temperature value. The task of fitting a function into the feature space that approximates all values known from the data points and allows for inter- and extrapolation is called *regression*. A third application occurs if no label is assigned to the data points at all, neither categorical nor real value. Then, the aim of the algorithm is to divide the data points into distinct sets of similar points that form one class, also known as cluster. This process is called clustering.

In general, all tasks that are solved using machine learning can be classified into one of the presented tasks: *representation learning*, *generative modeling*, *classification*, *regression* or *clustering*.

3.1.2 Agent Perspective

In general, each machine learning model can be interpreted as an agent. An agent interacts with its environment by receiving information about the environment via sensory inputs, called perception, and manipulates it using actuators, called action [116]. The core of the agent is the mapping from a perception to the action it should perform, i.e. the decision making. In general, this step is called the performing element, but to align the notation with current machine learning literature, it is also called model in the scope of this thesis. For some learning paradigms, the agent does not interact with the full environment during training, but only with a small subset of it. This subset is called training environment. One can imagine the training environment as a data set of finite size. A model can be trained on this data set, but the set cannot represent the full variety

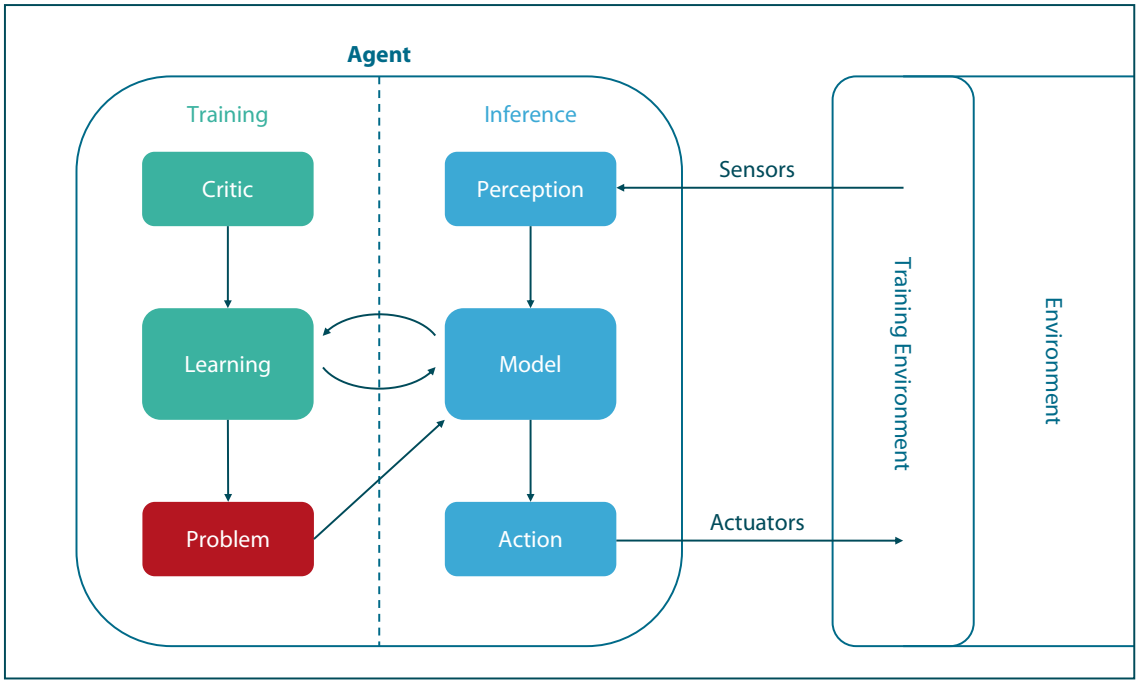

Figure 3.2: Sketch of an intelligent agent and its interaction with the environment.

of data samples that arise from the underlying real distribution. Fig. 3.2 visualizes an agent that interacts with its environment.

However, just selecting an action based on the current perception is not enough to call an agent intelligent. It should rather be capable of improving its own behaviour based on observations. It should be able to learn. Hence, each agent has an integrated learning routine that relies on some form of critic to evaluate pairs of perceptions and corresponding actions, aiming to learn whether this action was helpful given the perception. In general, all machine learning concepts can be fitted into this agent concept. However, different learning paradigms feature different settings of the agent's learning routine. In the following, the learning paradigms *supervised learning*, *unsupervised learning*, *semi-supervised learning*, *weakly supervised learning* and *reinforcement learning* are presented in the scope of the agent perspective.

In *supervised learning*, a finite training data set $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2) \ldots, (\mathbf{x}_n, y_n)$ of input vectors $\mathbf{x}_i$ and corresponding outputs $y_i,\ i = 1, \ldots, n$ is given. This can be interpreted as a training environment with which the agent interacts during training. The given y acts like a supervisor that tells the agent for each performed action whether it was the right one or not. Hence, the agent can learn to satisfy the supervisor on as many perceptions from the training environment as possible, aiming at mimicking the supervisor's decisions

as close as possible. The agent can be deployed to the full environment, i.e. it can decide on an action even for perceptions that were not part of the training environment. But a supervised agent is not able to improve in the full environment when no supervisor is present to rate its decisions.

The scenario is quite similar for *unsupervised learning*. Once again, the agent only observes a training environment, but in contrast to supervised learning, there is no external supervisor that rates the quality of decisions. This relates to a training environment where only a set of perceptions $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is given. In this setting, the agent has to define the critic functionality on its own. However, the training is still performed within a finite training environment. When the agent is deployed in a real environment, no further improvement is achieved.

There are also some hybrid approaches, e.g. *semi-supervised learning* and *weakly-supervised learning*. In the semi-supervised setting, the agent receives the critic from a supervisor only for a limited number of observations, while it has to use its own critic functionality on all other ones. For the *weakly-supervised learning*, the agent does receive critic from a supervisor, but this critic is not completely reliable. Hence, the agent expects noisy criticism and therefore can react to it. However, both approaches are specific or hybrid forms of the previously mentioned supervised and unsupervised agents.

In contrast, the agent directly interacts with the full environment in *reinforcement learning*. Once again, the agent defines the critic of performed perception-action-pairs on its own. However, it is capable of improving within the full environment, making open-end-learning possible. Additionally, the agent features another element within the learning part: a problem generator. This element adds curiousness by formulating new problems autonomously and proposing experiments for the model to perform. This allows the agent to assess whether novel actions provide a better reward for a given perception than already known ones. Hence, the reinforcement learning agent gives the highest flexibility and the most advanced learning methods, but it is also harder to define and the restrictions of other agent types, e.g. supervised methods, might be an advantage in safety-critical application scenarios like medical decision making.

In general, all machine learning algorithms can be classified into one of these learning paradigms and interpreted as an agent. Hence, each of the numerous machine learning models that have been proposed over the decades can be categorized and clustered within the framework of the task and the agent perspective.

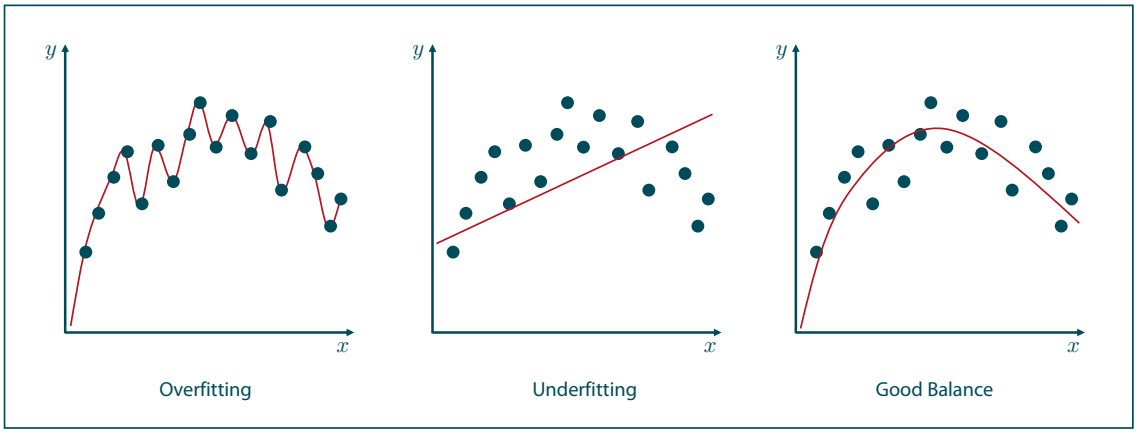

Figure 3.3: Visualization of the bias-variance-tradeoff. If the model's variance is too high, it is prone to overfit to the data (left). In contrast, if the variance is too low, the model is not capable of grasping relevant patterns (middle). An optimal fit can only be found when the tradeoff between bias and variance is sufficient (right).

3.2 Overfitting and Generalization

Typically, at least in the case of supervised and unsupervised learning, machine learning models are trained on a finite data set, the training data. The aim of machine learning is to adjust all model parameters in such a way that the model generalizes well [116]. This means that the model is not only capable of providing good performance on the training data but also on unseen samples. However, providing minimal error on the training data does not lead to the best generalization but actually might decrease the generalization performance as the model only focuses on learning the specific training data points without deriving general decision rules. This effect is known as overfitting. With increasing model capacity and flexibility, the risk of overfitting rises as the training data is almost always affected by noise or outliers. Hence, the model should be regularized to not capture each tiny bit of variance present in the training data but rather to focus on the general patterns. However, this implies the risk of ignoring quite important parts of the variance that is necessary for decision making on unseen data, which is called underfitting. The problem of finding a generalizing model that does neither overfit nor underfit is one of the biggest challenges in machine learning. This challenge is also known as the bias-variance-tradeoff. Fig. 3.3 visualizes this problem.

One possibility to asses the generalization performance of a machine learning model is *k-fold crossvalidation* [116]. Within a crossvalidation, the available data is divided into

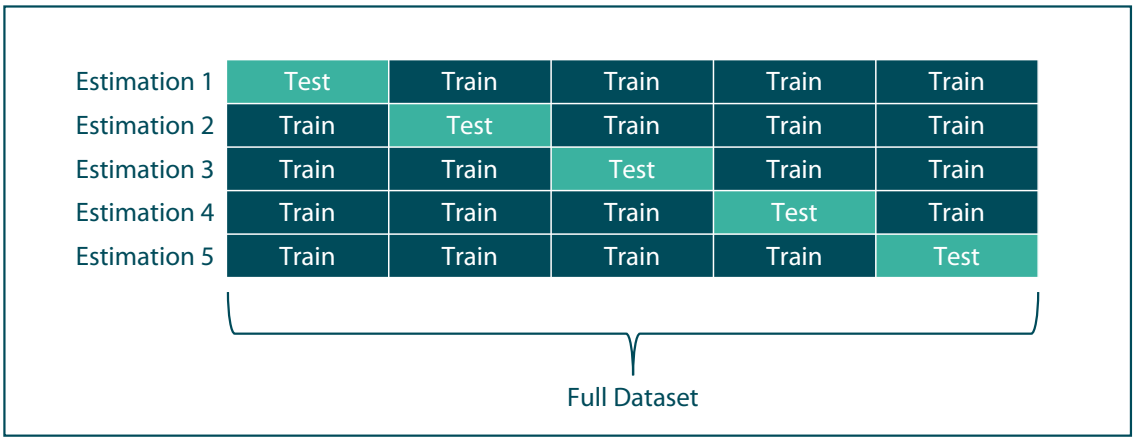

Figure 3.4: Visualization of a crossvalidation, exemplarily given for five folds. The data set is split up into five folds of different size. Then, the model is trained on four of them and evaluated on the fifth one. This procedure is repeated, each time using a different fold for testing.

three data sets: training data, validation data and test data. The model is trained on the training data, while the validation data can be used to optimize hyperparameters of the model. One typical example is stopping the training of a neural network after a specific number of epochs when the error on the validation set rises, indicating overfitting. Afterwards, the performance of the trained model is then assessed by applying it to the test data and comparing the predicted results with the ground truth. Hence, the generalization performance on unseen data can be approximated.

In a classical k-fold crossvalidation, this procedure is repeated k times so that each data point was part of the test data once. A typical value would be $k = 10$. Hence, $10\,\%$ of the available data would serve as test data while the other $90\,\%$ could be split up into training and test data. Note that a validation set for tuning hyperparameters can be drawn from the training data set within each fold. An alternative is the so-called *k-fold Monte Carlo crossvalidation*. Here, the data split within each fold is independent, i.e. it is not controlled in how many folds one data point serves as a training or a test data point. Instead, the training, validation and test sets are drawn randomly from the available data within each fold. Compared to classical crossvalidation, this approach might give rise to statistical errors, but it allows to assess the model's performance on an arbitrary number of folds.

3.3 Support Vector Machines

A Support Vector Machine (SVM) is a machine learning model for classification. The main idea is to optimize the model for generalization capability by finding the separating hyperplane with the largest distance to the observed training data points. This hyperplane is called the maximum-margin separating plane. In addition, SVMs utilize the kernel trick to deal with nonlinear classification problems. The first version of SVM was presented in [17]. The explanation in this thesis roughly follows [41].

3.3.1 Classification

The SVM is designed for supervised training, so it can be assumed that a data set with $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2) \ldots, (\mathbf{x}_n, y_n)$ is given. The data set consists of n data points with m features, respectively. Hence, $\mathbf{x}_i \in \mathbb{R}^m, i = 1, \ldots, n$ contains the features of the i-th data point while $y_i \in \{-1, 1\}, i = 1, \ldots, n$ is the known label. It is important to note that this label notation implies a binary classification with two possible output classes, 1 and -1. Further information on mutliclass SVM, i.e. SVM classification with more than two classes, can be found in [70].

The aim is to find a hyperplane that separates the points into the two classes, i.e. where all points of label 1 lie above the hyperplane and all points of label -1 lie below the hyperplane. Note that this is only possible if the data set is linearly separable. The kernel trick for dealing with non-linearly separable data sets will be presented after explaining the linearly seperable case. The desired hyperplane follows the general hyperplane equation

$$\mathbf{w}^\mathsf{T}\mathbf{x} + d = \mathbf{0}, \tag{3.1}$$

where $\mathbf{w} \in \mathbb{R}^m$ and $d \in \mathbb{R}$ are the parameters of the hyperplane and $\mathbf{0}$ represents a vector of dimension m filled with zeros. It is important to note that the term $\mathbf{w}^\mathsf{T}\mathbf{x} + d$ describes the distance of the point $\mathbf{x}$ to the hyperplane defined by $\mathbf{w}$ and d, while the sign of the result describes on which side of the hyperplane the point lies. By setting the distance to zero, 3.1 describes all points that lie on the hyperplane.

The key idea of the SVM is to define the optimal separating hyperplane as the separating plane that has the largest distance to the closest data point. This assumption directly follows from the concept of generalization where it can be assumed that the training data set only shows a specific part of the real variance. To be as robust as possible to test data that is more spreaded than the training data, the distance between the closest training data point and the hyperplane, also referred to as margin, should be maximal. This always leads to a symmetric margin, i.e. the margin is identical for both classes. The

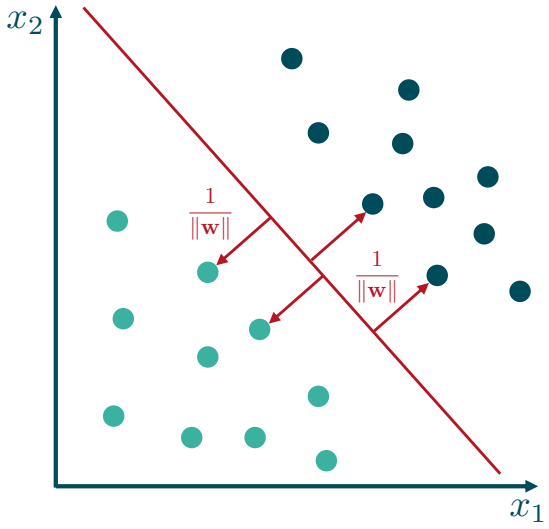

Figure 3.5: A visualization of the classification hyperplane (red) identified by maximizing the margin.

maximization of the margin can be formulated as an optimization problem as described in the following paragraphs. Fig. 3.5 visualizes the classification hyperplane and the margin.

It can be shown that the distance between the hyperplane and the closest data point is given by $\frac{1}{\|\mathbf{w}\|}$. Hence, the maximum margin can be found by minimizing the denominator of this fraction. However, not only the maximum margin should be taken into account. It is important to make sure that the training data is separated and classified correctly, i.e. that all training samples lie on the right side of the hyperplane. Hence, each datapoint poses a constraint on the optimization problem:

$$\begin{aligned} &\underset{\mathbf{w}}{\text{minimize}} && \|\mathbf{w}\| \\ &\text{subject to} && y_i(\mathbf{w}\mathbf{x}_i + d) \geq 1,\ i = 1, \ldots, n. \end{aligned} \tag{3.2}$$

The definition of the constraints arises from the canonical representation of the margin (see [41]. This means that if the hyperplane seperates the data correctly, $\mathbf{w}$ can be scaled in a way that the margin is exactly 1 or -1, respectively. Hence, the data points of class 1 that are closest to the hyperplane have the distance 1 and all other points have a higher positive distance. Vice versa, all data points of class -1 have -1 or a higher negative

distance. To enforce that all points are classified correctly, all points should fulfill the following conditions:

$$\begin{aligned} \mathbf{w}^\mathsf{T}\mathbf{x}_i + d \geq 1, &\forall \mathbf{x}_i | y_i = 1 \\ \mathbf{w}^\mathsf{T}\mathbf{x}_j + d \leq -1, &\forall \mathbf{x}_j | y_j = -1. \end{aligned} \tag{3.3}$$

As y is defined to have either the value 1 or -1, these conditions can be combined to the constraint given in 3.2.

Even though the optimization problem stated in 3.2 could be solved to find the maximum margin separating plane, a transformation to the dual optimization problem brings several advantages, including the support vector extension, the applicability of kernels and the possibility to define efficient solvers like the SMO algorithm [129]. Hence, the problem is reformulated to match the definition of a quadratic program (QP):

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \quad & \frac{1}{2}\mathbf{w}^\mathsf{T}\mathbf{w} \\ \text{subject to} \quad & y_i(\mathbf{w}\mathbf{x}_i + d) \geq 1,\ i = 1, \ldots, n. \end{aligned} \tag{3.4}$$

Note that both objective functions share the same minimum and hence, both problems have the same solution. But by formulating the problem as a QP, it is possible to utilize Lagrange-duality to derive the dual problem:

$$\begin{aligned} \underset{\alpha}{\text{maximize}} \quad & \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} y_i y_j \alpha_i \alpha_j \mathbf{x}_i^\mathsf{T} \mathbf{x}_j \\ \text{subject to} \quad & \sum_{i=1}^{n} y_i \alpha_i = 0, \\ & \alpha_i \geq 0,\ i = 1, \ldots, n. \end{aligned} \tag{3.5}$$

The variable of this dual problem is the vector $\alpha \in \mathbb{R}^n$ that contains the Lagrange multipliers $\alpha_i,\ i = 1, \ldots, n$. Each of these Lagrange multipliers corresponds to one data point and has a special property: if the datapoint matches the constraint with slack, i.e. if the distance to the separating hyperplane is not 1 or -1, the corresponding α is zero. Hence, the only nonzero α_i correspond to the data points that are closest to the hyperplane which is typically only a small subset of the training data set. These data points are called support vectors and gave the SVM its name.

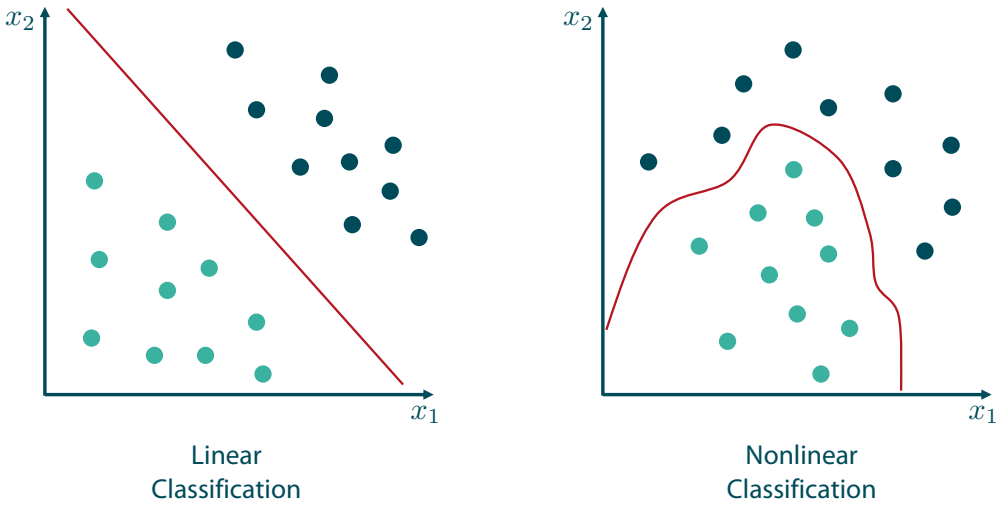

Figure 3.6: Using the kernel trick, the SVM cannot only be applied to linear classification (left) but also to nonlinear classification problems (right).

After solving the optimization problem 3.5 and retreiving the optimal Lagrange multipliers α^*, the parameters of the hyperplane can be computed as

$$\begin{aligned} \mathbf{w}^* &= \sum_{i=1}^{n} y_i \alpha_i^* \mathbf{x}_i \\ d^* &= \frac{1}{y_j} + \sum_{i=1}^{n} y_i \alpha_i^* \mathbf{x}_i^\mathsf{T} \mathbf{x}_j, \; j \in \{1, \ldots, n\}. \end{aligned} \tag{3.6}$$

Until now, it is only possible to classify linearly separable data sets correctly by optimizing the parameters of a hyperplane. However, it is also possible to learn nonlinear classifications using SVM by utilizing kernel functions (see Fig. 3.6). The basis of this idea, also referred to as the kernel trick, is that for every data set, there is a high-dimensional space in which this data set is linearly separable [116]. Hence, after mapping the data to this high-dimensional space, the problem is solvable with the linear SVM. But the only thing that needs to be computed in the high-dimensional space is the scalar product between data points $\mathbf{x}_i^\mathsf{T} \mathbf{x}_j$. And this scalar product can be calculated by a kernel function.

A kernel function k is a function that computes the scalar product between two points $\mathbf{x}, \mathbf{z}$ from the input space X after mapping the data to a higher-dimensional space using the mapping F:

$$k : X \times X \mapsto \mathbb{R} \tag{3.7}$$

$$k(\mathbf{x}, \mathbf{z}) = \langle F(\mathbf{x}), F(\mathbf{z}). \rangle \tag{3.8}$$

Whether a function is a kernel can be validated using Mercer's Theorem [41]. By using a potential kernel function k and a set of data points $\mathbf{x}_1, \ldots, \mathbf{x}_n$, it is possible to compute the kernel matrix K, also called Gram matrix, as

$$K := (K_{i,j})_{i,j=1,\ldots,n)} \tag{3.9}$$

with

$$K_{i,j} := k(\mathbf{x}_i, \mathbf{x}_j). \tag{3.10}$$

If this kernel matrix is symmetric and positive semi-definite, k is a kernel and can be used to compute the scalar product in a high-dimensional space. It is important to note that the mapping F remains unknown and is not even unique. Hence, the kernel trick offers an extremely efficient way of solving the problem in a higher-dimensional space. Using kernels, the SVM problem becomes

$$\begin{aligned} \underset{\alpha}{\text{maximize}} \quad & \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} y_i y_j \alpha_i \alpha_j k(\mathbf{x}_i, \mathbf{x}_j) \\ \text{subject to} \quad & \sum_{i=1}^{n} y_i \alpha_i = 0, \\ & \alpha_i \geq 0, \, i = 1, \ldots, n \end{aligned} \tag{3.11}$$

and the resulting hyperplane can be computed as

$$\begin{aligned} \mathbf{w}^* &= \sum_{i=1}^{n} y_i \alpha_i^* \mathbf{x}_i \\ d^* &= \frac{1}{y_j} + \sum_{i=1}^{n} y_i \alpha_i^* k(\mathbf{x}_i, \mathbf{x}_j), \; j \in \{1, \ldots, n\}. \end{aligned} \tag{3.12}$$

Over the years, many kernels were presented for applying SVMs to different tasks and data sets. In the scope of this thesis, two kernels are relevant: The linear kernel k_{linear} and the gaussian radial basis function (RBF) kernel k_{RBF}. The linear kernel does not induce a mapping and computes the scalar product in the original feature space:

$$k_{linear}(\mathbf{x}, \mathbf{z}) = \mathbf{x}^\mathsf{T} \mathbf{z} \tag{3.13}$$

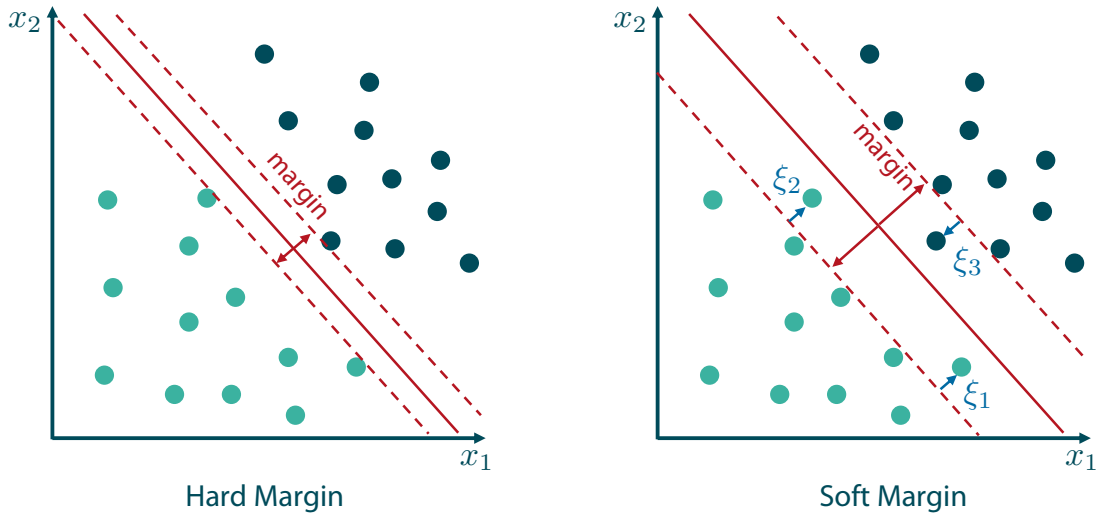

Figure 3.7: SVM classification using hard margin (left) and soft margin (right). By allowing misclassifications, a higher margin can be reached even though outlier are present in the data set.

Hence, using the linear kernel, the SVM learns a linear classification utilizing a hyperplane. In contrast, the RBF kernel is capable of highly nonlinear classification [129]:

$$k_{RBF}(\mathbf{x}, \mathbf{z}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{z}\|^2}{2\sigma^2}\right) = \exp\left(-\gamma\|\mathbf{x} - \mathbf{z}\|^2\right) \tag{3.14}$$

with

$$\gamma = \frac{1}{2\sigma^2}. \tag{3.15}$$

The hyperparameter γ gives control over the flexibility of the model. A high value leads to a very flexible model that is prone to overfitting, a small value tends to underfit the data. The optimal choice of γ is highly depending on the data set and is one challenge in training an SVM in practice.

3.3.2 Soft-Margin SVM

One important disadvantage of the SVM formulation in 3.11 is that the model is not capable of ignoring outliers. Even few outliers might lead to very small margins or models that tend to overfit as the constraints must not be violated. Hence, the method is also called hard-margin SVM. In contrast, in the so-called soft-margin SVM, slack variables

$\xi_i,\ i = 1, ..., n$ are introduced to the primal problem 3.4 to allow for violation of the constraints:

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \quad & \frac{1}{2}\mathbf{w}^\mathsf{T}\mathbf{w} + C\sum_{i=1}^{n}\xi_i \\ \text{subject to} \quad & y_i(\mathbf{w}\mathbf{x}_i + d) + \xi_i \geq 1, i = 1, \ldots, n, \\ & \xi_i \geq 0. \end{aligned} \tag{3.16}$$

In the constraints, the slack variables allow for distances smaller than 1 and hence allow that points lie within the margin or even on the wrong side of the hyperplane. However, the sum over all violations is added to the objective function. Hence, the optimizer aims at finding a good balance between a large margin and a smooth, generalizing model. Fig. 3.7 visualizes the difference between hard- and soft-margin SVM. The hyperparameter C gives control over this balance. For high values of C, smaller margins are accepted to ensure that no point violates the constraint. For small values, the model tries to find a large margin even if some points are misclassified. The optimal choice of C is highly task-specific and presents one challenge in applying SVMs to a practical problem.

In the dual formulation of the soft-margin SVM, C becomes an upper bound of the Lagrange multipliers:

$$\begin{aligned} \underset{\alpha}{\text{maximize}} \quad & \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{n} y_i y_j \alpha_i \alpha_j k(\mathbf{x}_i, \mathbf{x}_j) \\ \text{subject to} \quad & \sum_{i=1}^{n} y_i \alpha_i = 0, \\ & 0 \leq \alpha_i \leq C, i = 1, \ldots, n. \end{aligned} \tag{3.17}$$

Similar to the hard-margin case, kernels can be used to achieve nonlinear mappings.

3.3.3 Regression

The concept of SVMs cannot only be used for classification but also for regression problems. This extension is called Support Vector Regression (SVR) and is based on similar principles. However, the label $y \in \mathbb{R}$ is now a continous real value. Hence, the aim of SVR is to find a line that optimally fits the given data tuples $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2) \ldots, (\mathbf{x}_n, y_n)$. Once again, this problem can be formulated as a QP:

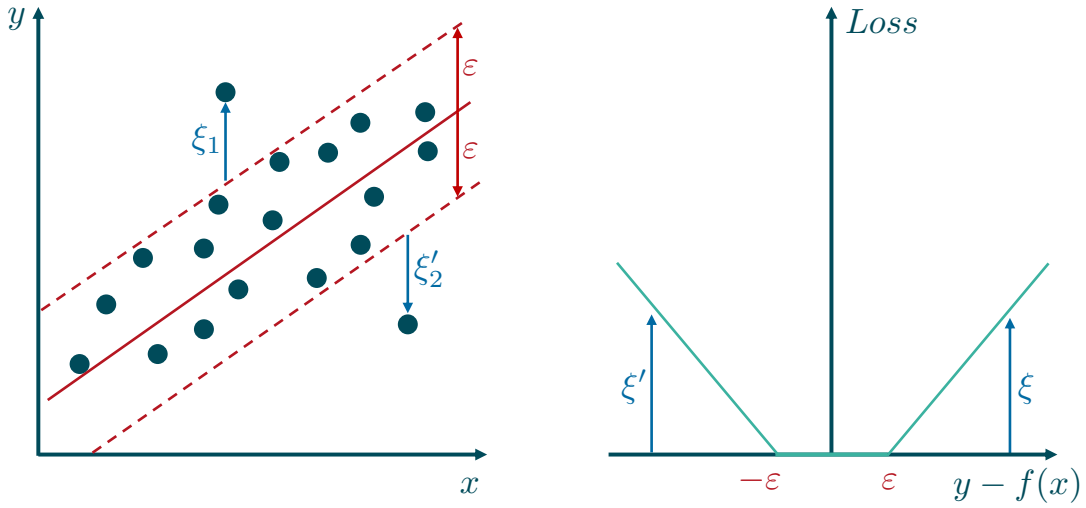

Figure 3.8: Visualization of a Support Vector Regression (left). All data points laying within the ε-tube are counted as perfectly fitted without errors. Data points outside the tube are penalized depending on their distance to the tube ξ or ξ'. This leads to the SVR's ε-insensitive loss function (right).

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \quad & \frac{1}{2}\mathbf{w}^{\mathsf{T}}\mathbf{w} + C\sum_{i=1}^{n}\left(\xi_i + \xi_i'\right) \\ \text{subject to} \quad & y_i - (\mathbf{w}\mathbf{x}_i + d) \leq \varepsilon + \xi_i, \\ & (\mathbf{w}\mathbf{x}_i + d) - y_i \leq \varepsilon + \xi_i', \\ & \xi_i, \xi_i' \geq 0, \, i = 1, \ldots, n. \end{aligned} \tag{3.18}$$

The constraints imply that the difference between the prediction $\mathbf{w}\mathbf{x}_i + d$ and the ground truth y_i is smaller than a specified value ε. As errors can occur in positive and negative direction, i.e. over- and underestimating the value y_i, both cases are included as a constraint for each data point. Similar to the soft-margin SVM, the slack variables ξ_i and ξ_i' are integrated to allow for violations of this constraint, while the number of violations is controlled by the parameter C. The hyperparameters ε and C should be optimized according to the task and the data set. Their choice presents a typical challenge in applying SVR to real-world problems. Fig. 3.8 visualizes the penalization in Support Vector Regression in dependence of the hyperparameters.

Similar to the SVM, the SVR problem stated in 3.18 can be transformed to its dual formulation using Lagrange duality:

$$\begin{aligned} \underset{\alpha,\alpha'}{\text{maximize}} \quad & -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i-\alpha_i')(\alpha_j-\alpha_j')\mathbf{x}_i^\mathsf{T}\mathbf{x}_j \\ & -\varepsilon\sum_{i=1}^{n}(\alpha_i+\alpha_i')+\sum_{i=1}^{n}y_i(\alpha_i-\alpha_i') \\ \text{subject to} \quad & \sum_{i=1}^{n}(\alpha_i-\alpha_i')=0, \\ & 0\le\alpha_i,\alpha_i')\le C, \qquad\qquad i=1,\dots,n. \end{aligned} \tag{3.19}$$

As each data point poses two constraints, there are Lagrange multipliers α' in addition to α. After solving the optimization problem, the corresponding parameters of the line can be computed as

$$\mathbf{w}^* = \sum_{i=1}^{n}(\alpha_i^*-\alpha_i'^*)\mathbf{x}_i \tag{3.20}$$

$$d^* = \begin{cases} y_j+\sum_{i=1}^{n}y_i(\alpha_i^*-\alpha_i'^*)\mathbf{x}_i^\mathsf{T}\mathbf{x}_j+\varepsilon, & \alpha_j\in(0,C),\ j\in\{1,\dots,n\} \\ y_j+\sum_{i=1}^{n}y_i(\alpha_i^*-\alpha_i'^*)\mathbf{x}_i^\mathsf{T}\mathbf{x}_j-\varepsilon, & \alpha_j'\in(0,C),\ j\in\{1,\dots,n\}. \end{cases} \tag{3.21}$$

Once again, the only computation that needs to be done in the feature space is the scalar product between data points. Hence, kernels can be used:

$$\begin{aligned} \underset{\alpha,\alpha'}{\text{maximize}} \quad & -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i-\alpha_i')(\alpha_j-\alpha_j')k(\mathbf{x}_i,\mathbf{x}_j) \\ & -\varepsilon\sum_{i=1}^{n}(\alpha_i+\alpha_i')+\sum_{i=1}^{n}y_i(\alpha_i-\alpha_i') \\ \text{subject to} \quad & \sum_{i=1}^{n}(\alpha_i-\alpha_i')=0, \\ & 0\le\alpha_i,\alpha_i')\le C, \qquad\qquad i=1,\dots,n. \end{aligned} \tag{3.22}$$

This allows the SVR model to learn nonlinear functions from the training data.

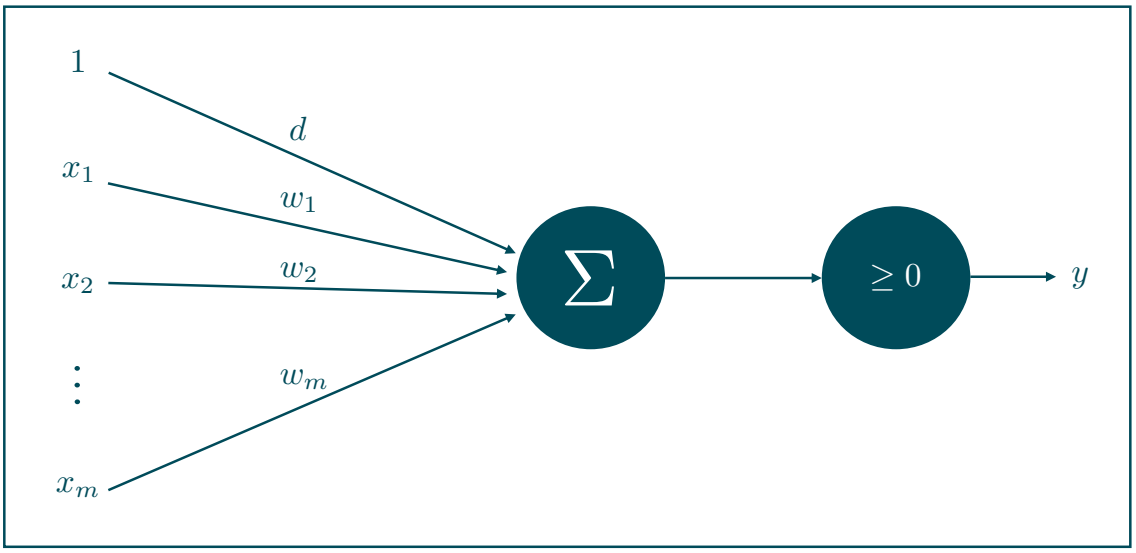

Figure 3.9: Visualization of a perceptron. The inputs are weighted by their individual weight and the sum of the result is computed. Then, the value of this weighted sum is compared to the threshold 0, leading to a binary classification. Note that the comparison to the threshold is a special kind of activation function σ.

3.4 Neural Networks

Artificial Neural Networks (NN) follow the approach to mimic biological information processing in human brains. As the name suggests, a neural network consists of many small units, the neurons, that are connected to a network-like structure. Even though a single neuron's information processing capabilities are limited, neural networks showed impressive performance in pattern recognition and learning from high-dimensional data [86]. In contrast to biological neural systems that consist of physical objects, NN are mathematical approximations of these systems and consist of blocks of program code, forming virtual networks of virtual neurons. The key idea behind NN training is to adjust the parameters in such a way that the prediction error for a given training data set gets lower. In this chapter, the theoretical framework of NN is presented. After describing the basic structure as well as the general learning strategy, specific types of NN architectures that are relevant in the scope of this thesis are presented in detail. An introductory overview can be found in [86], while [48] provides a deeper description of neural networks.

3.4.1 From Perceptrons to Feed-Forward Neural Networks

The explanations in this section mainly follow [100] and [106]. The smallest unit of a neural network is the mathematical model of a neuron, also called perceptron. The perceptron has several inputs $x_1, \ldots, x_m, x \in \mathbb{R}$ and one weight $w_1, \ldots, w_m, w \in \mathbb{R}$ assigned to each of these inputs. The information processing is performed by taking the weighted sum over all inputs and apply an activation function σ to the result. The computed value of the activation function is the output y of the perceptron. For classification, the activation function could be a step function with the step at 0. Then, the perceptron output can be computed as

$$y = \sigma\left(\sum_{i=1}^{m} w_i x_i + d\right) = \begin{cases} 1 & w_i x_i + d > 0 \\ 0 & w_i x_i + d \leq 0. \end{cases} \tag{3.23}$$

Fig. 3.9 shows a sketch of the perceptron's structure. If the output values 1 and 0 correspond to two different classes, a binary classification is possible using the above formula. It is interesting to note that this formulation directly relates to a separating hyperplane: The weighted sum with the added bias d defines a line like $\mathbf{w}^\mathsf{T}\mathbf{x} + d = 0$. Using the step function with threshold 0, the perceptron computes whether a data point $\mathbf{x} = (x_1, \ldots, x_m)$ lies on one side of this hyperplane or the other[1].

It is possible to adjust these parameters in such a way that all samples from a given data set are classified correctly. This is achieved via an iterative update of the weights, also known as the perceptron algorithm.

The perceptron itself is quite limited in its data processing capabilities. Specifically, the formulation in 3.23 constraints it to learning a linear classification using a separating hyperplane. Hence, the perceptron is not capable of dealing with non-linearly separable data. However, it is possible to link multiple perceptrons together to improve the data processing capabilities. This leads to a network of perceptrons, a neural network. In this context, the perceptrons are also called neurons.

The single neurons can be connected in different ways. One possibility is to stack multiple neurons in parallel to form a layer. Then, multiple layers can be stacked sequentially to form a Multi-Layer Perceptron (MLP). Fig. 3.10 exemplarily shows the architecure of an MLP with two hidden layers. In the MLP, the first layer is called input layer, the last layer is referred to as output layer and the intermediate layers are the hidden layers. Each neuron in the hidden layers and the output layer is connected to the output of every neuron in the

[1] There is plenty of literature that introduces the bias term d as a threshold S, where S is the step position in the activation function. The formulation in 3.23 was chosen due to its nice analogy to separating hyperplanes and SVM (see 3.3). Both notations can be transferred into each other by setting $d = -S$ [106].

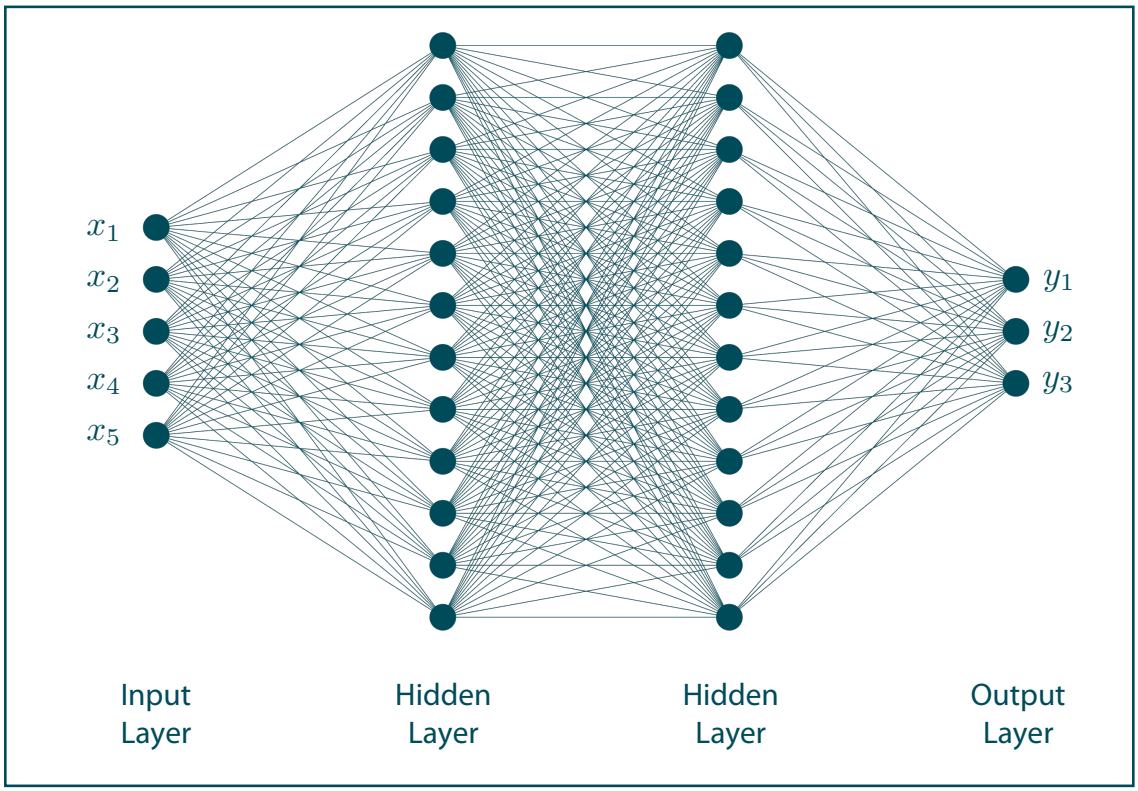

Figure 3.10: Architecture of a multilayer perceptron (MLP) with five inputs, two hidden layers with 12 neurons each and three outputs.

preceding layer, while each neuron in the input layer is connected to all input variables $x_1, \ldots, x_m$. The output of the MLP is the output of the last layer. Hence, the number of neurons in the output layer is the desired dimensionality of the target y. Due to the type of connections, the MLP is also called fully-connected and it is counted to the class of Feed-Forward Neural Networks. This means that the information within the network flows constantly from the input to the output layer without any back-loops or cycles. Note that there are plenty of network architectures that utilize such kind of backward connections, for example Hopfield Networks or Recurrent Neural Networks (RNN). For details, see [116].

The neurons in an MLP rather use a different activation function σ than a single perceptron. Besides the higher informative value of continuous outputs instead of the binary perceptron output, a step function is not convenient for training the MLP as described in the following. Additionally, nonlinear information processing can be integrated by using a

nonlinear σ. Typical activation functions that are also relevant in the scope of this thesis are *linear*, *sigmoid* and *Rectified Linear Unit (ReLU)* activation:

$$\begin{aligned} \sigma_{\text{linear}}(\nu) &= \nu \\ \sigma_{\text{sigmoid}}(\nu) &= \frac{1}{1+e-\nu} \\ \sigma_{\text{ReLU}}(\nu) &= \max(0,\nu). \end{aligned} \tag{3.24}$$

Another important activation function is the *softmax* activation. In contrast to the ones presented in 3.24, the softmax function is not applied within one neuron but across all K neurons in one layer, typically the last layer of the network, and is defined as:

$$\sigma_{\text{softmax}}(\mathbf{z})_i = \frac{\exp z_i}{\sum_{j=1}^{K} \exp z_j}, \; i = 1, \ldots, K. \tag{3.25}$$

Thus, the outputs of all neurons in the layer are transferred into a probability distribution where the sum over all outputs is one. This is a useful property to solve classification tasks as softmax output of each neuron corresponds to the probability of one class. By applying an *argmax* operation, the most probable class can be identified. It should be noted that an MLP in this form is not a bayesian system per se and hence, the values only represent an approximation of a true probability.

The basic idea of training an MLP is to adjust the weights of the network in such a way that the prediction error decreases. Thus, the accuracy of the prediction $\tilde{y}$ compared to the ground truth y is computed using an error function $\mathfrak{L}$, also called Loss Function. As the weights of all neurons influence this loss value, the loss function depends on each weight w and each bias d in the network. Hence, the network training can be formulated as a gradient descent on the loss function regarding the individual weights.

To compute the gradient of the loss function regarding each individual weight and bias, backpropagation is used. Even though first versions of this algorithm were around before, it was made popular in 1986 by [114]. As the name suggests, the key idea is to propagate the prediction error backwards through the network. Hence, the algorithm consists of a forward pass to compute the activations in each neuron and the value of the loss function and a backward pass to propagate the gradient backwards using the chain rule.

Let $w^l_{j,k}$ denote the weight of the connection from the k-th neuron in the $(l-1)$-th layer to the j-th neuron in the l-th layer (see Fig. 3.11). Similarly, d^l_j and a^l_j refer to the bias and the activation of the j-th neuron in the l-th layer. Furthermore, a weight matrix $\mathbf{w}^l$ and an

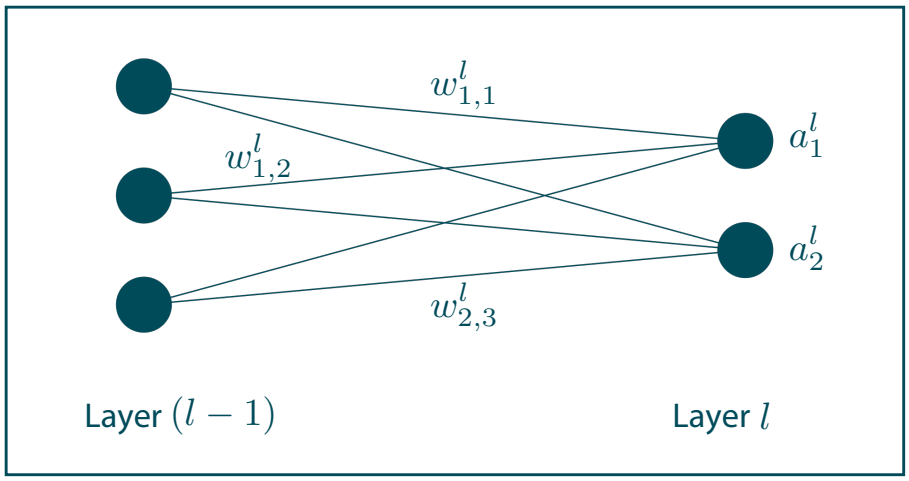

Figure 3.11: Nomenclature within an MLP, given for two subsequent layers $(l-1)$ and l.

activation vector $\mathbf{a}^l$ are introduced that contain all weights or activations of the neurons in the l-th layer:

$$(\mathbf{w}^l)_{j,k} = w^l_{j,k} \tag{3.26}$$

$$(\mathbf{a}^l)_j = a^l_j. \tag{3.27}$$

Then, the activation of the j-th neuron in layer l that contains K neurons is

$$a^l_j = \sigma\left(\sum_{k=1}^{K} w^l_{j,k} a^{l-1}_k + d^l_j\right) \tag{3.28}$$

while the activation of the whole layer can be computed as[2]

$$\mathbf{a}^l = \sigma(\mathbf{w}^l \mathbf{a}^{l-1} + d^l) \tag{3.29}$$

with $\mathbf{a}^0 \equiv x$, where x is the input data.

Given a data tupel (x, y), the activation in each layer of the MLP can be computed using 3.29 iteratively, starting with the first layer. This is called forward propagation and represents the first step of the backpropagation algorithm. After propagating the activations through the network, the output of the last layer is the predicted value $\tilde{y}$ that can be compared to the ground truth using the loss function $\mathfrak{L}(\tilde{y}, y)$. Hence, the error in each neuron of the last layer L can be computed as

$$\delta^L_j = \frac{\partial \mathfrak{L}}{\partial a^L_j} \sigma'(r^L_j) \tag{3.30}$$

[2]Note that the notation in 3.29 implies a vectorized computation of the activation function σ, see [106] for details.

with

$$r_j^L = \sum_{k=1}^{K} w_{j,k}^l a_k^{l-1} + d_j^l. \tag{3.31}$$

In matrix form, the error δ^L of the last layer is

$$\delta^L = \nabla_{\mathbf{a}^L} \mathfrak{L} \odot \sigma'(\mathbf{r}^L) \tag{3.32}$$

with

$$\mathbf{r}^L = \mathbf{w}^l \mathbf{a}^{l-1} + d^l, \tag{3.33}$$

where $\odot$ is the Hadamard-product. The vector $\mathbf{r}^l$ is called the weighted input of the l-th layer.

Based on this error term in the last layer, the error in the preceding layer can be computed recursively using the chain rule:

$$\delta_j^l = \frac{\partial \mathfrak{L}}{\partial r_j^l} \tag{3.34}$$

$$= \sum_{k=1}^{K} \frac{\partial \mathfrak{L}}{\partial r_k^{l+1}} \frac{\partial r_k^{l+1}}{\partial r_j^l} \tag{3.35}$$

$$= \sum_{k=1}^{K} \frac{\partial r_k^{l+1}}{\partial r_j^l} \delta_k^{l+1}. \tag{3.36}$$

Hence, the influence of each neuron on the prediction error can be calculated throughout the whole network. After backpropagating these errors, the partial derivatives of the loss function regarding the weights and biases of the network are given as

$$\frac{\partial \mathfrak{L}}{\partial w_{j,k}^l} = a_k^{l-1} \delta_j^l \tag{3.37}$$

$$\frac{\partial \mathfrak{L}}{\partial d_j^l} = \delta_j^l. \tag{3.38}$$

With these gradients, it is possible to perform a gradient descent to adapt the weights and biases in order to minimize the loss function value. Hence, one important condition on the loss function is that it is differentiable.

Typically, the gradient descent is performed in a stochastic way, i.e. the gradients are not computed for one data pair $(\mathbf{x}, y)$ but for a set of data points, the so-called batch or mini-batch. This improves the robustness of the optimizer and speeds up the training process.

3.4.2 Convolutional Neural Networks

As the name suggests, Convolutional Neural Networks (CNNs) follow the idea to integrate a convolution of the input data with kernels into the structure of a neural network. As the convolution of a signal with a kernel is a well-known approach to extract features from the signal and the values of the kernel are weights within the network, CNNs are capable of autonomously extracting features that are relevant for the given problem. While the concept was already presented in 1989 by [87], the breakthrough of CNN began in 2012 with the impressive performance of AlexNet in the ImageNet image classification challenge [83]. The explanations in this section follow [48].

As mentioned above, the key idea is to perform a convolution operation within a neural network. A convolution of a signal $x(t)$ with a Kernel K is defined as

$$s(t) = \int x(a)K(t-a)\mathrm{d}a = (x \star a)(t). \tag{3.39}$$

The result of the convolution $s(t)$ is also referred to as feature map. If the input is a two dimensional image I with $P \times Q$ pixels, the discrete convolution becomes

$$S'(i,j) = (I \star K)(i,j) = \sum_{p=1}^{P}\sum_{q=1}^{Q} I(p,q)K(i-p, j-q). \tag{3.40}$$

Fig. 3.12 shows the convolution of an input image with a kernel to receive a feature map. It is important to note that most programming libraries implement CNNs using cross-correlation, which is equal to a convolution with a flipped kernel:

$$S(i,j) = (I * K)(i,j) = \sum_{p=1}^{P}\sum_{q=1}^{Q} I(p,q)K(i+p, j+q). \tag{3.41}$$

To align with these libraries and the literature, 3.41 will be denoted as convolution in the scope of this thesis.

In classical signal processing or image analysis, the convolution of the input signal or image with a filter kernel is widely used to extract features. The values in the kernel relate to the feature that should be found: The value of the resulting feature map at a specific position gives information on the similarity between the corresponding image region and the kernel. However, designing feature kernels manually typically requires a lot of prior knowledge of the task and the given data. The key idea of CNNs is to introduce a type of layer into a feed-forward neural network that performs a convolution of the input with a specified kernel, while the values of the kernel are weights of the network. Hence, these

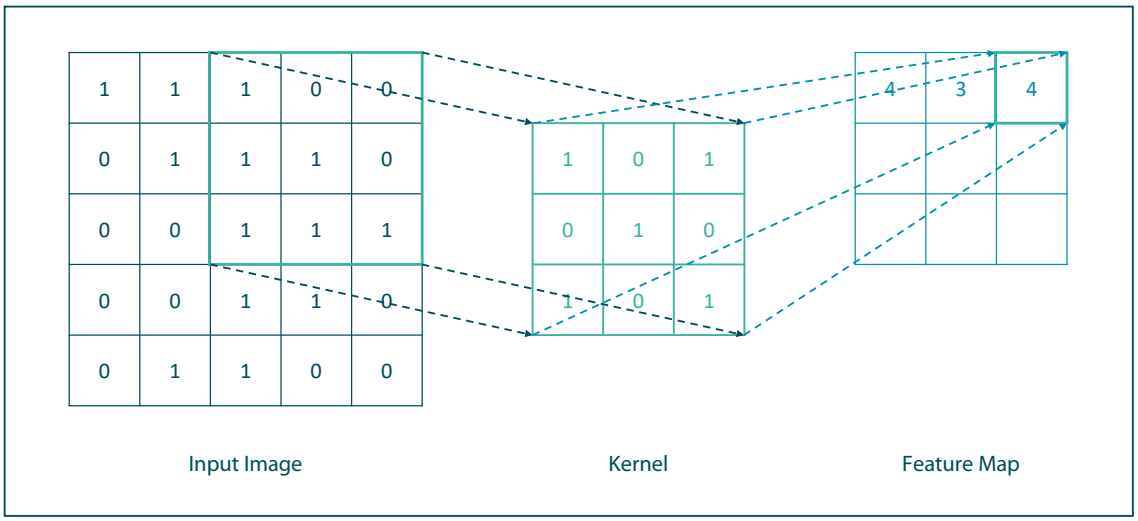

Figure 3.12: Convolution of an image with a filter kernel. The feature map is computed by calculating the weighted sum of the kernel with the corresponding image region.

weights, i.e. the kernel, is trainable according to the loss function. Therefore, CNNs are capable of extracting features that are relevant to solve the given problem completely data-driven.

In practice, this convolution within the structure of a neural network is achieved using three concepts: sparse interactions, parameter sharing and equivariant representation [48] which are explained in the following.

Sparse interaction means that in contrast to the MLP, neurons in a convolutional layer are not connected to all neurons of the previous layer, but only to a small part of them. In the case of two dimensional image analysis, one neuron would be connected to a small region of the input image, for example a 3×3 region. Note that the neuron performs the weighted sum of its inputs. If the weights of the neurons are interpreted as the values of a kernel, each neuron implements the computation of the weighted sum between the small image region and the kernel for one position.

Parameter Sharing connects this effect to a full convolution operation. Each neuron in one convolutional layer is connected to one input region and hence compares this region to a kernel specified by its individual weights. This is visualized in Fig. 3.13. Parameter sharing means that all these neurons are sharing the same weights. Hence, all image regions are compared to the same kernel and the layer performs a convolution of the input with a kernel defined by the shared weights of the neurons. The output of the layer is a feature map describing the activation of the filter kernel at a specific position of the input. Note that each neuron still applies an activation function σ to the result of the

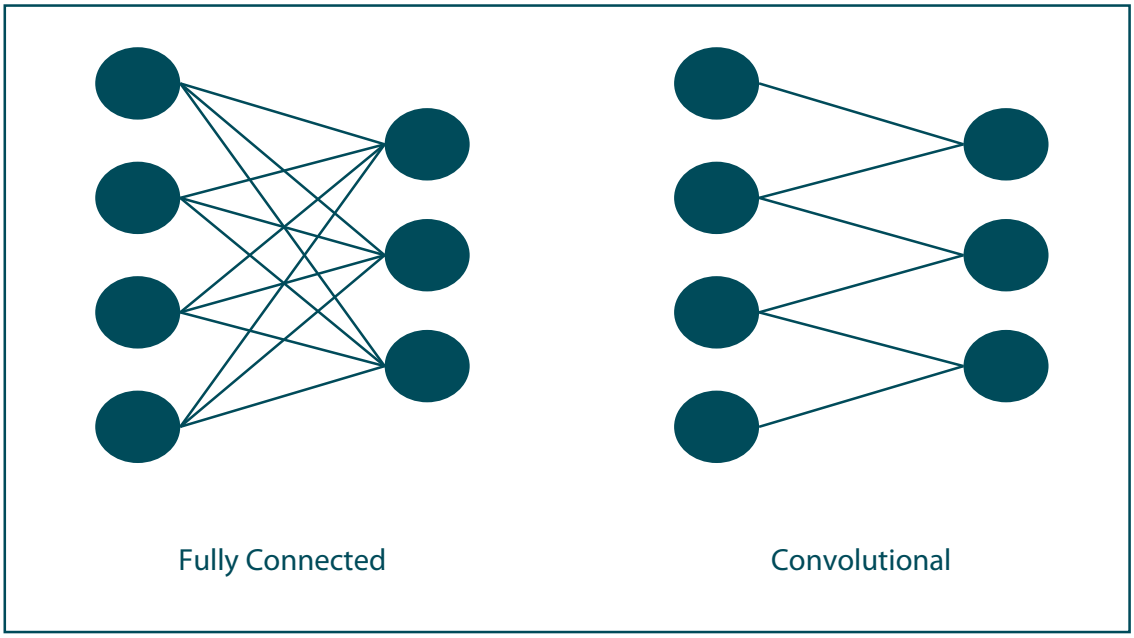

Figure 3.13: Connectivity between subsequent layers. While each neuron in a fully-connected layer is connected to each neuron from the precedent layer (left), each neuron in a convolutional layer only receives input values from a small region of the previous layer (right), exemplarily shown in 1D with a kernel size of two.

weighted sum. Hence, the output of a convolutional layer is the computed feature map after applying the potentially non-linear activation function to it.

Equivariance relates to the property of a function that a specific change that is applied to the input is similarly observed at the output. Convolutional layers are equivariant to translations, i.e. if the input is translated in a specific direction, the feature map is translated the same way. This property is very useful in image processing as the feature map does not only provide information on the feature activation but also on the location of it within the input image.

Architecture of Convolutional Neural Networks

Even though there are many possibilities to create a CNN, a typical CNN architecture emerged over the years [86]. One prominent example of this architecture is the *VGG-16* network [132]. Besides convolutional layers and fully connected layers as known from the MLP, two other kinds of layers are used: pooling layers and flatten layers. Pooling layers implement a size reduction by statistically summarizing nearby inputs. This means that a

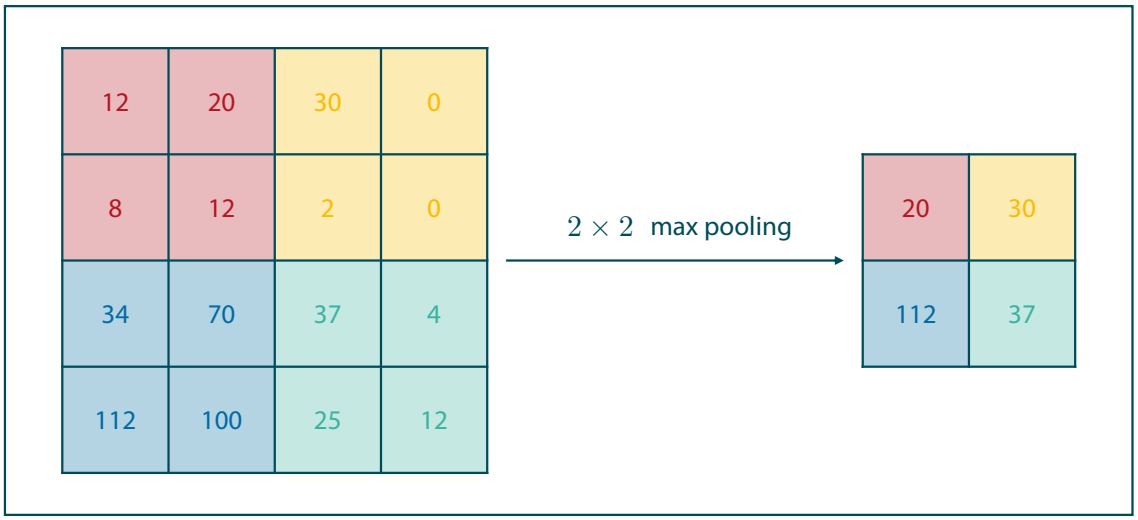

Figure 3.14: Principle of the max pooling operation, exemplarily shown for a pooling size of 2×2. From each 2×2 region, only the maximum value is kept, resulting in a dimensionality reduction.

certain region of the input is reduced to one value that summarizes all information in this region in a meaningful way. One prominent example is *max pooling* [150] that gives back the maximum of all values in the specific input region (see Fig. 3.14). Pooling operations have two advantages. First, they reduce the size of the data and, hence, the size of the following network layers. This is specifically important for high-dimensional data like high-resolution images or even image volumes. Second, pooling enforces invariance regarding small translations. Together with the general equivariance of convolutional layers, the network not only cares for the rough location of a feature to preserve relative positions of objects in the input image, but also cares whether a feature is there instead of giving the exact position on a finer scale. This combination provided impressive results in various applications of CNNs [86]. Typically, pooling applies to small input regions, e.g. 2×2.

Flatten layers perform a simple operation that simply converts a multi-dimensional input into a one-dimensional vector. Hence, a flatten layer does not change the input values but vectorizes it. A flatten layer can transform the output of a convolutional or pooling layer to a form that can serve as input for a fully-connected layer.

The typical architecture of a CNN for image classification can be separated into two parts: the feature learning part and the classification part. The first part is the feature learning part where features are extracted from the input image using convolution. Typically, multiple convolutions are stacked in parallel within one convolutional layer. This is

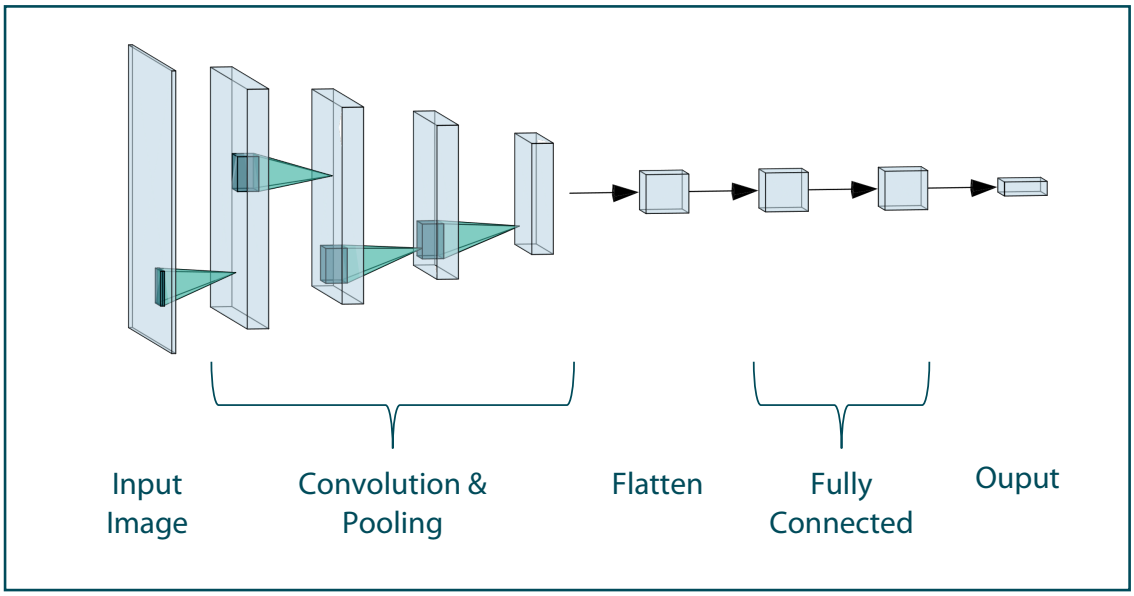

Figure 3.15: Typical architecture of a CNN. At first, meaningful features are extracted using convolutional layers and pooling. After a flatten operation, fully connected layers predict the output based on these features.

due to the fact that the aim is not to extract only one feature but a whole set, e.g. 64 features. Hence, the output of a convolutional layer is a stack of feature maps. After this convolutional layer, another convolutional layer might follow. Its inputs are the feature maps from the previous layer. The different feature maps are interpreted as further dimensions of the input data, also referred to as channels. Such a sequence of convolutional layers is typically followed by a pooling layer. It is also possible to apply pooling after each convolutional layer. The sequence of multiple convolutional layers and pooling operations can be interpreted as the feature learning part of the CNN architecture. At its end, a flatten layer marks the transition to the classification part. This part consists of a typical MLP that is connected to the flatten layer. Fig. 3.15 shows such a typical CNN architecture.

The structure of a CNN resembles the typical machine learning pipeline where meaningful features are extracted from the raw input data and based on these features, a classifier tries to distinguish the data into several classes. The huge advantage of CNNs is that the features do not have to be selected manually but can be learned implicitly by the model.

3.4.3 Autoencoders

Autoencoders are a special kind of architecture to achieve representation learning with a NN. The main goal is to find a low-dimensional representation $\mathbf{z} \in \mathbb{R}^d$ of the data $\mathbf{x} \in \mathbb{R}^m$ with $d \ll m$ that carries a lot of information of the original data. This is achieved by mapping the input to itself throughout the autoencoder while the architecture has a bottleneck with very few neurons. The activation in this bottleneck refers to $\mathbf{z}$ and by forcing the autoencoder to reconstruct the original input image from these few values, the bottleneck representation $\mathbf{z}$ becomes highly descriptive of the data $\mathbf{x}$. Hence, autoencoders are often used to identify sufficient features that can be used by other learning methods.

An autoencoder consists of two subnetworks: The encoder $enc : \mathbb{R}^m \mapsto \mathbb{R}^d$ and the decoder $dec : \mathbb{R}^d \mapsto \mathbb{R}^m$. The encoder maps from the input data to the low-dimensional representation while the decoder reconstructs the original data from this representation:

$$\mathbf{z} = enc(\mathbf{x}) \tag{3.42}$$

$$\tilde{\mathbf{x}} = dec(\mathbf{z}) = dec(enc(\mathbf{x})), \tag{3.43}$$

where $\tilde{\mathbf{x}}$ is the output of the autoencoder and, hence, an estimation of the input $\mathbf{x}$. Even though the encoder and the decoder network might appear quite different, the decoder network typically follows the mirrored architecture of the encoder. Fig. 3.16 shows the architecture of an autoencoder.

Note that $d \ll m$ is not a necessary condition for autoencoders in general. There are also variants with $d > m$, which is specifically important for learning sparse representations [48]. In the scope of this thesis, only the case $d \ll m$ is relevant, also referred to as undercomplete autoencoders. In image or signal processing, convolutional encoder architectures are common. In this case, the decoder performs a deconvolution operation [101].

3.4.4 Variational Autoencoders

A special form of autoencoders are variational autoencoders (VAE). The main difference to classical autoencoders is that the VAE assumes the autoencoder to be a variational model $p(\mathbf{z}, \mathbf{x})$, leading to a probabilistic view on the data. VAEs were introduced first by [81].

As the model is assumed to be variational, the encoder and decoder functions become the distributions $p_{enc}(\mathbf{z}|\mathbf{x})$ and $p_{dec} = p(\mathbf{x}|\mathbf{z})$. The representation $\mathbf{z}$ is also called latent representation or latent space. It is assumed that each data sample $\mathbf{x}_i, i = 1, \dots n$

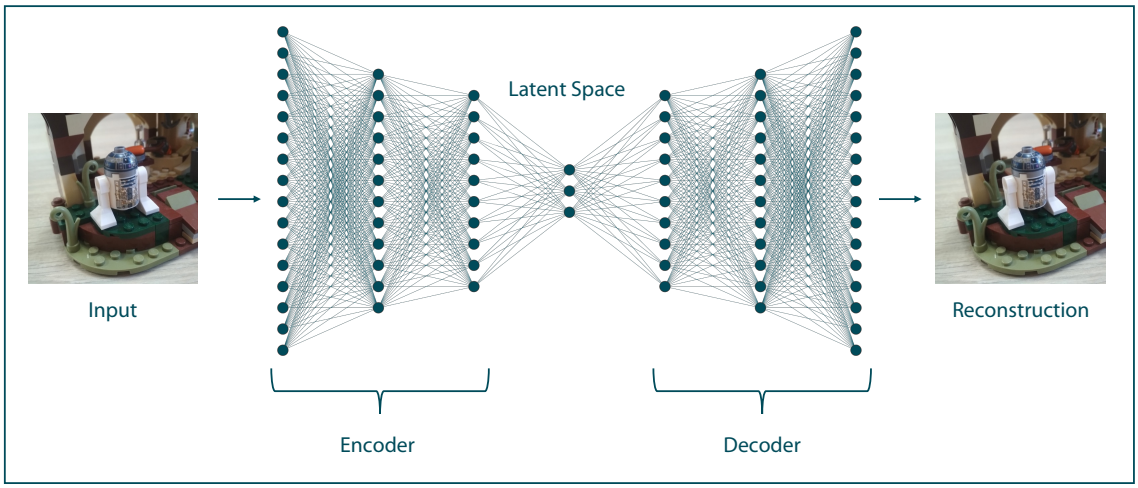

Figure 3.16: Architecture of an autoencoder. The input is propagated through the encoder, passes the low-dimensional bottleneck layer and is reconstructed throughout the decoder. The activation of the neurons in the bottleneck layer span the latent space, i.e. the identified representation.

is represented in the latent space as a gaussian with mean $\mu_i \in \mathbb{R}^d$ and variance $\mathbf{\Sigma}_i \in \mathbb{R}^d$. Hence, each data point has a probabilistic representation with a specific variance. To map from the latent space to the data space, the decoder samples from the gaussian distribution and reconstructs a data point based on this sample. Additionally, it is enforced that the mixture of all encoded data points in latent space approximates a gaussian distribution as well. Hence, the loss function of a VAE is given by

$$\mathfrak{L}_{vae} = \mathbb{E}(\log p_{dec}(X|z)) - D_{kl}(p_{enc}(z|X)||p_{dec}(z)), \tag{3.44}$$

where D_{kl} is the Kullback-Leibler-Divergance [81]. It is important to note that training a feed-forward neural network with integrated sampling from the latent sample distribution $\mathcal{N}(\mu_i, \mathbf{\Sigma}_i)$ is not possible in a straight-forward way. Hence, [81] introduced the reparameterization trick. The key idea is that the sampling is not done directly in latent space, but from a gaussian distribution $\mathcal{N}(\mathbf{0}, I)$, where $\mathbf{0} \in \mathbb{R}^d$ is a vector of zeros and $I \in \mathbb{R}^{d \times d}$ is the identity matrix. The result of this sampling is then multiplied with $\mathbf{\Sigma}_i$ and μ_i is added afterwards. This procedure provides the same result as sampling from $\mathcal{N}(\mu_i, \mathbf{\Sigma}_i)$ but allows for backpropagation through the whole VAE. Details can be found in [37]. Fig 3.17 visualizes the concept of the reparameterization trick.

In practice, VAE have the huge advantage of providing a smooth latent space. This is due to the fact that all data points form one cluster in latent space as they form a gaussian

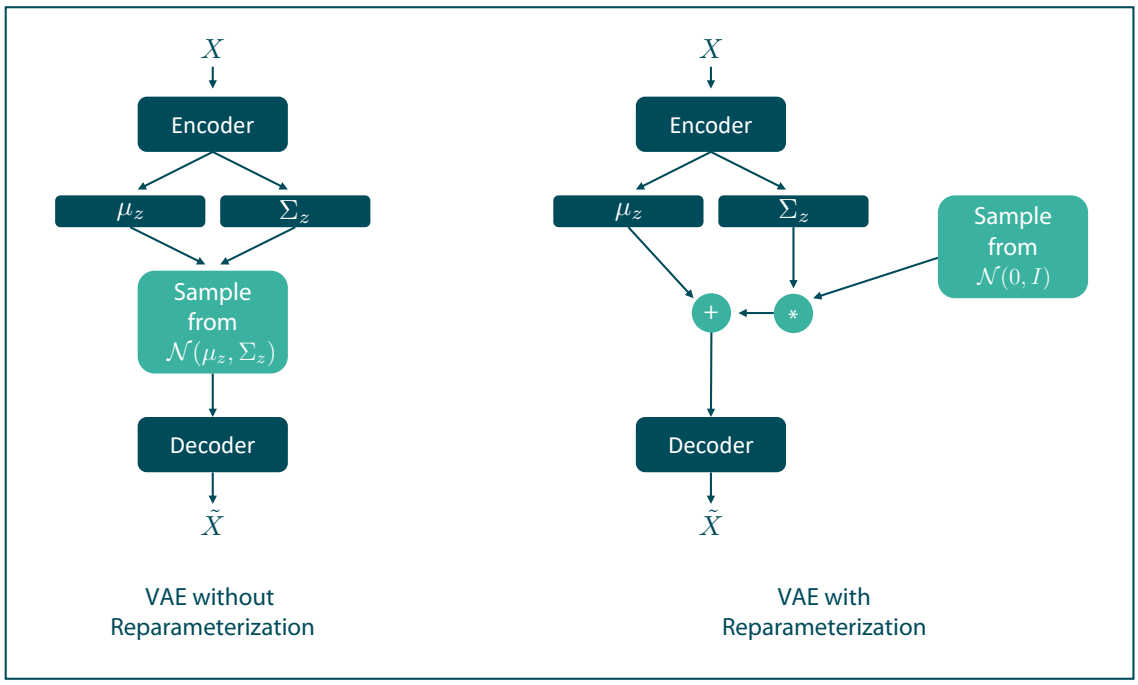

Figure 3.17: Visualization of the reparameterization trick. Instead of sampling from the distribution predicted by the encoder (left), the sampling is done from a gaussian distribution with a mean of zero and a variance of one. The sampled point is then scaled by the predicted variance and added to the predicted mean. This allows for applying backpropagation to train the model.

distribution and that during training samples around each data point are used for back-propagation. This implies an implicit constraint on the decoder that points that lie close together in latent space should lead to similar reconstructed data. Hence, interpolation in latent space becomes possible, making VAEs an often used approach for data synthesis ([71], [144], [122]).

3.5 Random Forests

Random Forests (RF) were introduced by [20] and are a machine learning method based on the concept of decision trees. Due to the small number of parameters and the strong statistical properties, they were shown to perform robustly in numerous applications. The following explanations mainly follow [64].

The Random Forest is based on decision trees. These are representations of the training data in a tree structure, allowing for a simple classification. Each node of the

tree describes a test on the input data, e.g. a threshold test. The branches that descend from this node represent the possible results of the test. The leafs represent the output value that can be assigned to the corresponding input data, i.e. the classification result. Hence, a decision tree implicitly partitions the feature space into regions with the same output, where all partition borders are lines orthogonal to one dimension. This is due to the threshold test on one feature within each node. The tree can be built using the training data, so that the threshold values of the individual nodes are adapted to the data. This corresponds to the training of a decision tree.

A Random Forest consists of a set of different, decorrelated decision trees. The individual trees are built on random subsets of the training data set. For this purpose, a random sample is drawn from the number of training data points to build a single tree, and the tree is trained based on this sample. Furthermore, only a random subset of the features is used during this training process, such that different trees are trained in different feature spaces, each of which represents a subspace of the entire feature space. This decorrelation procedure is also called bootstrapping [64].

Due to bootstrapping, the output of the trees differ for the same input data. The final prediction is determined as a majority vote, i.e. the output that is provided most frequently is the result of the classification.

Random Forests are not limited to classification but can also be applied to regression. Therefore, the majority vote is replaced by an averaging operation to combine the prediction results of all individual trees.

The advantage of Random Forests lies in their simplicity. There are only few parameters that have to be chosen by the user, which is why Random Forests often achieve good results without strong adaptation to the individual problem [64]. Moreover, both linear and nonlinear predictions are possible using the same method. On the other hand, a Random Forest's prediction can be computed very efficiently, since the construction of decision trees does not require a large amount of computational power.

3.6 Clustering

The term clustering refers to a problem class within unsupervised learning. Given data points $\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{x} \in \mathbb{R}^m$, the aim is to find clusters in the data set that divide the data into classes [116]. As the problem is unsupervised, no labels for the data points are known. Hence, the method's output should be a partitioning of the feature space into several clusters as well as an assigned cluster label for each data point. In the scope of this thesis, two typical clustering approaches will be relevant: k-means-clustering and

Gaussian Mixture Model (GMM) clustering. Both methods are presented in detail in the following paragraphs.

3.6.1 k-means Clustering

The idea of the k-means-algorithm is that each cluster of data points should have a center and that this center should lie in the centroid, i.e. the mean average, of all data points belonging to the respective cluster. Hence, the computation is straight-forward: the feature space is divided into a set of k clusters, defined by their cluster center. Each data point is assigned to the cluster with the closest distance. Then, the cluster centers are adjusted iteratively to approximate the centroid of their cluster. The below description follows [100].

First, the k cluster centers $\mu_1, \ldots, \mu_k, \mu \in \mathbb{R}^m$ are initialized randomly. These centers represent clusters where each one represents one class of data. After initialization, each datapoint $\mathbf{x}_i, ; : i = 1, \ldots, n$ is assigned a class label c_i depending on the closest cluster center:

$$c_i = \arg\min_j \|\mathbf{x}_i - \mu_j\|^2, \ i = 1, \ldots, n. \tag{3.45}$$

After assigning the data points to the cluster centers, each cluster center is moved to the center of its cluster points, respectively:

$$\mu_j = \frac{1}{|\Omega_j|} \sum_{\omega \in \Omega_j} \mathbf{x}_\omega, \ j = 1, \ldots, k, \tag{3.46}$$

where Ω_j is the set of indices of the data points assigned to the cluster j. The steps described in 3.45 and 3.46 are repeated until the cluster centers do not move anymore.

The k-means algorithm is an easy yet robust algorithm for clustering. However, the choice of the optimal value of k might be challenging for unknown data. Additionally, the distance metric in 3.45 is error prone in high-dimensional spaces or on non-flat manifolds [100].

3.6.2 Gaussian Mixture Models

Gaussian Mixture Models provide a probabilistic approach for clustering a data set. The key idea is to explain the data by a mixture of gaussian probability distributions. By maximizing the expectation, these gaussians can be adapted so that the resulting mixture has the highest probability to lead to the given data set when n data points are sampled from it. The description in this paragraph mainly follows [116].

In the GMM approach, the data is assumed to be sampled from a mixture distribution with k components where each of these components is a gaussian distribution. Hence, the mixture distribution is given by

$$p(\mathbf{x}) = \sum_{j=1}^{k} p(c=j)p(\mathbf{x}|c=j) \tag{3.47}$$

$$= \sum_{j=1}^{k} \pi_j \mathcal{N}(\mathbf{x}|\mu_j, \mathbf{\Sigma}_j), \tag{3.48}$$

where $\mathbf{x} \in \mathbb{R}^m$ is a data point, c denotes the component, μ_j and $\mathbf{\Sigma}_j$ are the mean and variance of the j-th component and π_j is the weight of this component. Hence, the parameters of the GMM are the component weights π_j, means μ_j and variances $\mathbf{\Sigma}_j,\ j = 1, \ldots, k$. These parameters can be optimized using the expectation maximization (EM) algorithm on given training samples $\mathbf{x_1}, \ldots, \mathbf{x_n}$. The algorithm consists of two steps: The E-step and the M-step [116]. In the E-step, the probability $p_{j,i} = p(c=j|\mathbf{x}_i)$ of the i-th data point being generated by the j-th component is computed for each data point. For the computation, bayes theorem is utilized with $p(c=j) = \pi_j$:

$$p_{j,i} = \pi_j p(\mathbf{x}_i|c=j). \tag{3.49}$$

Additionally, the number of data points assigned to component j is denoted as n_j with

$$n_j = \sum_{i=1}^{n} p_{j,i}. \tag{3.50}$$

Based on these probabilities, also referred to as expectation, the model parameters can be updated in the M-step. This update implicitly maximizes the log likelihood of the data. the update of the parameters is computed as

$$\mu_j \longleftarrow \sum_{i=1}^{n} \frac{p_{j,i}\mathbf{x}_i}{n_j} \tag{3.51}$$

$$\mathbf{\Sigma}_j \longleftarrow \sum_{i=1}^{n} \frac{p_{j,i}(\mathbf{x}_i - \mu_j)(\mathbf{x}_i - \mu_j)^\mathsf{T}}{n_j} \tag{3.52}$$

$$\pi_j \longleftarrow \frac{n_j}{n}. \tag{3.53}$$

The two steps of the EM algorithm are repeated iteratively until convergence is reached. Using this algorithm, GMMs become a clustering tool with a broad theoretical basis from bayesian theory. However, it is possible that one component only covers one data point, leading to a variance close to zero and a likelihood that goes to infinity. Another issue is the merging of two components that share the same mean and variance. Both problems challenge the clustering algorithm, especially in high-dimensional feature spaces.

4 Data Acquisition and Preprocessing

Training and evaluating methods for pseudohealthy synthesis require sufficient data sets. For the case of structural pathology, the model should estimate the shape of the individual valve leaflets based on the shape of the surrounding tissue, i.e. the aortic root. Hence, a setup for aortic root imaging as well as another one for leaflet imaging is necessary. For the case of morphological pathologies, the aortic root shape should be estimated based on a dilated root shape. Thus, the setup for aortic root imaging is sufficient. Fig. 4.1 visualizes the data sets necessary to develop and evaluate pseudohealthy synthesis methods for both pathological cases.

In this chapter, the acquisition of these data sets is described in detail. First, the general methodology of the data acquisition is presented in 4.1, including the processing of fresh hearts, the setups used for acquisition as well as data preprocessing methods. After that, the specific collection of the data sets for both cases using the presented methods is described in 4.2 and 4.3. Finally, possible sources of errors as well as limitations of the data set's validity are discussed in 4.4. A compact overview of the data set nomenclature can be found on page xxi.

Contribution

To the best of the author's knowledge, the data sets collected in the scope of this thesis are unique. In the case of a structural pathology, the collected data set presents the first and only one containing not only high resolution leaflet images but also 3D ultrasound images of the aortic root. Hence, it is the only data set available to evaluate pseudohealthy synthesis methods for personalized aortic valve prosthesis shaping. Comparable data sets only contain high-resolution images of the valve leaflets [118], [36], while the size of these data sets is still smaller than the on of the leaflets-only data set D_S^{aux} collected in the scope of this study. Regarding a morphological pathology, the data set presented in this thesis is the largest one available. The only comparable data set was collected by [1] using the same dilation method. However, only eight roots were examined in this study, while the data set used in the scope of this thesis contains 24 aortic roots.

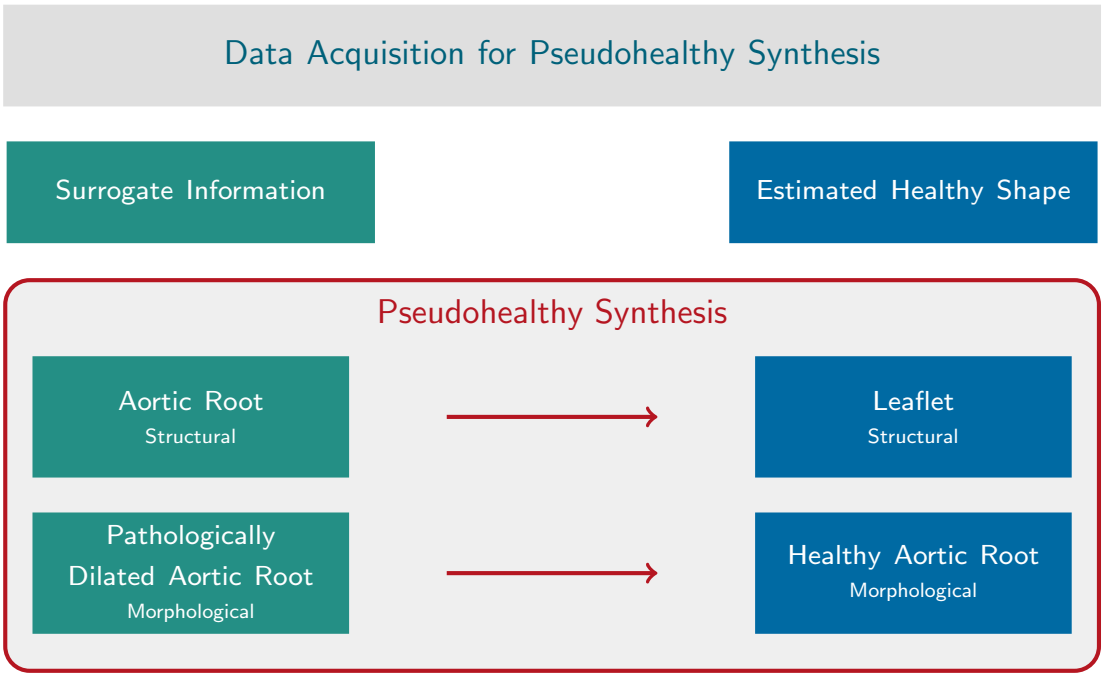

Figure 4.1: For both pathological cases, it is necessary to know the surrogate information as well as the desired healthy shape of each data sample. Hence, for the structural pathology, the aortic root shape as well as the shape of its leaflets are required. For the morphological pathology, the aortic root shape has to be known in a pathologically dilated as well as in its native, healthy state.

4.1 Methodology of Data Acquisition

In this thesis, pig hearts were used as a well-established animal model instead of human hearts. The cardiac anatomy of a pig is very similar to the human one [30]. This specifically holds for the aortic root region as well as the aortic valve apparatus which is of interest in this study. This is supported by the fact that porcine aortic valves are even used as xenoprosthetics for aortic valve replacement in humans [39]. Therefore, the porcine aortic root and valve are commonly used for cardiovascular research [118], [1], [10].

The fresh porcine hearts were bought at the slaughterhouse, so there are no ethical issues to declare. Directly after the purchase, the hearts were cleaned and frozen for later use to reduce biomechanical changes to a minimum. For each data acquisition experiment, one heart was defrosted in a water bath of room temperature over night.

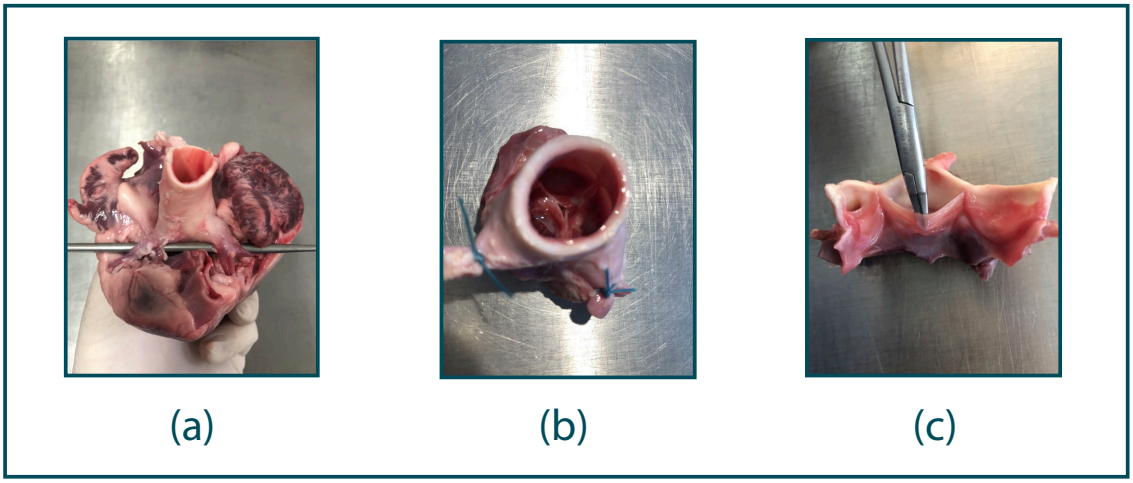

Figure 4.2: Aortic root extraction. At first, the aorta is cut below the aortic arch and the coronary arteries are exposed (a). Then, the aortic root is cut out of the left ventricle and the coronary arteries were tied up (b). To extract the leaflets, the root was cut vertically in between the right-coronary and non-coronary sinus (c).

4.1.1 Aortic Root Imaging

As the aortic valve apparatus consists of very thin structures that move quite fast in the blood stream, imaging the aortic valve and root is challenging in classical volumetric modalities like CT or MRI. Due to its high resolution, the low costs and the wide applicability, transesophageal echocardiography (TEE) became the gold standard for assessing the aortic valve apparatus [39]. As the name suggests, the small ultrasound probe is located in the tip of an endoscope that is conducted into the esophagus of the patient. The aim of the aortic root image acquisition was to mimic this examination while allowing for all necessary manipulations of the aortic root shape.

Aortic Root Extraction At first, the aortic root was extracted from the porcine heart. Thus, the aorta was exposed and cut away right caudal of the aortic arch. Then, the coronary arteries that originate from the left- and right-coronary sinus were exposed for about $2\,cm$ and cut there. Afterwards, the aortic root base was cut out of the left ventricle and the root was extracted from the heart. After cleaning and cutting away unneeded material, the coronary arteries were tied up closely to avoid flow through them. The result was an extracted aortic root with a functioning aortic valve inside. Fig. 4.2 (a-b) shows photos of the extraction steps.

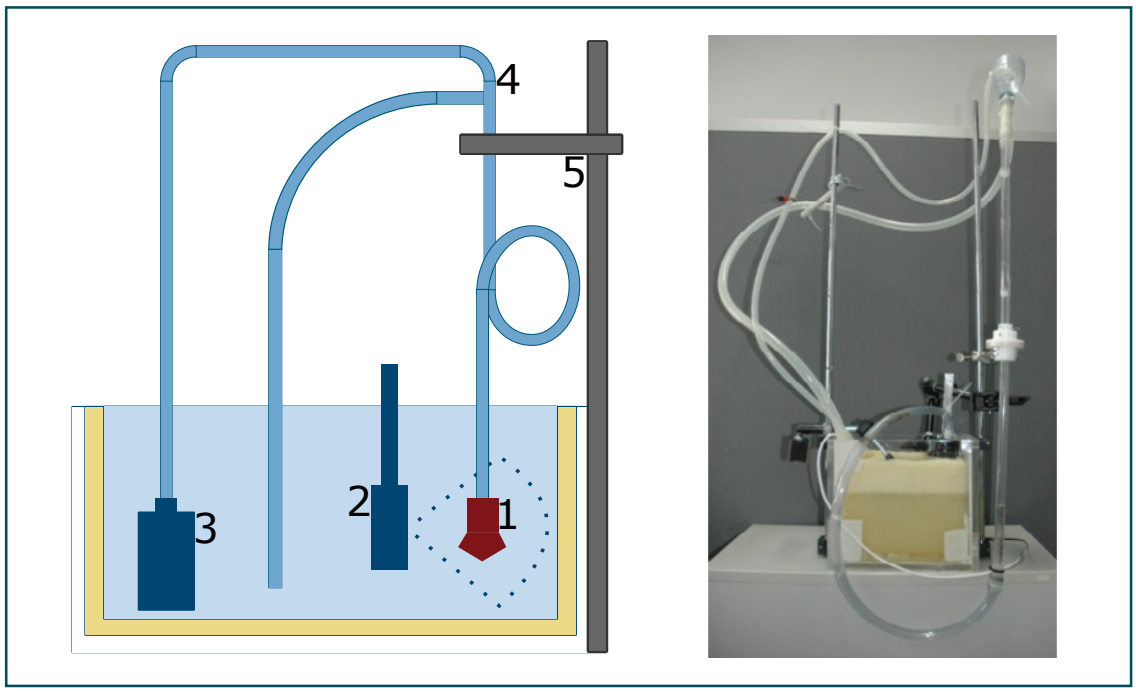

Figure 4.3: Setup for ultrasound imaging. Left: Sketch of the setup with mounted aortic root (1),TEE probe (2), water pump (3), an overflow for a constant water level (4) and adjustable height of the overflow to control the pressure (5). Right: Photograph of the setup. Adapted from [62] © 2016 IEEE.

Setup [1] As mentioned above, the aim was to mimic a TEE examination. This implies three key features the setup had to provide: The aortic root should be put under physiologically realistic pressure to assure a normal morphology, the TEE probe should be mounted roughly in parallel to the ascending aorta and both, the root and the probe, should be located in a medium that allows for ultrasound imaging.

Hence, the setup consisted of a water bassin in which the aortic root and the probe are located. The walls of the bassin were lined with foam to reduce sound echo artifacts. The aortic root was attached to a vertical tube filled with water. Hence, the column of water inside the tube created a pressure on the aortic valve, forcing it to close. The tube was mounted on a holder with adjustable height and featured an overflow back to the bassin. A pump constantly filled the vertical tube with water from the bassin, leading to a constant water level at the overflow even in the case of leakage in the aortic root. By adjusting the

[1] The setup described in this paragraph has been developed by the author in the scope of his master thesis [53].

height of the overflow, the pressure could be controlled. To reduce the amount of air bubbles in the ultrasound image, the pump and overflow tube were separated from the rest of the bassin using foam. Additionally, the bassin walls were lined with foam to reduce sound echos. Fig. 4.3 shows a sketch as well as a photograph of the setup.

The setup mimics a static scenario, i.e. there is no simulation of a pumping heart. Instead, it is assumed that the imaging takes place during the diastole when the aortic valve is closed. Besides allowing for a less complex setup, this assures that important features of the valves functionality can be assessed, e.g. the coaptation height or the height difference between the coaptation and the commissure plane. The conversion between the pressure units $mmHg$ and mmH_2O is given by

$$1\,mmHg \approx 13.595\,mmH_2O. \tag{4.1}$$

Hence, to achieve the typical diastolic blood pressure of around $80\,mmHg$, the height difference between the overflow and the aortic roots coaptation plane was adjusted to $108.7\,cm$.

3D Imaging For 3D imaging of the aortic root, a TEE probe was mounted in the bassin such that the aortic root was within its field of view, mimicking a realistic TEE examination (see Fig. 4.3). Throughout all measurements, the ultrasound images were taken in a aperture angle of $40°$ and a depth of $10\,cm$. Two different ultrasound setups were utilized: a *GE vivid 7* station with a *6T* TEE probe and a *GE vivid E95* station with a *6VT-D* TEE probe. While the the *6VT-D* is a matrix array probe and is capable of providing 3D volumetric images, the *6T* probe is limited to provide 2D images but features a rotatable transducer. Hence, in the latter case, a further postprocessing step was necessary to generate volumetric images. Using the $6T$ probe with the *GE vivid 7* station, a sequence of 2D images was acquired while continuously rotating the imaging plane from $0°$ to $180°$. Assuming a constant rotational velocity and hence angle steps between consecutive images, the current angle of each image within the sequence could be assigned by linearly interpolating the angle over the image indices. Given the 2D images and their corresponding acquisition angle, the data was interpolated on a 3D cartesian grid, resulting in a 3D volumetric image of the aortic root. It should be noted that this procedure results in a spatial dependency of the image resolution. Pixels in the outer regions of the volume have a lower resolution as the ones close to the rotational axis due to higher interpolation errors. Therefore, the aortic root was placed in the middle of the 2D image to avoid inaccuracies in the relevant image regions.

4.1.2 Leaflet Imaging

With an average thickness of $1.75 - 150\,\mu m$ [102], the aortic valve leaflets and specifically their inner structure of collagen fibers is barely visible in classical medical imaging modalities. Hence, the leaflets are extracted from the heart and assessed in an ex-vivo state.

Leaflet Extraction At first, the aortic root was extracted from the heart as described in Sec. 4.1.1. Then, the root was cut vertically in between the right-coronary and the non-coronary sinus without harming the leaflets. Hence, the root could be unrolled and the three leaflets were assessible (see Fig. 4.2). Then, the individual leaflets were cut off the aortic root wall along the commissure line. Special care was taken to follow the natural commissure line as close as possible. To prevent the extracted leaflets from drying out, they were preserved in individual containers in an isotonic saline solution ($9\,NaCl\,g/l$).

Setup To obtain high resolution images of the valve leaflets, a setup was utilized that has previously been developed in the Clinic for Thoracic Surgery, University Hospital Schleswig-Holstein, Germany, in the group of Dr. Michael Scharfschwerdt [123]. The setup was initially designed to examine bovine pericardium based on blue light transmission. This thesis presents the first approach to examine the aortic valve leaflets with such a technique.

The three leaflets of one valve were spread on an optical diffusion plate. Special attention has been paid to the preservation of the natural shape with minimal deformation. The plate was backlit with blue light (470 nm) using LEDs and a camera (*Canon DS 6041*) was mounted in a fixed distance above this plate, allowing to take a photograph of the scene. The diffusing plate provides homogeneous illumination. The blue light is absorbed in collagen tissue, resulting in a high contrast of the leaflet and its inner structure. Specifically, the prominent collagen fibers on the leaflets are clearly visible. Fig. 4.4 shows a sketch of the setup as well as a raw leaflet image exemplarily.

Preprocessing The data acquisition was followed by a preprocessing pipeline. Aiming at a suitable, unified representation of the leaflets, the steps of this pipeline are described in the following. First, the images were cropped manually to extract images of the single leaflets. Each image was converted to grayscale (ranging from 0 to 255) and inverted, resulting in an image corresponding to a thickness profile. To avoid noise around the leaflet, a mask was created. For this purpose, the background pixels were set to 0 using

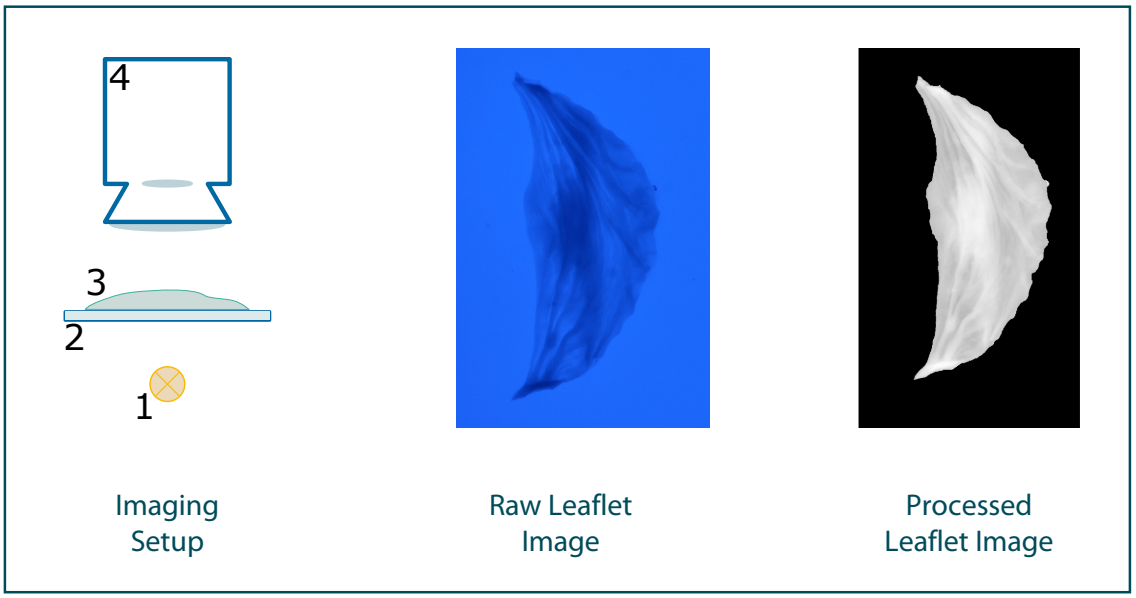

Figure 4.4: Leaflet imaging. Left: Sketch of the imaging setup with blue light source (1), diffusing plate (2), leaflet (3) and camera (4). Middle: Raw leaflet image. Right: Leaflet image after preprocessing. Adapted from [60].

thresholding. To avoid holes in the segmentation due to very thin areas of the leaflets, the segmentation thresholds were adjusted manually for each image, ranging from 158 to 168. Afterwards, the leaflets were centered by translating the center of mass to the image's mid point and by manually rotating the leaflet so that the commissure points were vertically aligned. Finally, the images were downsampled to a size of 128×64 pixels with a resolution of $0.34\ \frac{mm}{pixel}$. An example of a preprocessed leaflet image can be found in Fig. 4.4.

4.2 Data Acquisition for Structural Pathology

To examine the case of structural pathology, several data sets were acquired in the scope of this thesis. To answer the research questions, it is necessary to collect the aortic root shape as well as the shape of the individual leaflets from each examined heart. Additionally, to ensure the synthesis of realistic leaflet images, a broad data set of healthy leaflet shapes is helpful to train generative models. In total, three data sets were collected, where two contained the aortic root shape as well as the leaflet shapes and one consisted of leaflet shapes only.

4.2.1 Acquisition Workflow

As mentioned above, three data sets were collected. The acquisition workflow slightly differed between them, hence the workflows are presented individually for each data set in the following paragraphs.

Initial Data set with Root and Leaflet: $\tilde{D}_S$ The first data set contains the shape of the aortic root as well as the shape of the corresponding leaflets of ten porcine hearts. At first, the root was extracted and 2D ultrasound images were acquired using the *GE vivid7* ultrasound station with the *6T* probe and 3D volumes were computed as described in Sec. 4.1.1. Then, the leaflets were extracted from the root and images of them were acquired. This data set was used to validate that the prediction of leaflet features is possible based on the aortic root shape in the scope of answering research question *Q1*.

Broad Data set with Root and Leaflet: D_S To address the research questions *Q2* and *Q3*, the data set $\tilde{D}_S$ is limited in its expressiveness. Not only does it consist of ten hearts only, but also does the *6T* ultrasound probe provide a lower image quality compared to the *6VT-D* probe, especially in combination with the 3D volume computation that might lead to further distortions (see Sec. 4.1.1). Hence, a second data set was collected using the *6VT-D* probe and the *GE vivid E95* station. This set contains the aortic root shape as well as the leaflet shapes of 30 hearts. Thus, the roots were extracted and 3D ultrasound volumes were acquired. Then, the leaflets were extracted and images of them were taken. This data set was used to address the research questions *Q2* and *Q3*.

Data set with Leaflet only: D_S^{aux} In the scope of pseudohealthy synthesis, generative models have to be developed that ensure the synthesis of realistic healthy shapes. As this thesis follows a data-driven approach using machine learning, a vast data set of healthy leaflet shapes is desired to cover a maximum of realistic shape variances. Hence, another data set was collected that only contains leaflet shape images of 56 hearts. In this data set, no information on the aortic root shape is present. Thus, the aortic roots were extracted from the hearts and the leaflets were directly cut out, followed by imaging of them. This data set was used in the scope of answering the research questions *Q2* and *Q3*. As the data set can only be used for learning representations of the leaflets, it s also referred to as auxiliary data set.

4.3 Data Acquisition for Morphological Pathology[2]

To examine the case of morphological pathology, the aortic root shape is of interest in two states, namely the healthy and the pathological state. The evaluation of different pseudohealthy synthesis methods requires that both states are known for each heart sample so that the healthy ground truth is available. Hence, a data set was used that contains 24 porcine aortic roots in the healthy as well as an artificially dilated, i.e. a pathological, state.

4.3.1 Acquisition Workflow

First, the aortic root was extracted from the heart and 2D ultrasound images were taken using the *GE vivid7* ultrasound system and the *6T* probe. From these images, a 3D volume was computed as described in Sec. 4.1.1. This image represents the healthy state. Then, the aortic root was artificially enlarged to simulate a pathological dilation due to an aneurysm. Thus, a vertical cut was performed in each sinus, respectively. In each of these cuts, a diamond-shaped patch was sued in. The patches consisted of porcine aortic material to ensure realistic biomechanical properties. The sewing procedure was performed by trained personel from the Department of Cardiac Surgery, University Hospital Schleswig-Holstein, Germany. During the procedure, the root was regularly moisturized using isotonic saline solution ($9\,\mathrm{NaCl\,g/l}$) to prevent the tissue from drying out. Due to this additional tissue, an enlargement of the three sinuses and hence, a dilation of the root was achieved. After this procedure, another ultrasound image was acquired, representing the pathological state. Hence, both states are known for each heart. This workflow was applied to 24 hearts, resulting in 48 ultrasound images that form the data set D_M. Fig. 4.5 exemplarily shows one root in the healthy and the manually dilated, i.e. pathological, state.

Training generative models on 3D data is computationally expensive. Hence, a 2D subset of this data set was created. Thus, from each volume, the 2D slice image showing the commissure plane of the root was extracted manually. The commissure plane refers to the horizontal cut through the aortic root that shows all three commissure points. Previous studies showed the high information content of this image plane on the aortic root shape [56]. Accordingly, the data set $\tilde{D}_M$ contains 2D slice images of all 24 valves in the healthy and the pathological state, respectively, in total 48 2D images.

[2]The data set described in this paragraph was collected by the author in the scope of his master thesis [53].

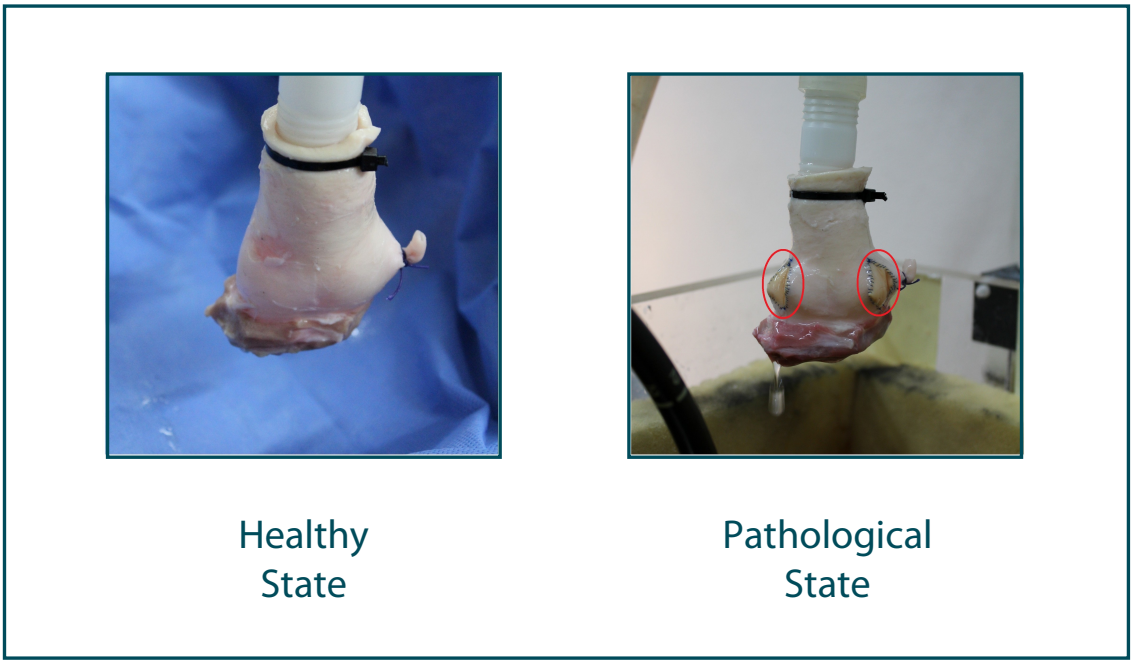

Figure 4.5: Aortic root in the native, healthy state (left) and after manual dilation to simulate a morphological pathology (right). The diamond-shaped patches are encircled in red. Adapted from [62] ©2016 IEEE.

4.4 Discussion

A data set is always just a sample and, hence, can only grasp a part of the variance present in the real world. In this section, limitations of the presented data set's expressiveness will be discussed. First, limitations of the general methodology, including the setups, are discussed, followed by reviewing the ones arising from the acquisition workflows.

In addition to the error-prone 3D reconstruction from 2D images, the ultrasound imaging setup involves several potential sources of errors. By extracting the aortic root from the heart and pressurize it isolatedly, it might take a slightly different shape than in its native surrounding. Additionally, the speed of sound within the water in the bassin differs from the average one in the human body that is used during the ultrasound image computation process, leading to depth-dependent distortions. However, as all images and measured geometric features are compliant with literature values, these effects can be neglected. Another issue arising during ultrasound imaging were air bubbles in the field of view that were challenging for further image analysis. Especially during the assess-

ment of manually dilated roots, many bubbles were present due to insufficiencies. While an insufficiency of the valve validates the success of the manual dilation to simulate a pathology, many air bubbles were induced by the stream through the valve. By applying image filtering methods, e.g. thresholding, the influence of these air bubbles was reduced to a minimum.

During the process of leaflet imaging, cutting the leaflet off the aortic root wall was challenging as the junction between leaflet and wall was not always easy to identify. Additionally, the manual spreading of the leaflets on the illuminated plate might be a source of error regarding the shape assessment. Although special care was taken during cutting and spreading, errors cannot be ruled out. However, it can be assumed that these errors are normal distributed and no bias towards a specific leaflet shape is introduced.

While the process of adding diamond-shapes patches into the aortic root wall is a common method for simulating pathological aortic root dilation [1], the resulting deformation should not be entitled as realistic. Even though the patch material was taken from the same aorta, a natural aneurysm with a year-long progression will result in a different shape.

Even though the data collection procedure is prone to errors at different stages, the collected data sets are unique and present the first or largest data sets of its kind, respectively. Hence, the procedures used within this thesis present the state-of-the-art for ex-vivo scenarios. Additionally, it is important to not that the focus of this work is on methodological development in the field of pseudohealthy synthesis. Therefore, the data sets should be realistic enough to allow for reliable proof-of-concept studies. An assessment of clinically relevant accuracies with a translational value is out of scope for this thesis. Therefore, the data sets do not have to meet the claim of reproducing the real clinical conditions as close as possible. Hence, the collected data sets are sufficient for this thesis.

5 Predicting Individual Prosthesis Features[1]

In this chapter, research question *Q1* is investigated aiming at the estimation of geometric key features of the prosthesis based on the surrogate:

Q1: Is it possible to predict individual prosthesis features?

The goal is to show that the information given by the surrogate is sufficient to predict features of the desired shape. Hence, a positive answer to this question could serve as a proof-of-concept for the general concept of pseudohealthy synthesis for cardiovascular prosthesis shaping. Fig. 5.1 visualizes the approach for investigating the research question *Q1*. The research question is investigated separately for both pathological cases. The investigation of *Q1* for structural pathologies is described in Sec. 5.2, for morphological pathologies it can be found in 5.3. For the latter case, two different approaches are followed to address *Q1* in a paired as well as in an unpaired data scenario. These sections are followed by a conclusion to answer the research question based on the results of the experiments performed for both cases.

[1] Parts of this chapter were published in:

[55] J. Hagenah, T. Evers, M. Scharfschwerdt, A. Schweikard and F. Ernst, A Support Vector Regression-Based Data-Driven Leaflet Modeling Approach for Personalized Aortic Valve Prosthesis Development. 2018 Computing in Cardiology Conference (CinC), 2018:1-4, (2018)

[62] J. Hagenah, E. Werrmann, M. Scharfschwerdt, F. Ernst and C. Metzner. Prediction of individual prosthesis size for valve-sparing aortic root reconstruction based on geometric features. Annu Int Conf IEEE Eng Med Biol Soc. 2016. 2016:3273-3276. (2016) ©2016 IEEE

[61] J. Hagenah, M. Scharfschwerdt, A. Schweikard and C. Metzner. Combining Deformation Modeling and Machine Learning for Personalized Prosthesis Size Prediction in Valve-Sparing Aortic Root Reconstruction. Functional Imaging and Modelling of the Heart. FIMH 2017. Lecture Notes in Computer Science, 10263:461-470. Springer, Cham. (2017)

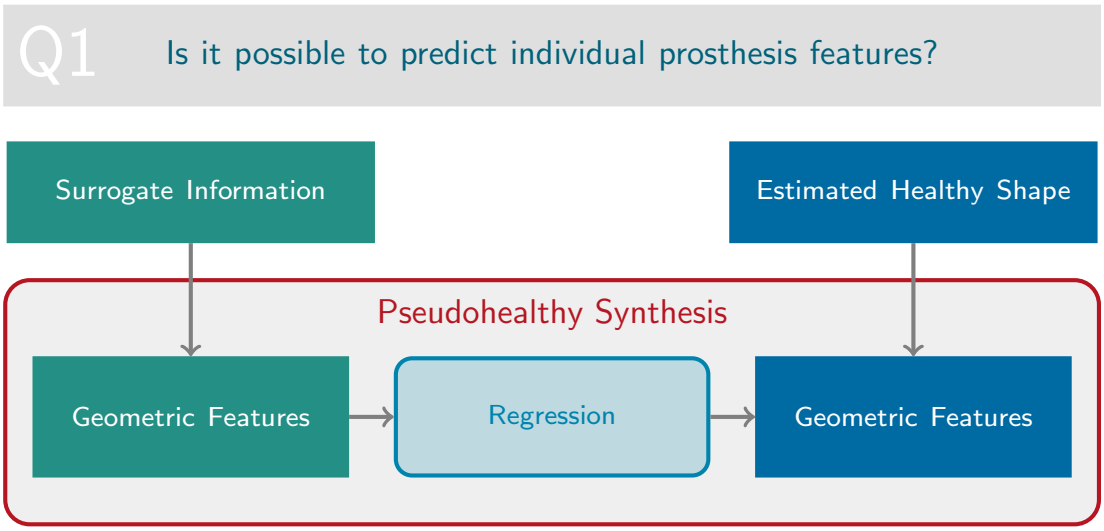

Figure 5.1: General approach investigated to answer research question *Q1*. Geometric features are extracted from both, the surrogate data as well as the desired healthy shape. Then, a regression model is trained to predict geometric key features of the prosthesis.

5.1 Related Work

Most approaches for cardiovascular prosthesis surgery planning are based on biomechanical simulations, typically achieved using the finite elements method (FEM) [152]. The core idea of these approaches is to model the blood flow as well as the mechanical behaviour of the aortic root and valve as realistic as possible, including the fluid-structure-interaction (FSI). Using such a model, it is possible to simulate the behaviour of the aortic valve apparatus for different prosthesis shapes that are integrated into the model. Hence, it allows for pre-operative assessment of the hemodynamics and closing behaviour of the valve regarding different kinds of prostheses. However, modelling the aortic valve apparatus is highly challenging as all components are mutually dependent and influence each other [119]. Additionally, the biomechanical properties are complex to describe. While the aortic tissue model of [67] provides adequate results for the aortic root, the leaflets biomechanics are highly nonlinear and are not fully understood yet. This is mainly due to the strong influence of the prominent collagen fibers at the surface of the leaflets on its biomechanical characteristics. Even though [63] could show the importance of modeling these characteristics, the estimation of the fiber distribution remains very rough and coarse.

For planning valve-sparing aortic root reconstruction, i.e. the morphological pathology case, one example of such a simulation-based approach is the work of [99] that focuses on the behaviour of specific geometric features during pathological dilation of the aortic root to identify an optimal prosthesis shape match. In contrast, [51] aimed at simulating the stress and strain within the leaflets during the cardiac cycle while [85] simulated the deformation of the aortic root.

In the case of structural pathologies, the aim of a biophysical simulation is to assess the characteristics of the system, e.g. the blood flow pattern, for different kinds of valve prostheses or surgical decisions in a virtual intervention. One example is [16] where the optimal orientation of a prosthetic valve was evaluated using simulation results. In recent years, several approaches were proposed to plan transcatheter aortic valve implantation (TAVI) procedures using FEM simulations [52], [18]. As this minimally invasive technique requires a higher degree of navigation, the planning tools mainly focus on the optimal location and orientation of the prosthesis.

The few approaches that utilize machine learning aim at extracting features from the image that are visible. Hence, they can be interpreted as image processing tools to reduce manual annotation steps rather than shape prediction methods. For example, the authors of [5] extracted the diameter of the aortic annulus from CT images using deep learning to identify the individually optimal diameter of a TAVI prostheses.

5.2 Structural Pathology

In the case of a structural pathology, the research question *Q1* aims on the question whether it is possible to predict geometric features of the individual aortic valve leaflet shapes based on information given by ultrasound images of the aortic root. One prominent key feature of the leaflet is its commissure line, i.e. the curved line where it is attached to the aortic root wall. The commissure line already contains much information about the whole leaflet shape. Additionally, in contrast to the coaptation, it is barely visible in the ultrasound images, so a prediction only based on information of the surrounding tissue can be assured. Fig. 5.2 visualizes the investigation of research question *Q1* for the case of a structural pathology. In addition to to providing a proof-of-concept for pseudohealthy synthesis, the results of this section also have a clinical value. As described above, modeling the biomechanical properties of the aortic valve apparatus in a personalized way is very challenging and many key properties, like the stiffness of the leaflets, remains uncertain. Therefore, from a the machine learning approach could present an interesting and model-free alternative to simulation-based approaches.

Contribution

In this section, to the best of the author's knowledge, the first approach for data-driven prediction of personalized prosthesis features in the scope of a structural prosthesis is presented. Together with Sec. 5.3, it also presents the first application of machine learning for aortic valve shape estimation and the first attempt of personalized aortic valve prosthesis shaping in general.

5.2.1 Leaflet Contour Line Prediction

In this study, geometric features are extracted from the ultrasound images and a Support Vector Regression (SVR) model is trained to find a mapping between these features and the commissure line of the corresponding leaflets. As the commissure line is extracted from the planar leaflet images as a part of the leaflet's contour, it is also called contour line in the following. Fig. 5.3 visualizes the basic idea of the contour line prediction approach.

Feature Extraction In this study, the data set $\tilde{D}_S$ is used. It consists of aortic root images of ten hearts as well as their corresponding valve leaflets. As described above, the leaflet shape feature that should be predicted is the commissure line. This leaflet

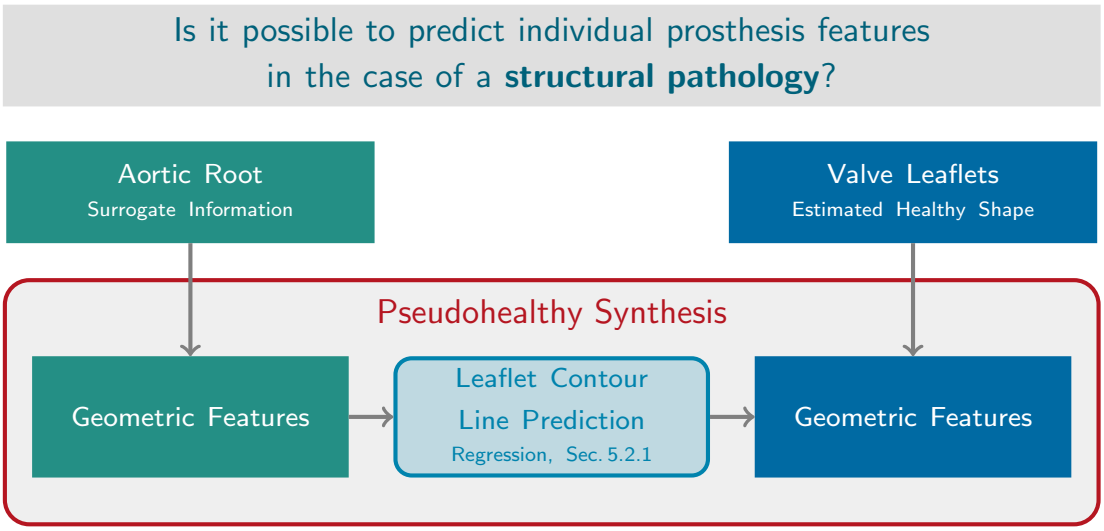

Figure 5.2: Approach to investigate research question *Q1* in the case of a structural pathology.

contour line was extracted using an automatic workflow. First, the leaflet was segmented using thresholding ($t = 0.63$, corresponds to a gray value of 160), followed by a closing operation to improve the segmentation result. The structural element was circular with a radius of 30 pixels. Finally, the leaflet contour was extracted using the Canny algorithm for edge detection [24]. For a more general description of the contour line, it was transformed to polar coordinates with its origin in the nodulus arantius, i.e. the point on the leaflet's free edge where it touches both other leaflets (see Fig. 5.4). Hence, the commissure line can be represented as a set of discrete points in the 2D image.

The data set $\tilde{D}_S$ allows for the derivation of an intelligent system to estimate the mapping of the geometric key features extracted from the ultrasound images to this commissure shape. Thus, the aim is to find a mapping F_i for each leaflet $i \in \{1, 2, 3\}$ with

$$F_i(\phi, x) \rightarrow \rho, \tag{5.1}$$

where ϕ is the angle of the polar contour coordinate system, ρ is the radius in this system and $x \in \{X_1(i), X_2(i), X_3(i), X_4(i)\}$ is a vector containing a set of individual geometric key features extracted from the ultrasound images. Thus, the leaflet contour line can be estimated for a discretized grid of angles depending on the curved valve shape represented by x. The parameter i is the leaflet index and indicates whether the right-coronary ($i = 1$), non-coronary ($i = 2$) or left-coronary ($i = 3$) leaflet shape is predicted. In this study, four different sets of key feature descriptions were examined to analyse the

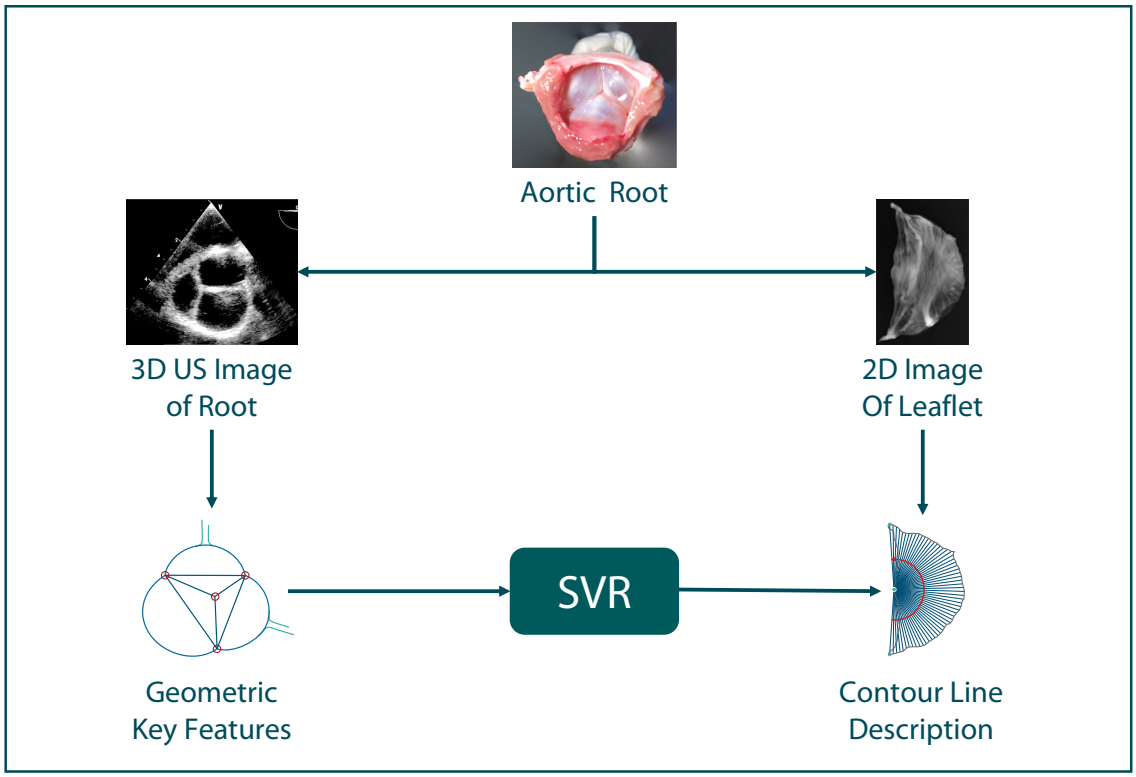

Figure 5.3: Approach for leaflet contour line prediction. The data set contains 3D ultrasound images as well as 2D leaflet images from ten porcine hearts. While the leaflet contour line is described in polar coordinates to allow for easy parameterization, geometric key features are extracted from the aortic root. Then, an SVR model is trained to learn a mapping from these geometric key features to the corresponding contour line. Adapted from [55].

influence of different features on the problem of personalization. The new features S_t and K_t were defined as

$$\begin{aligned} S_t &= S_1 + S_2 + S_3 \\ K_t &= K_1 + K_2 + K_3. \end{aligned} \tag{5.2}$$

S_t is the sum of all free edge lengths while K_t is the sum over the commissure distances. Hence, these new features allow an estimation of the valve's overall size. By division through them, a feature can be scaled to this size approximation. Accordingly, the four examined sets of geometric key features were defined as

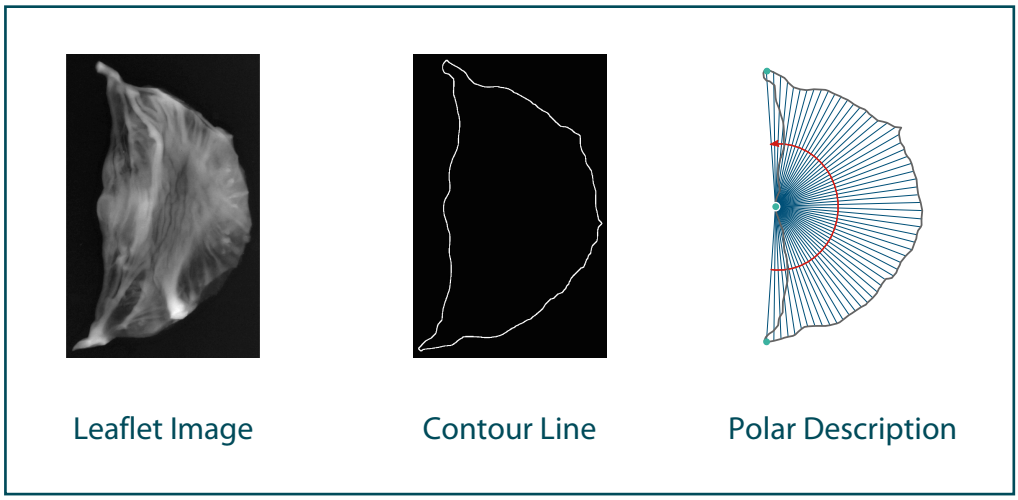

Figure 5.4: Leaflet contour line extraction. The contour is identified using Canny edge detection. Then, the points in this line are transformed to a polar coordinate system, where the origin is the nodulus arantius. Adapted from [55].

$$
\begin{aligned}
X_1(i) &= \left\{ \frac{\theta(i,1)}{S_t} \right\} \\
X_2(i) &= \left\{ \frac{\theta(i,1)}{S_t}, \frac{\theta(i,2)}{S_t} \right\} \\
X_3(i) &= \left\{ \frac{\theta(i,1)}{S_t}, \frac{\theta(i,2)}{S_t}, K_i \right\} \\
X_4(i) &= \left\{ \frac{\theta(i,1)}{S_t}, \frac{\theta(i,2)}{S_t}, \frac{K_i}{K_t} \right\},
\end{aligned}
\tag{5.3}
$$

where i is the leaflet index, K_i is this leaflet's commissure distance and $\theta(i,j)$, $j \in \{1,2\}$, is the j-th free edge length of the leaflet i. Fig. 5.5 visualizes the extracted geometric key features as well as the leaflet notation.

The four key feature descriptions differ in their complexity. Due to the small number of samples, low dimensional descriptions were tested. While $X_1(i)$ only takes one of the leaflet's free edge lengths into account, $X_2(i)$ considers both free edge lengths to ensure a better representation of the leaflets' shearing. By division through S_t, the key features were scaled according to an estimate of the valve's size. In addition, the commissure distance of the leaflet is included in $X_3(i)$ and $X_4(i)$ to measure the general span of the leaflet. In this case, the absolute distance ($X_3(i)$) as well as the distance scaled to the sum of commissure distances ($X_4(i)$) were examined.

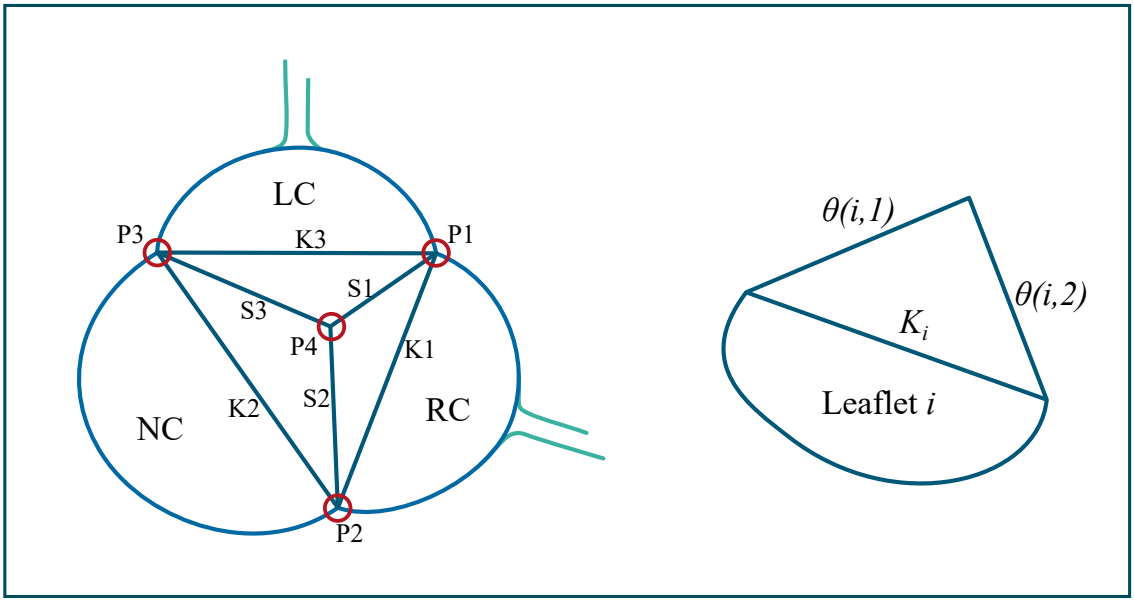

Figure 5.5: Geometric feature extraction. Based on the three commissure points P_1, P_2 and P_3 and the coaptation point P_4, the key features can be calculated for each leaflet (left). The nomenclature of θ follows a general leaflet nomenclature (right). Adapted from [55].

Model Definition and Training The mapping F_i was estimated using Support Vector Regression (SVR) [133] with an RBF kernel. The implementation was done utilizing the *LIBSVM* library [25]. To ensure that asymmetric geometries can be estimated, three different models were trained predicting the shape of one specific leaflet, i.e. the left-coronary leaflet, the right-coronary leaflet and the non-coronary leaflet, respectively. The SVR parameters C and ε as well as the kernel parameter γ have a strong impact on the model's accuracy and its generalization performance. Hence, they were optimized using a grid search method ($C \in [0.01, 1000], \varepsilon \in [0.01, 1], \gamma \in [1, 1100]$, discretization: 100 steps, respectively). 10-fold-crossvalidation was performed, i.e. the model was trained on nine valves and evaluated on the last one by predicting its shape and comparing it to the ground truth. This leave-one-out approach was chosen to address the small size of the data set to maximize the number of training samples.

Comparison to Existing Model For a better interpretation of the model accuracy, the geometric leaflet shape model of Sievers et al. [131] was evaluated to analyse its capability of reproducing personalized leaflet shapes. In this model, the contour line is modelled by adding three circle sections with different radii. The ratio of these radii was derived

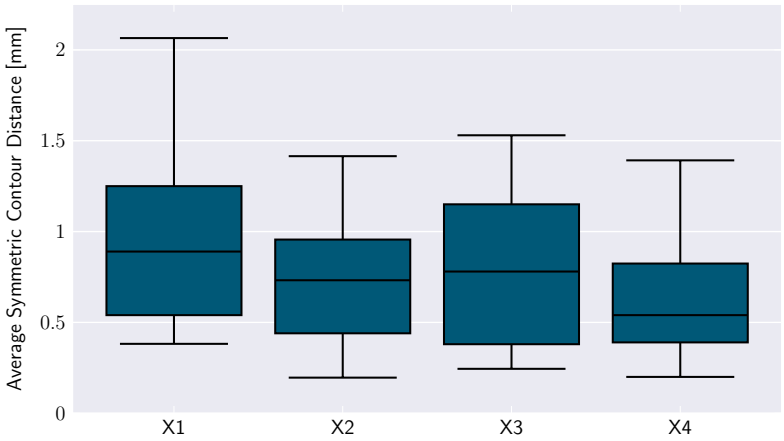

Figure 5.6: Accuracy of predicting the leaflet contour line for all four feature descriptions, given as the average symmetric contour distance (ASCD) between the prediction and the ground truth, averaged over all folds. Adapted from [55].

empirically [2]. The model creates symmetric leaflets, hence, each leaflet is meant to span around the same compartment of the aortic root, which is one third of the root circumference. To reach personalization, the radii were scaled in the scope of this study according to the commissure distance K_i of the specific leaflet $i, i \in 1, 2, 3$. This results in three individually shaped leaflets that span the entire circumference. It is important to note that the geometric model of Sievers et al. does not aim for personalization in its native form and that the individual scaling was introduced in this study. However, due to the lack of personalized prediction models in general, the comparison gives a hint on the accuracy of the proposed SVR-based approach.

Results and Discussion The data-driven model was evaluated for the four different key feature descriptions on all ten heart valves for each leaflet, respectively, resulting in 30 predictions per model. Fig. 5.6 shows the average symmetric contour distance (ASCD) for each examined feature space. The scaling of the features to measures of the valve's size (X_2, X_4) clearly improves the prediction accuracy. A combined geometry description using the leaflets' free edge lengths and the commissure distance in this scaled representation (X_4) provided the smallest errors. Therefore, the following predictions were based on this feature space.

[2] The exact model from [131] is confidential due to commercialization. Hence, no further details can be given in the scope of this thesis. Arising questions should be directly addressed to the authors of [131].

Table 5.1: Prediction accuracy of the two evaluated models for each leaflet type. All values given as the average symmetric contour distance in mm.

	Sievers et al.	Data-driven Model
Left-coronary	2.77 ± 0.84	0.69 ± 0.23
Right-coronary	2.96 ± 0.73	0.64 ± 0.31
Non-coronary	2.65 ± 0.70	0.41 ± 0.48
Mean	2.21 ± 0.56	0.61 ± 0.33

Additionally, the data-driven model with the optimal feature space and the model of Sievers et al. were evaluated on the data set. Table 5.1 shows the prediction accuracy of both models for all three leaflets as well as the mean prediction accuracy. While the error of Sievers et al.'s approach lies above 2 mm, the data-driven model can predict each leaflet's geometry with an error of less than 1 mm. Fig. 5.7 shows sample predictions for both models.

The results show that the geometric model of Sievers et al. is not capable of reproducing patient-specific leaflet shapes. This indicates that current models for heart valve prosthesis are still far from optimal. In contrast, the data-driven model provides good approximations and clearly outperforms the existing model. Hence, it could be shown that estimating the individual leaflet shape based on ultrasound image data is possible. Additionally, a feasible representation of the valve's shape based on four anatomical landmarks was found and it could be shown that personalized estimation of the leaflet shapes can be performed even based on this simple geometric description. In both approaches, the non-coronary leaflet can be modelled with the highest accuracy. This could be due to the fact that this leaflet is more likely to be symmetric, while the right- and left-coronary leaflets are more likely to appear sheared, i.e. the nodulus does not lie in the middle of the leaflet's free edge.

Due to several manual interaction steps, it is possible that the data set is biased (see Sec. 4.2). The first possible reason could be inaccuracies in the manual landmark identification in the ultrasound volume images, though the results indicate that the manual geometry description worked accurately. Automatic identification based on image recognition could eliminate these errors. A second source of error in the data set is due to the manual spreading of the single leaflets on the illuminated plate. Positioning errors will affect the assessed leaflet contour shape. Hence, the sample size should be increased to statistically minimize this error. Additionally, this bigger data set could increase the ac-

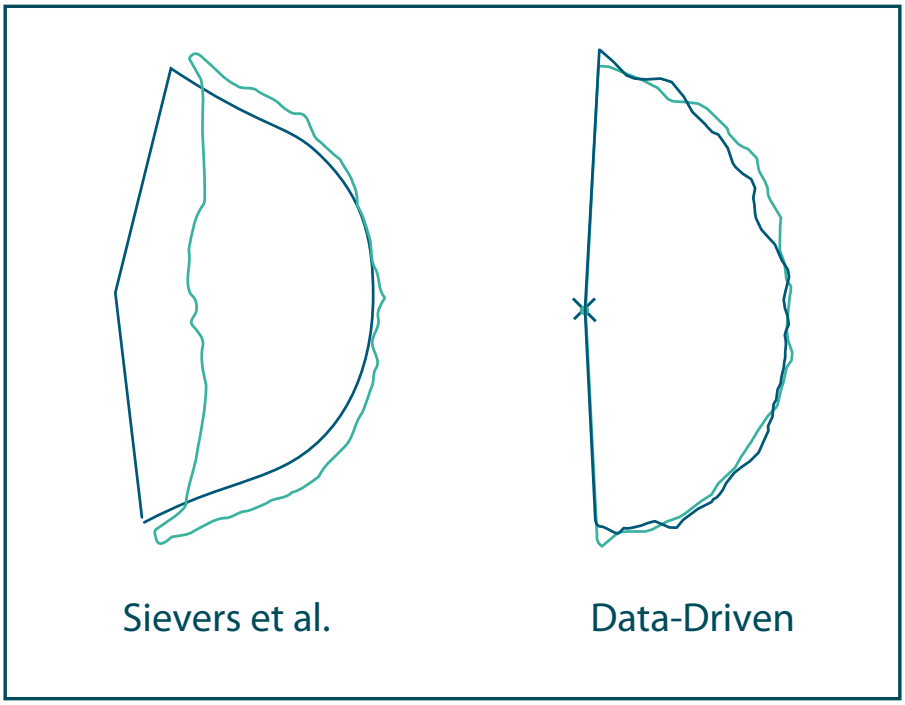

Figure 5.7: Example shape prediction (dark green) of the method from Sievers et al. (left) and the proposed data-driven model (right), together with the ground truth (light green). Adapted from [55].

curacy of the learning algorithm. However, it could be shown that even with such a small data set, sufficiently accurate prediction of the leaflet's contour line are possible. Another focus of future work could be the optimization of the learning method, aiming at an implied smoothing of the contour. Furthermore, a correlation of the estimations of the three different leaflet shapes could increase the accuracy due to the geometric dependency between the leaflets.

This study presents the first method to assess features of the planar shape of the aortic valve leaflets based on ultrasound imaging of the surrounding aortic root. It was shown that the developed model outperformed a modified, geometric model from Sievers et al. regarding the accurate prediction of the leaflets contour line and that acceptable prediction accuracies can be reached. Hence, the results indicate that the aortic root carries a sufficient amount of information on the shape of the leaflets and can therefore serve as surrogate information in a pseudohealthy synthesis algorithm for personalized aortic valve shaping.

5.3 Morphological Pathology

In the case of a morphological pathology, it is necessary to predict geometric key features of the healthy root based on the dilated state of it. For this proof-of-concept study, the aim was to estimate the healthy root diameter. This feature has a huge clinical impact as the most common aortic root prostheses are tube-shaped. Hence, the diameter of the tube is the only parameter that can be estimated patient-individually. If this mapping is possible, it can be assumed that the dilated state still carries enough information to estimate the original, healthy shape.

This section is divided into two approaches, *Direct Feature Estimation* in Sec. 5.3.1 and *Deformation Modeling* in Sec. 5.3.2. The *Direct Feature Estimation* approach uses feature based regression that is derived in a supervised way. However, this requires a paired data scenario, i.e. a set of corresponding pathological and healthy samples for model training. With clinical data, this is hard to achieve and hence, the significance of this study is limited. To overcome this issue and evaluate the capability of predicting healthy features in an unpaired setting, the *Deformation Modeling* approach is presented that does not rely on paired data. Fig. 5.8 visualizes the approaches to investigate research question *Q1* for the case of a morphological pathology.

Contribution

To the best of the author's knowledge, this section presents the first approaches for data-driven prediction of personalized prosthesis features in the scope of a morphological pathology. The *direct feature estimation* approach proposed in Sec. 5.3.1 is the first application of machine learning for aortic root shape prediction as well as for its personalization. Furthermore, Sec. 5.3.2 presents the first approach to model the aortic valve's deformation during aortic root dilation in a parameterized way. Additionally, it presents a novel concept for fusing this hand-crafted modeling with machine learning for cardiovascular surgery planning.

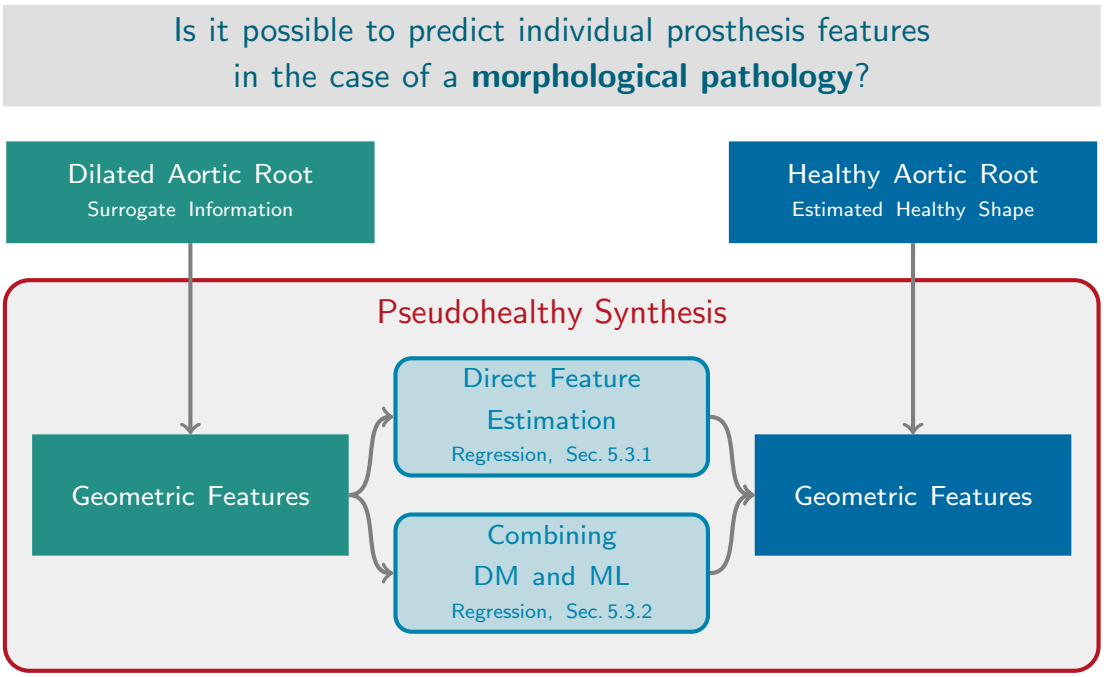

Figure 5.8: Approaches to investigate research question *Q1* for the case of a morphological pathology.

5.3.1 Direct Feature Estimation

The main idea of this approach is to describe the root and valve geometry by a small set of geometrical properties, called features. The pathological dilation of the aortic root is a change in these features, as well as the reconstruction of the original aortic root shape during surgery. Hence, a functional relationship between the features of one valve in its dilated and in its healthy state can be assumed. By estimating this mapping, it is possible to predict the healthy features of a valve based on its dilated features. Thus, the surgery planning can be described as a learning problem, which can be solved using supervised machine learning algorithms.

The full 3D data set D_M was used for assessing the aortic root geometry. In this initial study, the features are the three commissure distances K_1, K_2 and K_3, where K_1 corresponds to the right-coronary, K_2 to the left-coronary and K_3 to the non-coronary leaflet (see Fig. 5.5). The commissure distances are the distances between the commissure points, respectively, while the commissure points are the highest points where the leaflets are attached to the aortic root wall. This feature space definition was chosen for two reasons: First, the commissure points are easy to identify in the ultrasound im-

ages. Secondly, the diameter of the aortic root can be approximated by the diameter of the circumcircle of the triangle formed by the three commissure distances. Hence, given the predicted healthy commissure distances, it is easy to calculate the optimal prosthesis size based on these features.

The commissure points were manually identified in the 3D image of each valve in the healthy and the dilated state, respectively. Additionally, the corresponding aortic root diameter was calculated for the valves in the healthy state. The result is a data base with the commissure distances in healthy and dilated state of 32 valves and the optimal prosthesis diameter of each valve as ground truth. Fig. 5.9 shows the distribution of commissure distances within the data set for healthy and pathologically dilated valves.

Prediction Model As mentioned before, the surgery planning can be described as a learning problem. Its goal is to learn the mapping from dilated to healthy features of one valve based on the data base described above. This learning problem can be solved using *Support Vector Regression* (SVR). A detailed description can be found in 3.3.

To identify the optimal regression model, two different kernels were examined: the *linear* kernel as a robust, linear model and the *radial basis function* (RBF) kernel as a flexible, non-linear one. The kernel functions are given by:

$$\begin{aligned} \text{linear: } & K(x, x') = x * x' \\ \text{RBF: } & K(x, x') = e^{-\gamma ||x - x'||^2}. \end{aligned} \tag{5.4}$$

Additionally, different dimensions of the feature space were tested, i.e. 1D and 3D. In the 1d approach, each healthy commissure distance is estimated based on its corresponding dilated commissure distance. This provides an easy model. On the contrary in the 3D approach, the estimation of each healthy commissure distance is based on all three dilated commissure distances. Hence, dependencies between the different features can be considered in the model. The different models were evaluated using a 4-fold cross validation. The implementation was done using the library *libsvm* [25].

Obviously, there are some parameters of the SVR method, namely ϵ, C and, if the RBF kernel is used, γ. ϵ defines the weighting of the distance between every known point and the fitted curve. The regularization parameter C controls the trade-off between a flat curve and an exact fit. If the RBF kernel is used, γ regulates the flexibility of the function. The parameters are optimized using a grid search approach [13]. For this purpose, the

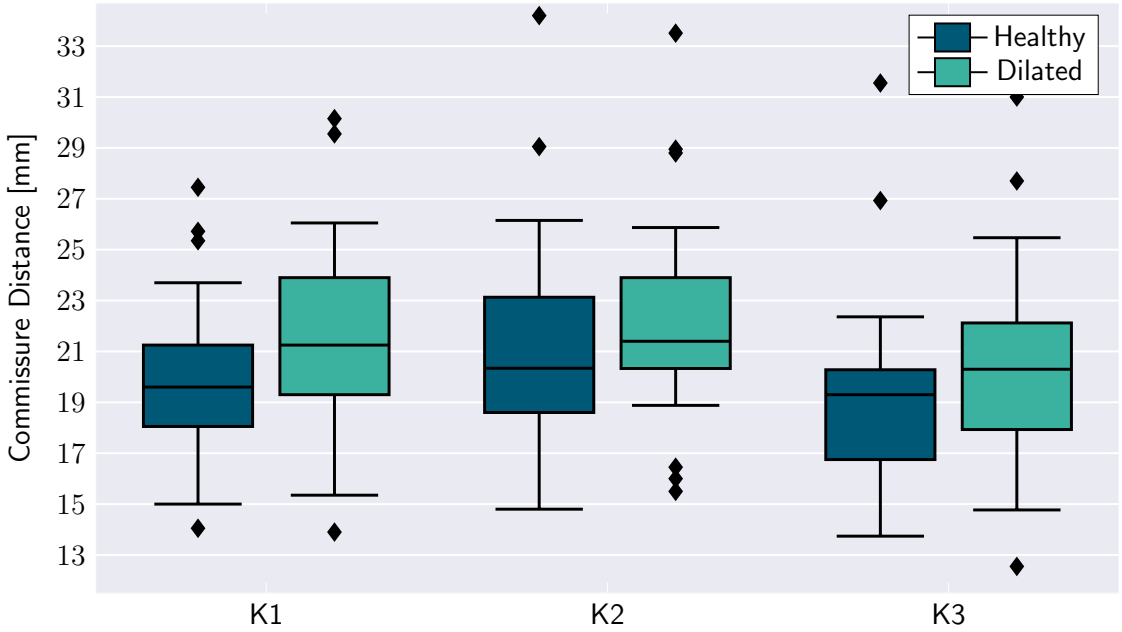

Figure 5.9: Distribution of commissure distances for all three leaflet types, given for the healthy and pathologically dilated valves. Adapted from [62].

SVR is performed on a set of different parameter combinations. The parameter set that provides the minimal prediction error is treated as the optimal parameter set.

In total, four different models are tested and compared, distinguishable by their kernel and their feature space dimension. The parameters are optimized for every model. Two kinds of errors are obtained: The mean error of the predicted prosthesis diameter and the matches. The continuous metric matches lies between 0 and 1 and is a relative value describing the number of valves which prosthesis diameter was predicted correctly divided by the number of all valves in the test set. As the commonly used prostheses are produced with even-numbered diameter values, the diameters were rounded appropriately.

Robustness Test Aortic root prosthesis size prediction suffers from high interpatient variability. Therefore, the optimal model was tested for robustness against high feature variance in the training set and the prediction set, respectively. For this purpose, the data base was divided into three subsets, T, P and P_{out}. P_{out} was filled with outliers, i.e. one quarter of the valves in the data set that show the highest feature difference to the median. The set P contains another quarter, the set T a half of the data base. Both of them were filled randomly from the cleaned data base. On these sets, two experiments

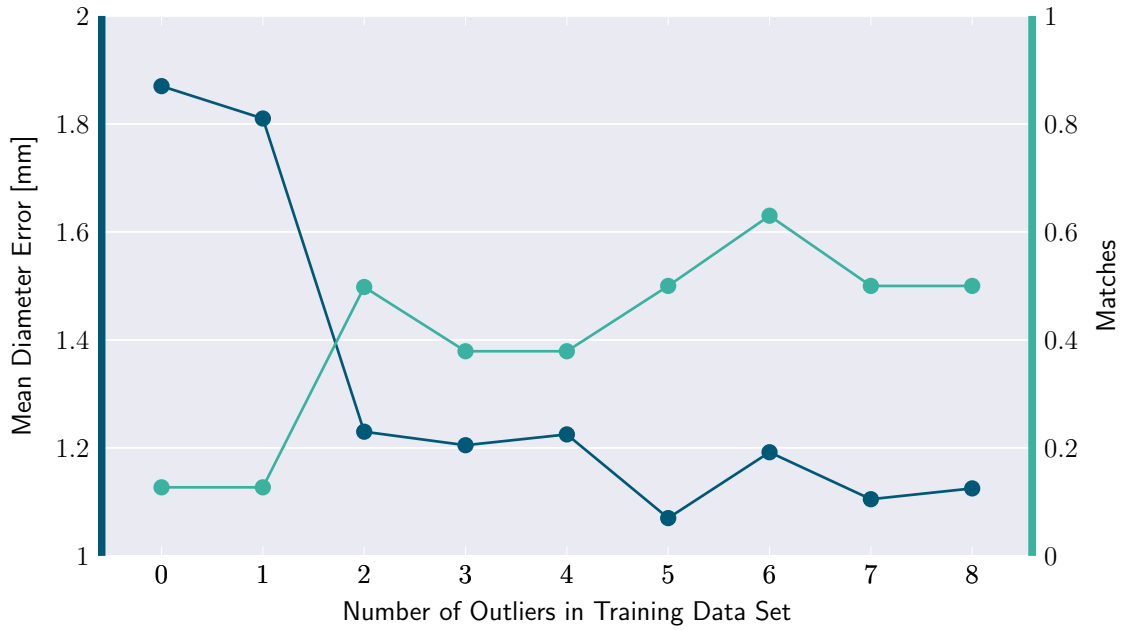

Figure 5.10: Accuracy of predicting the optimal prosthesis diameter regarding the number of outliers in the trainings data set, i.e. depending on the training data set's variance. The mean diameter error is given in mm while matches describes the relative amount of diameters predicted correctly within a $2mm$ error tolerance. This relates to the available prosthesis shapes that are manufactured in diameter steps of $2mm$. Adapted from [62].

were performed. To examine the ability to predict diameters of valves with quite different geometry than the valves in the test set, a model was trained on set T. With this model, the valve diameters of P and the valve diameters of P_{out} were predicted. To investigate the influence of the feature variance in the training set, the model was trained on T and the valve diameters of P were predicted. After that, single samples in T were replaced by samples of P_{out}, inducing a gradually increase of the feature variance in the training set. Each replacement was followed by a prediction of the valve diameters of P.

Results and Discussion Four different models were compared using cross validation. Table 5.2 shows the results of this comparison. The 1D feature space provides better results than the 3D feature space. The RBF kernel with 1D feature space achieves the smallest diameter error and the best matches value. For robustness examination, two experiments were performed. The comparison between the prediction of P and P_{out} based on the training set T showed that the mean diameter error of P is smaller than the

Table 5.2: Comparison of different models. The mean diameter error is given in mm, matches is given as relative value between 0 and 1.

Kernel	Features	Mean Diameter Error	Matches
linear	1d	1.03	0.47
	3d	10.19	0
RBF	1d	**0.95**	**0.63**
	3d	2.76	0.34

error of P_{out} by a factor of 2.4. Fig. 5.10 depicts the increasing prediction accuracy in dependency of the increasing feature variance of the training set.

The constructed data base shows large variances across all three features, which means that individual aortic root geometries drastically differ. This confirms the assumption that a satisfactory estimation of prosthesis sizes has to be based on individual parameters and highlights the value of this thesis. Furthermore, it can be seen that in the dilated state on average commissure distances increase for all three features. This affirms that the approach of artificially dilating the aortic roots as proposed in 4.3 reproduces aortic root dilation as it might be found in human patients.

Comparing results from the four different models shows that calculating a prediction model for each dimension separately, clearly outperforms the three-dimensional approach, regardless of the kernel function (see Table 5.2). Furthermore, it was observed that the RBF kernel produces slightly more accurate predictions in the 1D scenario and drastically better predictions in the 3D scenario. This is probably due to the fact that, the RBF kernel is more flexible than the linear kernel, which is especially important when a 3D prediction model is calculated. As expected, the differences in prediction accuracy are also reflected in the ratio of matches, which is also highest for the 1D model using an RBF kernel. Overall, the majority of aortic root diameters can be predicted correctly using the proposed approach, despite the high variability of the commissure distances within the data set.

As explained before, two tests were performed in order to investigate the robustness of the approach. First, it could be observed that the prediction accuracy dramatically decreased with an increase in feature variance in the test set. This is expected and probably due to the fact that, if the training data set is too homogeneous in comparison to the test set, the model overfits. This might imply that the prediction model needs to be trained on a data base that sufficiently represents the individual variability of aortic root

geometries. Second, it could be observed that the prediction accuracy can be improved by introducing more variability into the training data set. This shows that, indeed, the prediction accuracy of the approach can be improved by using a training data set that is more representative of the true distribution of aortic root geometries. A downside of the proposed approach is the need for paired training data as the SVR is trained in a supervised way.

Conclusion In this section, a novel SVR approach was presented for the prediction of the individual prosthesis size for valve-sparing aortic root reconstruction based on ultrasound images. It could be shown that this approach can accurately predict geometric features and prosthesis diameters from only three geometric features extracted from the pathological ultrasound data. Furthermore, the investigations show that special emphasis should be given to the creation of the training data set, which should reflect the inter-subject variability seen in aortic root geometries. These results highlight the high variability of aortic root shapes and emphasizes the need for personalized prosthetics.

Hence, it could be shown that the pathological aortic root carries a sufficient amount of information to estimate the individual healthy shape of the root. However, the prediction method proposed in this section is limited to a paired data scenario.

5.3.2 Combining Deformation Modeling and Machine Learning

As mentioned above, the *Direct Feature Estimation* approach relies on paired data samples of pathological and healthy shapes. This limits the clinical applicability as paired data is typically harder to collect. Hence, another approach was developed to investigate the sufficiency of information in the pathological image for pseudohealthy synthesis in an unpaired scenario.

The purpose is the introduction of a hybrid approach combining simplified aortic root deformation modeling and machine learning. It is inspired by the surgeon's decision making process, avoids biomechanical uncertainties and minimizes the demand on the underlying data base. Once again, the full 3D data set D_M was used for evaluating this approach.

The proposed method is inspired by the surgeon's decision making process. This process can be separated into four steps:

1) Obtaining the patient's individual dilated aortic root geometry
2) Estimation of the valve geometry if it would be attached to a prosthesis
3) Rating the correspondence between the resulting geometry and a healthy behaviour
4) If the rating was good enough, the optimal prosthesis size is found. Otherwise, repeat with decreasing prosthesis size until geometry is rated as healthy

The proposed approach implements the single steps of this decision making pipeline (cf. Fig. 5.11). In the following, these steps are explained in more detail.

Aortic Valve Deformation Modeling The aim of this step is to model the deformation of the aortic valve geometry in correspondence to a change in the root diameter. The proposed method utilizes strong simplifications of the aortic valve geometry and behaviour. Thus, it is assumed that the valve geometry can be described meaningful enough based on four landmarks: the three commissure points P_1, P_2, P_3 and the coaptation point P_{coap}. The commissure points are the points where the leaflets are attached to the root wall while the coaptation point is defined as the point were all three leaflets are touching each other (cf. Fig 5.12). Accordingly, the individual valve geometry is given by three vectors $L_1, L_2, L_3 \in \mathbb{R}^3$ pointing from P_{coap} to one of the coaptation points, respectively. After transformation to spherical coordinates with P_{coap} as the origin, the valve geometry $\mathbf{V}$ is given by

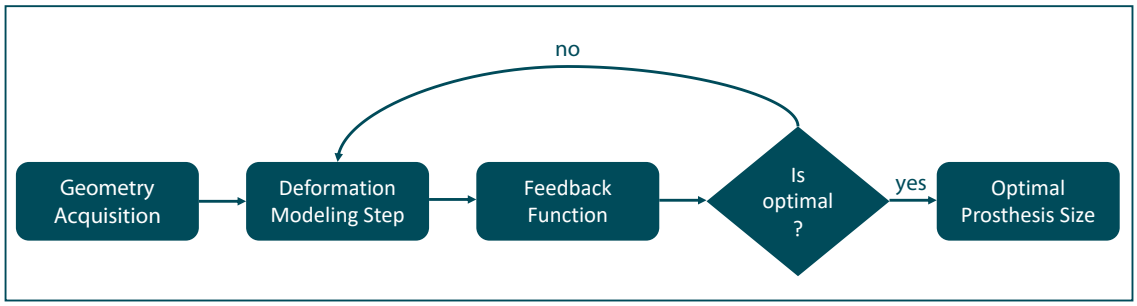

Figure 5.11: Workflow of the proposed method. The acquired morphological pathology is deformed incrementally according to a deformation model. The feedback function rates the resulting geometry regarding its likelihood of corresponding to a healthy valve. While the feedback function rates the current geometry as pathological, further deformation steps are performed until the geometry is rated as healthy. Based on this resulting geometry, the optimal prosthesis diameter can be computed. Adapted from [61].

$$\mathbf{V} = \begin{pmatrix} L_{1,r} & L_{2,r} & L_{3,r} \\ L_{1,\theta} & L_{2,\theta} & L_{3,\theta} \\ L_{1,\varphi} & L_{2,\varphi} & L_{3,\varphi} \end{pmatrix} , \tag{5.5}$$

where r, θ and φ are the spherical coordinates. r describes the distance to the origin, θ is the zenith angle and φ is the azimuthal angle.

While the aortic root wall dilates, the leaflet size is not affected by the deformation [126]. Hence, it can be assumed that the length of the vectors L_1, L_2 and L_3 stays constant over all relevant deformation steps. Following this assumption, the influence of the dilation on the valve geometry can be described as a change of the orientation of L_1, L_2 and L_3. This can be reached by varying two parameters of the spherical geometry description of every vector: θ, i.e. the angle to the vertical axis, and φ, i.e. the angle of the rotation around the vertical axis (cf. Fig 5.13). Accordingly, the deformation of the valve geometry $\mathbf{V}$ can be modeled as a transformation T given as

$$\begin{aligned} \mathbf{V}' =& T(\mathbf{V}, \Delta\theta_1, \Delta\theta_2, \Delta\theta_3, \Delta\varphi_1, \Delta\varphi_2, \Delta\varphi_3) \\ =& \begin{pmatrix} L_{1,r} & L_{2,r} & L_{3,r} \\ L_{1,\theta} - \Delta\theta_1 & L_{2,\theta} - \Delta\theta_2 & L_{3,\theta} - \Delta\theta_3 \\ L_{1,\varphi} - \Delta\varphi_1 & L_{2,\varphi} - \Delta\varphi_2 & L_{3,\varphi} - \Delta\varphi_3 \end{pmatrix} . \end{aligned} \tag{5.6}$$

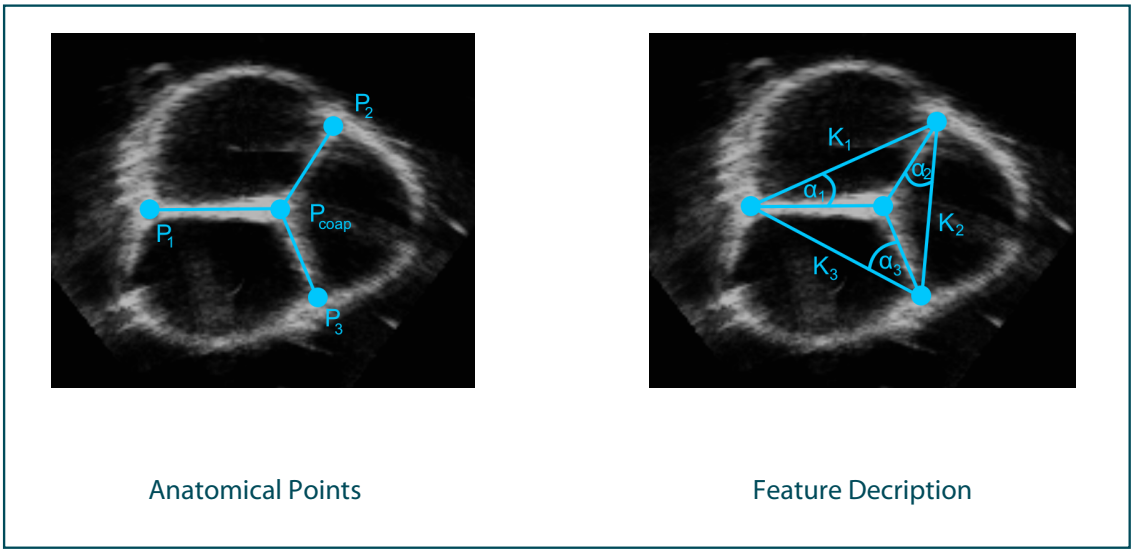

Figure 5.12: Anatomical points (left) and geometric features (right) extracted from ultrasound images. Note that the figure shows a two-dimensional projection of the three-dimensional geometry. Adapted from [61].

The definition of $\Delta\theta_i$ and $\Delta\varphi_i, i = 1, 2, 3$ induces different deformation models. In the following, three models of different complexity are presented.

Homogeneous Dilation (HD): In this model it is assumed that the orientation of the plane formed by the commissure points P_1, P_2, P_3 stays constant. The change of these points cartesian z coordinate is the same for all three points without allowing any rotation around the vertical axis. This corresponds to a homogenous dilatation of all three sinuses, i.e. the areas of the root wall enclosing one leaflet, and can be derived by setting

$$\Delta\theta_i = \arccos\left(\frac{L_{1,r}}{L_{i,r}}(\cos(L_{1,\theta} + \Delta\theta_1) - \cos L_{1,\theta} + \cos L_{i,\theta})\right) - L_{i,\theta},\ i = 2, 3. \tag{5.7}$$

The two values $\Delta\theta_2$ and $\Delta\theta_3$ are dependant of $\Delta\theta_1$, which is the only model parameter. $\Delta\varphi_j, j = 1, 2, 3$ are set to 0. Hence, the HD model has only one degree of freedom (DOF).

Angle-compensated Homogenous Dilation (AHD): As mentioned above, the HD model does not take different dilations of the sinuses into account. This can be compensated by allowing small rotations around the vertical axis. This results in three independent parameters $\Delta\varphi_i, i = 1, 2, 3$ within a range of $\pm 10°$. The parameters $\Delta\theta_i, i = 1, 2, 3$ are set as in the HD model. This yields a model with four DOF.

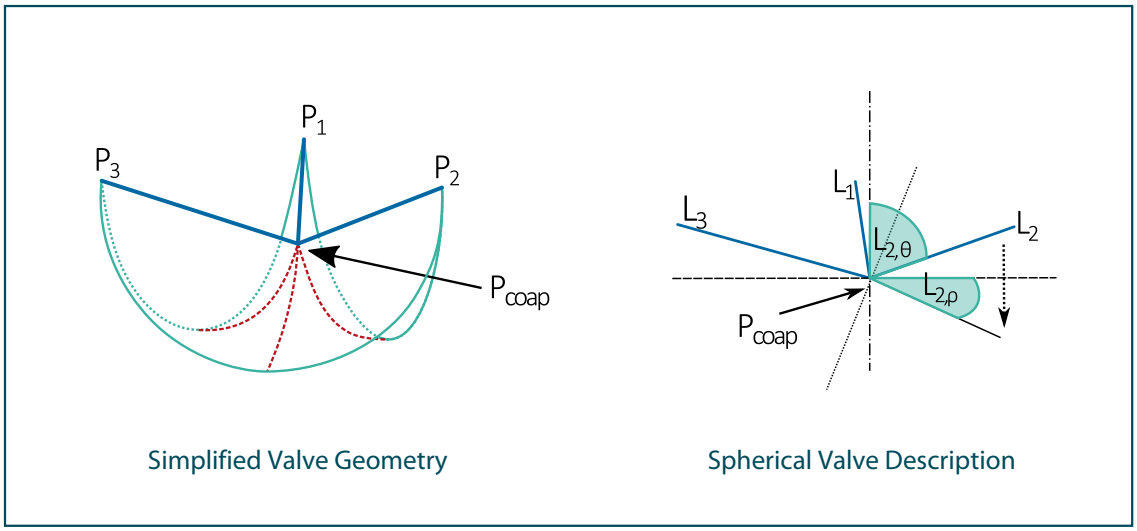

Figure 5.13: Left: Simplified valve geometry, described by the three commissure points P_1, P_2 and P_3 as well as the coaptation point P_{coap}. Right: Spherical coordinate description of the valve's geometry. Manipulating the polar angles, different kinds of deformation can be modeled. Adapted from [61].

Free Deformation (FD): In this model, the parameters $\Delta\theta_i, i = 1, 2, 3$ are decoupled. Hence, all orientation parameters $\Delta\theta_i$ and $\Delta\varphi_i, i = 1, 2, 3$ can change independently within given ranges. With six DOF, the FD model offers the highest complexity of the presented deformation models.

With this simplified valve geometry description and the defined transformation models, parameterized, realistic deformations of arbitrary aortic valve geometries can be calculated efficiently. Biomechanical uncertainties are avoided by making simplifying assumptions. To evaluate the models, the dilated geometries from the data base were deformed to match the corresponding healthy geometries as close as possible, followed by a comparison of the resulting root diameter with the healthy reference.

Feedback Function While the developed deformation models allow the prediction of the valve's geometry after a prosthesis induced deformation, there is a need for a feedback function that rates the healthiness of an arbitrary aortic valve geometry. Since no reliable model for such a rating function exists due to the high variability of the valve's shape, the use of machine learning is proposed to achieve a data-driven estimation of this complex relationship.

For that purpose, a suitable feature space has to be defined. Due to the description of the valve geometry based on four anatomical landmarks, all features should be extractable from these points. Instead of using the raw point coordinates, meta-descriptors were defined and used as features. This allows for a feature space dimensionality reduction, position- and orientation-invariant features and the integration of prior knowledge, i.e. geometrical descriptors known to have a strong relationship to the valve's functionality.

In this study, two features were defined: $\frac{K}{L_\mathrm{f}}$ and $\frac{\alpha}{L_\mathrm{f}}$. K is the mean of the three commissure point distances, while α is the mean over the three angles $\alpha_i, i = 1, 2, 3$ that describe the angle between the vector pointing from the i-th commissure point to its clockwise neighbor and the vector from the i-th commissure point to the coaptation point (cf. Fig. 5.12). L_f is the leaflets free edge length, calculated as the sum over the lengths of the three vectors L_1, L_2 and L_3. Due to the assumption that L_f stays constant, the defined features measure K and α relative to the valves specific leaflets free edge length.

Two different feedback function concepts were investigated: a binary feedback and a continuous feedback. While the binary feedback function gives back a distinct binary flag whether the geometry is healthy or not (classification problem), the continuous feedback functions output is a continuous score to rate the healthiness of the valve (regression problem).

The binary feedback function was derived from the data using a Support Vector Machine (SVM) utilizing the open-source library *LIBSVM* [25]. Non-linear relationships in the data were addressed by choosing a radial basis function (RBF) kernel. The SVM-parameter C and the kernel-parameter γ were optimized utilizing a leave-one-out-method on the training data set with an iterative gridsearch on multiple grids, ranging steadily from coarse to fine (initial grid: 5 values within [1,1000] and [0.001,1] for C and γ, respectively, finer grids centered around optimal parameters of coarser grid from last iteration). This SVM model learns a separation of valve geometries in healthy and dilated ones. To provide reasonable classification results for this purpose, the training set needs to include healthy and dilated valve geometry samples.

Since a binary separation of valve geometries is not very realistic, a continuous feedback function was used additionally. This leads to a continuous healthiness score. To this end, the feedback function was learned using a Gaussian Mixture Model (GMM) [98]:

1) For every training sample point $j \in [1, n]$, define a Gaussian distribution $\mathcal{N}(\mu_\mathbf{j}, \mathbf{\Sigma_j})$
2) Set μ_j to the samples feature space coordinates and $\mathbf{\Sigma}_j$ to a diagonal matrix with the estimated measurement noise of each feature at the corresponding diagonal entry (10% of the feature's mean value)
3) Derive feedback function F by superimposing these distributions with equal weighting:

$$F(\mathbf{x}) = \frac{1}{n} \sum_{j=1}^{n} \mathcal{N}(\mu_j, \mathbf{\Sigma}_j) \; . \tag{5.8}$$

This completely data-driven feedback function combines low computation times with a simpler and much more practical data base. The score predicted from the GMM model corresponds to the probability of an arbitrary valve geometry to be healthy. The probability to be dilated is included implicitly in the low score values. Hence, a training data set only consisting of healthy geometry samples is sufficient and no dilated samples are required.

Surgery Planning The surgery planning can be interpreted as an optimization problem, where the parameters of the deformation models should be optimized to provide the most healthy valve geometry, i.e. the highest value of the feedback function. Hence, the objective function is given by the feedback function and the parameters of the deformation model are the variables. The solution space is constrained by the geometries that can be reached by deforming the individual dilated valve geometry according to a deformation model. The deformation model defines a trajectory in the feature space with its origin in the observed dilated valve geometry. The optimal prosthesis size corresponds to the point on this trajectory with maximal feedback output (cf. Fig. 5.14). The relationship between a valve geometry and the corresponding prosthesis size is approximated by the diameter of the circumcircle of the triangle formed by the three commissure points. In the case of the binary feedback function, the optimal solution is defined as the valve geometry classified as healthy with the smallest deformation, i.e. the intersection between the trajectory and the class border.

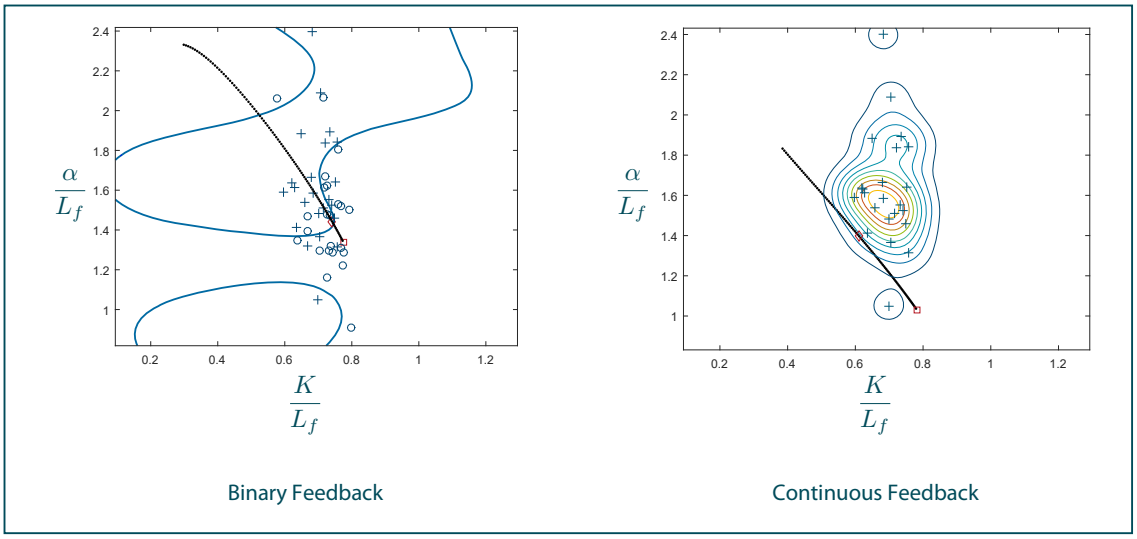

Figure 5.14: Visualization of surgery planning in the feature space using a binary feedback (left) and a continuous feedback (right). In the binary case, the feedback function is visualized as a decision boundary, while a contour plot of the feedback landscape is given for the continuous case. The trajectory defined by the incremental deformation steps is shown as in black. Adapted from [61].

In this initial study, the optimization problem was solved using a gridsearch method (discretization: $0.006°$ for $\Delta\varphi_i$ and $\frac{\mathbf{L}_{i,\theta}}{100}{}^{\circ}$ for $\Delta\theta_i, i = 1, 2, 3$). This corresponds to a valve diameter change (i.e. change of prosthesis size) of about $0.073\ mm$ for each step of θ and about $5.2 \cdot 10^{-5}\ mm$ for each step of φ. This fine discretization provides an acceptable avoidance of local optima.

The evaluation of the surgery planning method was performed on all 24 aortic roots using a 10-fold cross validation. During this method, optimal prosthesis diameters were predicted for the dilated valve geometries and compared to the natural root diameter received from the corresponding healthy geometry, serving as a reference.

Results and Discussion A fundamental assumption of the method is that the length of the vectors $L_i, i = 1, 2, 3$ stays constant during the dilation. To confirm this statement, a two-sample t-test (95% level of significance) on the L_i derived from the healthy and the dilated data was performed. We found no significant differences ($L_1 : \ t = -0.4978, \ p > 0.05$; $L_2 : \ t = -1.1455, \ p > 0.05$; $L_3 : \ t = -0.0405, \ p > 0.05$). To evaluate the consistency of the presented deformation models, they were applied to reproduce the

Table 5.3: Evaluation of the *deformation modeling* approach for all combinations of the two feedback functions (binary and continous) and the three deformation models (HD, AHD and FD). For each combination of feedback function and deformation model the mean prosthesis size prediction error in mm and the number of matches in %.

	HD	**AHD**	**FD**
Binary	21 %	25 %	21 %
	3.36 mm	2.81 mm	4.38 mm
Continous	63 %	50 %	13 %
	1.64 mm	1.64 mm	7.60 mm

deformation observed in the experiments. The mean errors of the predicted diameters for all valves were $0.37\ mm$ for HD, $0.09\ mm$ for AHD and $0.25\ mm$ for the FD model.

The presented *deformation modeling* method was evaluated on the data set aiming on the prediction accuracy of the optimal root diameter for each dilated valve geometry with the corresponding healthy geometry as a reference. Table 5.3 shows the mean prediction error and the number of matches, i.e. the relative number of predictions with an error of less than 1 mm, for all combinations of deformation models and feedback functions. The continuous feedback combined with the HD model reaches the highest prediction accuracy.

A comparison of the previously proposed *direct feature estimation* method and the *deformation modeling* method with the binary as well as the continuous feedback function is shown in Table 5.4. Using the continuous feedback function, the *deformation modeling* approach provides the same prediction accuracy while only healthy samples are requested in the training phase.

The evaluation of the three deformation models shows that all models reach a mean diameter prediction error of less than 0.5 mm. This indicates that all three models are capable of mimicking realistic aortic root deformations. Table 5.3 shows that for the HD and the AHD deformation model, the continuous feedback function clearly outperforms the binary feedback. This indicates that the optimal solution does not have to be close to the binary class border. Combining the HD deformation model with a continuous feedback function provides the best prediction results. This might be related to the fact that the information about the individual dilated valve geometry disappears with growing size of

Table 5.4: Comparison of the *direct feature estimation* (DE) method as described in 5.3.1, the *deformation modeling* (DM) with binary feedback function (AHD model) and the *deformation modeling* with continuous feedback function (HD model) subject to the prediction accuracy and the needed training data. The deformation models were chosen to provide optimal results.

Method	Matches	Training Data Set
DE	63 %	Dilated and healthy valve geometry of each subject in training set
DM, binary	25 %	Dilated and healthy valve geometries, not necessarily from the same subject
DM, continuous	63 %	Healthy valve geometries

the search space, which is the case with increasing DOF of the model and once more emphasizes the advantage of simplification.

The comparison with the *direct feature estimation* approach shows that using a continuous feedback function, the *deformation modeling* approach reaches comparable results based on training data set that is simpler to collect. The collection of clinical paired data for the *direct feature estimation* approach is time consuming and challenging. In contrast, the *deformation modeling* method utilizing a continuous feedback function that only relies on healthy valve geometries, which can be easily extracted from clinical TEE examinations. Hence, the application of simplified *deformation modeling* and GMMs could push the pre-operative planning of valve-sparing aortic root reconstruction towards clinical application.

However, the prediction accuracy still has to be improved. One reason for the relatively low maximal accuracy could be a biased training data set. Due to the manual identification of the landmarks in the ultrasound images, significant observer-specific errors are possible. Hence, an automatic valve geometry extraction could improve the method. A more detailed study of the influence of observer errors on the predictor performance would be interesting and relevant for future work. Additionally, the approach could be enhanced by a larger training data set. This would provide a more realistic representation of the underlying distribution of valve geometries and a higher dimensional feature space could be explored.

Conclusion In this work, a new approach for pre-operative planning of valve-sparing aortic root reconstruction was presented. The approach combines simplified *deformation modeling* with machine learning to mimic the surgeon's decision making process and represents a first alternative to the *direct feature estimation* approach. Three deformation models and two feedback functions were defined, evaluated and compared to the previously described *direct feature estimation* approach. The results show that the presented method provides comparable results while the demand on the training data is dramatically decreased, taking a step towards clinical application for the prediction of personalized aortic root prosthesis sizes. Hence, it could be shown that predicting features of the healthy shape only based on pathological shape information as a surrogate is not only possible with paired data but also in scenarios where only unpaired data is available.

5.4 Conclusion

The aim of this chapter was to answer the research question

Q1: Is it possible to predict individual prosthesis features?

Therefore, feature-based regression approaches were developed to predict individual features of the desired healthy shape only based on the available surrogate information. The prediction models were trained data-drivenly by implementing SVR models with various hyperparameters. This concept was adapted for both cases, i.e. structural pathology and morphological pathology. In the latter case, it was proposed to interpret the prediction as a deformation of the pathological shape until a model rates the resulting shape as sufficiently healthy. This approach doe not only open up the capability of dealing with unpaired training data but it also lays the basis for the concept of pseudohealthy synthesis via manipulation in latent space which will be presented in Sec. 6.3.1.

In both pathological cases, it was possible to predict meaningful features of the desired healthy shape only based on geometric features extracted from the surrogate information. The accuracy of the prediction was acceptable in both cases and, hence, can be seen as a proof-of-concept. Therefore, the research question can be answered with yes. In both cases, the structural pathology and the morphological pathology, the respective surrogate data, i.e. the surrounding tissue and the pathologically dilated state, carry enough information to infer on the individual. This proves the basic assumptions of this thesis and lays the foundation for the development of pseudohealthy synthesis methods.

Regarding the translational value of this chapter, both approaches, i.e. for personalized valve shaping and for personalized root shaping, are limited by only predicting features of the desired shape. In the structural pathology case, the commissure line of the three leaflets will barely be enough to design personalized valve prostheses and hence, the study should be seen as a proof-of-concept that the surrogate information is sufficient. In contrast, in the case of a morphological pathology, predicting the optimal diameter of a tube-shaped aortic root prosthesis could be of high interest as a surgery planning tool since estimating the individually optimal diameter remains a challenging task in clinical practice. Therefore, a further study on a human data set should be part of future research, where the deformation modelling approach should be followed since it is capable of dealing with unpaired data.

6 Synthesizing Personalized Healthy Shapes[1]

In this chapter, research question *Q2* will be addressed:

Q2: Is it possible to synthesize personalized healthy shapes?

In the previous chapter, it could be shown that predicting geometric key features of the desired healthy state is possible based on the surrogate information. Hence, the assumption that the surrogate information carries a sufficient amount of information about the individual healthy shape can be considered as true. Research question *Q2* addresses the question how this information can be used to not only predict geometric features but to synthesize complete healthy shapes. In general, this poses two core challenges: Identifying adequate shape representations and estimating the right shape within this representation based on the surrogate information. The approaches to solve both challenges differ for the two types of pathology considered in this work.

In case of a structural pathology, the pseudohealthy synthesis can be interpreted as bridging a domain gap as the surrogate images as well as the target shapes are just different kinds of information about the same anatomical structure. However, this domain gap is substantial compared to classical domain adaptation methods as not only a different part of the anatomical structure is shown, e.g. the aortic root instead of the valve's leaflets, but the surrogate information might also differ in its appearance as a different imaging modality is utilized. In the case of a morphological pathology, the target images as well as the surrogate information are in the same domain, but the distribution spans

[1] Parts of this section are published in:

[58] J. Hagenah, M. Mehdi and F. Ernst. Generating Healthy Aortic Root Geometries From Ultrasound Images of the Individual Pathological Morphology Using Deep Convolutional Autoencoders. 2019 Computing in Cardiology (CinC), 2019:1-4, (2019)

[59] J. Hagenah, M. Mehdi and F. Ernst. Fully Data-Driven Pseudohealthy Synthesis for Planning Valve-Sparing Aortic Root Reconstruction using Conditional Variational Autoencoders. Current Directions in Biomedical Engineering 6(3):284-287. (2020)

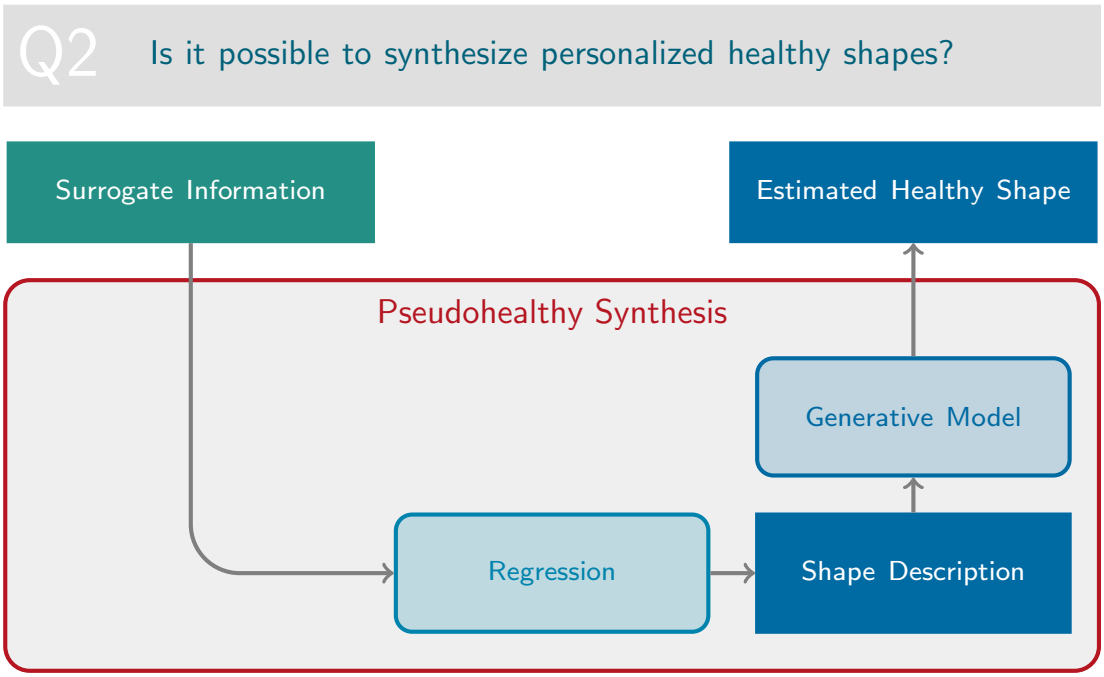

Figure 6.1: Approach to investigate research question *Q2*.

across healthy as well as pathological shapes. Fig. 6.1 visualizes the general approach of investigating research question *Q2*.

To address these challenges, new concepts for pseudohealthy synthesis are presented for both pathological cases, i.e. structural pathologies in Sec. 6.2 and morphological pathologies in Sec. 6.3. After describing the investigation of the research question in these section, the results are discussed and the question is answered in a conclusion.

6.1 Related Work

In the following, the related work will be described separately for both pathological cases. First, studies related to structural pathologies are presented, followed by a paragraph on relevant literature on morphological pathologies.

Structural Pathology In general, dealing with data from different domains is referred to as domain adaptation [82]. Numerous studies aimed at generalizing across different domains [45], [72], [78]. The typical focus of domain adaptation is dealing with input data from different domains, not to transfer an image to another domain [137].

The problem of estimating an image based on another image is called image-to-image translation and was introduced in [73]. Typically, generative adversarial networks (GAN) [49] are utilized to learn a distribution over the target images, conditioned by the input image [73]. Such approaches have been applied to a wide range of topics, for example semantic image synthesis [108], image segmentation [92], style transfer [151] or image inpainting [110]. However, most of these methods are used in settings with a moderate domain gap while the general objects in the input image remain unchanged in their shape.

Several studies focused on image-to-image-translation for shape deformation, aiming at bridging a wider domain gap [107]. However, all these approaches were evaluated on large data sets that are typically not available in clinical applications. Even though [94] proposed a few-shot approach for image-to-image translation, the model has to be pretrained on a large data set that is relatively similar to the small one. In [93],a framework based on a cascaded GAN structure is presented to learn image-to-image translation from one image pair only. However, such a model is not capable of generating the full variance of anatomies present across patients as it is trained to fit one anatomy as close as possible. In contrast, the framework presented in 6.2 directly models this variance by encoding all observed anatomies into latent space descriptions.

Morphological Pathology In the field of pseudohealthy synthesis, the usage of autoencoders is a common approach to learn a latent description [127], [9], [141], [28]. The idea is to train the model only with healthy data so that the decoder will ignore the pathological part of the information given in a pathological input image and synthesizes a pseudohealthy state. This preserves the individuality of the healthy shape, also referred to as subject identity [146]. However, these approaches were criticized recently as pathological out-of-bag samples with a very different shape compared to the healthy ones may lead to unrealistic predictions [105]. In contrast, the pathological samples are explicitly included into the representation learning method presented in 6.3 to disentangle the pathological variance and the subject identity.

In recent years, GANs were applied to pseudohealthy synthesis [2], [143], [136]. Even though very realistic images could be synthesized, these approaches suffered from a poor preservation of the subject identity as the GAN was trained to synthesize any image inside the observed healthy distribution [146]. Even though different methods were pro-

posed to overcome this problem [8], [73], preserving the subject identity in GAN-based approaches remains a challenging task. In [145], a new GAN-based architecture for pseudohealthy synthesis in a brain lesion setting was proposed. The authors used a segmentation of the pathological area as additional information to disentangle the healthy and pathological latent representation and thus preserve subject identity. However, this approach is not applicable to a pathological dilation process, e.g. in the aortic root, where the entire structure is affected and not just a specific region.

The concept of shape primitives presented in Sec. 6.3.1 is related to movement primitives which are used in robotics to describe complex trajectories [125]. Additionally, the mapping process to the mixture of shape primitives is reminiscent of coherent point drift for point cloud matching, but in the method proposed in this section, the deformation is not constrained to be affine [104].

6.2 Structural Pathology

Despite ongoing advancements of medical imaging techniques, there are structures in the human body that are difficult to visualize using typical medical imaging modalities like Computed Tomography (CT), Magnetic Resonance Imaging (MRI) or Ultrasound Imaging (US). However, knowledge about these anatomical shapes is highly relevant for clinical decision making and intervention, e.g. for biomechanical simulations or for personalized prosthesis shaping. One example of such a structure is the aortic valve. This valve consists of three thin leaflets embedded in the aortic root that are pressed together during diastole to prevent the blood in the aorta from flowing back into the left ventricle [39]. The geometry of the aortic valve, i.e. the shape of its three leaflets, shows a substantial inter-patient variability [85]. As the aortic valve and root form a complex biomechanical system with close interaction between different anatomical structures [148], this individual geometry is crucial for the correct functionality of the valve. The valve's leaflets are extremely thin and flutter in the blood stream, hence, imaging using typical medical imaging modalities is challenging. However, knowledge about the individual geometry of the aortic valve is necessary for many applications, ranging from heart modeling and simulation to the development of personalized prostheses. For both applications, the leaflet shape should be assessed in an unpressurized state to avoid stress-related deformations. Due to the high impact of the individual valve geometry on a procedure's outcome, a personalization of this prosthesis would be beneficial but is currently out of scope as the desired prosthesis shape is not assessible using typical in-vivo imaging modalities. Hence, the synthesis of personalized leaflet shapes presents a promising approach for solving this problem.

In an ex-vivo setting, it is possible to extract the aortic valve leaflets and collect high-resolution images in a planar state [118]. Such a data set could serve as a training set for a generative model aiming at leaflet shape synthesis. An additional advantage of this approach is that the synthesized leaflet shapes are already in a planar shape, allowing for direct manufacturing of a prosthesis out of typical material, e.g. pericardium. However, the shape synthesis should be tailored to the individual patient. Thus, the generative model should receive a prior consisting of information about relevant surrounding tissue to estimate the patient's individual leaflet shapes. In this case, the three-dimensional (3D) shape of the aortic root might be a sufficient prior as its geometry should be closely related to the shape of the individual leaflet. Additionally, the aortic root is clearly visible in ultrasound images acquired using transesophageal echocardiography (TEE).

However, the domain gap between a 3D ultrasound image of the aortic root and an RGB image of the leaflet shape in its planar state is substantial. Typical methods for

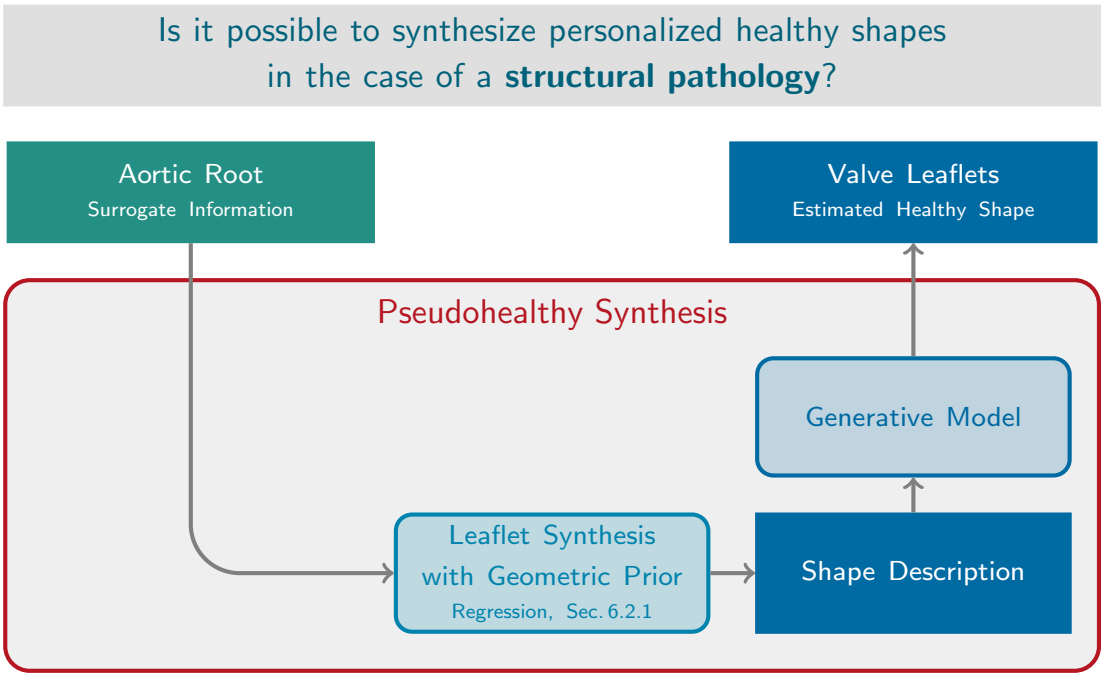

Figure 6.2: Approach to investigate research question *Q2* in the case of a structural pathology.

bridging domain gaps utilize adversarial training, i.e. generative adversarial networks, to transfer an image from one domain to the other [73]. However, GANs need vast amounts of training data to converge that is typically not available in medical imaging. Especially for structure shape synthesis based on surrounding tissue, data collection is very time-consuming and requires additional effort due to ex-vivo experiments. Additionally, GANs are prone to sometimes synthesizing unrealistic images [91], which should be avoided in the scope of medical decision support systems or prosthesis manufacturing.

In this section, a robust approach for synthesizing aortic valve leaflet shapes with the individual aortic root shape as geometric prior is presented based on shape encoding with autoencoders. The given problem is formalized and different ways to solve it are presented. All these approaches are evaluated on the collected ex-vivo data set, including a hyperparameter analysis and a comparison of the proposed approaches. Fig. 6.1 visualizes the approach for investigating research question *Q2* for the case of a structural pathology.

Contribution

The contribution of this section is twofold. First, a novel methodology for bridging big domain gaps that works robustly even on small data is described. As this method is not limited to aortic valve leaflet synthesis, the methodological contribution holds the potential to be of great interest in the medical image analysis community as well as for general computer vision researchers. Second, the proposed method can be directly applied in the scope of personalized aortic valve modeling, not only for prosthesis development but also for personalized cardiac simulations.

6.2.1 Leaflet Synthesis with Geometric Prior

The key idea is to identify a generative model that is capable of synthesizing aortic valve leaflet shapes using representation learning. Specifically, a convolutional autoencoder is utilized to encode the whole variety of leaflet shapes into one unified representation. Autoencoders allow for a high flexibility during the learning process as well as for identifying abstract representations. In comparison to conventional representation learning methods like e.g. principal component analysis (PCA), autoencoders were shown to be more accurate for a data set size of greater than 20 in the medical image analysis domain [140]. Accordingly, it could be shown that a PCA-based representation of aortic valve leaflets only has a limited descriptive power [60].

After identifying an adequate representation, a mapping can be learned between the surrogate information, i.e. a volumetric ultrasound image of the aortic root, and the latent representation of the individual leaflet shapes. This mapping can be interpreted as a mapping between the two different domains, i.e. surrogate data and target data.

Data Sets In the scope of this study, high-resolution images of the aortic valve leaflets as well as an image of the surrounding tissue is necessary to develop suitable solutions. Here, a volumetric ultrasound image of the aortic root serves as geometric prior. Thus, the data set D_S is used that includes such paired data from 30 porcine hearts. The size of the volumetric ultrasound images was $84 \times 202 \times 84$ with a voxel size of $0.71\,mm \times 0.49\,mm \times 0.71\,mm$. From each of these ultrasound volumes, a subvolume was cropped out that shows the aortic root, respectively. Therefore, the commissure plane, i.e. the horizontal slice through the aortic root that shows all three commissure points was identified manually. The subvolume was extracted starting from the commissure plane and spanning the 31 layers below. Hence, the subvolume covered $22.72\,mm$ of the aortic root's height. To remove imaging artifacts and background noise, a thresholding

was applied, setting all gray values smaller or equal to $t = 80$ to 0. Finally, the ultrasound images were scaled with the factor $\frac{1}{255}$ to be in the range $[0, 1]$. The preprocessing of the leaflet images followed Sec. 4.1.2. In addition, the data set D_S^{aux} which only contains leaflet image data is utilized as an auxiliary data set for learning representations of the aortic valve leaflets. However, it is only used for training as the corresponding geometric prior from surrounding tissue remains unknown.

Problem Formulation As described above, the data lies in two different spaces: The volumetric ultrasound data space $\mathcal{D}_{US} \subseteq \mathbb{R}^{32\times128\times128}$ and the leaflet data spaces $\mathcal{D}_{Lf}^{l} \subseteq \mathbb{R}^{128\times64},\ l \in \{rc, lc, nc\}$ for the right-coronary, left-coronary and non-coronary leaflets, respectively. In addition, a latent representation $\mathcal{Z}$ is defined for each data space, leading to the latent spaces $\mathcal{Z}_{US} \subseteq \mathbb{R}^{m_l}$ and $\mathcal{Z}_{Lf} \subseteq \mathbb{R}^{n_l}$. Note that the existence of a shared latent space representation for all three leaflet types, i.e. right-coronary, left-coronary and non-coronary, is assumed. The latent space descriptions can be derived from the data space using representation learning on the corresponding data set, respectively. Figure 6.3 shows the different coordinate spaces and their connection.

The general goal is to synthesize the unknown personalized leaflet images $I^{rc} \in \mathcal{D}_{Lf}^{rc}, I^{lc} \in \mathcal{D}_{Lf}^{lc}$ and $I^{nc} \in \mathcal{D}_{Lf}^{nc}$ of one patient based on the information contained in the patient's individual aortic root geometry given by $V \in \mathcal{D}_{US}$. The direct mapping from V to I^{rc}, I^{lc} and I^{nc} is known as *image-to-image translation* and typically requires a vast amount of training data as it is usually solved using GANs (see section 6.1). First encoding V to its latent representation z_{US} and then mapping to I^{rc}, I^{lc} and I^{nc} is still interpreted as image-to-image translation as the only difference is an unsupervised pre-training of a subnetwork, which is a well-known technique in deep learning on small data [12]. By encoding I^{rc}, I^{lc} and I^{nc} to their latent representations $z_{Lf}^{rc}, z_{Lf}^{lc}, z_{Lf}^{nc} \in \mathcal{Z}_{Lf}$ using the encoder q_{Lf}, the image-to-image translation problem can be reformulated as a regression problem where z_{Lf}^{rc}, z_{Lf}^{lc} and z_{Lf}^{nc} should be estimated based on the ultrasound image data. If a decoder p_{Lf} is known, the corresponding shape leaflet images can be synthesized. Due to intense preprocessing of the leaflet image data, $\mathcal{Z}_{Lf}$ mainly contains shape information. Hence, predicting this latent information based on the volumetric ultrasound data V is referred to as *shape estimation*. Another option is to not only encode the leaflet images but also the ultrasound data. It is assumed that there is a connection between both latent spaces as they represent different parts and states of the same organ. However, the difference between both latent spaces is the domain gap between the data sets. One way to overcome this domain gap is to train a machine learning model to

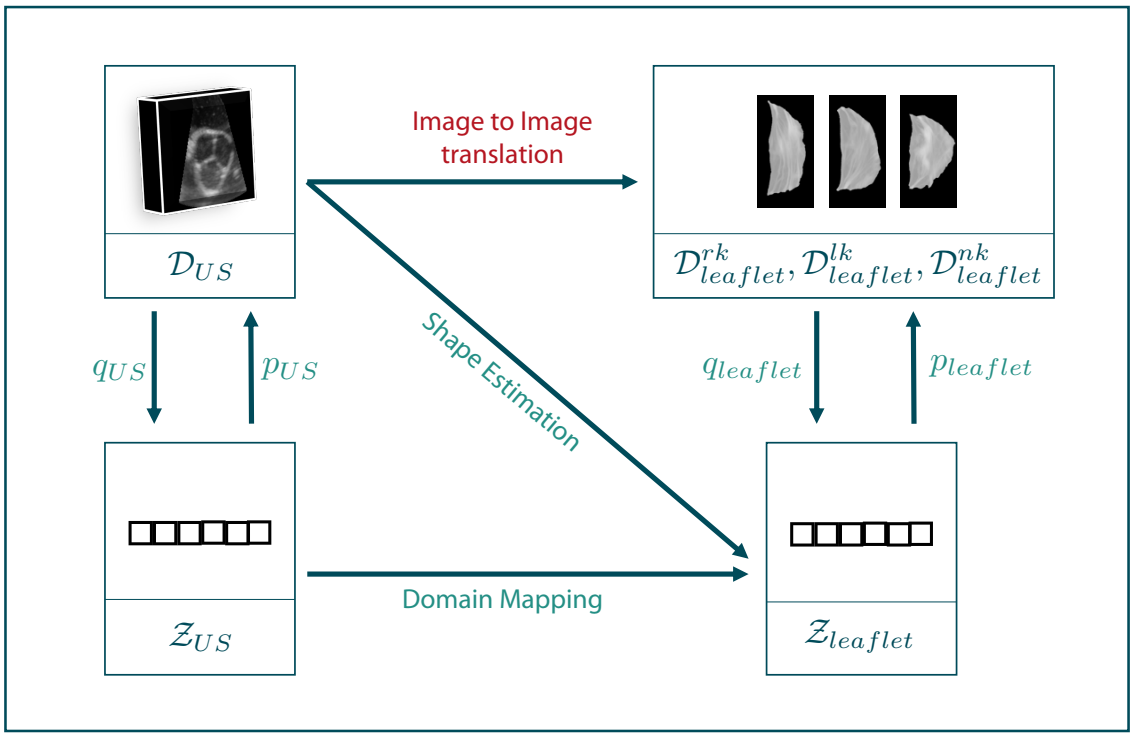

Figure 6.3: Sketch of the proposed approaches. As a direct mapping between the two data domains, also called *image-to-image translation*, is barely possible in the given scenario, it is proposed to encode the data to latent space descriptions using representation learning. Then, the latent leaflet description can be predicted based on the 3D ultrasound image, called *shape estimation*, or based on the latent description of the ultrasound volume, referred to as *domain mapping*.

map from $\mathcal{Z}_{US}$ to $\mathcal{Z}_{Lf}$. Adding the encoder q_{US} and the decoder p_{Lf}, a full path from an ultrasound volume to synthesized leaflet images is given. This approach is called *domain mapping*.

As described above, training an image-to-image translation model on very limited data until convergence is a challenging task. In addition, it is likely that an adversarial model learns the full distribution of leaflets without taking the prior information V into account, making the model fail in personalized shaping. Therefore, it is proposed two follow the other two approaches, namely *shape estimation* and *domain mapping* (see Fig. 6.3). Details on the implementation of both are presented in the following sections.

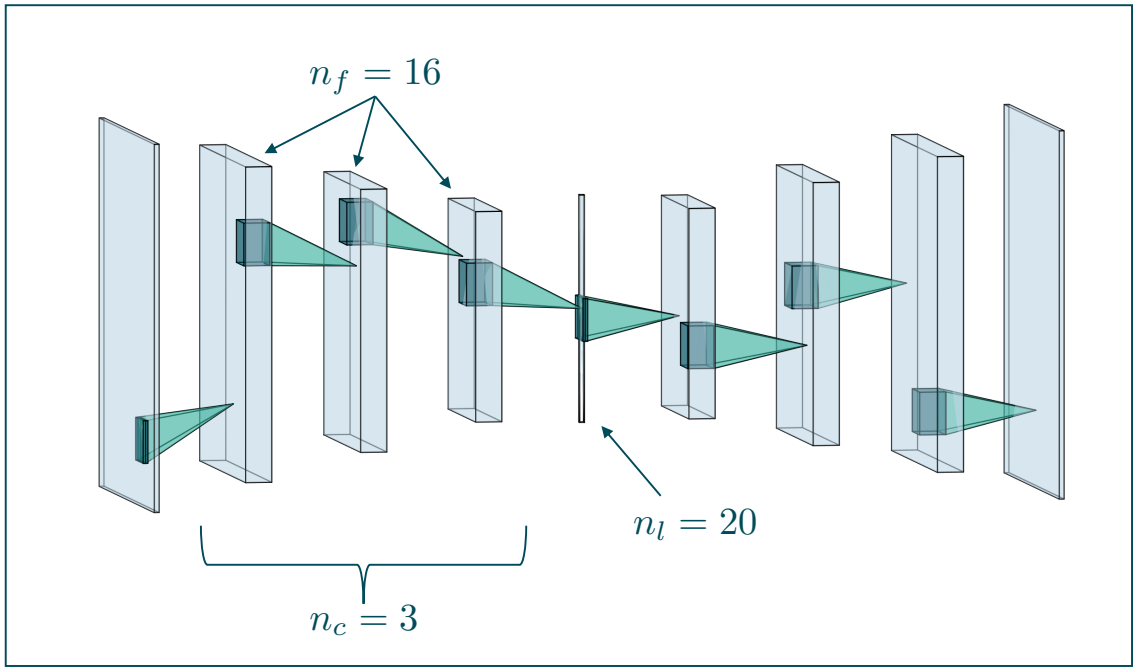

Figure 6.4: Architecture of the leaflet autoencoder, exemplarily shown for $n_c = 3$ convolutional layers, $n_f = 16$ filters per convolutional layer and $n_l = 20$ neurons in the bottleneck layer.

Shape Estimation Approach The goal of the *shape estimation* approach is to encode the leaflet shape, given as an image, into a latent space description and predict the individually optimal latent shape based on a 3D US image of the aortic root. Thus, the approach consists of two models: the leaflet autoencoder for shape encoding and the regression network for mapping from the US image to the latent representation.

For the leaflet autoencoder, a parameterized architecture is used as shown in Fig. 6.4. The encoder q_{Lf} consists of n_c convolutional layers with n_f filters and *ReLU* activation, followed by a 2×2 average pooling each. Following this convolutional part, a flattening operation and fully connected layer (*ReLU* activation) with as many neurons as outputs of the last pooling layer connects the encoder with the bottleneck layer. The bottleneck layer has n_l neurons, featuring *linear* activation. The decoder p_{Lf} follows the mirrored encoder architecture utilizing upconvolution and upsampling. The autoencoder was trained using the *adam* optimizer, mean squared error loss and a batch size of 32 for 100 epochs.

It is assumed that there is a shared latent space description over all leaflet types, i.e. right-coronary, left-coronary and non-coronary leaflets. Hence, the autoencoder is trained on all three kinds of leaflets, leading to the data set $\mathcal{D}_{Lf} = \mathcal{D}_{Lf}^{rc} \cup \mathcal{D}_{Lf}^{lc} \cup \mathcal{D}_{Lf}^{nc}$. To

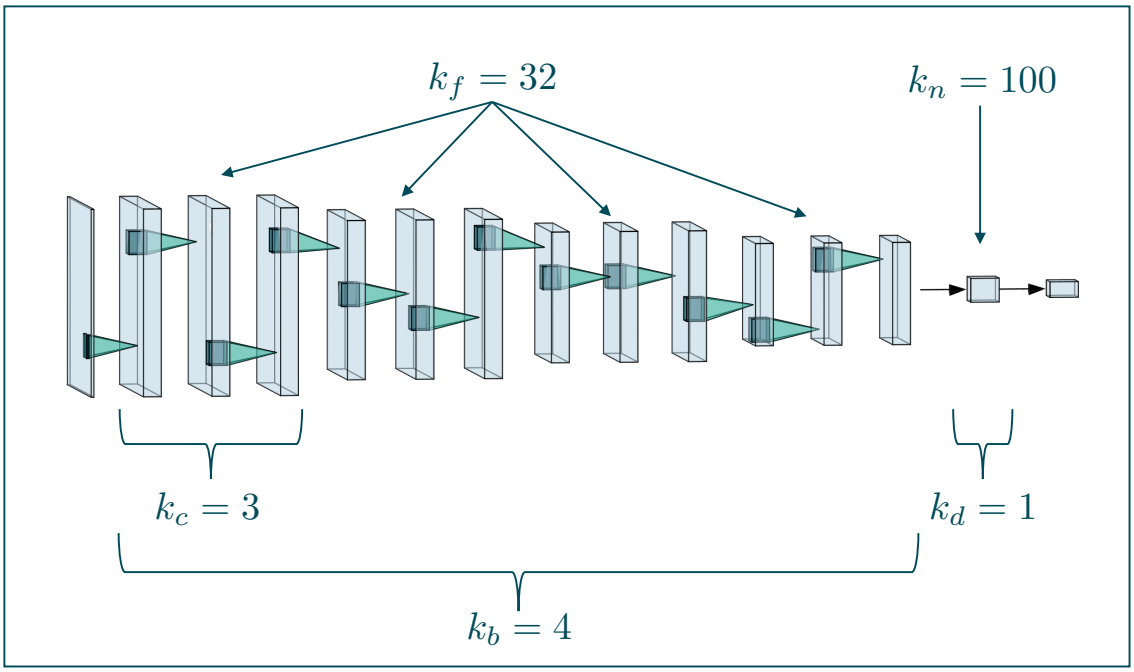

Figure 6.5: Architecture of the regression network, exemplarily shown for $k_b = 4$ convolutional blocks, $k_c = 3$ convolutional layers per block, $k_f = 32$ filters per convolutional layer, $k_d = 1$ fully connected layer and $k_n = 100$ neurons per fully connected layer.

ensure a reliable representation, the data set D_S^{aux} was added to the training data as an auxiliary data set $\mathcal{D}_{Lf}^{\text{aux}}$, containing 168 images of leaflets from 56 valves.

For the regression network, it is proposed to use a 3D convolutional neural network (CNN) with a VGG-like architecture ([132]). Thus, the network consists of k_b convolutional blocks. Each of these blocks consists of k_c 3D convolutional layers with k_f filters and *ReLU* activation. Each block is followed by a 2×2 average pooling layer. Behind the last convolutional block and a flattening operation, k_d fully connected layers with *ReLU* activation and k_n neurons each are attached, followed by the output layer with k_l neurons and *linear* activation. The architecture is illustrated in Fig. 6.5. It is assumed that the prediction of the individual shapes of the right-coronary, left-coronary and non-coronary leaflets are independent. Hence, three regression models are created and trained, predicting one leaflet type each. The three models share the same architecture but are trained independently.

To identify an optimal autoencoder architecture, i.e. an optimal combination of the hyperparameters n_c, n_f and n_l, different combinations were assessed regarding the

Table 6.1: Combinations of hyperparameters of the leaflet autoencoder assessed during hyperparameter analysis. The optimal combination is marked in bold.

Parameter	Values
n_c	**3**, 4, 5
n_f	**16**, 32
n_l	10, **20**, 30

model's ability to reconstruct the input image after propagating it through the model in a grid-search approach. Thus, a 10-fold Monte-Carlo crossvalidation (80% train, 20% test) was performed on $\mathcal{D}_{Lf}$. After training the autoencoder on the training data, the test data was propagated through the model and the resulting reconstruction was compared to the original image using the root mean square error (RMSE). The auxiliary data set $\mathcal{D}_{Lf}^{\text{aux}}$ was added to the training data within each fold. Table 6.1 shows all combinations of hyperparameters tested.

After identifying an optimal autoencoder architecture, the hyperparameters of the regression network were optimized in a similar way. For multiple combinations of the hyperprameters k_b, k_c, k_f, k_d and k_n, the model's performance on predicting the latent shape representation of the corresponding leaflets for unseen ultrasound images is assessed using a 10-fold Monte Carlo crossvalidation (80% train, 20% test) on $\mathcal{D}_{US}$ and $\mathcal{D}_{Lf}$. The autoencoder was trained on the training leaflet images using the optimal hyperparameters. Afterwards, all training leaflet images were encoded to the latent space and the three regression networks were trained to predict the latent space representation of the respective leaflet based on the corresponding 3D ultrasound image. The accuracy was assessed by predicting the shape representation for the test US images, reconstructing images of the predicted leaflet shapes using the decoder and comparing these predicted images to the ground truth leaflet images, once again using RMSE. Table 6.2 shows all hyperparameter combinations assessed in this study.

After identifying optimal sets of hyperparameters for the autoencoder and the regression networks, the performance of the *shape estimation* approach was analyzed. The autoencoder and the three regression networks were trained on the training data, using the respective optimal hyperparameters, and were applied to predict the leaflet shapes for unseen test data. This was performed using a 10-fold Monte Carlo crossvalidation (80% train, 20% test) on $\mathcal{D}_{US}$ and $\mathcal{D}_{Lf}$, while again $\mathcal{D}_{Lf}^{\text{aux}}$ was added to the training data for training the autoencoder. The predicted leaflet images were compared to their corre-

Table 6.2: Combinations of hyperparameters of the regression network assessed during hyperparameter analysis. The optimal combination is marked in bold.

Parameter	Values
k_b	3, **4**, 5
k_c	1, 2, **3**
k_f	16, **32**
k_d	**1**, 2, 3
k_n	50, **100**, 200

sponding ground truth using four metrics: Jaccard similarity, Hausdorff distance, average symmetric contour distance (ASCD) and RMSE. To compute the Jaccard similarity, the Hausdorff distance and the ASCD, the leaflets were segmented in the predicted images utilizing thresholding ($t = 0.45$, corresponding to a grayscale value of 115).

Domain Mapping Approach The key idea of the *domain mapping* approach is to encode both, the leaflet images as well as the volumetric US data, into a latent space description, respectively, and train a model to map from one latent space to the other. One important advantage is that the dimensionality of the mapping is much smaller comparison to the *shape estimation* approach. Hence, it is possible to learn this mapping using classical machine learning methods like Random Forests (RF) or multi-layer perceptrons (MLP) ([116]).

The autoencoder for the leaflet images is the same as presented in Sec. 6.2.1. To find a representation of the US data, a similar architecture is proposed but featuring 3D convolution and pooling. Hence, the encoder q_{US} consists of m_c 3D convolutional layers with m_f filters, *ReLU* activation and followed by a $2\times2\times2$ average pooling each. Following this convolutional part, a flattening operation and fully connected layer (*ReLU* activation) with as many neurons as outputs of the last pooling layer connect to the bottleneck layer with m_l neurons, featuring *linear* activation. Once again, the decoder p_{US} follows the mirrored encoder architecture. Fig. 6.6 shows the architecture of the full autoencoder.

To find a mapping between the two latent spaces $\mathcal{Z}_{US}$ and $\mathcal{Z}_{Lf}$, an RF with t decision trees and an MLP with l_h hidden layers and l_n neurons in each hidden layer were evaluated. For the MLP, all hidden layers used *ReLU* activation, while the output layer utilized *linear* activation.

The optimal hyperparameters of the US autoencoder were identified in a similar way as for the leaflet autoencoder. A 10-fold Monte Carlo crossvalidation (80% train, 20%

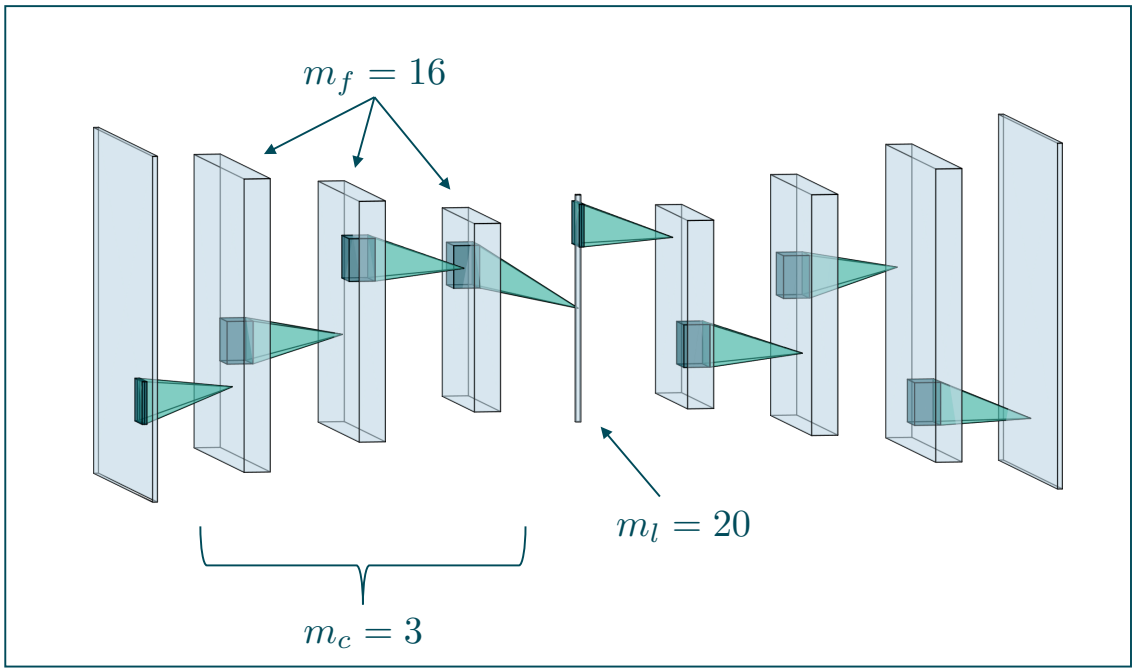

Figure 6.6: Architecture of the ultrasound autoencoder, exemplarily shown for $m_c = 3$, $m_f = 16$ and $m_l = 20$.

test) was performed on $\mathcal{D}_{US}$ and the reconstruction accuracy on the test data was assessed using RMSE. Table 6.3 shows all evaluated combinations of hyperparameters. For finding the mapping between the latent spaces, different values of the hyperparameters of both learning methods were evaluated. For the RF approach, the values $50, 100, 150, 200, 250$ were analyzed for t. For the MLP approach, the values $1, 2, 3, 4, 5$ for l_h and $50, 100, 150, 200$ for l_n were assessed.

After identifying optimal sets of hyperparameters, the performance of the *domain mapping* approach on predicting the individual leaflet shapes based on a 3D US image was evaluated on unseen data. A 10-fold Monte Carlo crossvalidation (80% train, 20% test) was performed on $\mathcal{D}_{US}$ and $\mathcal{D}_{Lf}$. Both autoencoders were trained on the training data (including $\mathcal{D}_{Lf}^{\text{aux}}$ for the leaflet autoencoder) and the training samples were encoded to their latent space representation. Then the models to predict z_{Lf} for a given z_{US} were trained. As described for the *shape estimation* approach, it is proposed to use three different models, one for each leaflet type, i.e. right-coronary, left-coronary and non-coronary. RFs as well as MLPs were evaluated for this regression, each of them using the optimal hyperparameters. After training the models, the test US images were encoded using q_{US}, the latent leaflet representations were predicted and corresponding images were

Table 6.3: Combinations of hyperparameters of the ultrasound autoencoder assessed during hyperparameter analysis. The optimal combination is marked in bold.

Parameter	Values
m_c	**3**, 4, 5
m_f	**16**, 32
m_l	**20**, 100, 200

synthesized using the decoder p_{Lf}. These predicted leaflet images were compared to the ground truth using Jaccard similarity, Hausdorff distance, ASCD and RMSE, with a thresholding for Jaccard similarity, Hausdorff distance and ASCD ($t = 0.45$, corresponding to a grayscale value of 115).

Results and Discussion At first, the results of the hyperparameter analysis are presented, followed by the performance analysis of both approaches. Afterwards, the results are discussed in detail and their impact on future research is given in an outlook paragraph. All hyperparameters were optimized regarding a minimal RMSE. For both autoencoders, an architecture featuring three convolutional blocks with 16 filters each and a latent dimension size of 20 was identified as optimal, with an RMSE of 0.0617 ± 0.0106 for the leaflet autoencoder and 0.0678 ± 0.0200 for the ultrasound autoencoder. For assessing the hyperparameters of the *shape estimation* approach, the optimal leaflet autoencoder hyperparameters were used and the parameter combinations of the CNN mapping from the ultrasound volume space $\mathcal{D}_{US}$ to the latent space of the leaflet $\mathcal{Z}_{Lf}$ were evaluated. The optimal architecture features four convolutional blocks with three convolutional layers using 16 filters each, followed by a single fully connected layer with 100 neurons. This indicates that feature extraction needs a certain degree of abstraction, but the identified features are meaningful and can be processed with a simple classification model, i.e. a single fully connected layer. The RMSE between the true leaflet samples in the latent space description and the predicted ones with this architecture was 0.1331 ± 0.0392. For the *domain mapping* approach, the previously identified, optimal hyperparameters were used for both autoencoders and the hyperparameter influence on the model mapping from $\mathcal{Z}_{US}$ to $\mathcal{Z}_{leaflet}$ was evaluated. In case of an RF regression, 200 trees were found to perform best with an RMSE of 0.1365 ± 0.0278, measured in the leaflet's latent space between the predicted and the true latent representation of the leaflet. Using an MLP for *domain mapping*, the optimal architecture consisted of four hidden layers with 100 neu-

Table 6.4: Accuracy of leaflet shape prediction for all methods, regarding RMSE, Jaccard similarity, Hausdorff distance and ASCD. The values for the Hausdorff distance and the ASCD are given in pixels (resolution $0.34\frac{mm}{pixel}$). Results are given as the mean value over all cross validation folds, averaged over all three leaflet types, and the standard deviation.

Method	RMSE	Jaccard	Hausdorff	ASCD
Shape Estimation	$\mathbf{0.1383 \pm 0.0434}$	$\mathbf{0.8310 \pm 0.0679}$	$\mathbf{6.10 \pm 3.39}$	4.03 ± 9.87
Domain Mapping with RF	0.1417 ± 0.0479	0.8230 ± 0.0737	6.45 ± 4.29	3.10 ± 6.81
Domain Mapping with MLP	0.1482 ± 0.0482	0.8060 ± 0.0796	7.02 ± 4.46	$\mathbf{2.84 \pm 5.01}$

rons each, reaching an RMSE of 0.1386 ± 0.0409. All hyperparameters that were identified to be optimal are marked in bold in Tables 6.1, 6.2 and 6.3.

To compare the different approaches for leaflet shape synthesis, they were evaluated in a 10-fold crossvalidation. Fig. 6.7 shows qualitative synthesis results exemplarily for four valves drawn from a test set within one fold of the crossvalidation. In general, all three synthesis methods are capable of providing realistic leaflet shapes in all cases. The *shape estimation* approach shows vanishing inter-patient variance and tends to predict an average leaflet for each individual. While the *domain mapping* approach with RF provides a slightly higher variance, *domain mapping* with MLP shows the highest variance and hence individualization performance. The latter approach is capable of following individual shapes with relatively high accuracy in some cases, e.g. the right-coronary leaflet of valve 1. However, all presented methods struggle in predicting atypical leaflets, e.g. the right-coronary leaflet of valve 2.

Table 6.4 shows the results of this study regarding Jaccard similarity, Hausdorff distance, ASCD and RMSE, averaged over all folds and all leaflet types. All three methods are capable of achieving a Jaccard similarity of over 80%, while the *shape estimation* approach reaches the maximal value of around 83% coverage. This is most likely due to the high overlap between the test leaflets and an average one. However, regarding the contour line accuracy, the *domain mapping* approach provides lower ASCD values, indicating that it outperforms the *shape estimation* approach regarding finer details of the leaflet shape. As the image resolution was $0.34\frac{mm}{pixel}$, the *domain mapping* with an MLP achieves an ASCD of $0.97\,mm$ and hence is shown to predict the leaflet contour with sub-millimeter accuracy, with a Hausdorff distance of $2.39\,mm$.

In Fig. 6.8, the ASCD is given for each method regarding the accuracy in predicting the different leaflet types, i.e. right-coronary, left-coronary and non-coronary. The results show that all methods perform much better when predicting the non-coronary leaflet

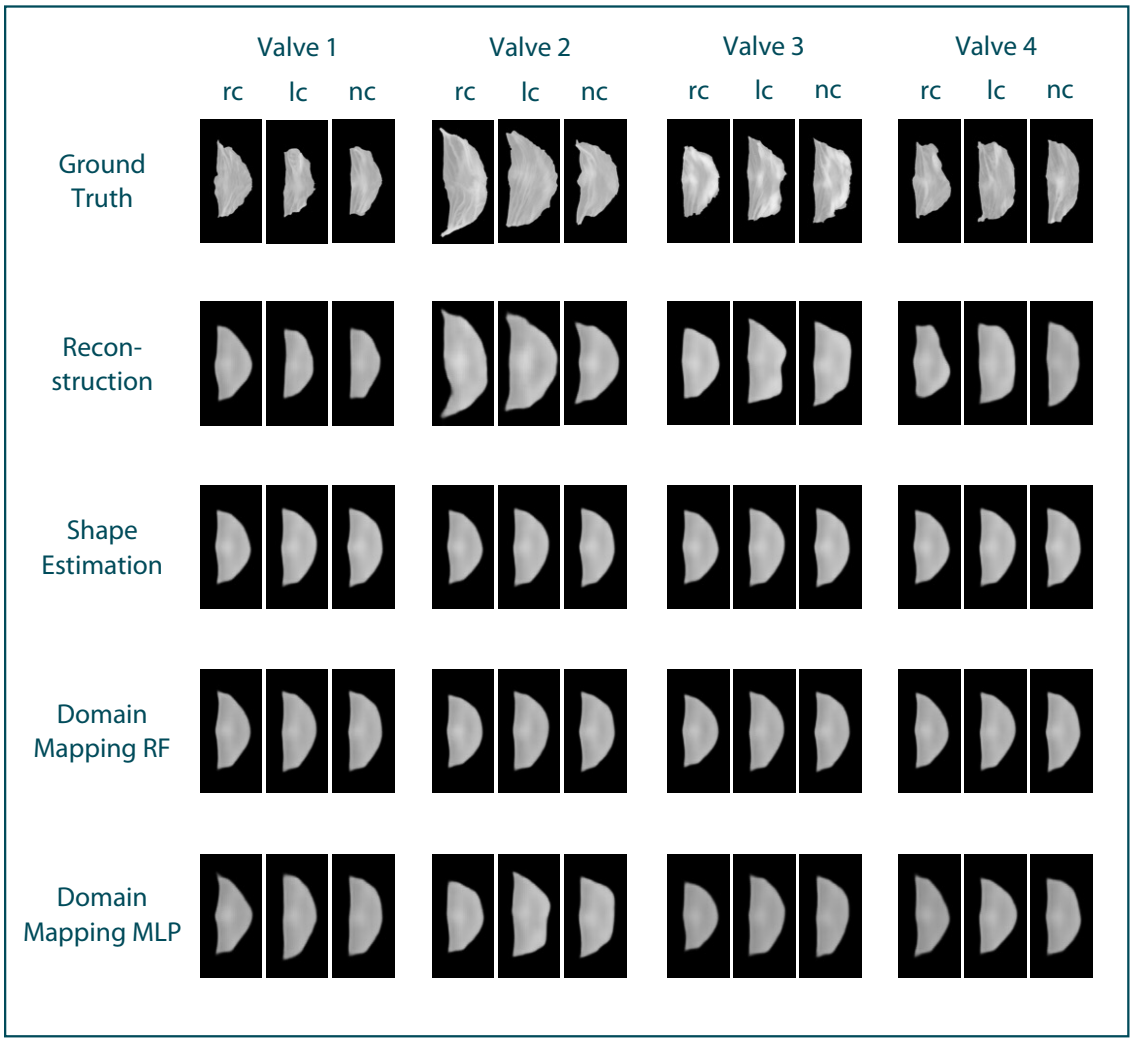

Figure 6.7: Qualitative results for leaflet shape synthesis with geometric prior, exemplarily given for four valves (columns) with its right-coronary (rc), left-coronary (lc) and non-coronary (nc) leaflet, respectively. The rows show the ground truth, the reconstruction of the ground truth after propagating it through the whole autoencoder, as well as the synthetic leaflets produced by the different approaches, i.e. *shape estimation*, *domain mapping* with Random Forests and *domain mapping* with multi-layer perceptrons. All four valves were drawn from the test set of one fold of the crossvalidation.

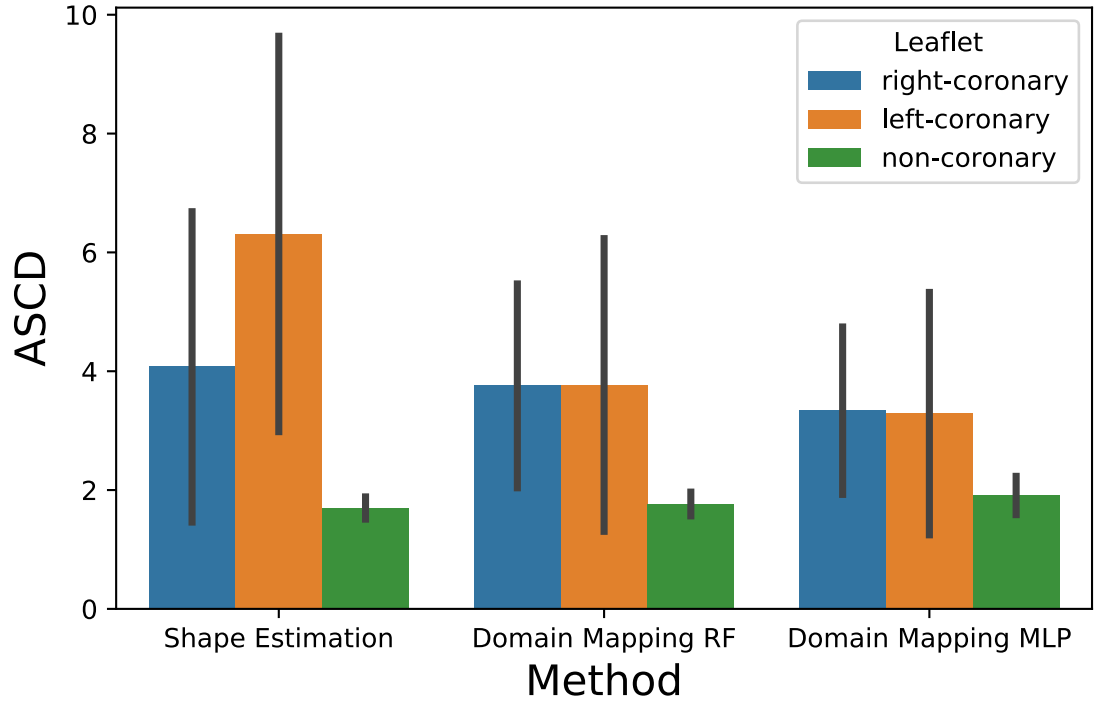

Figure 6.8: Accuracy of all methods depending on the predicted leaflet type, given as the ASCD in pixels. While the left- and non-coronary leaflet can be sufficiently predicted, the accuracy drops for right-coronary leaflets.

shape compared to the other two leaflet shapes. This might be due to the fact that the inter-patient shape variance is higher for the right- and left-coronary leaflets than for the non-coronary one. Additionally, the *shape estimation* performs worse in predicting the left-coronary leaflet than the right-coronary one, while both accuracy values are comparable in the case of *domain mapping*, regardless of the regression method. This indicates a higher robustness of the latter approach by utilizing meaningful features that describe the aortic root geometry identified during representation learning.

Outlook The results indicate that bridging the large domain gap between volumetric ultrasound images of the aortic root and the planar shape of its three leaflets is possible even using a small amount of data. As the *domain mapping* with MLP provided personalized synthetic shapes, the ultrasound image carries enough information for this mapping when it is combined with a proper leaflet representation. However, the relevant

information seems to be hidden within the ultrasound image and hard to extract as the *shape estimation* approach was not capable of providing adequately personalized synthetic images. By encoding the images to a latent space representation, it is possible to extract the relevant information, but still a relatively high level of abstraction is needed to make use of it as the optimal MLP featured four hidden layers. Accordingly, even though the RF grasps some parts of this information, its performance lags behind the MLP-based approach. Hence, in this scenario, the *domain mapping* approach with MLP clearly outperformed the other methods. However, this might be highly task-dependent and it is conceivable that the *shape estimation* method might perform better for tasks with lower-dimensional surrounding tissue data, i.e. 2D images. A broader analysis of the performance of the proposed methods for different applications and scenarios could provide important insights.

Since the influence of personalized leaflet shaping on cardiac simulations or prosthesis performance has barely been studied, it is not clear which quantitative value for the accuracy of leaflet synthesis can be considered as clinically acceptable. The methods proposed in this work present the first model to personalize leaflet shapes in a data-driven way and hence open up possibilities for the asssessment of the influence of individual leaflet shaping, for example in a whole-heart simulation or, after fabricating a personalized prosthesis, in a left-heart simulator setup.

This study focused solely on predicting the leaflets' shapes. However, regarding personalized biomechanical modeling or prosthesis development, a prediction of the leaflet's inner structure, i.e. the distribution of the prominent collagen fibers, would also be of high interest. The autoencoder used in this study smoothes the inner structure and focuses on representing the general shape, which is the desired behaviour in the scope of this work. However, the leaflet representation could be extended to also contain information about the inner structure. Using this adaptation, the approaches described here would be capable of predicting the full leaflet structure, i.e. the shape and the inner structure. This highlights the flexibility of the presented framework.

Conclusion In this section, a new framework to synthesize the shape of unknown anatomical structures based on the geometry of surrounding tissue was presented by solving a domain mapping problem. The problem was formalized and two general approaches were proposed to solve it. In an evaluation of this framework regarding the application of synthesizing aortic valve leaflet shapes based on volumetric ultrasound images of the aortic root, it could be shown that the method is capable of reliably synthesizing realistic leaflet shapes and that the geometric prior carries enough information to synthesize personalized leaflet shapes when both domains are encoded into a latent space and an MLP is used to learn a mapping between both latent spaces. This proof-of-concept study does not only open up numerous novel applications of personalized aortic valve modeling but also presents a transferable approach for anatomical shape synthesis with geometric prior in a more general context.

6.3 Morphological Pathology

Personalized prosthesis shaping often suffers from a lack of information on the optimal desired shape. If the shape of an organ is deformed due to a pathology, the prosthesis should ideally remodel the original, healthy shape that is usually unknown. Hence, the choice of the optimal shape is an estimation problem where the prior information is the individual pathological shape. For the surgeon, this estimation presents a challenging task. One example is the optimal graft prosthesis shaping in valve-sparing aortic root reconstruction. During this surgery, the pathologically dilated aortic root is remodeled using a graft prosthesis, while the patient's native aortic valve is spared and installed in this prosthesis [32].

Consequently, the success of the surgery highly depends on the optimal choice as the prosthesis' shape strongly affects the functionality of the aortic valve. A patient-prosthesis-mismatch can lead to severe complications, ranging from aortic valve insufficiency to the patient's death [126].

However, the desired healthy shape is typically unknown. To assess this shape, pseudohealthy synthesis is necessary. The term pseudohealthy synthesis was introduced in [19] and refers to the synthesis of images that show an estimated pseudohealthy state based on pathological information. The typical approach is to find an appropriate low-dimensional representation of the healthy images and find a mapping from the pathological information to the corresponding healthy image's representation. The pseudohealthy image is then synthesized using a decoder structure. Previously published methods proposed end-to-end models where the mapping between the pathological information and the pseudohealthy image is learned jointly together with the low-dimensional representation. For example, [28] utilized the decoder subnetwork of an autoencoder to perform this mapping, while the mapping is completely integrated into the representation learning in the GAN-based approach of [136].

However, this approach is not sufficient for personalized prosthesis shaping. Pathological out-of-bag samples with a very different shape compared to the healthy state may lead to unrealistic predictions [105]. As inner organs like the aortic valve apparatus show a substantial inter-patient variability not only in the healthy but also in a pathologically dilated state [62], the restriction to homogeneous input data is unrealistic. [146] proposed to integrate a segmentation of the pathological area into the method to overcome this problem. However, as a pathological dilation affects the whole aortic root, a segmentation does not benefit the method in this context.

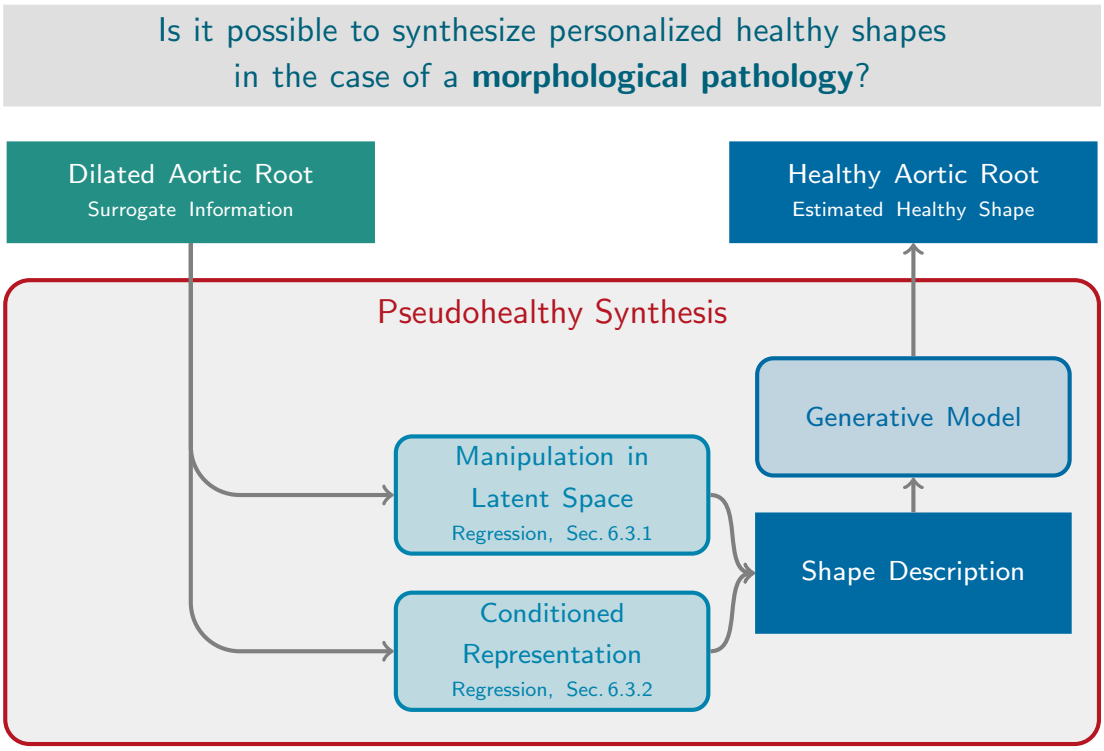

Figure 6.9: Approaches to investigate research question *Q2* in the case of a morphological pathology.

In a previous study, a GAN-based approach was assessed to solve the mapping from pathologically dilated to healthy aortic root shapes using the data set D_M [34]. Even though several architectures, hyperparameters and model types were evaluated, none of the GANs adequately converged to a state where a realistic aortic root shape could be synthesized. This supports and emphasized the need for a novel, accurate and consistent method for this case.

In Sec. 6.3.1, a novel concept for pseudohealthy synthesis in the context of personalized prosthesis shaping is presented. The problem of pseudohealthy synthesis is reformulated as a manipulation, i.e. a translation, of the organ shape in a latent space description. Following this approach, the representation learning and the mapping from pathological to pseudohealthy samples are disentangled, allowing for direct control of the model complexity and the integration of consistency constraints. In Sec. 6.3.2, an end-to-end method using conditioned representation learning is evaluated to benchmark the manipulation-based approach against this end-to-end one. Fig. 6.9 visualized the approaches to investigate research question *Q2* for the case of a morphological pathology.

Contribution

This section presents multiple contributions to the field. First, this study demonstrates the first application of pseudohealthy synthesis for aortic root shaping and, hence, represents the basis for personalized modeling and prosthesis shaping in the context of cardiovascular implants. Second, a novel framework to reduce the complex image transformation to a simple translation in latent space is presented, being the first of its kind. Third, the concept of shape primitive described in Sec. 6.3.1 is the first method for pseudohealthy synthesis that is implicitly constrained to avoid out-of-cluster predictions, making it highly relevant for clinical application. Fourth, for the *Fixed Translation* agent as well as for probabilistic prediction with *Shape Primitives*, both presented in Sec. 6.3.1, the method does not provide a single solution but a set of solution candidates, opening up the possibility to suggest multiple options to the surgeon to pick the optimal one. Hence, the method takes the step from a decision system to a decision support systemholding the potential to highly increase the acceptance among surgeons substantially. Additionally, this novel concept is not constrained to aortic root shapes but is applicable to all fields of pseudohealthy synthesis for computer-assisted interventions.

6.3.1 Pseudohealthy Synthesis via Manipulation in Latent Space

In this section, a novel concept of interpreting pseudohealthy synthesis as a manipulation in a latent space description is presented. As each point in this space represents one anatomical shape, a translation within the latent space corresponds to a manipulation of the anatomical shape in image space. Thus, this complex deformation becomes a simple translation that can be learned data-drivenly.

The core idea is to use a semi-supervised variational autoencoder to identify an appropriate representation with distinct clusters of pathological and healthy samples. For pseudohealthy synthesis, a new concept is introduced where deformation agents in this latent space estimate a corresponding healthy point for each pathological input sample. The image synthesis can be performed using the decoder network. Three different approaches are developed and proposed to define this deformation agent, ranging from linear, fixed translation over flexible, nonlinear models to probabilistic prediction. These approaches not only feature different levels of complexity of the mapping but also open up dedicated possibilities for including the surgeon into the final decision-making pipeline, bridging the gap from end-to-end decision systems to decision support systems with the surgeon in the loop. In the following, the model construction as well as the inference workflow is described in detail and all proposed approaches are evaluated in the setting of planning valve-sparing aortic root reconstruction.

Methods In this paragraph, a detailed description of the proposed method is provided. The key idea is to find a latent space representation of shapes where pathological and healthy samples form two distinct clusters using a semi-supervised variational autoencoder. Then, a deformation agent is trained to predict the individually optimal healthy shape based on the latent space description of the pathological state. Using the decoder, an image of the predicted pseudohealthy shape can be reconstructed. It is important to note that in this concept, the deformation agent only receives input from the latent space while the image space remains unkown. This allows the agent to use the learned features of the autoencoder that are known to be descriptive. In the proposed concept, the autoencoder and the deformation agent are trained individually and one after another instead of in an end-to-end-manner. By disentangling representation learning and agent training, the model can be implicitly constrained to consistently predict realistic healthy shapes. After explaining the representation learning approach, the concept of deformation agents in latent space for pseudohealthy synthesis is presented. Three different approaches are proposed to define this agent, *Fixed Translation*, *Direct Estimation* and *Shape Primitives*.

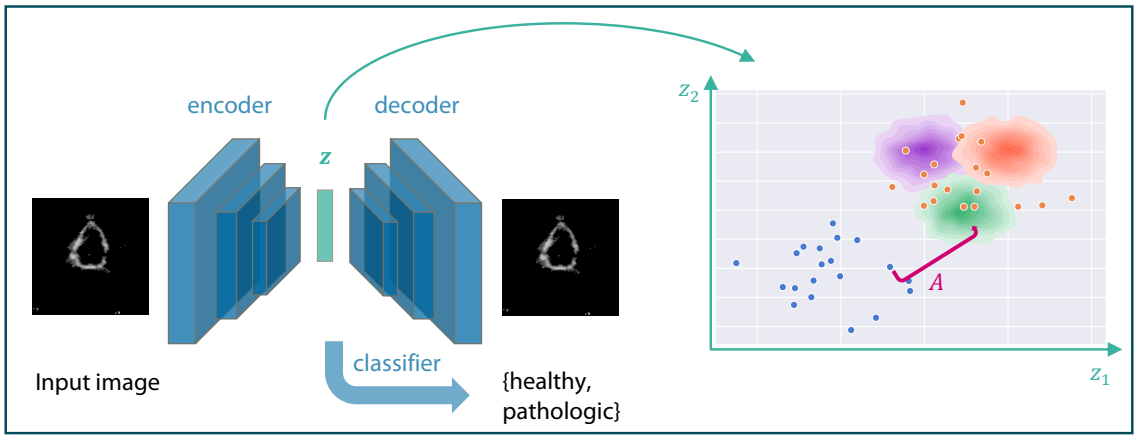

Figure 6.10: Illustration of the representation learning approach for pseudohealthy synthesis. The image data is encoded to a latent space description z using a semi-supervised variational autoencoder. In the latent space, the pathological (blue) and healthy (orange) samples are separable. A deformation agent $\mathcal{A}$ (magenta) is trained to map from a pathological sample to the corresponding healthy sample. If the concept of *Shape Primitives* is used, the healthy samples are described using radial basis functions (violet, red and green distributions) and the agent maps to their activation, respectively.

For the latter, a novel concept of shape primitives in latent space is introduced to represent shapes in a probabilistic way. In the case of a *Fixed Translation* or *Shape Primitives* agent, it is possible to design concepts for candidate prediction to include the surgeon into the decision making pipeline.

Representation Learning First, a suitable representation of the anatomical shapes must be identified. A sufficient representation should contain unique mappings from image space to the latent space and back, the generative model should be smooth in the latent space to ensure realistic synthetic images and the pathological and healthy subjects should lie in distinguishable clusters. To fulfill these requirements, a semi-supervised variational autoencoder inspired by [80] is proposed for implicitly learning this representation. The encoder subnetwork is denoted as q and the decoder as p. In addition to the classical encoder-decoder-architecture, another branch δ is added, arising from the bottle-neck layer and in parallel to the decoder. It consists of a single classification layer (fully-connected, sigmoid) to predict the right label, i.e. *healthy* or *pathologic* in one-hot-encoding, for a specific point in latent space. This branch forces the representation to

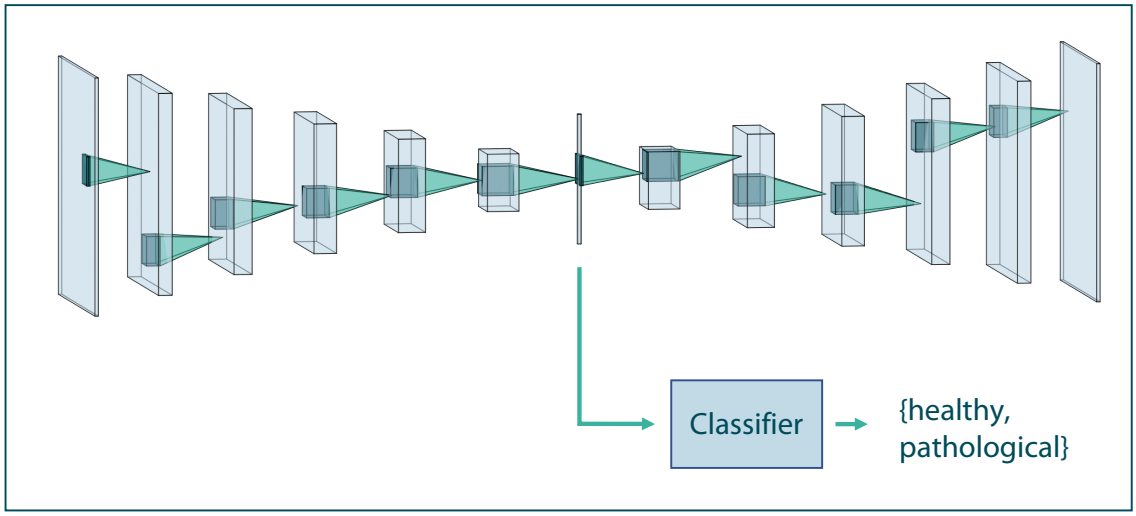

Figure 6.11: Architecture of the proposed autoencoder. In addition to the decoder, a second branch is integrated that classifies the datapoints within the latent space into the categories *healthy* or *pathological*. Note that the latent space is sketched as a single vector to improve clarity. In the implementation, the latent space is a variational model described by μ and σ utilizing reparameterization (see Sec. 3.4 for details).

make healthy and pathological samples distinguishable, resulting in two separate clusters (see Fig. 6.10). The architecture of the autoencoder is shown in Fig. 6.11.

The loss function L is designed as the weighted sum between the classical variational autoencoder loss L_{vae} as described in [81] and the classification loss in the discriminative branch L_δ as

$$\begin{aligned} L =& \alpha L_{vae} + (1-\alpha) L_\delta \\ =& \alpha(\mathbb{E}(\log p(X|z)) - D_{\text{kl}}(q(z|X)||p(z))) \\ &+ (1-\alpha) D_{\text{mse}}(\delta(\mathbb{E}(q(z|X))), y_{\text{true}}), \end{aligned} \tag{6.1}$$

where X is the input image, y_{true} is the true binary label of this image (*healthy* or *pathological*), D_{kl} is the Kullback-Leibler-Divergence and D_{mse} is the mean squared error. The parameter α controls the trade-off between a smooth representation (L_{vae}) and the separability of the two clusters (L_δ).

Deformation Agents in Latent Space After finding a sufficient representation with two distinct clusters, the complex 3D deformation of the organ shape can be formulated as a translation in this latent space. To this end, the concept of deformation agents in latent space is introduced. The aim is to train an agent to estimate an optimal healthy state based on the latent representation of the individual pathological one. Please note that in the scope of this chapter, the term agent relies on the very broad and general definition of an agent described in [116]. Specifically, the agent's training routine is not fixed, so supervised, unsupervised or even reinforcement- or adversarial-based training is possible. This open definition is important as in clinical data, the pathological as well as the healthy shape from each patient is not necessarily known and such an unpaired data set might make a different training routine necessary. But even in this case, the framework stays the same. In the following, three approaches are presented to define the deformation agent: *Fixed Translation*, *Direct Estimation* and probabilistic prediction with *Shape Primitives*. All three approaches feature different levels of complexity and are applicable either in a paired or an unpaired data scenario.

Fixed Translation This naive approach assumes that a fixed translation vector, added to the pathological sample, is sufficient as long as the vector is guaranteed to point towards the healthy cluster. As this vector can be computed distribution-based as the difference between the cluster center points, no point-to-point correspondences are needed for model construction. This makes the *Fixed Translation* applicable in an unpaired data setting. Additionally, the subject identity is spared and the prediction result is implicitly constrained to lie in the healthy cluster as long as the pathological shape lies within the cluster of observed pathological shapes. However, the complexity is limited as the translation is not patient-dependent.

For model construction, the translation vector t is computed based on the given training data. Thus, all training images X_{healthy} and X_{patho} are encoded to compute their d-dimensional latent representation z_{healthy} and z_{patho} using the encoder q. The vector $t \in \mathcal{R}^d$ is the difference between the center points of the healthy and the pathological cluster. Its components can be calculated as

$$t^m = \frac{1}{N_{\text{healthy}}} \sum_{i}^{N_{\text{healthy}}} z_{\text{healthy}}^{i,m} - \frac{1}{N_{\text{patho}}} \sum_{l}^{N_{\text{patho}}} z_{\text{patho}}^{l,m},\ m = 1, \ldots, d, \tag{6.2}$$

where N_{healthy} and N_{patho} are the numbers of healthy and pathological training samples, respectively, and $z_{\text{healthy}}^{i,m}$ describes the m-th component of the latent representation of the i-th healthy sample. The resulting t serves as translation vector during inference.

Inference on an unseen pathological image X^j_{patho} is performed after encoding this image to its latent space representation z^j_{patho}. Then, the translation is a simple vector addition:

$$\tilde{z}^j_{\text{healthy}} = \mathcal{A}^{\text{Fixed Translation}}(z^j_{\text{patho}}) = z^j_{\text{patho}} + t. \tag{6.3}$$

The predicted pseudohealthy image $\tilde{X}^j_{\text{healthy}}$ is synthesized using the decoder p.

One important property of the agent is that it is possible to divide the translation into S steps where after each step, a pseudohealthy candidate solution $z^{j,s}_{\text{healthy}}, s = 1, \ldots, S$ is predicted:

$$\tilde{z}^{j,s}_{\text{healthy}} = z^j_{\text{patho}} + \frac{s}{S}t. \tag{6.4}$$

This way, multiple candidate solutions with different levels of deformation can be predicted. This allows the surgeon to pick the optimal candidate and thus facilitates human-in-the-loop decision making.

Direct Estimation One downside of the *Fixed Translation* approach is that the translation is independent of the individual pathological shape, leading to a linear deformation model. To overcome this problem, the *Direct Estimation* approach is introduced to learn a sample-dependent mapping using supervised learning. This leads to a highly flexible model that is capable of learning complex deformations. However, the model is not implicitly constrained and unreal predictions might be possible. Additionally, due to the supervised training routine, the approach is only applicable in paired data scenarios.

As the name suggests, the *Direct Estimation* agent maps directly from the pathologic to the healthy sample in the latent space:

$$\tilde{z}_{\text{healthy}} = \mathcal{A}^{\text{Direct Estimation}}(z_{\text{patho}}). \tag{6.5}$$

In this study, supervised learning is utilized to derive this mapping in a data-driven way. Hence, the construction step mainly consists of training the model based on all given training samples, mapping from the pathological to the known, healthy ground truth by minimizing the mean absolute error in latent space. A fully-connected neural network model is proposed for the agent $\mathcal{A}^{\text{Direct Estimation}}$.

After encoding the pathologic image X^j_{patho} to retrieve z^j_{patho}, the pseudohealthy state can be estimated using the already trained agent as

$$\tilde{z}^j_{\text{healthy}} = \mathcal{A}^{\text{Direct Estimation}}(z^j_{\text{patho}}) \tag{6.6}$$

and the corresponding pseudohealthy image $\tilde{X}^j_{\text{healthy}}$ is computed using the decoder.

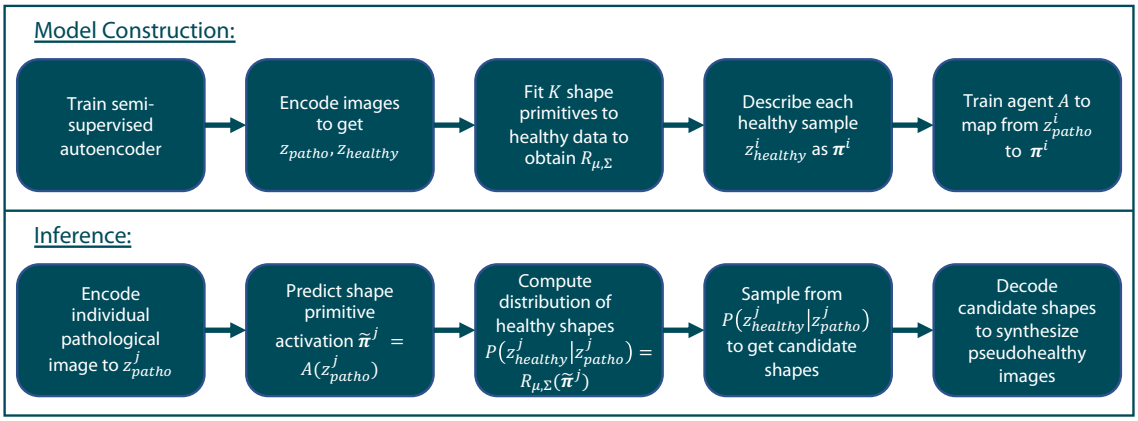

Figure 6.12: Flowchart of the proposed method for pseudohealthy synthesis with shape primitives, divided into the model construction step including the training process (top) and the inference step for unseen data (bottom).

Probabilistic Prediction with Shape Primitives Even though the *Direct Estimation* agent is capable of learning highly nonlinear mappings in latent space, the model is not constrained to avoid unrealistic prediction results. Hence, it is not guaranteed that the pseudohealthy image resembles a realistic shape as observed in the healthy training data. To overcome this problem, the concept of shape primitives in latent space is introduced.

The key idea of this method is to approximate the set of healthy samples in a latent space description with a set of radial basis functions, i.e. gaussians in latent space, called shape primitives. It is assumed that each organ shape can be constructed by a linear combination of these typical shapes, the shape primitives. This implies a probabilistic view on anatomical shape characteristics as the linear combination of shape primitives does not lead to a discrete point but rather to a distribution, described by the mixture of these gaussians. For pseudohealthy synthesis, a mapping from an individual pathological sample to this distribution can be identified using a trainable agent instead of predicting a discrete point. From the estimated distribution, healthy candidate shapes can be sampled. Fig. 6.12 illustrates the workflow for the model construction step as well as for inference on new data. Both steps are described in detail in the following paragraphs.

After training the autoencoder and encoding all samples to the latent space, the shape primitives can be computed. The K shape primitives are defined as gaussians $\mathcal{N}(\mu_k, \Sigma_k)$

in the d-dimensional latent space. Hence, the likelihood of the i-th sample $z^i \in \mathbb{R}^d$ can be completely represented by the weight vector $\boldsymbol{\pi}^i \in \mathbb{R}^K$ of the mixture of these gaussians:

$$P(z^i) = \sum_{k=1}^{K} \pi_k^i \mathcal{N}(\mu_k, \Sigma_k) = \mathcal{R}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\boldsymbol{\pi}^i), \tag{6.7}$$

where μ_k and Σ_k are the mean and the covariance matrix of the k-th shape primitive and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the sets of all K means and covariances, respectively.

To identify $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ for the given data set, all healthy shape samples of the training set are encoded to the latent space and a Gaussian Mixture Model (GMM) with K components was fitted to this data using Expectation Maximization, in direct analogy of GMM clustering [103]. It is assumed that the shape primitives are radial. Thus, the method can rely on the broad and very well-established theoretical foundation of radial basis functions as universal function approximators, including their robustness in high dimensional spaces with a small number of observations [111]. Following this assumption, Σ_k is a diagonal matrix where all diagonal entries have the identical value of v_k and the representation of the likelihood can be simplified to $\mathcal{R}_{\boldsymbol{\mu},\boldsymbol{v}}(\boldsymbol{\pi}^i)$, where $\boldsymbol{v}$ is the vector of the variances of the K shape primitives.

To compute $\boldsymbol{\pi}^i$ for a specific data point z^i, the idea of radial basis functions is used to calculate the activation of this data point for each shape primitive as

$$\pi_k^i = \frac{\hat{\pi}_k^i}{\sum_{j=1}^{K} \hat{\pi}_j^i} \tag{6.8}$$

with

$$\hat{\pi}_k^i = \exp\left(-\theta \frac{\|z^i - \mu_k\|}{2v_k}\right), \tag{6.9}$$

where θ is an additional scaling factor to control the width of the shape primitives. The normalization of the activations ensures that all activations for one data point sum up to one.

For mapping a pathological sample point z^i_{patho} to the distribution of corresponding healthy shapes $P(z^i_{\text{healthy}})$, the method relies on an deformation agent $\mathcal{A}$ that estimates the weight vector $\boldsymbol{\pi}^i$ with

$$\boldsymbol{\pi}^i = \mathcal{A}(z^i_{\text{patho}}). \tag{6.10}$$

The agent can be trained in a supervised way on tuples $\langle z^i_{\text{patho}}, \boldsymbol{\pi}^i \rangle$ known from the training data set. It is proposed to train the autoencoder and the deformation agent individually and one after another instead of an end-to-end-manner as the shape primitives

should be already known when the deformation agent is trained. If the shape primitives could be varied during the training process, the regularizing effect of using them would vanish, the risk of overfitting would increase and out-of-bag prediction would be more probable. After training the agent, the whole model is ready for inference on unseen data.

For performing pseudohealthy synthesis, the individual image of the pathological shape is encoded to its latent space representation $z^j_{\text{pathologic}}$ and the agent is used to predict the estimated weight vector $\tilde{\boldsymbol{\pi}}^j$. Then, given the individual pathological shape, the distribution of healthy shapes can be computed as

$$P\left(z^j_{\text{healthy}}|z^j_{\text{patho}}\right) = \mathcal{R}_{\boldsymbol{\mu},\boldsymbol{v}}\left(\tilde{\boldsymbol{\pi}}^j\right) = \mathcal{R}_{\boldsymbol{\mu},\boldsymbol{v}}\left(\mathcal{A}\left(z^j_{\text{patho}}\right)\right). \tag{6.11}$$

From this distribution, possible candidate shapes can be sampled. Using the decoder network, a pseudohealthy image can be synthesized for each of these candidate shapes. These images could be shown to the medical staff for final decision making.

Application to Aortic Root Reconstruction In this section, it is described how the concept of pseudohealthy synthesis via manipulation in latent space can be applied to assist the surgeon's decision making in planning valve-sparing aortic root reconstruction surgery. For this proof-of-concept study, the data set $\tilde{D}_M$ was utilized, containing 24 roots. For each aortic root in the data set, a pathologically dilated state as well as the corresponding healthy state are known, resulting in 48 images.

The 2D ultrasound images used in this study show a horizontal slice through the aortic root at the height of the commissure plane as this imaging plane provided optimal results in previous studies [56].

Model Definition In this paragraph, the definition of the network architectures as well as the choice of the parameters used in this study will be described.

The autoencoder used in this study follows the architecture proposed in [58] and consists of five convolutional layers in the encoder q (8/16/32/64/128 filters, respectively, filter size 3×3, ReLU activation), each followed by a pooling layer (average pooling, filter size 2×2). The decoder p was the mirrored encoder architecture where the pooling layers were replaced by upsampling layers. The discriminative branch consists of a single classification layer (fully-connected, softmax activation). In this study, the loss weighting parameter α was set to $\alpha = 0.5$, so both parts of the loss function were weighted equally. The number of latent space dimensions d was set to 12. Data augmentation was

intentionally not applied even though the data set is small since all artificial variance introduced by the augmentation would have to be encoded in the latent space. Additionally, the definition of realistic image transformations in ultrasound images is not trivial due to the physics of this modality.

As the *Direct Estimation* and the *Shape Primitive* approach feature a supervised agent for learning the mapping in latent space, a multi-layer perceptron (MLP) is used for this purpose. An MLP with h hidden layers and n neurons in each hidden layer will be denoted as $A_{h,n}$. The output layer has d neurons for *Direct Estimation* and K outputs for *Shape Primitives*, respectively. While all layers use ReLU activation in general, sigmoid activation was applied in the output layer for the *Shape Primitive* approach to force the predicted weights to lie in the interval $[0, 1]$. For *Direct Estimation*, the output layer has linear activation. The *adam* optimizer was used for training with mean squared error loss [86].

Evaluation Methodology The proposed method was evaluated on the given data set. Monte-Carlo cross validation was performed with ten runs and splits of 80% training and 20% test data, where the training data was used for model construction and the test data for inference. It is important to note that cross validation was not performed on images, but on image pairs of the healthy and dilated state of the same aortic root. To analyze the influence of hyperparameters of the deformation agents for *Direct Estimation* and *Shape Primitives* on the model performance, all combinations of $h \in \{1, 2, 3, 4, 5\}$ and $n \in \{50, 100\}$ were evaluated. To further assess the influence of the number of shape primitives in the *Shape Primitives* approach, each agent was evaluated for $K \in \{2, 4, 6, 8, 10, 12, 14, 16, 18, 19\}$, as 19 is the number of healthy samples in each training set. To assess the accuracy quantitatively, errors were calculated in image space. A predicted shape was decoded to synthesize an image and the root shape within this synthetic image was compared to the one in the ground truth image. To achieve this, the aortic root was segmented manually in all healthy images. For the synthesized images, automatic segmentation was applied using thresholding (pixel grayscale value > 80 defined as part of the root). Then, the segmentation of the predicted root shape could be compared to the ground truth using the Dice coefficient [153].

For the *Shape Primitive* approach, the candidate prediction requires a selection step. Since the resulting prediction is not a single shape but a set of sampled shapes from the distribution, the decision process of a human expert using the proposed system for surgery was simulated. Assuming an idealized expert, five candidate shapes were sampled, the similarity to the ground truth was computed using the Dice coefficient and the

best sample was taken as the final prediction. The implementation was done using *Tensorflow* with the *Keras* framework [29].

Results and Discussion This section presents the results of the described experiments and discusses advantages, limitations and future work regarding candidate prediction, the transfer of the method to similar problems as well as the comparison of the three agents in different scenarios.

Fig. 6.13 shows the synthesized root images for each method as well as the pathological image and the healthy ground truth. All agents were capable of synthesizing realistic aortic root shapes, indicating that the autoencoder performs well. Specifically, the decoder appears to be smooth over the latent space. It even tends to predict closed root shapes, eradicating imaging artifacts like holes in the root, e.g. in root 1 or root 6. The *Shape Primitive* approach seems to provide the highest variance of predicted shapes, while the *Fixed Translation* and the *Direct Estimation* agent tend to predict more similar shapes for different input samples. All agents perform well in specific cases, e.g. root 2 for *Fixed Translation*, root 4 for *Direct Estimation*, root 5 for *Shape Primitives* with $k = 2$ or root 4 for *Shape Primitives* with $k = 16$. However, each of the agents also shows poor performance on some samples.

The three approaches for defining the deformation agent were evaluated for various agent architectures in a 10-fold cross validation. Table 6.5 shows the accuracy of each method for the best agent and parameter choices, respectively. For *Direct Estimation* and *Shape Primitives*, the worst performing hyperparameters are listed in addition. As the *Fixed Translation* approach does not include any hyperparameters, only one result is given. The highest Dice coefficient was reached by the *Shape Primitives* approach with $k = 2$. This might be due to the symbiosis of a flexible prediction model and the implicit constraint to realistic shapes. Apart from that, the small number of shape primitives leads to a strong impact of the expert decision on the result as the predicted distribution spans a big area of the latent space with such a small number of shape primitives. Hence, the expert can choose from a wide range of different shapes. Assuming the expert to choose optimally, this might be a decisive advantage of this agent. In contrast, the *Shape Primitives* agent with $k = 16$ did not reach the same accuracy, but still outperformed *Fixed Translation* and *Direct Estimation* regarding the Dice coefficient. Even though realistic images could be synthesized and performance differences between the methods are observable, the overall quantitative performance is not optimal. This is highly related to the very simplistic segmentation of the aortic root in the synthesized images. A deeper

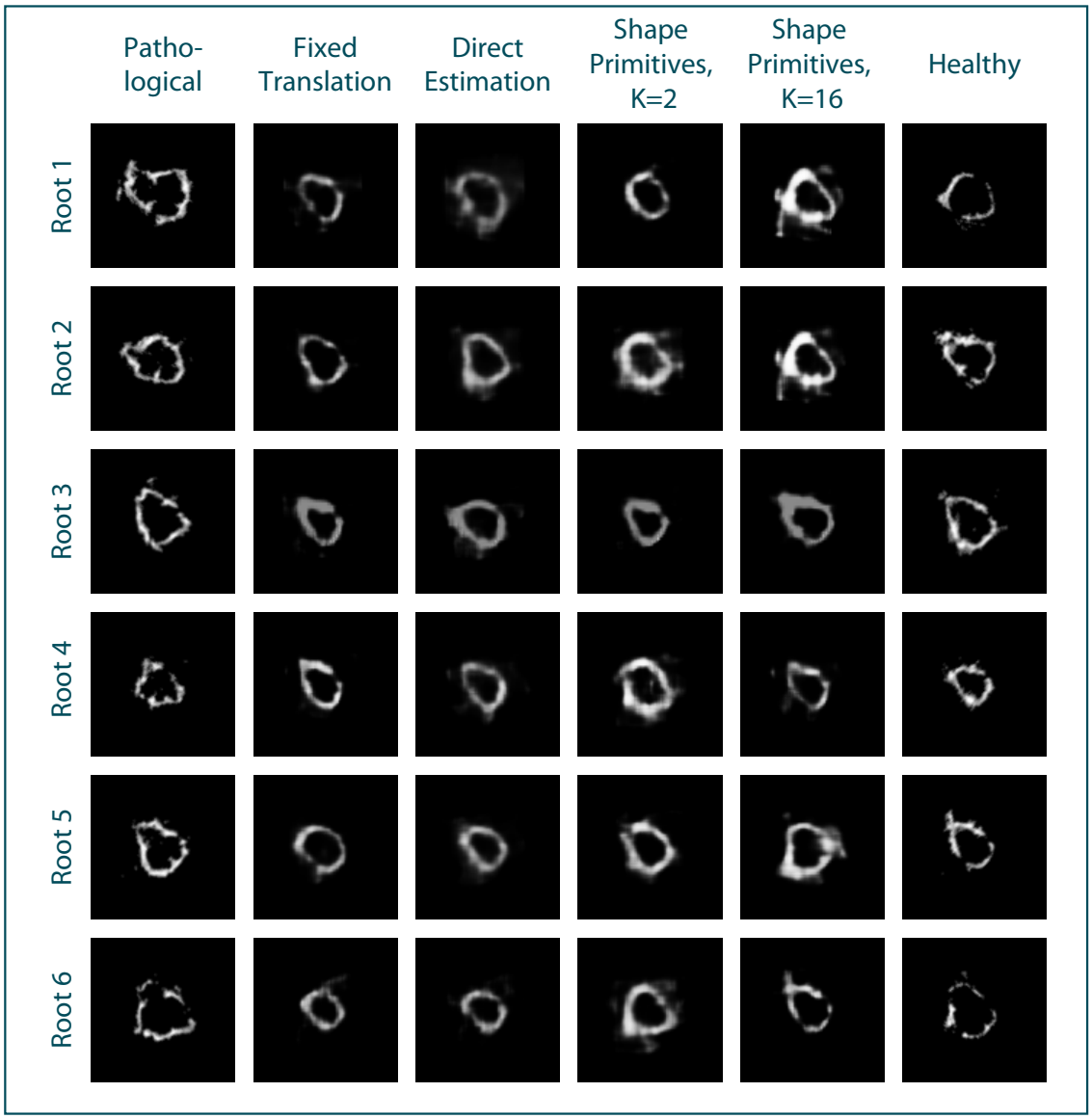

Figure 6.13: Example images synthesized using the different methods in comparison with the pathologic image (left column) and the healthy ground truth (right column). Each row exemplarily shows the predictions for a different aortic root from the data set, respectively. The agent hyperparameters for the methods were $A_{1,50}$ for *Direct Estimation*, $A_{1,100}$ for *Shape Primitives* with $k = 2$ and $A_{1,50}$ for *Shape Primitives* with $k = 16$.

Table 6.5: Quantitative Results of each agent as the mean over all crossvalidation folds. The results are given as the Dice coefficient between the segmented prediction and the segmented ground truth. For the *Direct Estimation* and *Shape Primitive* approaches, the optimal (above line) as well as the worst performing (below line) parameter combinations are given. In addition to the *Shape Primitive* method with $K = 2$, which provided the highest accuracy, $K = 16$ is also listed as an example with a higher number of shape primitives. For comparison, the Jaccard similarity is given additionally for each method.

Method	**Dice Coefficient**	**Jaccard Similarity**
Fixed Translation	0.48 ± 0.05	0.32 ± 0.03
Direct Estimation, $\mathcal{A}_{1,50}$	0.49 ± 0.14	0.32 ± 0.08
Shape Primitives, $K = 2$, $\mathcal{A}_{1,100}$	$\mathbf{0.58 \pm 0.08}$	$\mathbf{0.41 \pm 0.04}$
Shape Primitives, $K = 16$, $\mathcal{A}_{1,50}$	0.54 ± 0.07	0.37 ± 0.04
Direct Estimation, $\mathcal{A}_{3,100}$	0.46 ± 0.15	0.30 ± 0.08
Shape Primitives, $K = 14$, $\mathcal{A}_{5,100}$	0.42 ± 0.10	0.27 ± 0.05

quantitative analysis using advanced segmentation methods might provide more insights or even higher Dice scores.

One important observation is the performance gap between the *Direct Estimation* and the *Shape Primitives* approach, highlighting the implicit advantages of our probabilistic mapping. As the shape primitives' locations are fixed in the latent space and optimized to fit the healthy samples, the predicted distributions will always lie within the area of the observed healthy shapes, implicitly avoiding out-of-bound-predictions. This robustness is also observable in the lower standard deviation of the error. As the standard deviation is also lower for the *Fixed Translation* approach, the higher deviation of *Direct Estimation* is most probable due to the high flexibility of its prediction model.

For the *Direct Estimation* and the *Shape Primitives* approach, different combinations of agent's hyperparameters were evaluated. Fig. 6.14 shows the performance of all agents regarding their number of layers, their number of neurons per layer and, for the *Shape Primitives* approach, their number of shape primitives, respectively. In general, shallow agents performed better than deeper ones and a smaller number of neurons per layer provided higher Dice scores. This indicates that the complexity of the mapping is limited, highlighting the performance of the representation learning as well as the usefulness of the concept of manipulating a point in latent space for pseudohealthy synthesis. Regarding the number of shape primitives, the best performance was achieved with $K = 2$. The Dice score first decreases with an increasing K to a minimum at $K = 8$, but increases

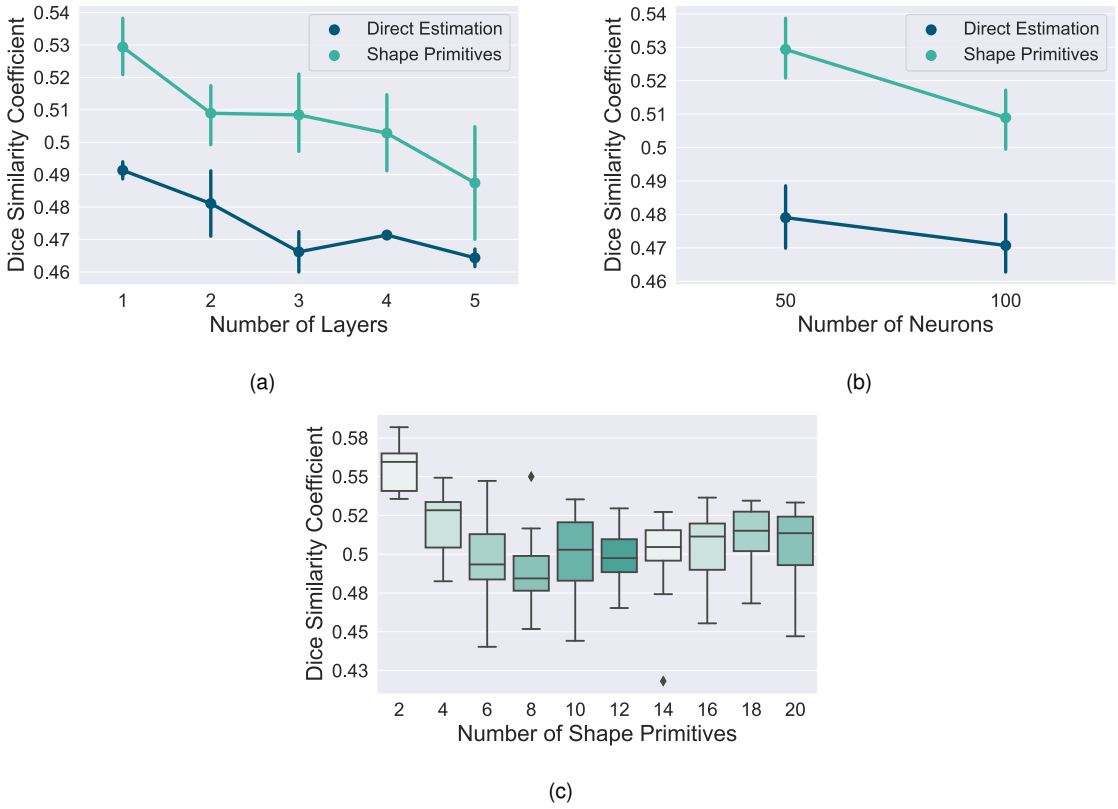

Figure 6.14: Results of the hyperparameter analysis. The results are given for all agents that share a specific number of layers (a) and a specific number of neurons per layer (b) for *Direct Estimation* and *Shape Primitives* approach, respectively. For the *Shape Primitives* approach, the influence of the number of shape primitives is given as a boxplot (c).

for higher K. This effect is most probable due to the assumption that the human expert chooses optimally. For small values of K, the shape primitives span across large areas of the latent space and hence, the candidate samples cover a big variance. As a result, the human expert has a strong impact on the prediction performance. This effect vanishes with an increasing K. However, the increase of the Dice coefficient for higher values of K indicates that the model's prediction works well even though the human expert can only choose from a set of candidate solutions with a smaller variance.

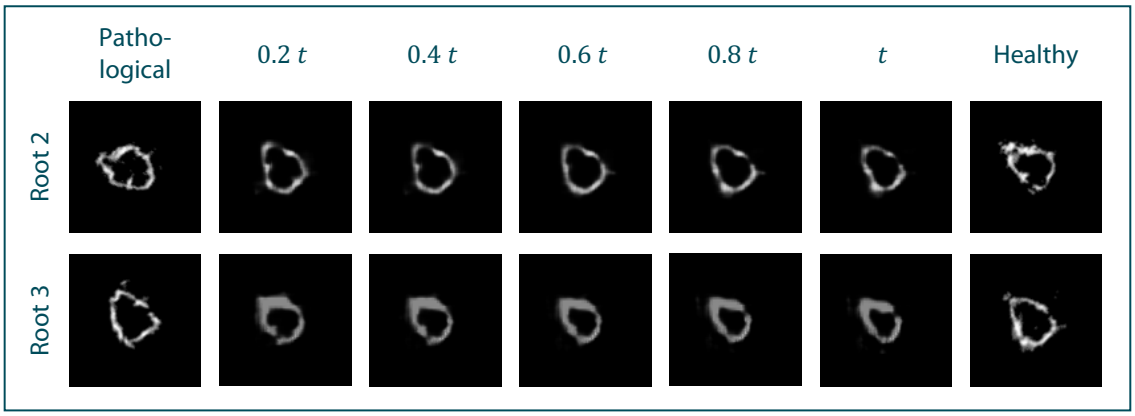

Figure 6.15: Candidate predictions for the *Fixed Translation* approach. By not adding the full translation vector t to the pathologic sample but only a specific portion of it, different steps along the trajectory can be synthesized from which the surgeon can choose the optimum. Here, four intermediate steps as well as the full translation vector (columns) are given exemplarily side by side with the pathological (left column) and healthy (right column) image for two aortic roots from the data set (rows).

Candidate Prediction The possibility to predict candidate solutions for a human expert to choose from is an important advantage of the *Shape Primitives* approach. Fig. 6.16 shows examples of predicted candidates. Through this procedure, the surgical planning tool takes the step from a decision system to a decision support system, integrating the surgeon into the loop. This might increase the acceptance of the method and could also simplify translation into the clinic by reducing legal and ethical implications.

While the candidate prediction in the *Shape Primitive* approach is done by sampling from the predicted distribution, it is possible to integrate candidate prediction in the *Fixed Translation* approach as well. Thus, instead of adding the full translation vector to the encoded pathological sample, a specific portion of it can be added and the corresponding image can be synthesized using the decoder p. By performing this with equidistant steps along the latent space trajectory given by the translation vector, the whole process of deforming the root can be assessed in image space. Therefore, the current "degree of healthiness" can be intuitively assessed by a human expert. Fig. 6.15 shows examples of this procedure. Hence, the human expert can decide on the "level of deformation" applied to the root by choosing the optimal step on the trajectory. While this is a completely different candidate selection procedure than for the *Shape Primitives*, both methods are capable of including human experts into the final decision while supporting them with data-driven deformation modeling. It should be noted that this interaction step might also

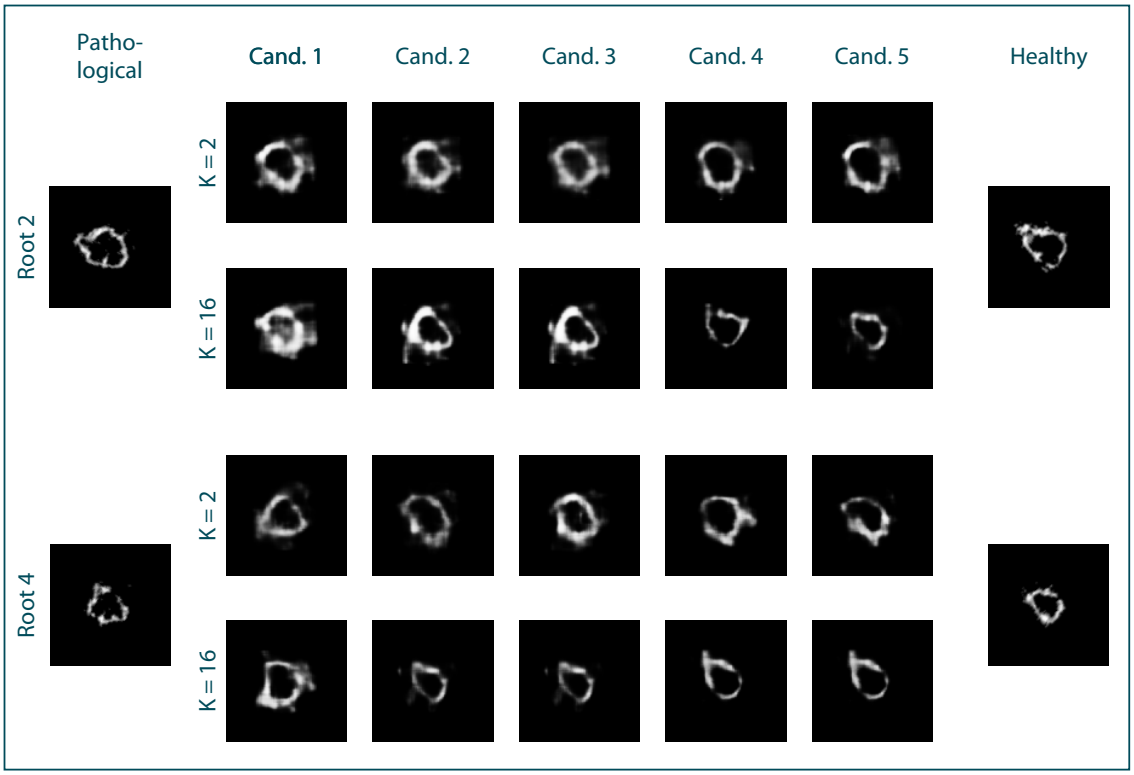

Figure 6.16: Candidate Predictions for the *Shape Primitives* approach. The agent does not map to a point but to a distribution in latent space. From this distribution, candidate solutions can be sampled. The figure shows candidate predictions for two roots (rows) for two different values of K, respectively, alongside with the pathological image (left column) and the healthy ground truth (right column).

introduce human errors as the surgeon might not always choose optimally. A further study of the surgeons' decision quality would be an interesting part of future work.

Transfer to Similar Problems It is obvious that the simulated pathology present in the data set does not perfectly resemble real dilation. However, the results show that the proposed method is capable of removing this pathological dilation in the images without explicitly modeling this pathology. The knowledge about the pathological deformation is extracted by the autoencoder on a completely data-driven basis. Given realistic training data, it can be assumed that the method is also capable of performing well in a clinical setting, i.e. dealing with real pathological dilation. Additionally, the proposed concept is not limited to planning valve-sparing aortic root reconstruction but might also present a

Table 6.6: Comparison of the three deformation agent concepts regarding mapping complexity, the implicit constraints, the possibility to predict multiple candidate solutions for human-in-the-loop applications and the capability of training the agent in an unpaired data scenario.

Agent Type	Flexible Mapping	Implicitly Constrained	Candidate Prediction	Unpaired Data
Fixed Translation	✗	✓	✓	✓
Direct Estimation	✓	✗	✗	✗
Shape Primitives	✓	✓	✓	✗

promising approach for shaping prostheses for other organs. Further studies should incorporate an in-depth analysis of the hyperparameters for a wider range of applications. Considering organ shapes, 3D point clouds could be an interesting alternative to the 2D images used in this study. The proposed method could be extended to work directly on point clouds using geometric deep learning methods [21]. As concepts for geometric autoencoders have been proposed during the last years [135], the deformation in latent space as well as the three developed agents can be directly transferred to 3D point clouds.

Comparison of Agents The three different agents proposed in this section have different features and advantages, providing options for different scenarios. Table 6.6 shows a summary of the pros and cons of each approach. Due to the MLP agent, *Direct Estimation* and *Shape Primitives* provide a higher flexibility.

However, this results in an unconstrained prediction for *Direct Estimation* while both of the other methods are implicitly constrained to predict a realistic, healthy shape. A substantial advantage of the *Shape Primitives* and *Fixed Translation* approaches is the capability of predicting candidate solutions to integrate the human expert into the decision loop. Another important feature is the capability to deal with unpaired data. This scenario is common in clinical data sets where no ground truth is known. In this case, the *Fixed Translation* approach provides a robust method to learn a prediction without the need for paired data as the translation vector can be derived without a healthy ground truth as long as some unrelated healthy and pathological data is available. On the other hand, it might be possible to use long-term screening studies to observe organ shapes of one patient in their native as well as a pathological shape. One example is the SHIP study [75] where long-term follow-up cardiac MRI scans are available for numerous patients, potentially including patients with specific pathologies. Hence, it might also be possible

that a paired data set becomes available and that the more flexible and more accurate *Shape Primitives* method proves to be a better choice.

In general, each of the three approaches has their characteristic strengths and weaknesses. However, the results indicate that if paired data is available, the *Shape Primitives* approach outperforms the *Direct Estimation*, making it the method of choice. In an unpaired data setting, *Fixed Translation* presents a good alternative. Both methods feature a candidate prediction and an implicit consistency constraint. Hence, it can be recommended to use either *Fixed Translation* or *Shape Primitives*, depending on the available data and the desired kind of predicted candidates.

Conclusion In this section, a novel approach for pseudohealthy synthesis in the context of personalized cardiovascular prosthesis shaping was presented. The problem was formulated as a translation in the latent space of an autoencoder and three different approaches were developed to learn this in a completely data-driven manner. These approaches were evaluated on a data set with a given ground truth, including an analysis of the hyperparameters. Furthermore, a comprehensive comparison of the different methods was provided as well as special features for their application, e.g. the prediction of candidate solutions for human-in-the-loop decision making or dealing with unpaired data. This study presents a new approach to personalized prosthesis shaping and, by introducing the concept of pseudohealthy synthesis via manipulation in latent space, opens up a new family of algorithms within the field of computer-assisted interventions.

6.3.2 Pseudohealthy Synthesis via Conditioned Representations

Even though the concept of performing pseudohealthy synthesis using agents for latent space manipulation gives plenty of opportunities for controlling the flexibility of the synthesis, designing such an agent requires a lot of engineering and domain knowledge. As a comparison, an end-to-end model that learns the complete pseudohealthy synthesis pipeline implicitly would allow for easier translation and application. In addition, training an agent in latent space in a supervised manner diminishes the possibilty of training the model on unpaired data. One challenge in designing such an end-to-end method is to disentangle the influence of the pathology on the organ shape and the individual anatomical appearance that should be preserved.

In this section, an alternative approach for pseudohealthy synthesis of the aortic root that does not require direct interaction with the latent space is presented. By including a condition on the image class to produce, it is possible to direct the autoencoder to synthesize reasonable images with specific attributes. Thus, the concept of Conditional Variational Autoencoders (CVAE) for planning valve-sparing aortic root reconstruction is proposed. The key idea is to condition the encoder as well as the decoder such that the predicted latent space representation or image is pathological or healthy, respectively. For pseudohealthy synthesis, a pathological image can be propagated through the autoencoder while conditioning the decoder to predict a healthy image. As training is performed on individual images with their corresponding label information, it is possible to train the CVAE on unpaired data, i.e. a data set of healthy and pathological images without correspondences.

Data Set As in the previous experiments, the data set $\tilde{D}_M$ is used. Thus, to keep the end-to-end approach comparable to the previously presented latent space manipulation method, only the 2D commissure slice is used from each volumetric image.

Model Architecture In this study, the Conditional Variational Autoencoder has a similar architecture to the VAE used in Sec. 6.3.1. However, there is no classification layer in the latent space. Hence, the encoder consists of five convolutional layers (kernel size 3×3) followed by average pooling (kernel size 2×2). The decoder is obtained by mirroring the architecture of the encoder. Furthermore, two new layers that represent the conditions c_q and c_p are imposed on the inputs of the encoder q and the decoder p [134]. Both conditions are defined in a one-hot encoding, where $(1, 0)$ represents a healthy and $(0, 1)$

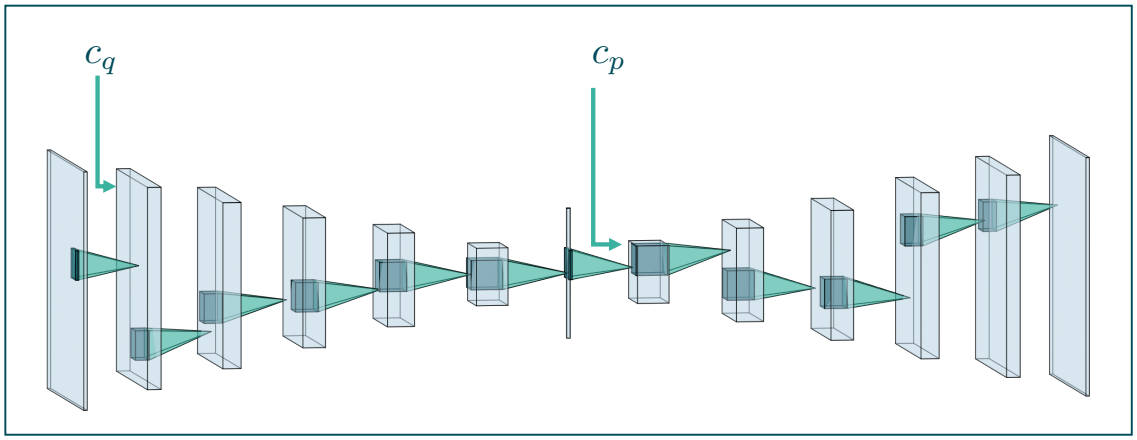

Figure 6.17: Schematic illustration of the proposed architecture conditioned by c_q and c_p. Note that the latent space is sketched as a single vector to improve clarity. In the implementation, the latent space is a variational model described by μ and σ utilizing reparameterization (see Sec. 3.4 for details).

a pathological state. Hence, the probability distributions learned by the model are now conditioned by c_q and c_p and the loss function $\mathfrak{L}_{\text{cvae}}$ becomes

$$\mathfrak{L}_{\text{cvae}} = \mathbb{E}(\log p(X|z, c_p)) - D_{\text{kl}}(q(z|X, c_q)||p(z|c_p)), \tag{6.12}$$

where X is the input image, z is its latent representation and D_{kl} is the Kullback-Leibler divergence. This conditioning pushes the CVAE towards disentangling individual anatomical traits from the pathological deformation of the shape. The information on whether the current sample is healthy or pathological is directly given in the decoder as a condition. In addition, the latent space representation z is also conditioned to encode two different states. Therefore, the representation becomes more independent from the pathological state and focuses on representing individual anatomical traits. Figure 6.17 illustrates the architecture of the proposed model.

ReLU activation was applied throughout the network except in the last layer, where *Sigmoid* activation was used. Similar to the architecture presented in Sec. 6.3.1, the dimension of the latent space was set to 12. For training, the *adam* optimizer was used.

Training and Inference All images in the data set were assigned their corresponding label, i.e. *healthy* or *pathological*. During training on one image of the training batch, the encoder as well as the decoder were conditioned by the true label of the image. Thus, the CVAE learned the desired conditioned distribution in the latent space as well as the conditioned generative model.

For inference, a pathological image is propagated through the model while conditioning the decoder to generate a healthy shape. As two different conditions c_q and c_p are defined for the encoder and the decoder, respectively, there are two approaches for the inference step: *Homogeneous conditioning* and *heterogeneous conditioning*.

In the *homogeneous conditioning* approach, the encoder as well as the decoder are conditioned the same. The conditions are set to

$$c_q = (1, 0) \tag{6.13}$$
$$c_p = (1, 0). \tag{6.14}$$

Thus, the image is already encoded as if it would be healthy. This concept is related to the idea of interpreting pseudohealthy synthesis as a denoising task, as it was e.g. suggested by [127], [9] or [28]. The basic assumption is that by learning the manifold of healthy anatomies, the model is capable of identifying the pathological deformation as an anomaly that is removed.

In contrast, the conditions for the encoder and decoder differ in the *heterogeneous conditioning* approach. While the decoder is conditioned to generate a healthy image, the encoder is conditioned to encode the input image as pathological:

$$c_q = (0, 1) \tag{6.15}$$
$$c_p = (1, 0). \tag{6.16}$$

This approach follows the assumption that the latent features identified by the CVAE carry sufficient information on the individual anatomy, independently of the pathology, such that simply flipping the condition provides an adequate approximation of the individual healthy state.

Both approaches were evaluated using a 10-fold Monte-Carlo crossvalidation ($80\,\%$ training, $20\,\%$ test). It is important to note that the data split was performed on the aortic roots, regardless of their state. Hence, the healthy and the pathological image of the same root will either both be in the training or both in the test set, providing a fair and reliable evaluation. Within each fold, the CVAE was trained on all images in the training

Table 6.7: Quantitative Results of both approaches as the average over all crossvalidation folds. The results are given as the Dice coefficient and the Jaccard similarity between the segmented prediction and the segmented ground truth.

Method	Dice Coefficient	Jaccard Similarity
Homogeneous Conditioning	0.45 ± 0.04	0.29 ± 0.02
Heterogeneous Conditioning	0.47 ± 0.05	0.31 ± 0.03

set using the true labels for c_q and c_p for 350 epochs and a batch size of 12. Then, the pathological images from the test set were propagated through the model using the conditions given in 6.14 and 6.16, respectively. The synthesized image was segmented using a threshold ($t = 80$ in a grayscale range $[0, 255]$) and compared to the manual segmentation of the corresponding healthy ground truth image using the Dice score.

Results and Discussion Table 6.7 shows the average Dice score over all folds. Both methods provide a Dice score below 0.5, with the *heterogeneous conditioning* performing slightly better. Qualitative results are shown in Fig. 6.18.

Overall, the CVAE synthesizes realistic and adequate aortic root images for both conditioning methods. Even though the model is trained in an unpaired way and no handcrafted agent design is needed, the synthesis provides consistent results. In comparison, the *heterogeneous conditioning* approach qualitatively approximates the individual healthy shape slightly better than the *homogeneous conditioning* approach, e.g. for root 3 or root 4. This is supported by the quantitative results, even though the difference between both average Dice scores is small.

In comparison to the concept of agents for latent space manipulation presented in Sec. 6.3.1, all three agent concepts outperform the end-to-end approach using a CVAE, regardless of the conditioning method. In addition, the CVAE does not provide a reasonable method for candidate prediction. However, the CVAE can be trained without further modeling steps and is capable of dealing with unpaired data. Hence, it might provide a promising alternative in settings where little to no domain knowledge is present and no paired data is available.

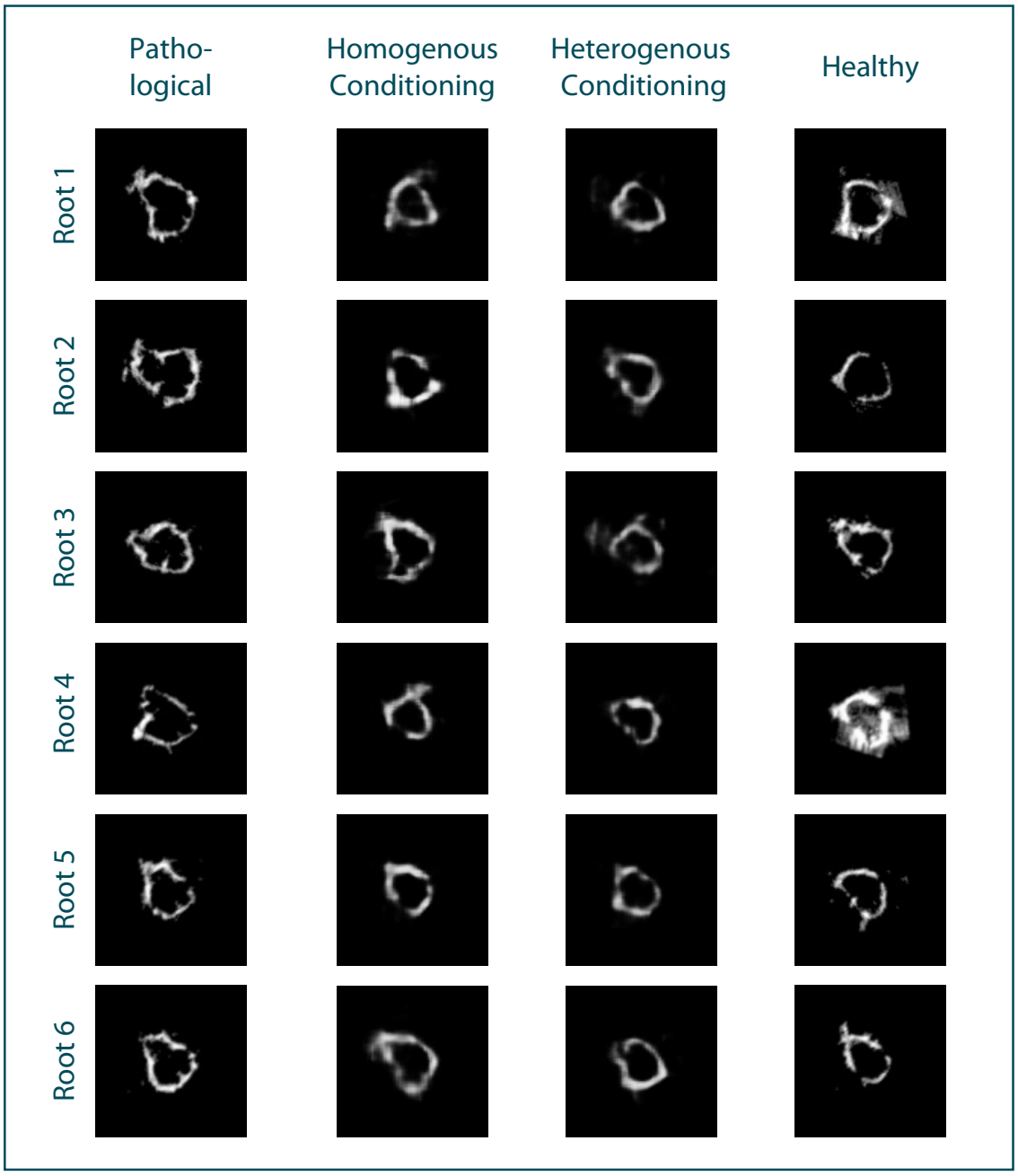

Figure 6.18: Qualitative results of an end-to-end CVAE using *homogeneous conditioning* and *heterogeneous conditioning* as well as the pathological input image and the healthy ground truth.

6.4 Conclusion

The aim of this chapter was to answer the research question

Q2: Is it possible to synthesize personalized healthy shapes?

Different frameworks were developed to achieve pseudohealthy synthesis in the case of a structural or a morphological pathology. Within both frameworks, multiple approaches were developed, optimized, evaluated and compared on the collected ex-vivo porcine data sets. Based on the results of both studies, the research question *Q2* can be answered with yes.

In both cases, the developed methods were capable of synthesizing consistently realistic leaflet or root shapes and it was possible to approximate the original, healthy shape only based on surrogate information. Even though the data sets were comparably small, the generative models performed robustly and no out-of-bag predictions were observed. While the developed frameworks differ for the two pathology cases, both of them have in common that they are based on representation learning using autoencoders. This chapter shows the power of representation learning and that a well-identified latent representation opens up possibilities for numerous applications. In case of a morphological pathology, it was possible to include implicit constraints on predicting realistic outputs as well as candidate prediction methods. Despite being a novel approach in a new field, the translational value for use in clinical applications is high. Additionally, these advantages of the latent space manipulation approach do not come at the cost of a lower accuracy since the method outperformed a classical end-to-end learning approach.

One important point to note is that the target accuracy necessary to ensure sufficient prosthesis functionality or an advantage of the personalization compared to state-of-the-art prosthetics remains unclear. Consequently, the interpretation of the quantitative results cannot be brought to a final stage based on the results presented in this chapter. However, this work presents the first approach for personalized cardiovascular prosthesis planning and hence lays the basis for personalized prosthesis manufacturing and testing in physical setups. Comparing different biomechanical parameters of the personalized prostheses and native aortic roots as well as benchmarking them against state-of-the-art prostheses might provide new insights in the impact of personalized prostheses shaping. The methods developed and presented in this chapter make such experiments possible for the first time.

It should also be considered that the data sets collected and used in the scope of this work are limited in their realism. For the structural pathology case, the ultrasound

volume images only show the extracted aortic root without surrounding tissue and in a static state. However, there are validated methods for detecting phases of the cardiac cycle [46] as well as segmenting the aortic root [26], both based on transesophageal ultrasound images. By utilizing these methods as a preprocessing, a realistic human data set can be transferred to a similar state as the data set used in this work. However, one challenge arises with collecting leaflet images from humans in an ex-vivo setting as proposed in Sec. 4.2. As ex-vivo studies can only be performed with organ donors, the individuals are typically elderly and the leaflets show a certain degree of calcification. This might harm the spreading of the leaflets as well as the imaging procedure. Hence, a data set should be collected with younger organ donors, which might interfere with the need for transplants. An alternative would be to collect the leaflet shape images not ex-vivo but in-vivo. Possible imaging modalities could be intravascular optical coherence tomography (OCT) or high-resolution CT with manual postprocessing [43]. Even though the leaflets would then be in their native state, i.e. curved and pressurized, it might be easier to collect a meaningful data set of a sufficient size. The methods developed and presented in this work can still be applied. Regarding the morphological pathology it is obvious that the manual dilation of the roots is only a rough approximation of a realistic pathological deformation. However, the deformation is not modeled within the method but learned by the agent in a completely data-driven way. As the agents developed in this chapter were capable of learning the deformation caused by manual dilation, it can be assumed that they are also capable of dealing with realistic deformation caused by pathology.

The methods developed in this chapter present the first approaches for personalized cardiovascular prosthesis shaping using pseudohealthy synthesis. They lay the groundwork for personalized shaping of aortic valve and aortic root prostheses. Furthermore, they can be applied to personalized cardiac and cardiovascular modeling and simulation where individual shaping plays a critical role as well. Furthermore, none of the methods is restricted to the application in the cardiovascular domain. As all representations and mappings are learned data-drivenly, the proposed frameworks present a promising approach for any kind of anatomical shape modeling based on surrogate information and hence, might be of high interest within the wide field of computational anatomy.

7 Identifying Optimal Sets of General Prostheses[1]

The aim of this chapter is to answer the research question

Q3: Is it possible to identify optimal sets of general prostheses?

In the previous chapter, it could be shown that personalized shaping of cardiovascular prostheses using pseudohealthy synthesis is possible for both types of pathologies. The presented methods allow for manufacturing prostheses specifically tailored to the patient's individual anatomy. This individual tailoring comes at high costs, logistical challenges and regulatory issues as the regulation of a product whose shape varies from patient to patient remains an unclear procedure. In addition, there is no data available on the impact of personalized prostheses on the patients' outcome. Therefore, the accuracy of resembling the individual anatomy that is necessary to ensure an increased functionality and patients' outcome is unknown and it is possible that providing a less accurate approximation of the individual anatomy is already sufficient. This leads to a new perspective on personalized prostheses in between full personalization and a "one-shape-fits-all" approach as given by state-of-the-art prostheses. In this chapter, the usage of a set of typical prosthesis shapes is proposed.

The shape of each anatomical structure shows a distribution over all individuals. The basic assumption of the proposed approach is that there is a set of typical shapes that

[1] Parts of this section were published in:

[57] J. Hagenah, K. Kühl, M. Scharfschwerdt and F. Ernst. Cluster Analysis in Latent Space: Identifying Personalized Aortic Valve Prosthesis Shapes using Deep Representations. Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning, in Proceedings of Machine Learning Research 102:236-249 (2019)

[54] J. Hagenah and F. Ernst. Discrete Pseudohealthy Synthesis: Aortic Root Shape Typification and Type Classification with Pathological Prior. Proceedings of The 4th International Conference on Medical Imaging with Deep Learning, in Proceedings of Machine Learning Research 143:252-267 (2021)

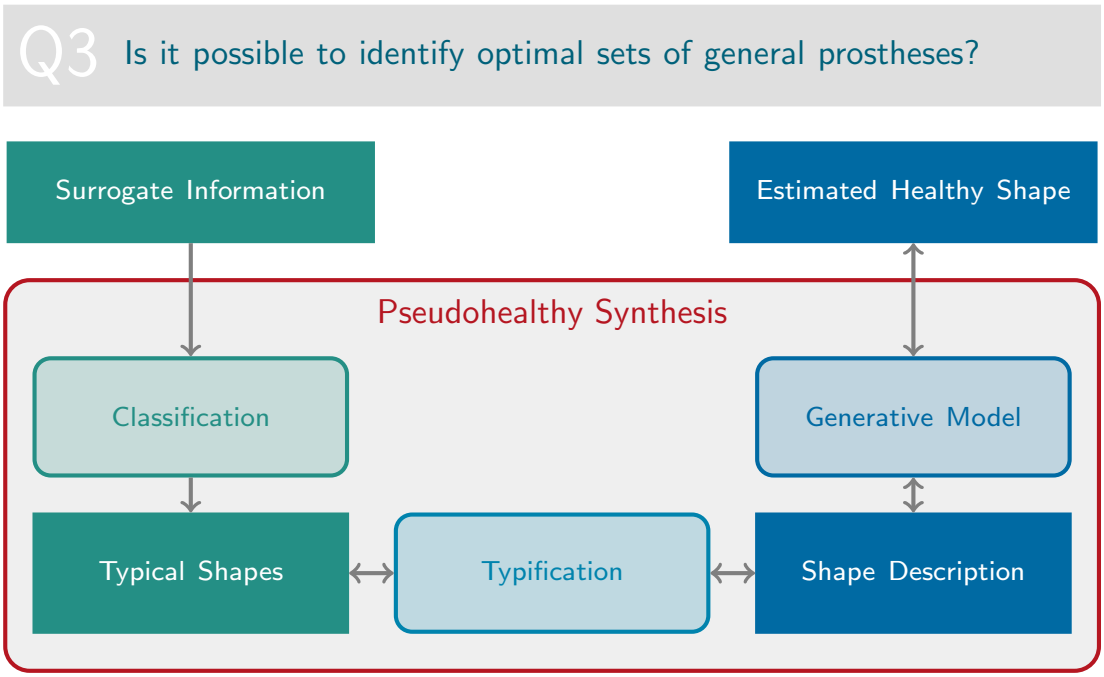

Figure 7.1: General approach to investigate research question *Q3*. In contrast to the previous chapter, typical shapes are identified in an automatic typification procedure. These typical shapes serve as prosthesis shape types out of which the individually optimal one can be classified based on the surrogate information.

represents this whole distribution adequately. Regarding personalized prostheses shaping, this means that each native shape can be sufficiently approximated by one of these typical shapes. Even though a finite set of shape types is used, the prostheses will mimic the patient's original anatomy more closely than state-of-the-art prostheses where only the diameter is personalized. This leads to a promising compromise between superior personalization and the possibilty for serial production. While the pseudohealthy synthesis methods in Chap. 6 featured continuous prediction, i.e. regression, the choice of the optimal prosthesis shape becomes a classification of the optimal shape type, leading to a discrete prediction.

Hence, regarding a clinical application scenario, prostheses of all typical shape types could be manufactured in mass production and with regulations. In the clinic, all types can be stored in a sufficient number and when a patient needs an implant, the optimal shape type is classified and the prosthesis can be inserted immediately. In comparison to the continuous approach presented in Chap. 6, the translation of this procedure to clinical application is way easier and more realistic.

In this chapter, a methodological framework for identifying a set of typical shape, referred to as typification, as well as type classification is developed and evaluated for both pathological cases, i.e. structural pathology and morphological pathology. At first, a method for automatically identifying a set of typical shapes, i.e. typification, is developed. Then, based on the typification results, discrete prediction models for classifying the optimal shape for an individual patient type based on surrogate information are presented. Due to its fundamental principles, the whole framework is also referred to as discrete pseudohealthy synthesis. Fig. 7.1 visualized this approach to investigate research question *Q3*.

7.1 Related Work

Most approaches for pseudohealthy synthesis use a form of representation learning, either utilizing autoencoders [127], [9], [141], [28] or by training Generative Adversarial Networks (GANs) [2], [143], [136], [146]. All these approaches have in common that they solve a regression problem, i.e. synthesizing the optimal image out of a continuous space. In contrast, in this chapter, it is proposed to only allow a finite set of discrete shapes, leading to a classification problem. This chapter presented the first studies on discrete pseudohealthy synthesis in general as well as the first approaches to identify typical aortic valve and root shapes using unsupervised clustering.

Clustering anatomical shapes has been used e.g. in the scope of image registration [149] or to identify and distinguish healthy and pathological patterns [22]. These approaches have in common that the clustering is performed in a manually identified feature space. In contrast, the methods presented in this chapter rely on clustering in the latent space of an autoencoder to identify typical shapes.

7.2 Structural Pathology

Even though imaging of the leaflets is a remaining challenge due to the thin structure and the high movement, it could be shown that an estimation of the individual planar leaflet shapes is possible just based on an ultrasound image of the aortic root, utilizing Support Vector Regression (see Sec. 5.2). Furthermore, in Sec. 6.2, a framework for predicting and synthesizing the full shape of the individual leaflet based on the shape of the surrounding tissue was presented. Hence, a personalization of the leaflet shapes in a prosthesis is possible in general. However, a completely individual manufacturing of personalized prostheses is unrealistic in the near future due to economical, logistical and regulatory issues. An alternative approach for a trade-off between these issues and a higher patient outcome would be to offer a number of specific prostheses types that approximate the realistic distribution of valve shapes in a data-driven manner. Then, each patient could be treated with the prosthesis type that matches his or her individual anatomy and physiology best. This leads to a classification problem, which should be easier to solve than the regression problem from Sec. 6.2 and presents a cost-efficient and hence realistic way of aortic valve prostheses personalization.

In this chapter, a method is presented to perform a cluster analysis in aortic valve leaflet shapes to identify these valve types. Thus, each cluster center is interpreted as a typical shape and therefore represents a prosthesis shape. On challenge is that clustering in image space where each pixel presents one dimension is not suitable due to an effect called the curse of dimensionality [77]. In high dimensional spaces, distance metrics are influenced by the high number of dimensions and variance in the data cannot be incorporated in a sufficient way. Additionally, even small translations or rotations of the valve might lead to big distances in image space, whereas the shape of the valve stays the same. Hence, it is proposed to perform the cluster analysis in a latent space description of the valves using representation learning. Utilizing generative modeling, the latent cluster centers, also referred to as prosthesis shape types or typical prosthesis shapes, can be transformed back to image space. Like this, the cluster analysis focuses on abstract, highly descriptive features instead of pixel-wise grayscale values, while the evaluation of the identified prosthesis shapes can be done using intuitive metrics in image space. After this automatic typification, a type classification pipeline is developed to predict the optimal valve prosthesis shape type based on the surrounding tissue, i.e. an ultrasound image of the aortic root. Fig. 7.2 shows the approaches to investigate research question *Q3* in the case of a structural pathology.

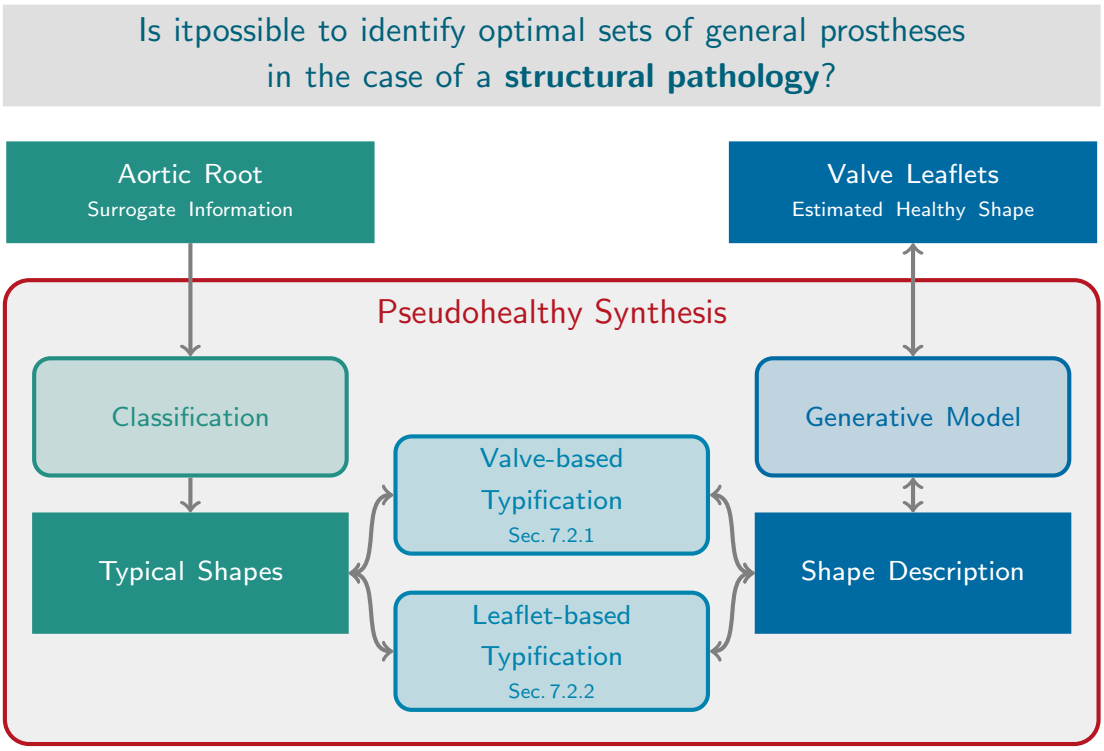

Figure 7.2: Approaches to investigate the research question *Q3* for the case of a structural pathology.

As described in Chap. 4, the leaflet images are distinct, i.e. each image shows one single leaflet. Each valve shape is given by three leaflet images, respectively. Therefore, typical valve shapes have to be identified on the basis of these three leaflet images, which can be either done *valve-based* or *leaflet-based*. Both concepts have in common that at first, a representation of one single leaflet image is learned utilizing an autoencoder. In the *valve-based* approach, the latent space representations of all three leaflets of one valve, i.e. the right-coronary, left-coronary and non-coronary one, are concatenated to achieve a latent representation of the whole valve. Then, the clustering is performed in the space of these concatenated latent representations, i.e. in a valve shape space. Hence, for pseudohealthy synthesis, such a valve representation is predicted and the shape representations of the individual leaflets can be extracted from it.

In contrast, in the *leaflet-based* approach, the clustering is directly performed in the leaflet's representation. Thus, the identified shape types are leaflet shapes, and one valve consists of three individual shape types. Hence, for pseudohealthy synthesis, three shape types are predicted for each valve that correspond to the shape of the right-coronary, left-

coronary and non-coronary leaflet, respectively. In the following, the typification as well as the type classification are presented for both approaches, the *valve-based* and the *leaflet-based* approach.

Contribution

There are several contributions of this section. First, the valve-based typification study described in Sec. 7.2.1 presents the first approach to identify typical organ shapes by performing a cluster analysis in a latent space description. To the best of the author's knowledge, it also presents the first study to identify a set of typical prosthesis shapes using clustering in general. Second, the type classification approaches presented in Sec. 7.2.1 and Sec. 7.2.2 are the first methods to solve the formulated domain gap problem in a discrete way. Hence, they present an important extension of the framework proposed in Sec. 6.2.1. Furthermore, this section presents the first study on identifying and predicting a set of aortic valve prostheses shapes and thus presents a big step towards clinical applicability of personalized aortic valve implants.

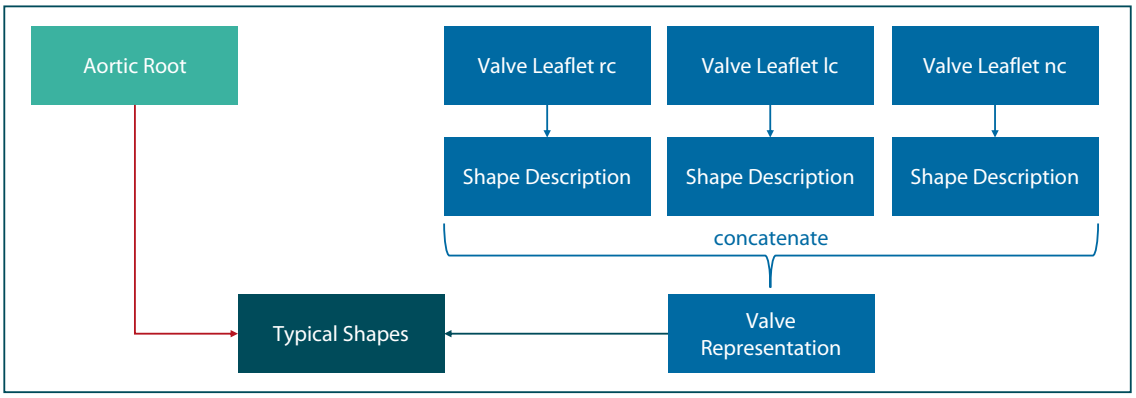

Figure 7.3: Workflow of the valve-based typification approach. The three leaflets of one valve are encoded and the three latent representation vectors are concatenated to achieve a latent description of the full valve. In this valve representation, typical shapes are identified. For a given aortic root shape, the individually optimal valve type can be classified.

7.2.1 Valve-based Typification and Type Classification

The aim of this approach is to identify typical shapes of the full valve. It is assumed that a valve's shape can be described by the shape of its three leaflets. Hence, each valve type consists of three leaflet shapes. The core idea of this approach is to introduce a valve representation that combines latent features of all three leaflets by concatenating their latent representation vectors. In this valve representation, a clustering can be performed to identify typical valve shapes. Fig. 7.3 visualizes the valve-based approach. The advantage of this approach is that the three leaflet shapes are assumed to be correlated, reflecting the natural interplay between them. Hence, it can be assumed that realistic and consistent valve shapes are identified and predicted. At first, the valve-based typification method is presented, including an evaluation, followed by describing the valve-based type classification.

Valve-based Typification

In this section, the identification of typical valve shapes is presented. The key idea is to find a latent representation of the individual leaflets and concatenate three of them to describe a full valve. In this abstract description of a valve's shape, typical shapes can be identified utilizing clustering.

Data Set In this study, two different data sets were used. The valve-based typification, i.e. the clustering of valves, was performed and evaluated on D_S^{aux} that contains leaflet images of 56 valves. The type classification experiments were performed on D_S were the corresponding ultrasound volume image of the aortic root is known. Thus, D_S^{aux} was used for representation learning and clustering and the resulting trained model was utilized to compute the ground truth classes of the valves from D_S.

Representation Learning At first, the shape of the valve, i.e. the shape of its three leaflets, has to be encoded into a latent representation. An artificial neural network was trained to find a compact yet meaningful representation of one leaflet image in an unsupervised manner. To this end, a convolutional autoencoder AE is used, which is known to be capable of encoding images with good representations in the bottleneck layer, i.e. the latent space [66]. The autoencoder is divided into two different subnetworks: the encoder network $q : I \in \mathcal{R}^{128\times64} \rightarrow z \in \mathcal{R}^{n_z}$ and the decoder network $p : z \in \mathcal{R}^{n_z} \rightarrow I \in \mathcal{R}^{128\times64}$ such that

$$AE(I) = p(q(I)) = I_{\text{reco}}, \tag{7.1}$$

where I is an image of one leaflet, n_z is the dimensionality of the latent space, i.e. the number of neurons in the bottleneck layer, and I_{reco} is the predicted reconstruction of I.

One important advantage of autoencoders is that generative modeling of images based on latent representations is possible using the decoder part. Like this, the clustering can be performed in the latent space while the evaluation of the clustering can be executed in image space by propagating the latent cluster centers through the decoder. To this end, a symmetric autoencoder with 3 convolutional layers (ConvL) with 32 filters, each followed by a maximum-pooling layer (maxPoolL), and a fully connected layer (FCL) was used for the encoder. In all layers, *ReLU*-Activation was applied. The decoder architecture was identical but mirrored. The autoencoder was trained with mean-squared-error loss using the *adam* optimizer. An illustration of the proposed architecture can be found in Fig. 7.4.

To evaluate the quality of the representation, the reconstruction accuracy of the autoencoder was analyzed by propagating the leaflet images through the encoder and generating an image of the resulting latent representation using the decoder. These predicted images were compared to their ground truth using two metrics: the Jaccard similarity d_J and the Hausdorff distance d_H. The Jaccard similarity measures the overlap of two leaflets and hence evaluates the overall shape similarity:

$$d_J(A, B) = \frac{A \cap B}{A \cup B}, \tag{7.2}$$

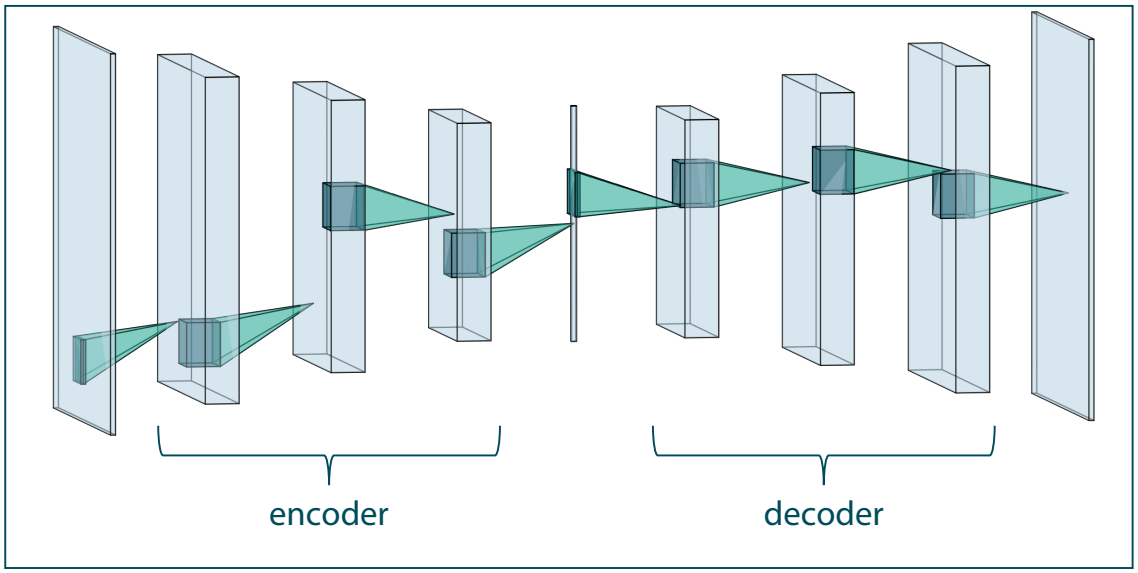

Figure 7.4: Architecture of the proposed autoencoder model.

where A and B are the sets of nonzero pixels in the leaflet images, respectively. The Hausdorff distance describes the maximum of the minimal distances of two contours, measuring the detailed accurateness of the leaflet contour in the reconstruction:

$$d_H(X,Y) = max\left(\sup_{x\in X}\inf_{y\in Y} d(x,y), \sup_{y\in Y}\inf_{x\in X} d(y,x)\right), \tag{7.3}$$

where d is the euclidian distance and X and Y are sets of contour pixels, respectively. Since the reconstruction accuracy is highly dependant on the network architecture, this analysis was also performed for a more shallow (1 ConvL followed by 1 maxPoolL, 1 FCL) and a deeper encoder architecture (5 ConvL followed by 1 maxPoolL, respectively, 1FCL) in addition. Once again, the decoder architecture was the mirrored encoder architecture. In this comparison, the number of latent dimensions n_z was set to 20.

Furthermore, the reconstruction accuracy of the autoencoder was evaluated for different choices of n_z, ranging from 2 to 50 dimensions.

However, not only the reconstruction accuracy is important to ensure a sufficient representation, but also a smooth decoder function is desired. Since the identified cluster center will barely match an observed valve geometry, the reconstructed image has to be interpolated. Smoothness of the decoder ensures that interpolations in the latent space provide realistic images, which is mandatory for the generation of valve images of the identified cluster centers. Two prove smoothness, the proposed architecture was compared to two alternative methods. First, the same autoencoder was used but data

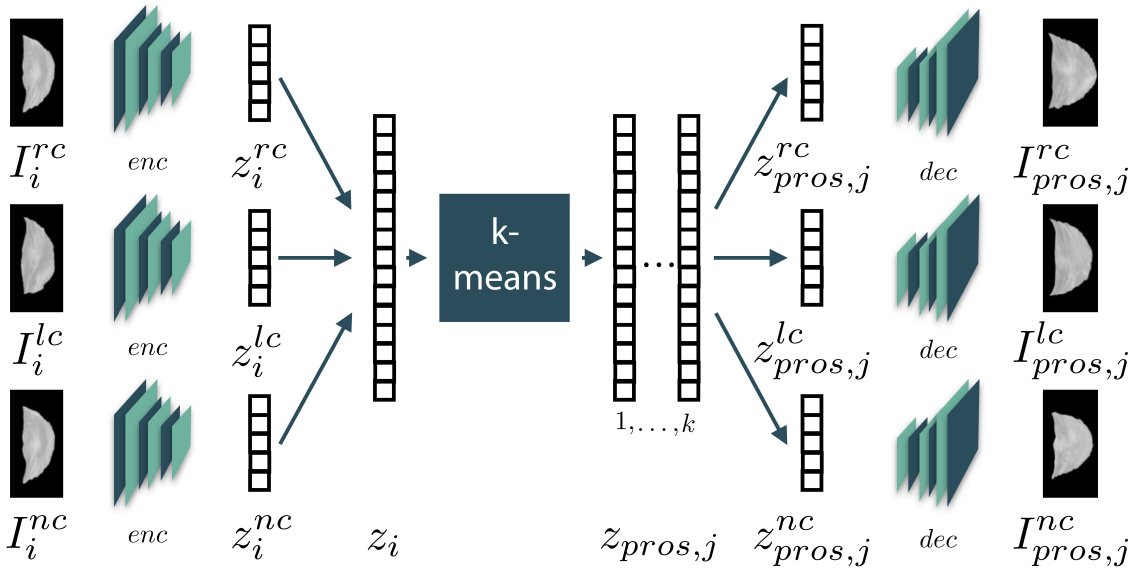

Figure 7.5: Sketch of the proposed method. The trained encoder model is used to receive the latent vector of each leaflet. These vectors are stitched together to represent a valve. In this space, the clustering is performed. Afterwards, the latent representation of the single leaflets of an identified prosthesis are extracted and the corresponding images are generated using the decoder. Adapted from [57].

augmentation (random shearing and zooming by up to $5\,\%$, translations in any direction with a maximum of 2.2 mm) was applied. Second, a variational autoencoder (VAE) of the same encoder/decoder architecture was evaluated as VAEs are optimized to identify a smooth latent space implicitly [81]. The comparison was done quantitatively regarding the reconstruction accuracy and qualitatively using linear interpolation between two points in the latent space and propagating the interpolated points through the decoder to produce leaflet images.

Clustering The aim is to perform a cluster analysis in latent space. To this end, it is assumed that the geometry of the whole valve can be described by the shape of its three leaflets. Thus, the three leaflets of one valve were propagated through the trained encoder network to get their latent representations $z_i^{rc} \in \mathcal{R}^{n_z}$, $z_i^{lc} \in \mathcal{R}^{n_z}$ and $z_i^{nc} \in \mathcal{R}^{n_z}$, where rc stands for the right-coronary, lc for the left-coronary and nc for the non-coronary leaflet, respectively, and $i = 1, \ldots, 56$ describes the current valve. By concatenating the latent vectors, a latent representation $z_i \in \mathcal{R}^{3n_z}$ of the whole valve is retrieved. In this representation, a k-means-clustering was performed, aiming at identifying k cluster centers

$z_{pros,j}, j = 1, \ldots, k$ that could serve as prosthesis types [95]. Even though a clustering in the latent space is convenient, the evaluation of the identified prosthesis types in latent space is not intuitive. Hence, the identified prostheses were transformed back to image space using the decoder. For this purpose, the latent prosthesis representation was split up into the three parts $z^{rc}_{pros,j} \in \mathcal{R}^{n_z}$, $z^{lc}_{pros,j} \in \mathcal{R}^{n_z}$ and $z^{nc}_{pros,j} \in \mathcal{R}^{n_z}$ corresponding to the three leaflets. These vectors were propagated through the decoder, resulting in images of the three leaflets of the prosthesis, i.e. $I^{rc}_{pros,j} \in \mathcal{R}^{128\times64}$, $I^{lc}_{pros,j} \in \mathcal{R}^{128\times64}$ and $I^{nc}_{pros,j} \in \mathcal{R}^{128\times64}$. These images could be compared to the images of all valves that correspond to this cluster using the metrics described above. Hence, the mean Jaccard similarity $\overline{d_{J,k}}$ and the mean Hausdorff distance $\overline{d_{H,k}}$ over all valves can be computed in dependency of the number of identified prosthesis types k as

$$\overline{d_{J,k}} = \frac{1}{56}\sum_{j=1}^{k}\sum_{I_c\in C_j}\frac{1}{3}\left(d_J\left(I^{rc}_c, I^{rc}_{pros,j}\right) + d_J\left(I^{lc}_c, I^{lc}_{pros,j}\right) + d_J\left(I^{nc}_c, I^{nc}_{pros,j}\right)\right) \tag{7.4}$$

$$\overline{d_{H,k}} = \frac{1}{56}\sum_{j=1}^{k}\sum_{I_c\in C_j}\frac{1}{3}\left(d_H\left(I^{rc}_c, I^{rc}_{pros,j}\right) + d_H\left(I^{lc}_c, I^{lc}_{pros,j}\right) + d_H\left(I^{nc}_c, I^{nc}_{pros,j}\right)\right), \tag{7.5}$$

where C_j is the set of all images I_c corresponding to the j-th cluster. Using these metrics a set of used prosthesis types can be evaluated regarding its capability of approximating each individual valve shape in the data set. To analyze the amount of valve prosthesis types needed to ensure a good shape approximation, the clustering was performed for different values of k, ranging from 1 to 56. Note that a clustering with $k = 1$ corresponds to the current clinical situation, where each patient is treated with the same prosthesis shape and only the diameter is tailored to the patient's anatomy. For the clustering study, the number of dimensions of the latent space n_z was set to 20.

Results of Representation Analysis The results of the experiment regarding the reconstruction accuracy of the proposed architecture as well as of alternative architectures is given in Table 7.1. Compared to a shallower and a deeper architecture, the proposed autoencoder achieves the highest Jaccard similarity and the smallest value of the Hausdorff distance, given as the mean over all leaflets. Fig. 7.6 shows the reconstruction accuracy in dependency of n_z, the number of latent dimensions. The accuracy increases at first, but saturates at a value of about 20 dimensions. Hence, additional dimensions have a neglectable benefit at the cost of higher challenges for the clustering algorithm. Therefore, 20 dimensions can be seen as a good tradeoff.

Table 7.1: Reconstruction accuracy of the proposed autoencoder (marked in bold), compared to a shallower and a deeper architecture regarding the Jaccard similarity d_J and the Hausdorff distance d_H. Additionally, the influence of data augmentation as well as the usage of a variational autoencoder is shown. The metrics are given as the mean over all leaflet images.

Architecture	$\mathbf{d_J}$	$\mathbf{d_H}$ **[mm]**
Shallower architecture	0.9272	3.17
Proposed architecture	**0.9471**	**2.60**
Deeper architecture	0.9283	3.18
Proposed AE with Data Augmentation	0.9424	2.70
VAE of same architecture	0.9423	2.74

The introduction of data augmentation as well as the use of a variational autoencoder provide comparable results to the proposed architecture (see Table 7.1). The evaluation of the smoothness of the different models was done qualitatively. While the VAE delivered slightly different shapes at the leaflet tips, no relevant differences between the proposed AE with or without data augmentation could be observed. Results are exemplarily shown in Fig. 7.7.

Results of Cluster Identification The clustering was performed for different numbers of clusters k. The capability of different numbers of prosthesis shape types to approximate the individual shapes given in the data set is shown in Fig. 7.8. For low values of k, the increase of accuracy is very steep, while it starts to flatten at a value of about 7. The resulting prostheses are exemplarily shown for $k = 3$ in Fig. 7.9 .

To critically question the whole approach of the proposed autoencoder pipeline, the clustering and evaluation routine was additionally performed regarding the image description using principal components. To this end, a Principal Component Analysis (PCA) was performed on the vectorized images and the 20 most relevant components were used as a representation. While the mean Jaccard similarity was substantially smaller compared to the autoencoder approach, outliers in the range of up to $20mm$ were observed in the mean Hausdorff distance. Detailed results of this comparison can be found in Fig. 7.10.

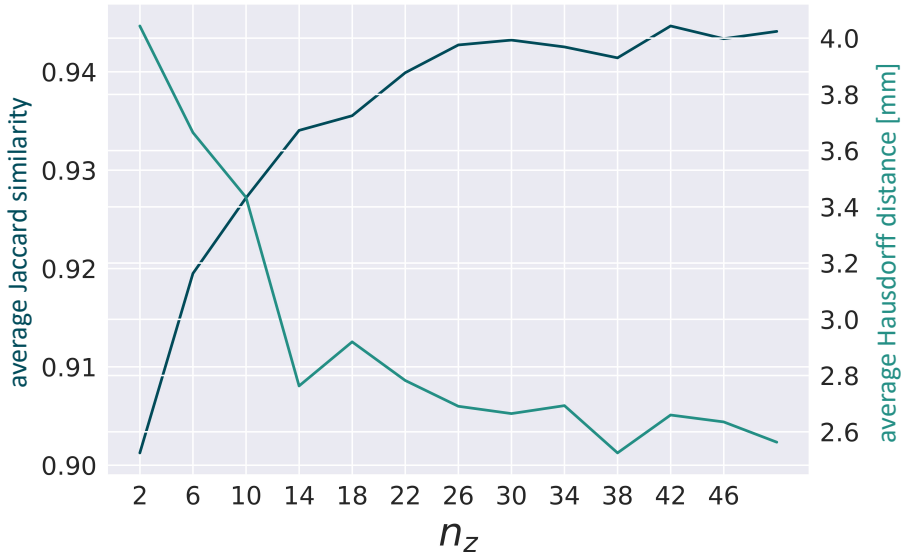

Figure 7.6: Reconstruction accuracy of the autoencoder in dependency of the number of latent dimensions n_z. Adapted from [57].

Discussion and Conclusion The results show that learning a feasible representation of the aortic valve leaflets is possible using convolutional autoencoders. Due to the saturation of the metrics with n_z greater than 20, a good trade-off between compactness of the representation and reconstruction accuracy can be achieved with 20 latent dimensions on this data set. The proposed architecture delivers the best results of all compared models. It is interesting that the usage of a deeper architecture did not lead to an increase in reconstruction accuracy. This might be related to the fact that autoencoders do not encode in general and that more complex decoders are more likely to just approximate the distribution of images without being conditioned by the latent space value [27]. Another possible explanation is the higher number of trainable parameters on the relatively small data set.

It could also be shown that no increase in reconstruction accuracy is achieved by introducing data augmentation. This might be due to the small variance in the data that is induced by the preprocessing pipeline. By accurate centering and unification, common advantages of data augmentation become irrelevant, such as translational or rotational invariance. Additionally, the general, rough shape of the leaflets is quite similar while the differences are mainly given by different contour lines or detailed shape differences, which also explains the overall high values of the Jaccard similarity. In the qualitative smoothness identification, slight differences between the proposed autoencoder and the

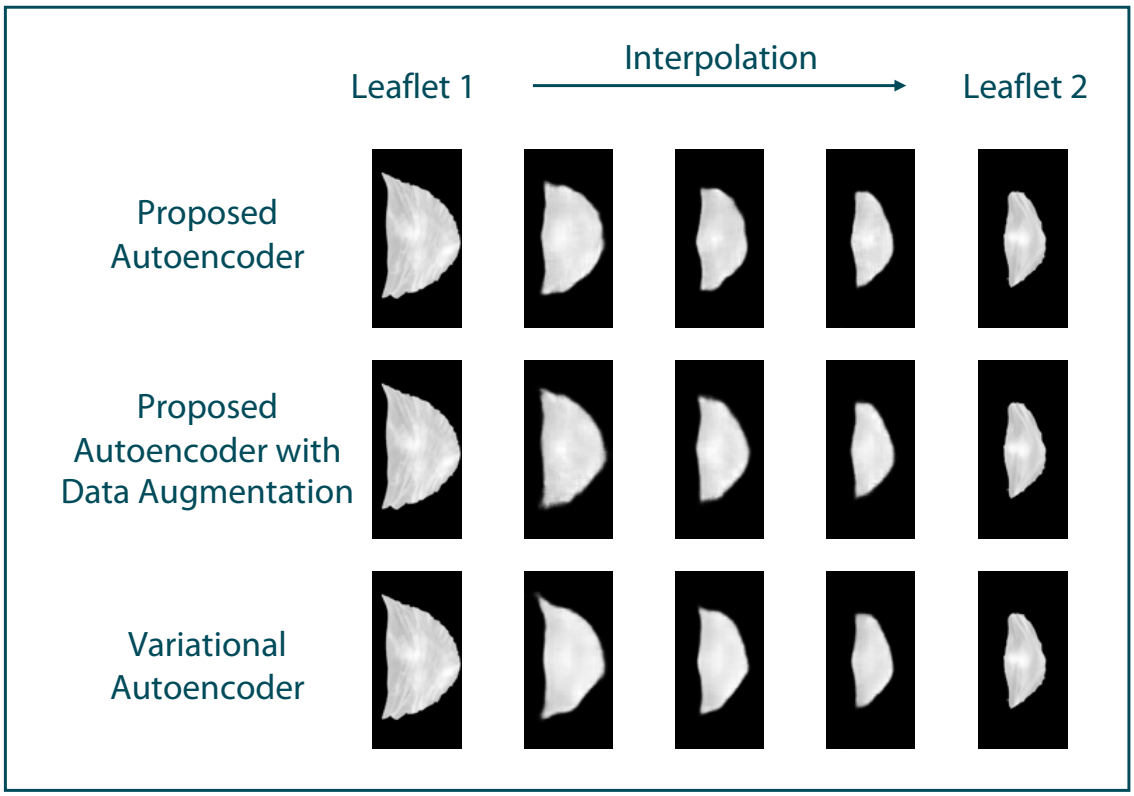

Figure 7.7: Smoothness of the autoencoder. Between leaflet 1 and leaflet 2, linear interpolation steps were acquired in latent space and the corresponding images were reconstructed. This was done for the proposed autoencoder without and with data augmentation as well as for the variational autoencoder. Adapted from [57].

VAE are visible. Hence, the VAE derives a different latent representation. However, an improvement of the reconstruction accuracy by using the VAE could not be observed.

The cluster analysis reveals the high potential of personalization. The clustering with $k = 1$ corresponds to the current situation where each patient is treated with the same valve geometry. By introducing more than one prosthesis type, the metrics for estimating how good the shape variance can be approximated deliver much better values. Note that the Jaccard similarity for $k = 1$ is already at 0.8. This means, that the whole potential of personalization lies within a range of a Jaccard similarity increase of 0.2. Just by introducing three different prosthesis types, about one quarter of this personalization gap can be closed. Likewise, the Hausdorff distance drops by $1.25mm$ when three prosthesis types are used. With more than three types, the increase or decrease of the metrics is slower, but still significant. However, this is reasonable, because a data set can always be

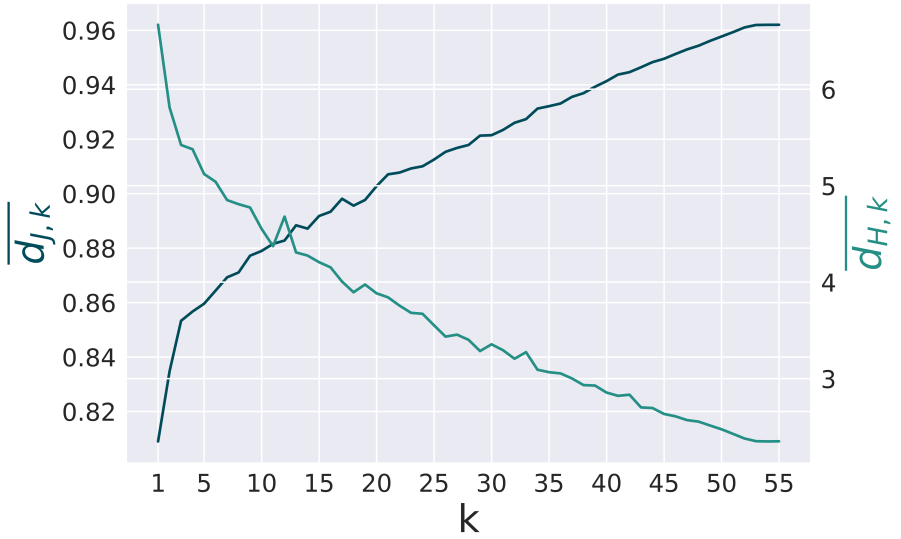

Figure 7.8: Capability of reproducing all individual valve shapes given in the data set in dependency of the number of prosthesis types. Adapted from [57].

approximated better the more cluster centers are used. At a number of about $k = 7$, the increase appears to be approximately linear. This leads to the expected conclusion that the biggest advantage in the trade-off between patient's outcome and economical issues can be achieved by introducing $3 - 5$ prosthesis shape types.

In this study, only the geometric shape of the leaflets could be evaluated. Remaining questions are the dynamical and biomechanical improvements of the valve prosthesis' functionality achieved by the proposed personalization technique. By manufacturing the identified prostheses, a comparison between state-of-the-art prosthesis and the personalized ones could be done in a left-heart simulator to analyze the impact of personalization. Due to the manual processing of the porcine heart valves, deformations of the leaflet shapes cannot be entirely avoided. However, it is assumed that these errors would be normally distributed, hence, no systematic bias towards specific shapes should be present in the data set. Another limitation of the study is the transferability to the human heart. Even though the aortic valves of pigs and humans appear to be quite similar and pig valves are already used as xenological prostheses, the direct usage of the identified prosthesis shapes in humans should be further investigated. However, the proposed autoencoder model could still present a sufficient basis using transfer learning, i.e. fine-tune the autoencoder with a very small set of human valves.

This study presents the first approach to derive prosthesis shapes via clustering in a latent space description. Like this, the clustering focuses on abstract meta-features given

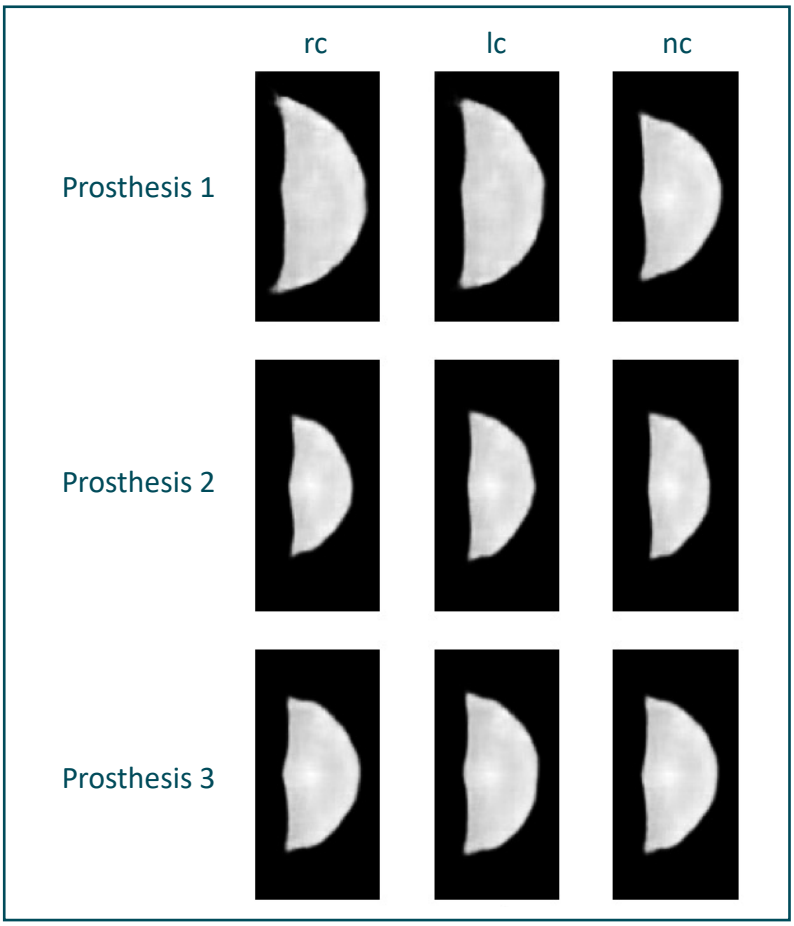

Figure 7.9: Resulting prosthesis shapes, exemplarily shown for $k = 3$. Adapted from [57].

by the deep representation rather than on pixel-wise differences. This makes the concept transferable to a lot of comparable problems in the area of personalized medicine. By adjusting the architecture of the autoencoder, it is possible to encode very different kinds of anatomies or images from different modalities, like 3D tomographic volumes. Hence, the proposed method could present an important step towards personalized prosthetics.

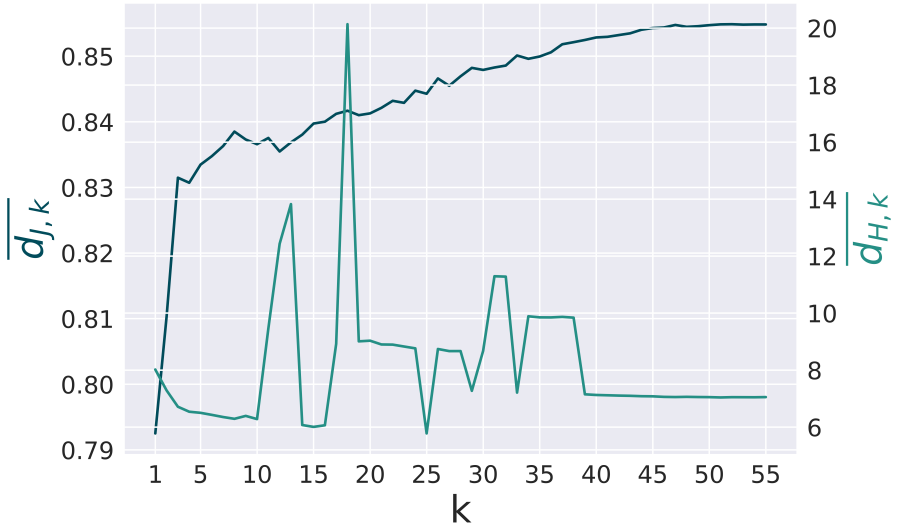

Figure 7.10: Capability of reproducing all individual valve shapes given in the data set in dependency of the number of prosthesis type when Principal Component Analysis (PCA) is used for deriving the latent space representation. For comparison, the number of principal components was set to $n_z = 20$. Adapted from [57].

Valve-based Type Classification

Based on the valve typification presented in section 7.2.1, it was examined whether it is possible to classify the optimal shape type only based on an ultrasound image of surrounding tissue, i.e. the aortic root. Thus, a leaflet representation was derived as described in Sec. 7.2.1 based on the data set D_S^{aux} and $k = 3$ typical valve shapes were identified. Then, all leaflet images contained in D_S were encoded using the trained autoencoder and the individual shape type was computed for each of them, respectively, using the previously calculated clustering. These clustering results serve as ground truth labels of the data set D_S.

To extract the relevant geometric prior information from the individual ultrasound image, three different approaches were assessed: 2D prediction, 2D transfer learning and 3D prediction. In all three approaches, a CNN is used to learn the classification based on the input image information. In the 2D prediction setting, only the horizontal slice through the ultrasound volume that shows the coaptation plane, i.e. the plane that shows the three noduli arantii in the closed state, is used as input for the model. Similarly, in the 2D transfer learning approach, the same slice is used. However, the image is replicated three times as simulated channels to achieve the shape of an RGB image. Then, the data can be processed using pretrained neural networks that learned meaningful features from

huge RGB data sets. In this study, a VGG16 architecture [132] pretrained on ImageNet [35] is used as its features were proven to carry relevant information on the aortic root geometry [56]. For fine-tuning, the last fully-connected layer was replaced by a freshly initialized one. For the 3D prediction approach, the full volume served as input of a 3D CNN. While the CNN architecture is predefined as VGG16 for the 2D transfer learning approach, three different architectures were assessed for the other cases.

Architecture A_1 consists of five convolutional layers (filter size 3×3 or $3 \times 3 \times 3$, depending on the input dimensionality) with an increasing number of filters $(4, 8, 16, 32, 64)$. Each of the convolutional layers was followed by a maximum pooling layer (size 2×2 or $2 \times 2 \times 2$) and three fully connected layers are attached to the last pooling layer with 512, 256 and 3 neurons, respectively. All layers use *ReLU* activation while the last layer utilizes *softmax* activation.

Architecture A_2 consists of four convolutional blocks, each followed by a maximum pooling (size 2×2 or $2 \times 2 \times 2$). Each of the blocks consists of three convolutional layers (filter size 3×3 or $3 \times 3 \times 3$), where the layers in the first two blocks have 16 filters and the ones in the last two blocks have 32 filters. Four fully connected layers are attached to the last pooling layer with 256, 128, 64 and 3 neurons, respectively. Once again, all layers use *ReLU* activation while the last layer utilizes *softmax* activation.

Similarly, architecture A_3 also features four convolutional blocks followed by a maximum pooling (size 2×2 or $2 \times 2 \times 2$) each. However, each block only consists of two convolutional layers (filter size 3×3 or $3 \times 3 \times 3$) where the number of filters increases with each block (8, 16, 32, 64). Similar to architecture A_1, three fully connected layers follow the last pooling layer with 512, 256 and 3 neurons, respectively. While the last layer uses *softmax* activation, all other layers utilize *ReLU* activation. Fig. 7.11 visualizes the three architectures A_1, A_2 and A_3.

All three approaches (2D prediction, 2D prediction with transfer learning and 3D prediction) were evaluated for all relevant architecture types (architecture 1, architecture 2, architecture 3, VGG16) on the data set D_S using 5-fold crossvalidation. To ensure a fair training process, the training data was balanced within each fold by oversampling the minority classes [50]. Within each fold, each of the models was trained on the current training data, i.e. the ultrasound image input and the corresponding shape type, represented as a one-hot encoded class vector. Then, the model was used to estimate the optimal shape types for all aortic roots in the current test set. Each model was evaluated by assessing the classification accuracy on the test data.

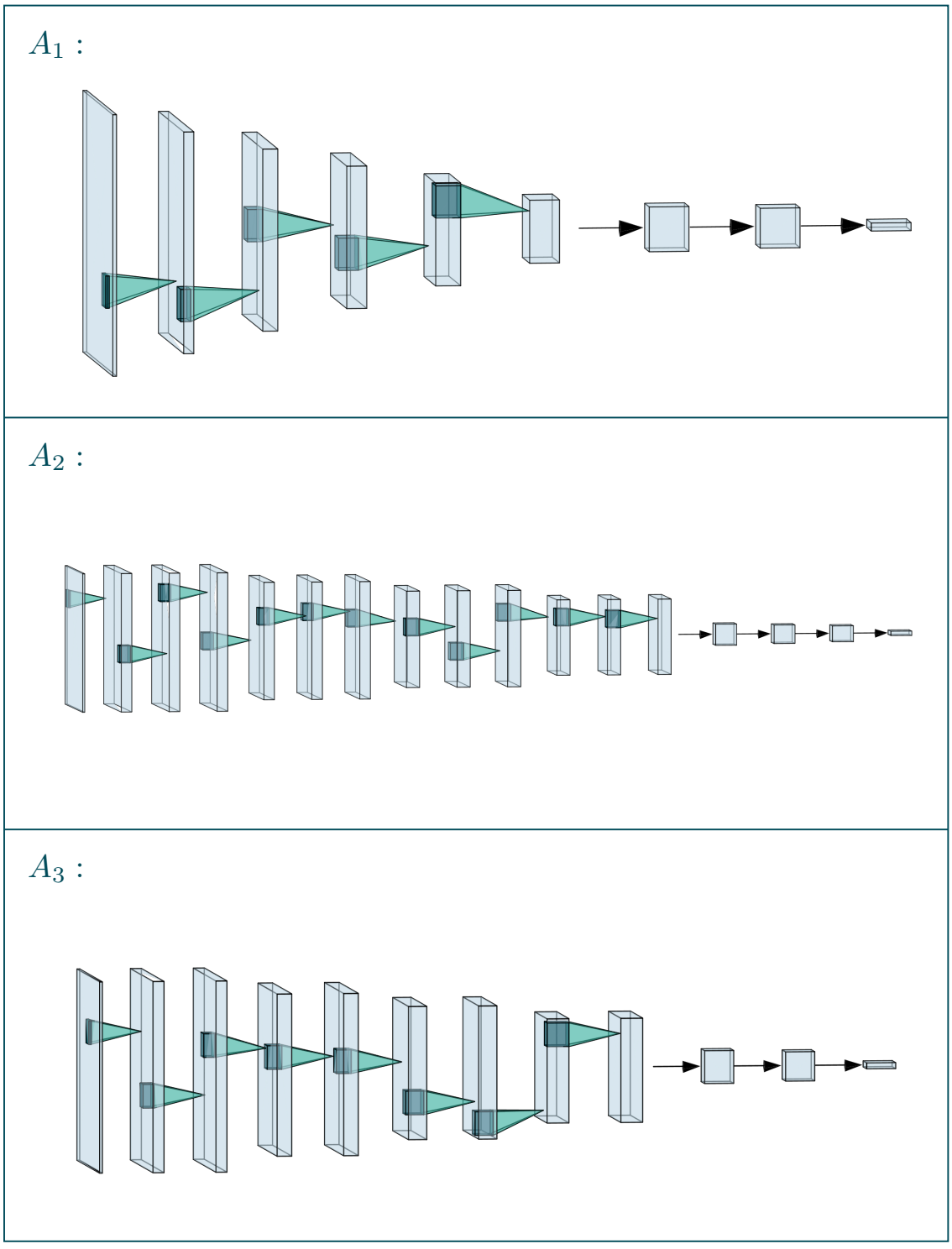

Figure 7.11: Proposed architectures for valve-based type classification.

Table 7.2: Prediction accuracy of all three approaches for each assessed architecture, respectively, given as the mean classification accuracy on the test set over all folds and its standard deviation.

Method	Architecture	Accuracy (mean $\pm$ std)
2D Prediction	A_1	$38.29 \pm 12.03\ \%$
	A_2	$40.14 \pm 8.26\ \%$
	A_3	$38.43 \pm 13.96\ \%$
2D Transfer Learning	VGG16	$28.29 \pm 17.62\ \%$
3D Prediction	A_1	$30.57 \pm 22.12\ \%$
	A_2	$36.10 \pm 8.12\ \%$
	A_3	$29.24 \pm 12.31\ \%$

Results and Discussion All three approaches were assessed for each relevant architecture using a 5-fold crossvalidation. Tab. 7.2 shows the classification accuracy for each model. As there are three potential classes, i.e. shape types, most models provide results in the area of guessing where an accuracy of around $33\,\%$ is expected. The maximum mean accuracy is $40.14\,\%$ for 2D prediction with architecture A_2. Neither the transfer learning approach using pre-trained features nor the integration of 3D shape information provides higher scores.

These results are not satisfying. As 2D and 3D input, very different architectures and even transfer learning were assessed in this study, it can be assumed that the poor performance is not due to manual hyperparameter choices but is rather systematic. Hence, the results indicate that predicting the individually optimal valve shape type only based on the geometry of the aortic root is not possible. This might be due to the high level of abstraction of the shape type as it does not only contain latent space information of one leaflet but of the combination of all three leaflets together. It should be noted that, following the nomenclature introduced in Sec. 6.2.1, all presented type classification approaches are *shape estimation* methods, while a *domain mapping* approach was not evaluated. However, as the classification accuracy for the valve-based type classification is that low and in the range of random guessing, a leaflet-based approach was followed rather than implementing *domain mapping* for valve-based type classification.

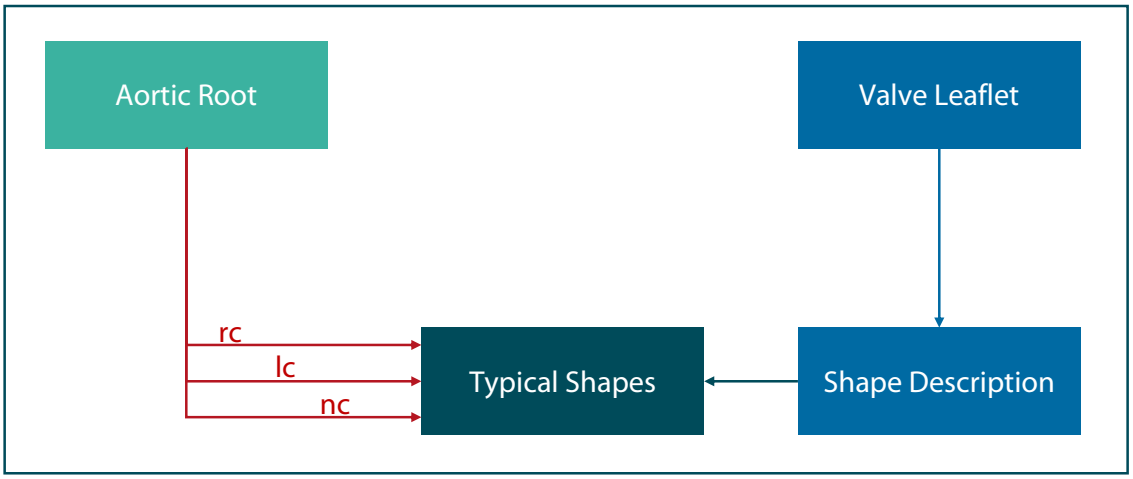

Figure 7.12: Workflow of the leaflet-based typification approach. After encoding all leaflet, typical leaflet shape types are identified. Note that these shape types include all three different kinds of leaflets, i.e. right-coronary, left-coronary and non-coronary ones. For a given aortic root shape, the individually optimal shape type can be classified for the right-coronary (rc), left-coronary (lc) and non-coronary (nc) leaflet separately.

7.2.2 Leaflet-based Typification and Classification

An alternative to the valve-based approach is to identify shape types of the single leaflets. It is assumed that leaflets from all types, i.e. right-coronary, left-coronary and non-coronary ones, can be described using a unified, type-agnostic shape typification model. Hence, each valve consists of three shape types, one for each leaflet. Based on the surrounding tissue information, i.e. the US volume image of the aortic root, each of these types can be classified independently, leading to three classification models, similar to the regression approach presented in Sec. 6.2.1. Since, a close anatomical relationship between the three leaflet shapes is expected, the independence assumption introduced by classifying each leaflet shape individually might appear unintuitive. However, it can be assumed that each of the classification models is capable of extracting relevant geometric information not only from one but from all three sinus of valsalvae. Since all three models receive the same input, each classification model should be able to take all relevant information from the US image into account, leading to correlated outputs of the three independent prediction models. Fig. 7.12 visualizes the general concept of the leaflet-based approach.

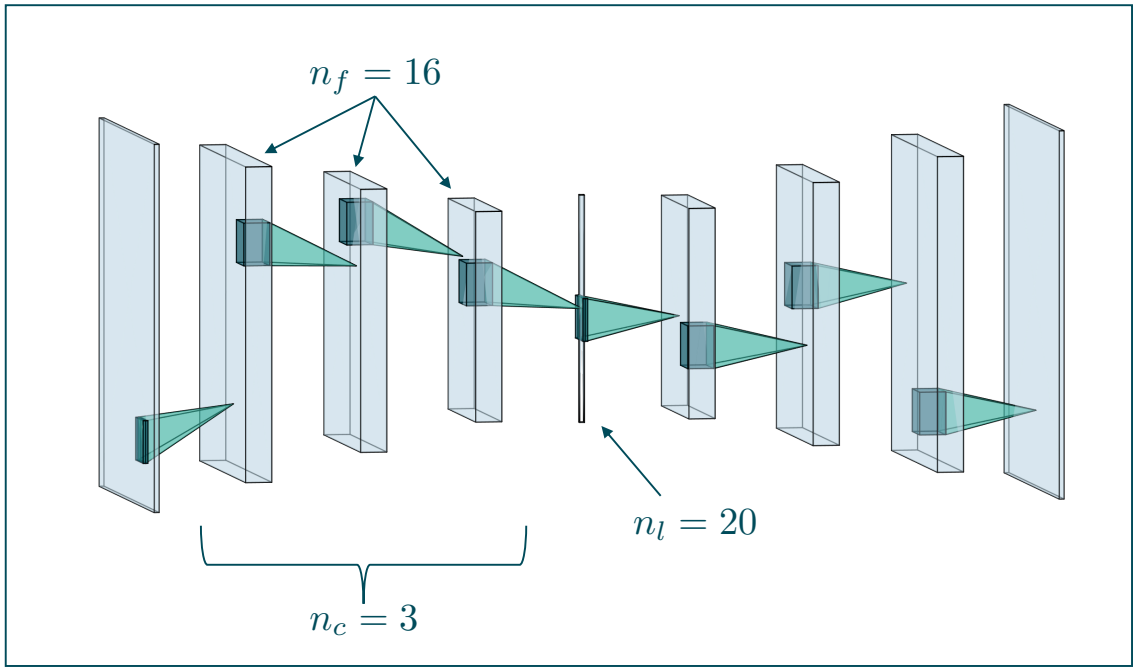

Figure 7.13: Architecture of the leaflet autoencoder. Due to the previously performed hyperparameter optimization on the same data set, the architecture is identical to the one proposed in 6.2.1.

Leaflet-based Typification

In the leaflet-based typification approach, typical shapes of individual leaflets are identified in the whole variety of observed leaflet shapes. Hence, there is no distinction between right-coronary, left-coronary and non-coronary leaflets. All three leaflet categories are merged together to identify typical leaflet shapes. Therefore, once again, a representation learning approach is used to perform clustering in a latent space description.

To make use of the maximal amount of data, the leaflet images from the data sets D_S^{aux} and D_S were used to identify typical leaflet shapes, resulting in 258 leaflets from 86 valves. As this data is identical to the data set used in the continuous approach described in Sec. 6.2.1 and the developed autoencoder architecture was proven to perform optimally on this data set, the same architecture and hyperparameters are used in this study. Hence, the encoder consisted of 3 convolutional layers (*ReLU* activation) with 16 filters and followed by a 2×2 average pooling layer each. The number of neurons in the bottleneck layer, i.e. the dimensionality of the latent space was 20. The decoder followed

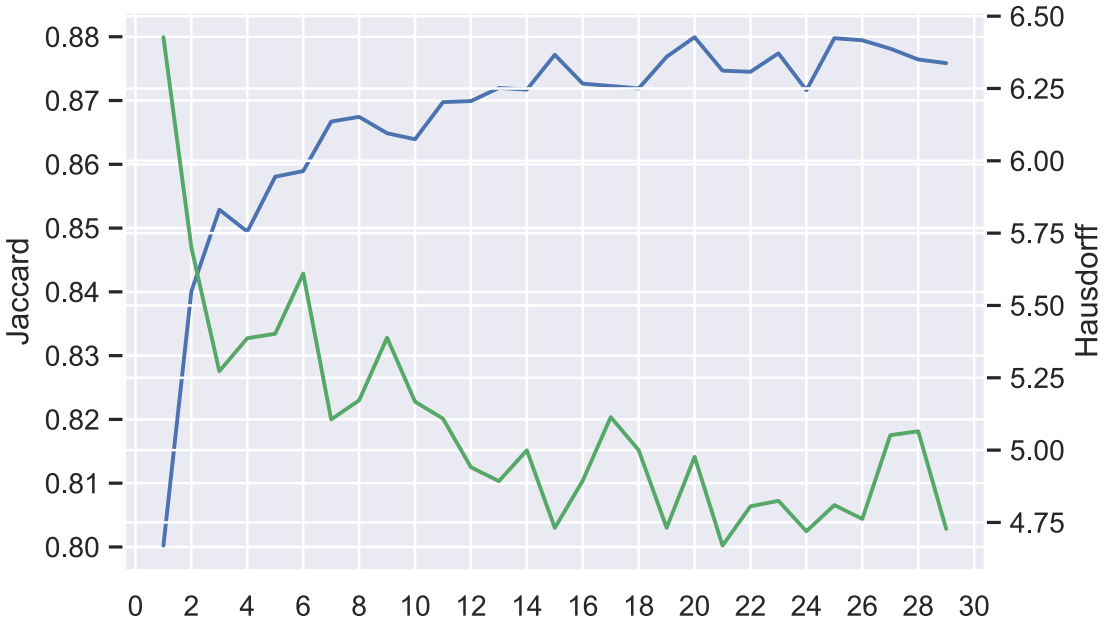

Figure 7.14: Jaccard similarity (blue) and Hausdorff distance (green) between test leaflet images and their corresponding leaflet type image, respectively, averaged over all test leaflets and folds.

the mirrored architecture. Fig. 7.13 visualizes the model's architecture. The autoencoder was trained for 100 epochs with a batch size of 32.

In order to analyze the capability of a set of prostheses shapes to represent the full variance observed in the data set, a 10-fold Monte-Carlo crossvalidation was performed. Thus, within each fold, the autoencoder was trained based on the training data, the training data was encoded to the latent space and k-means clustering was performed with k typical shapes, i.e. cluster centers. Then, the test images were encoded and their corresponding cluster centers were identified, respectively. These identified cluster centers, i.e. individually optimal discrete prostheses shapes, were compared to their corresponding leaflet images as described in 7.2.1 regarding the Jaccard similarity, Hausdorff distance, ASCD and RMSE. To assess prostheses shape sets of different sizes, this procedure was repeated for all different values of $k \in \{3, 4, 5, \ldots, 28, 29, 30\}$. Note that the minimal number of $k = 3$ was chosen according to the three native leaflet types, i.e. right-coronary, left-coronary and non-coronary.

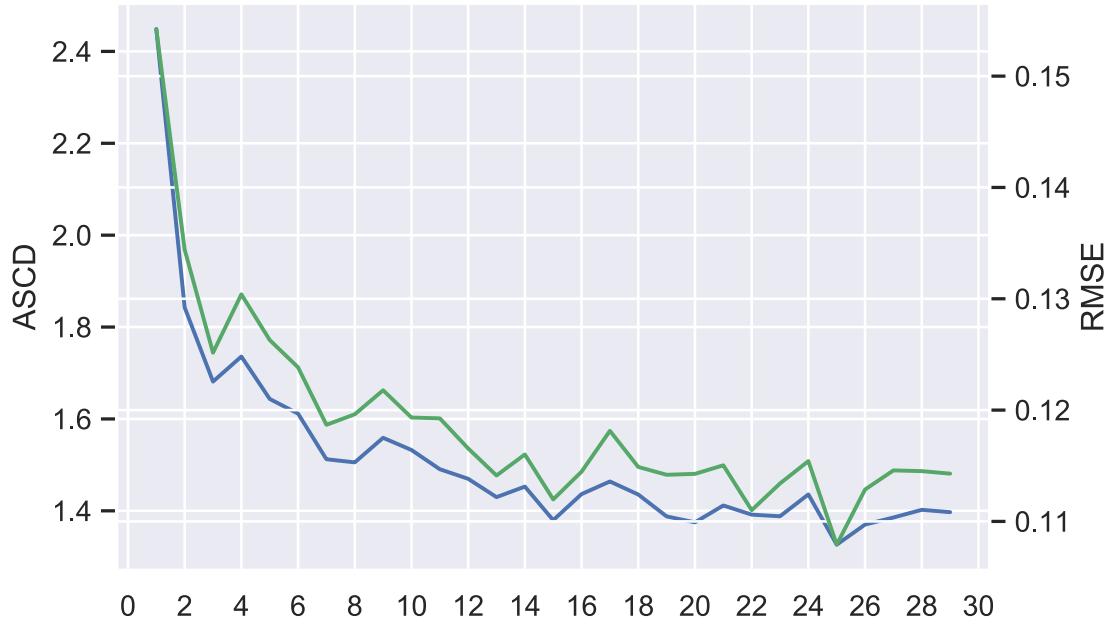

Figure 7.15: ASCD (blue) and RMSE (green) between test leaflet images and their corresponding leaflet type image, respectively, averaged over all test leaflets and folds.

Results of Typification The capability of approximating all observed leaflet shapes using a set of k prostheses was assessed in image space using a crossvalidation. Fig. 7.14 and 7.15 show the Jaccard similarity, Hausorff metric, ASCD and RMSE between the test leaflet images and their corresponding leaflet type image, respectively, averaged over all test leaflets and folds.

As expected, the Jaccard similarity increases while all other metrics decrease with an increasing number of typical shapes k. All metrics reach a certain degree of saturation. This verifies the assumption that a set of leaflet shapes is sufficient to approximate all observed leaflets adequately. Based on all four metrics, a value of $k = 15$ can be identified as optimal in terms of a trade-off between high accuracy and a small number of different leaflet shapes. As these 15 shapes span all three native leaflet categories, each of them are represented by five shape types on average.

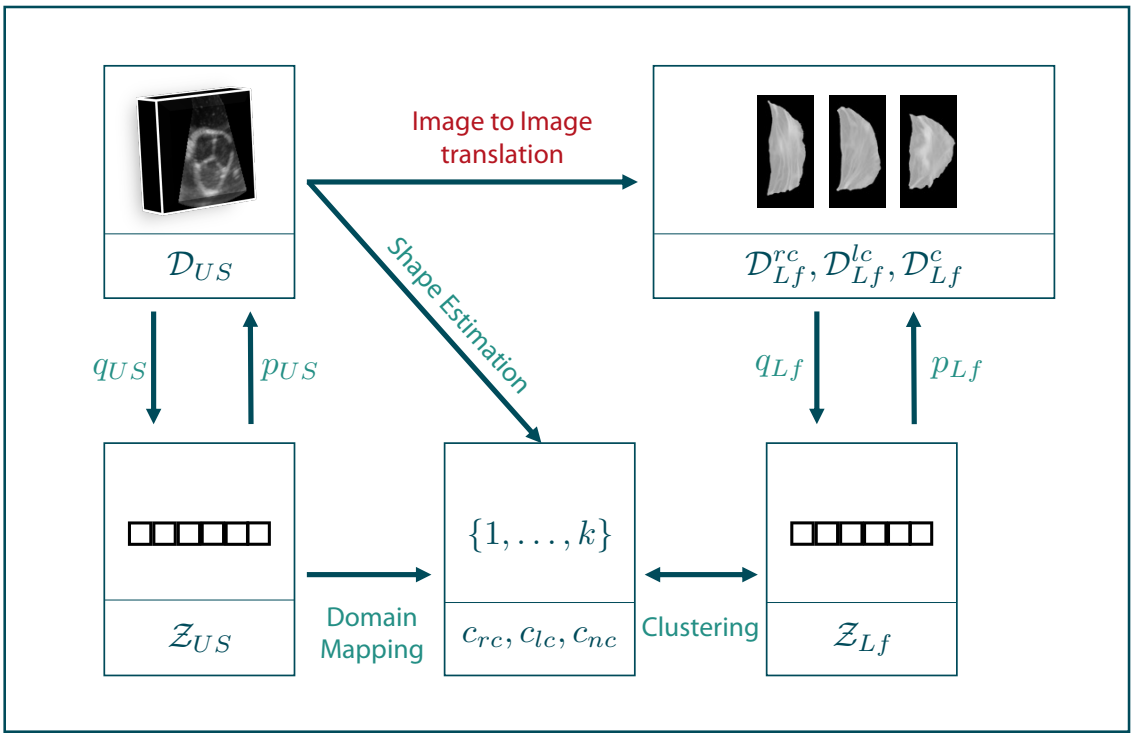

Figure 7.16: Sketch of *shape estimation* and *domain mapping* approach for the discrete case. In the latent respresentation of the leaflet images, a clustering is performed to identify typical shapes. Then, a mapping is learning to classify the optimal shape type of each leaflet based on the ultrasound image, either directly (*shape estimation*) or after encoding the ultrasound image to a latent space description (*domain mapping*).

Leaflet-based Type Classification

After identifying typical leaflet shapes, mapping is needed between the individual surrounding tissue geometry and the optimal prosthesis type. This is comparable to the pseudohealthy synthesis approach presented in Sec. 6.2.1. However, due to the discrete typification approach, the mapping becomes a classification problem. Thus, the individually optimal leaflet shape types should be predicted based on a volumetric US image of the aortic root. As mentioned above, each of the three leaflets is classified independently as one class label describes the shape type of one leaflet. Therefore, the presented method can be considered as the discrete version of the continuous pseudohealthy synthesis method presented in Sec. 6.2.1. Fig. 7.16 illustrates this concept. The shape type can be classified either directly based on the US image (*shape estimation*) or based on a latent representation of it (*domain mapping*).

Table 7.3: Combinations of hyperparameters of the regression networks assessed during hyperparameter analysis for the *shape estimation* approach. The optimal combination is marked in bold.

Parameter	Values
k_b	**3**, 4, 5
k_c	1, **2**, 3
k_f	**16**, 32
k_d	**1**, 2, 3
k_n	50, 100, **200**

To solve the classification task, three approaches were followed that are highly similar to the ones presented in Sec. 6.2.1: *shape estimation*, *domain mapping* with RF and *domain mapping* with MLP. However, in this setting, the decision models were designed to provide a classification instead of a continuous regression output. Details on the main concepts behind the three approaches can be found in Sec. 6.2.1.

As mentioned in Sec. 7.2.2, the data sets used in this study are identical to the ones used in Sec. 6.2.1. Hence, the optimal autoencoder architectures can be adopted from the hyperparameter analysis performed in Sec. 6.2.1. Thus, both autoencoders were parameterized by three convolutional layers with 16 filters each and a latent space dimensionality of 20.

Hyperparameter Analysis For the two variants of the *domain mapping* approach, an analysis of the hyperparameters of the classification models was done by performing a 10-fold Monte Carlo crossvalidation (80% training, 20% test) for each hyperparameter combination. Within each fold, the leaflet autoencoder was trained on the training data and the auxiliary data set D_S^{aux}. To obtain ground truth labels, all leaflet images of these two data sets were encoded and k-means clustering was performed in the latent space to identify typical leaflet shapes with $k = 15$. This number was identified as optimal in Sec. 7.2.2. Then, all leaflet images from the training and test sets were encoded and assigned to their cluster to obtain ground truth labels. Afterwards, the ultrasound autoencoder was trained on the ultrasound volumes from the training set. Finally, the three classification models were trained to predict the class label of their respective leaflet based on the latent representation of the ultrasound volume image. Note that three classification models were trained, one for each leaflet category, i.e. right-coronary, left-coronary and non-coronary. To assess the accuracy, the trained models were used to predict class labels

Table 7.4: Accuracy of discrete leaflet shape prediction for all methods, regarding RMSE, Jaccard similarity, Hausdorff distance and ASCD. The values for the Hausdorff distance and the ASCD are given in pixels, while the resolution was $0.34\frac{mm}{pixel}$. The table shows the mean value over all cross validation folds, averaged over all three leaflet types, and the standard deviation.

Method	RMSE	Jaccard	Hausdorff	ASCD
Shape Estimation	0.1658 ± 0.0525	0.7828 ± 0.0831	7.49 ± 3.99	2.72 ± 1.42
Domain Mapping with RF	$\mathbf{0.1616 \pm 0.0506}$	$\mathbf{0.7917 \pm 0.0839}$	$\mathbf{7.21 \pm 4.04}$	$\mathbf{2.58 \pm 1.42}$
Domain Mapping with MLP	0.1711 ± 0.0562	0.7744 ± 0.0920	7.92 ± 4.77	2.86 ± 1.64

for each ultrasound volume in the test set. Then, the predicted cluster centers were decoded and compared to the ground truth leaflet images using RMSE.

Thus, for the Random Forest with t trees, all values of $t \in \{50, 100, 150, 200, 250\}$ were assessed. For the MLP approach, the network consisted of l_h hidden layers with l_n neurons each. While the hidden layers used *ReLU* activation, the output layer features *softmax* activation to allow discrete classifications. In the scope of the hyperparameter analysis, all combinations of $l_h \in \{1, 2, 3, 4, 5\}$ and $l_n \in \{50, 100, 150, 200\}$ were assessed.

For the *shape estimation* approach, the parameterized architecture presented in 6.2.1 was used. Thus, the convolutional part of the VGG-like architecture consists of k_b convolutional blocks. Each of these blocks consists of k_c 3D convolutional layers with k_f filters and *ReLU* activation. Each block is followed by a 2×2 average pooling layer. Behind the last convolutional block and a flattening operation, k_d fully connected layers with *ReLU* activation and k_n neurons each are attached, followed by the output layer with k_l neurons and *linear* activation. Tab. 7.3 lists the assessed values for each parameter. To identify the optimal architecture, each combination of them was evaluated.

Results and Discussion The optimal number of trees in the *domain mapping* with RF approach identified in the hyperparameter analysis was $t = 150$ with an average RMSE over all folds of 0.1556 ± 0.0473. The optimal average RMSE with an MLP was achieved using $l_h = 1$ hidden layer with $l_n = 150$ neurons, with a value of 0.1590 ± 0.0479. For the *shape estimation* approach, the optimal architecture featured $k_b = 3$ convolutional blocks of $k_c = 2$ convolutional layers with $k_f = 16$ filters each, followed by $k_d = 1$ fully connected layer with $k_n = 200$ neurons, reaching an RMSE over all folds of 0.1534 ± 0.0471.

To assess the accuracy of the methods, a 10-fold Monte Carlo crossvalidation (80% training, 20% test) was performed for each of the three approaches using the optimal hyperparameters, respectively. The accuracy was measured using the Jaccard similarity, Hausdorff distance, ASCD and RMSE. Table 7.4 shows the quantitative results of all

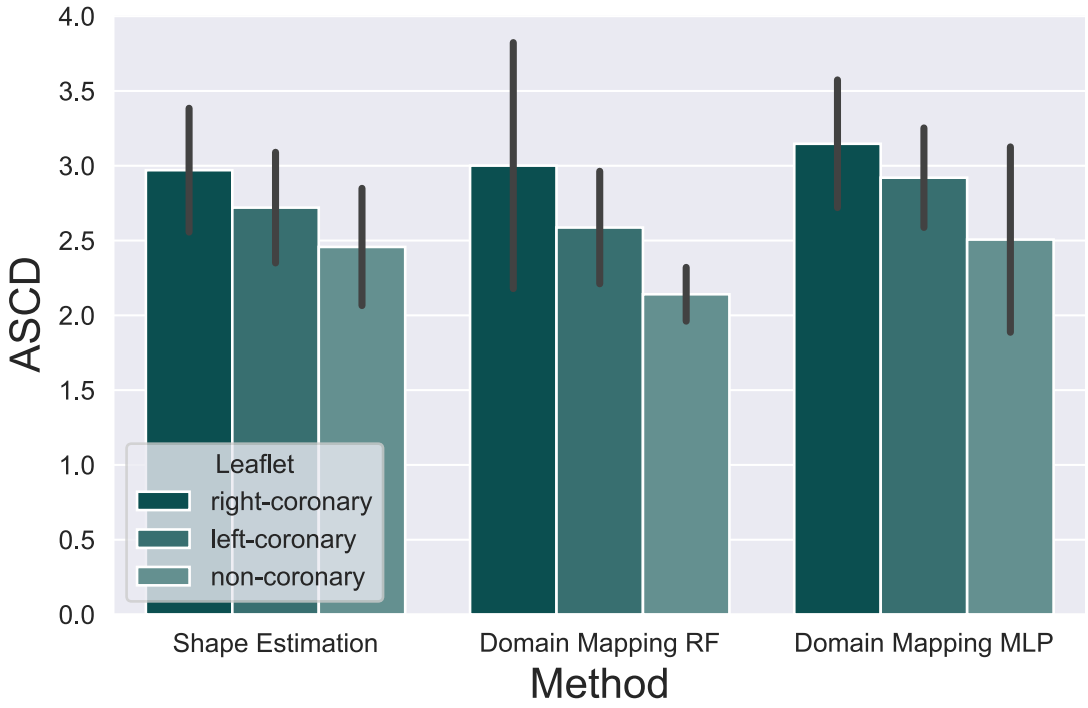

Figure 7.17: Average ASCD of each method, given for each of the three leaflet categories right-coronary, left-coronary and non-coronary. The ASCD is given in pixels, the resolution was $0.34 \frac{mm}{pixel}$.

methods, averaged over all folds. The average classification accuracy, i.e. the relative number of correctly classified leaflet shape types in the test set, was $18.9 \pm 12.7\%$ for *shape estimation*, $15.0 \pm 15.1\%$ for *domain mapping with RF* and $12.2 \pm 13.6\%$ for *domain mapping with MLP*. Fig. 7.17 shows the average ASCD of each method regarding each of the three leaflet categories.

As the method is implicitly constrained to predict typical shapes, all synthesized images look realistic. In general, the predicted leaflet types resemble the native, healthy ones well. Example predictions for the three approaches are shown in the Figures 7.18, 7.19 and 7.20, respectively. All in all, each method provides reasonable results with Jaccard similarity values over $77\,\%$ and an ASCD below $1\,mm$. While the *domain mapping* approach was more accurate than the *shape estimation* approach, using a Random Forest for *domain mapping* provided the best quantitative results throughout all metrics. Hence, in contrast to the continuous regression approach presented in 6.2.1, the Ran-

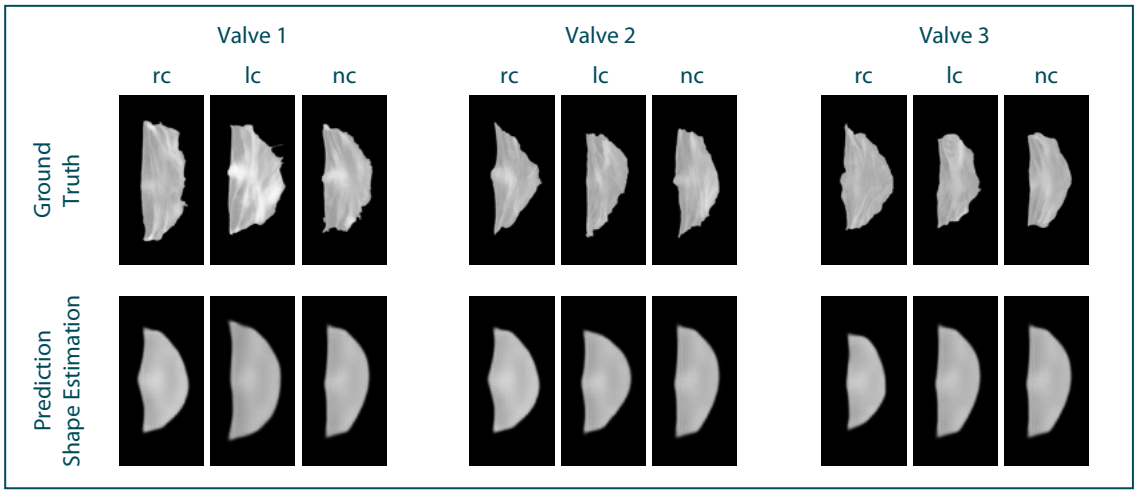

Figure 7.18: Qualitative prediction results of the discrete *shape estimation* approach, exemplarily shown for three valves from the test set.

dom Forest seems to be superior to an MLP in the discrete type classification case. The non-coronary leaflet is approximated with the smallest ASCD, which is probably due to a smaller variance of its shape. In the classification scenario, the *shape estimation* approach performs better than in the continuous approach. It outperformed the MLP-based *domain mapping* approach in all metrics.

In comparison to the continuous approach from Sec. 6.2.1, the optimal MLP architecture for the discrete case is quite shallow with only one hidden layer, while four hidden layers provided optimal results in the continuous approach. This indicates that the classification problem addressed in this study is less abstract and hence, easier to solve than the regression problem that occurs with continuous pseudohealthy synthesis. Overall, the classification accuracy on the test set is quite low for all three approaches with about $12 - 15\,\%$. However, it is important to note that there are 15 possible classes, so all methods perform better than guessing randomly. Additionally, there might be ordinal relationships between the classes as some shape types might be more similar than others. Hence, by predicting a wrong but similar shape type, the prediction accuracy decreases while the approximation of the individual leaflet shape is still fair.

In this study, it could be shown that by following a leaflet-based typification approach, leaflet type classification is possible. The superior classification results compared to the valve-based approach might be due to the lower level of abstraction. By training individual classification models for each leaflet category, each model can focus on extracting the in-

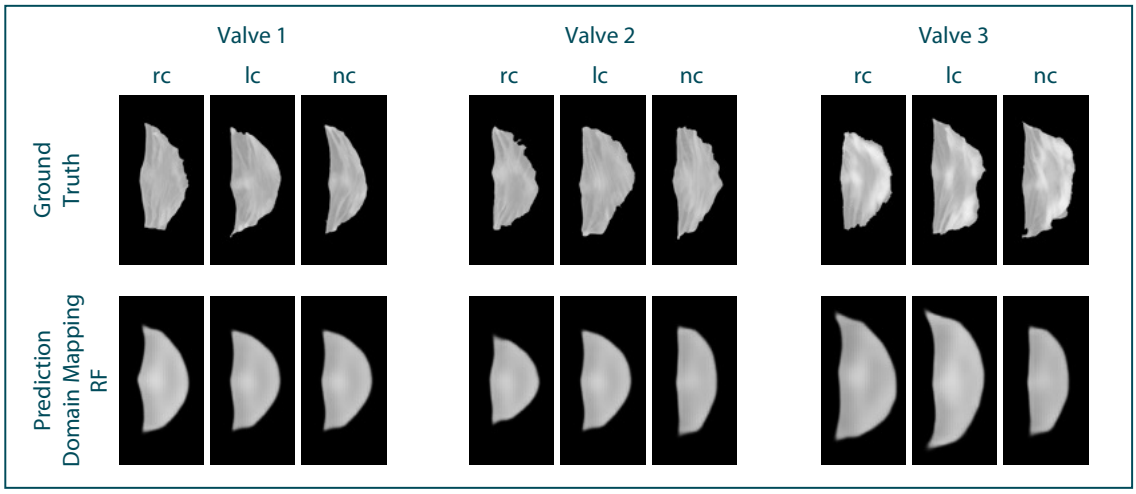

Figure 7.19: Qualitative prediction results of the discrete *domain mapping* approach using RF, exemplarily shown for three valves from the test set.

formation necessary to predict the single corresponding leaflet only, while one model has to collect all necessary information about all three leaflets in the valve-based approach. In addition, the leaflet shapes are encoded in the class labels in a more abstract way in the latter approach.

Conclusion In this section, a concept for automatic typification of aortic valve shapes based on individual leaflets was developed and presented by introducing clustering in latent space. Following this typification concept, models for valve type classification based on the individual aortic root shape were proposed and evaluated. It could be shown that it is possible to identify a meaningful representation of the aortic valve leaflets using convolutional autoencoders. Furthermore, the typification approach revealed typical leaflet and valve shapes and allowed for a quantitative analysis of the number of shape types used regarding the approximation of the whole variety of shapes using this finite set.

For the *leaflet-based* approach, it could be shown that it is possible to classify the shape type only based on a geometric prior from surrounding tissue. Qualitatively, the shape types are realistic, consistent and complementary, indicating that the typification method performs well. The quantitative analysis of predicting the individually optimal prosthesis revealed an adequate approximation with acceptable ranges of all metrics. Even though the classification accuracy on the test set is comparably low, the metrics for shape comparison are in an acceptable range. This might be due to the ordinal nature of

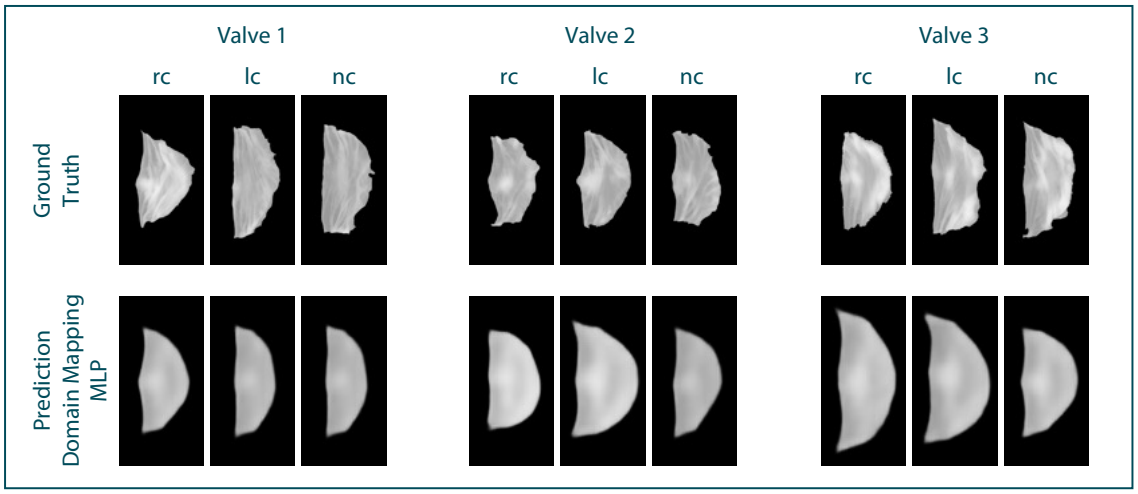

Figure 7.20: Qualitative prediction results of the discrete *domain mapping* approach using MLP, exemplarily shown for three valves from the test set.

this classification task: If the model predicts a wrong type with a shape similar to the correct one, the classification will be counted as false even though the shape approximation might be sufficient.

Since type classification was not possible in the *valve-based* approach, it can be stated that the *leaflet-based* approach outperformed the *valve-based* one. This might be due to the additional layer of abstraction that is introduced in the *valve-based* approach by concatenating three latent descriptions, challenging the classifier to learn the relationship between the three leaflets implicitly. I contrast, in the *leaflet-based* approach, the classification models can focus on predicting one single leaflet, which seems to be a less complex problem. It should be noted that the *domain mapping* approach was not evaluated in the scope of *valve-based* typification (see Sec. 7.2.1). However, as the *shape estimation* models did nit perform better than random guessing in the *valve-based* approach, which was the case in the *leaflet-based* one, it can be assumed that *domain mapping* with a *valve-based* approach will not outperform the *leaflet-based* method.

Overall, it could be shown that discrete pseudohealthy synthesis presents a promising alternative to continuous shape synthesis for personalizing aortic valve prostheses. It should be noted that the valve shape is reduced to the shape of the three leaflets, while the stent shape is not regarded in the scope of this study. However, the stent design might also benefit from the discrete approach as the finite number of leaflet shape types also imply a finite number of stent shapes that can be manufactured in serial production.

7.3 Morphological Pathology

As described above, the individual tailoring of personalized prostheses comes with high financial, logistical and regulatory issues. In order to achieve a balance between the benefits of personalization and a minimization of costs and regulatory issues, providing a set of general prosthesis shape types, which cover the anatomical diversity best, presents a promising trade off. For each patient, the individually optimal one can be predicted from the provided set for each patient. Regarding personalized aortic root prostheses, this might present an interesting alternative to the continuous pseudohealthy synthesis approach presented in Sec. 6.3. Fig. 7.21 visualizes the approach to investigate research question *Q3* in the case of a morphological pathology.

In this section, a framework for fully automatic shape typification as well as type classification based on a pathological prior is presented. This concept of discrete pseudohealthy synthesis is applied and evaluated for personalizing aortic root prostheses for the case of a morphological pathology. After describing the typification of healthy aortic root shapes in Sec. 7.3.1, the classification of the individually optimal shape type based on the pathological shape is presented in Sec. 7.3.1.

Contribution of this Section

The contribution of this section is twofold. First, to the best of the author's knowledge, this study presents the first approach to formulate pseudohealthy synthesis in a discretized way aimed at aortic root prosthesis shaping. As the developed framework is not limited to aortic root prosthesis shaping, it could be applied to a wide range of organ shape synthesis problems from a pathological prior and hence, provides a new method for pseudohealthy synthesis in general. Second, there is a high clinical value of this study for the development of personalized aortic root prostheses. The usage of a set of prosthesis types has a high potential for translation to clinical application.

7.3.1 Aortic Root Typification and Type Classification

In this section, approaches for aortic root typification as well as for type classification based on the individual pathologically dilated aortic root shape are presented. In contrast to the latent space manipulation approach presented in Sec. 6.3.1, the estimation problem becomes a classification and cannot be interpreted as a deformation of the pathological shape. Therefore, the classification approach presented in this section directly

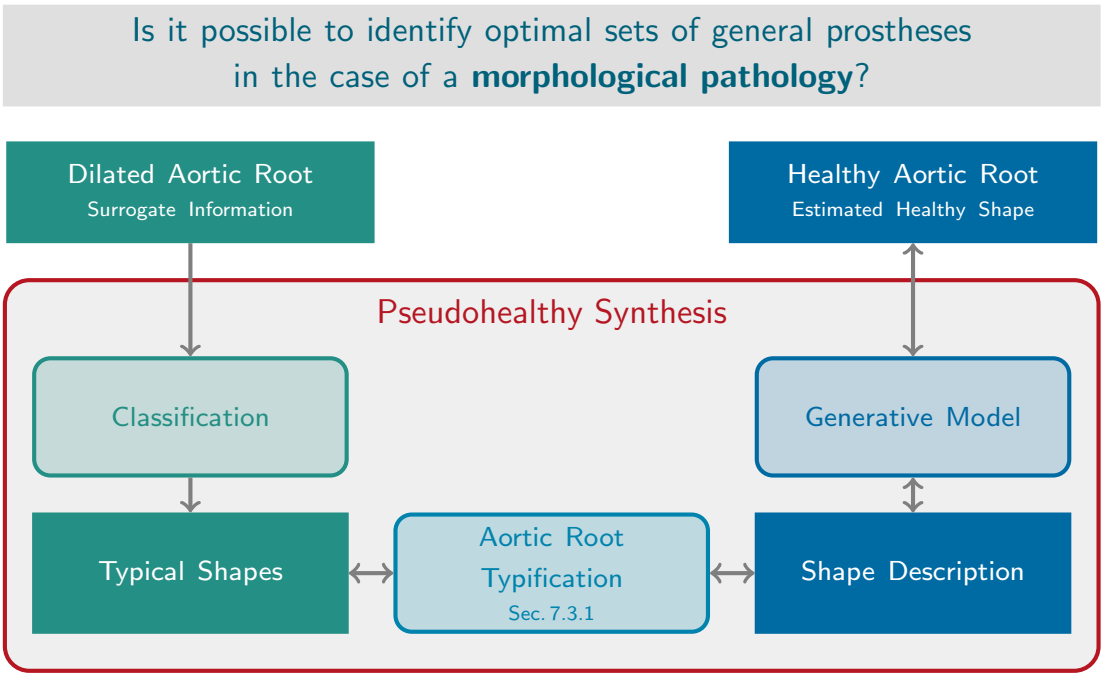

Figure 7.21: Approach to investigate research question *Q3* in the case of a morphological pathology.

estimated the individually optimal shape based on the pathologically dilated aortic root shape. Thus, it is closely related to the *shape estimation* approach presented in Sec. 6.2.1. Fig. 7.22 shows an overview of this proposed framework.

Aortic Root Typification

The goal of the typification step is to identify a set of typical aortic root shapes that can serve as prostheses shapes. To automate this task, unsupervised clustering of the healthy roots could reveal these shape types. However, clustering directly in image space is insufficient as images are comparably high-dimensional and the clustering algorithm faces the curse of dimensionality [4]. Thus, the approach of clustering in a latent space description is followed as presented in Sec. 7.2. Consequently, a latent representation of the healthy aortic root images is learned using a convolutional autoencoder. After encoding all healthy images into this low-dimensional representation, a clustering can be performed to identify typical shapes, i.e. the cluster centers, within the latent space. The desired shape of these potential prostheses can be synthesized using the decoder

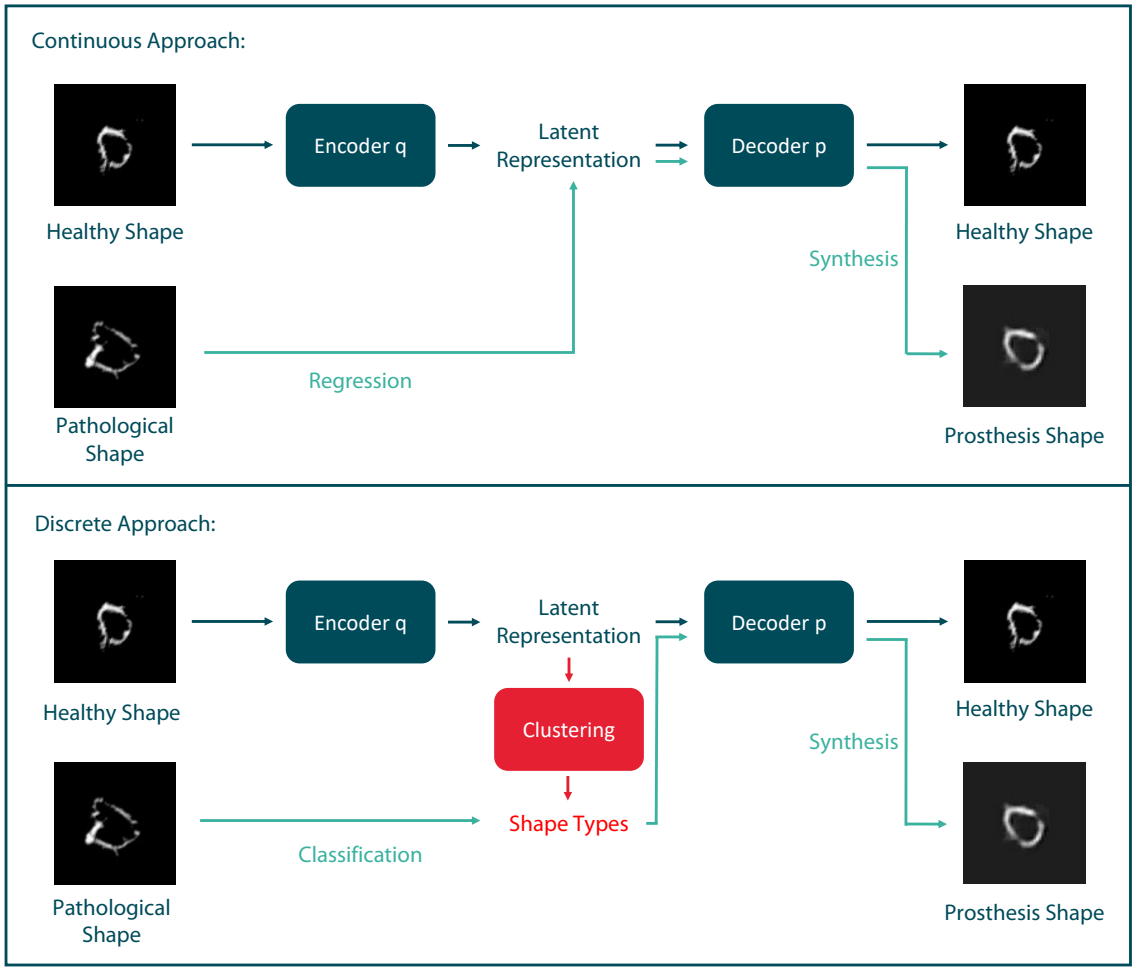

Figure 7.22: A sketch of the proposed method. The healthy images are encoded to a latent space representation. Previously presented methods (cf. 6 predict this latent representation continuously (top). In contrast, we propose to perform clustering in latent space to identify typical aortic root shapes (bottom). To compute the optimal prosthesis type for an individual patient, the shape type is classified based on an image of the pathological state and the corresponding image of the prosthesis can be synthesized.

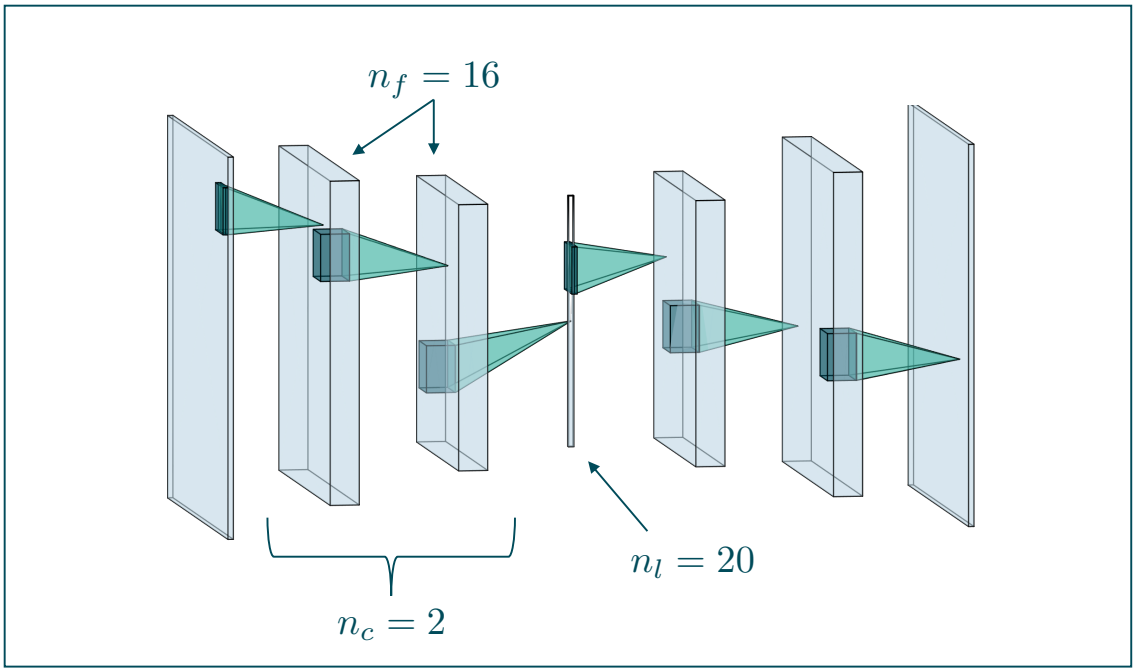

Figure 7.23: Architecture of the autoencoder, exemplarily shown for $n_c = 2$, $n_f = 16$ and $n_l = 20$.

network. Thus, clustering can be performed in a meaningful representation while the similarity between a specific prosthesis and a real aortic root can be directly assessed in image space using typical metrics.

Data Set The data set used here was $\tilde{D}_M$. Thus, in the scope of this study, only the horizontal slice image through the root that shows the commissure plane, i.e. the image plane that shows all three commissure points, is used. For these 2D images, a manual segmentation of the aortic root shape is available for the healthy roots, making a quantitative analysis more reliable. The images were resized to 96×96 pixels with a resolution of $1.27 \frac{mm}{pixel}$. The gray values were scaled to lie in the interval $[0, 1]$. These images will be denoted as I_i and the pathological ones as $\tilde{I}_i$ with $i = 1, \ldots, 24$, respectively. The data set was split into $N_{test} = 5$ aortic roots that served as a hold-out test set and 19 aortic roots for training and validation, referred to as training images.

Table 7.5: Examined hyperparameter values for the autoencoder. All combinations were assessed using 10-fold crossvalidation on the training images. The optimal combination regarding the RMSE between the original and the reconstructed image is marked in bold.

Parameter	Values
n_c	$\mathbf{2}, 3, 4$
n_f	$\mathbf{16}, 32$
n_l	$\mathbf{20}, 40, 60, 80, 100, 120, 140$

Representation Learning and Clustering The cluster centers, i.e. the prosthesis types, in latent space are denoted as $z_j^c,\ j = 1, \ldots, k$, where k is the number of types. After decoding, the synthesized image of the j-th shape type is called $\hat{I}_j^c$. To this end, a general, parameterized autoencoder architecture is proposed. The encoder $q(I),\ I \in \mathbb{R}^{96\times96}$ with $z_i = q(I_i)$, consists of n_c convolutional layers, each with n_f filters, *ReLU* activation and followed by a 2×2 average pooling. After a flattening operation, a dense layer follows where the number of neurons is the number of outputs of the last pooling layer. Then, a dense layer with n_l and linear activation forms the bottleneck layer that outputs the latent representation $z \in \mathbb{R}^{n_d}$. The decoder $p(z)$ with $\hat{I}_i = p(z_i)$ follows the mirrored encoder architecture using upconvolution and upsampling, where $\hat{I}_i$ is the reconstructed healthy image. Fig. 7.23 shows the architecture of the autoencoder. For training, the *adam* optimizer [79], mean squared error loss and a batch size of 12 was used while training was performed for 100 epochs.

To assess the influence of the hyperparameters n_c, n_f and n_l, several combinations of them were evaluated regarding the performance of the resulting architecture on encoding and reconstructing unknown aortic root images. Thus, a 10-fold Monte-Carlo

Table 7.6: Examined hyperparameter values for the classification CNN. All combinations were assessed using 10-fold crossvalidation on the training images. The optimal combination regarding the RMSE between the original and the reconstructed image is marked in bold.

Parameter	Values
m_b	$3, \mathbf{4}, 5$
m_c	$\mathbf{1}, 2, 3$
m_f	$\mathbf{16}, 32$
m_d	$1, 2, \mathbf{3}$
m_n	$50, 100, \mathbf{150}$

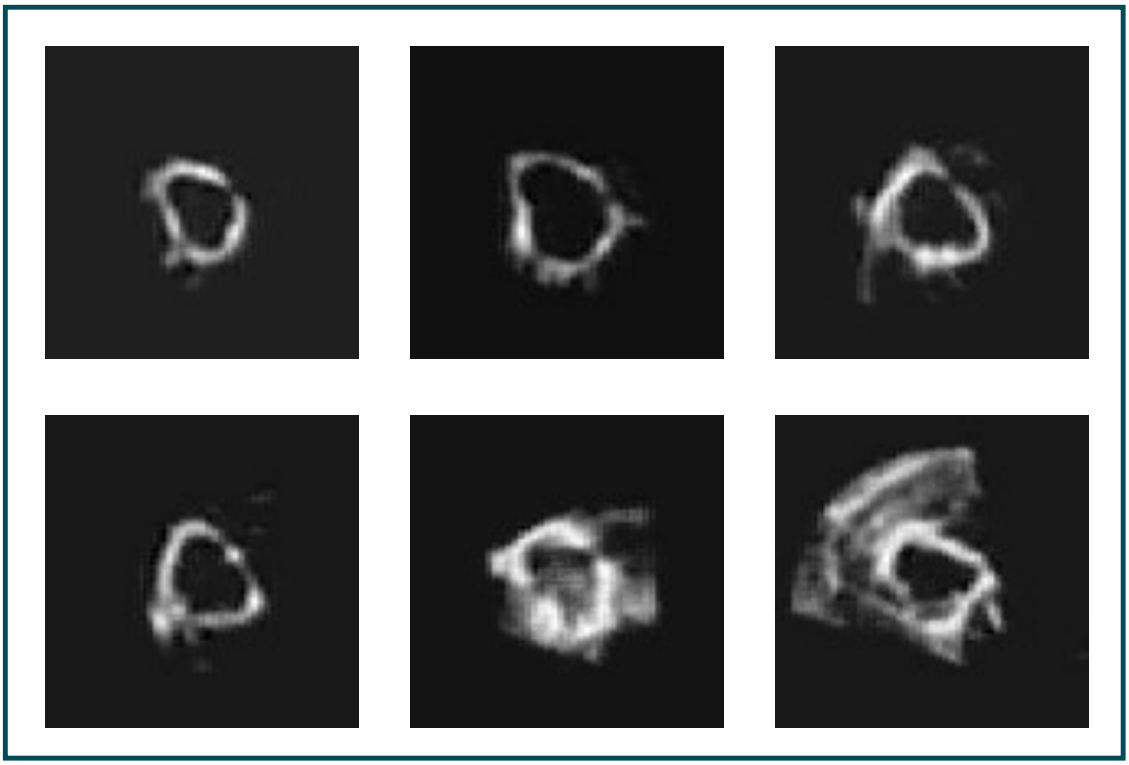

Figure 7.24: Example of a set of identified shape types for $k = 6$.

crossvalidation was performed on the training images (80% training, 20% validation) for each combination of hyperparameters examined, given in Table 7.5. The autoencoder was trained on the training data, the test data was propagated through the full network and the average root mean square error (RMSE) between the output and the original test images was computed in each fold.

After identifying the optimal architecture, the influence of the number of cluster centers, i.e. prosthesis types, on the capability of covering the full variance of observed root shapes by a small number of shape types was examined. Following the approach presented in Sec. 7.2, k-means clustering in the latent space was utilized. Different values of k were evaluated in a 10-fold Monte-Carlo crossvalidation (80% training, 20% validation) on the healthy training images. Within each fold, the autoencoder was trained on the training data using optimal hyperparameters, the training data was encoded to the latent space and k-means clustering was performed to identify k shape types. Then, the N_{test} test images were encoded and assigned to their respective clusters $z_j^c, j = 1, \ldots, k$. Finally, images of the prosthesis types $\hat{I}^c$ were synthesized by propagating the cluster

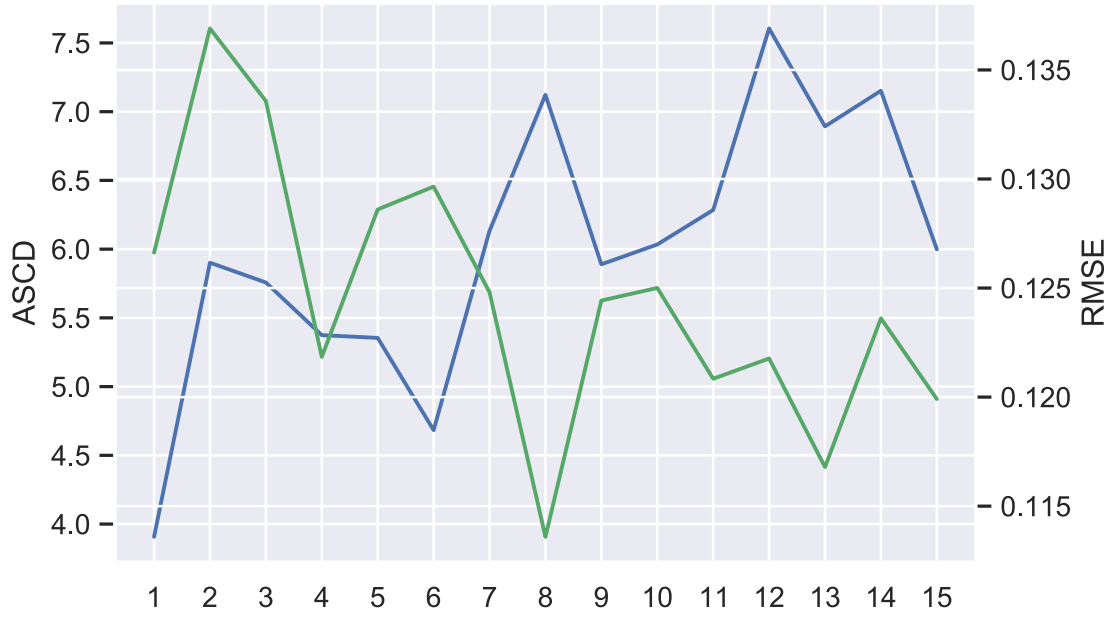

Figure 7.25: Mean Similarity of healthy root shapes and their assigned prosthesis shapes for different values of k, given as the ASCD (blue, left axis) and the RMSE (green, right axis).

centers through the decoder and each test image $I_i, i = 1, \ldots, N_{test}$ was compared to its corresponding prosthesis type image. This was performed for $k = 1, \ldots, 20$ while four different metrics for the image comparison were used: The Jaccard similarity, the Hausdorff distance, the RMSE and the average symmetric contour distance (ASCD). To compute Jaccard similarity, Hausdorff distance and ASCD, the synthesized prosthesis image was segmented using thresholding ($t = 0.31$, corresponds to a grayscale value of 80) and compared to the segmentation of the test image, respectively. This holds for all calculations of these metrics within this work.

Results of Typification and Discussion As the data set used in this study contains paired data, a quantitative analysis is presented in addition to a qualitative analysis of the predicted and reconstructed images of the prosthesis shapes. The presentation of the results and their discussion is divided into typification and type classification, followed by an outlook focusing on practical challenges aiming at clinical application.

The hyperparameter analysis revealed that an architecture with $n_c = 2$, $n_f = 16$ and $n_l = 20$ provides the best image reconstruction accuracy with an RMSE of 0.09 ± 0.03. Fig. 7.24 exemplarily shows the synthesized images of a set of identified prosthesis shapes

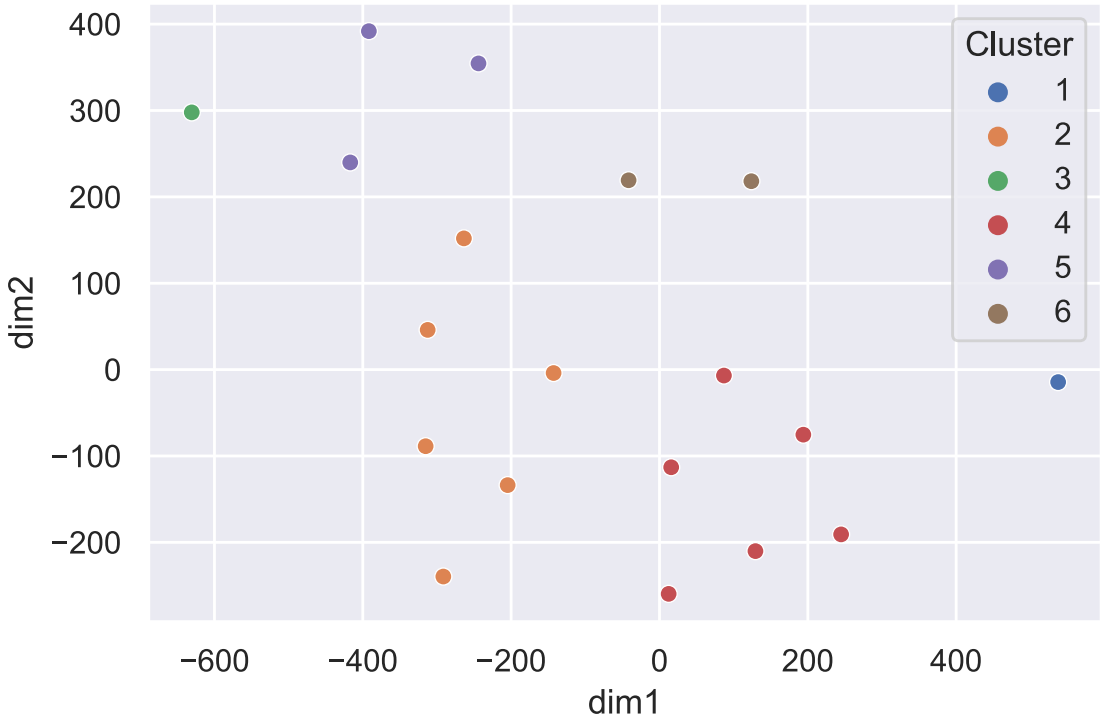

Figure 7.26: tSNE embedding of the healthy training images, separated by their respective cluster, i.e. shape type.

with $k = 3$. All of them look realistic and they complement each other well, indicating that an automatic typification of aortic root shapes using clustering in latent space is possible. Given the optimal architecture of the autoencoder, Fig. 7.25 shows the capability of approximating all healthy test images by their respective prosthesis in dependency of the number of prosthesis shapes k. As expected, the RMSE decreases with increasing k. However, the ASCD increases, indicating that, regarding the shape, a smaller number of prosthesis types leads to better results. $k = 6$ can be identified as an optimal value to assure reasonable values in both metrics. Average results for all metrics are given in Tab. 7.7. Fig. 7.26 shows the tSNE embedding [142] of the healthy training images as well as their identified shape type, respectively, for $k = 6$. In general, the clusters are consistent and homogeneous. However, some shape types, e.g. cluster 1 or cluster 3, only appear very rarely, which might be due to the small data set.

Aortic Root Type Classification

For discrete pseudohealthy synthesis, it is proposed to train a convolutional neural network (CNN) to predict the individually optimal prosthesis type based on an image of the pathological aortic root shape. As the decoder is known, it is possible to synthesize an image of the predicted prosthesis shape. In this study, the CNN was trained in a supervised way, where the ground truth label is retrieved by assigning each aortic root in the training set its corresponding cluster in latent space. Hence, in its current form, the method assumes that paired data is available, i.e. that for each patient in the training data set, the healthy as well as the pathological shape is known.

Similar to previous experiments, a parameterized architecture is proposed. It follows a VGG-like structure [132] and consists of m_b convolutional blocks, each consisting of m_c convolutional layers with m_f and *ReLU* activation followed by a 2×2 average pooling layer. After a flattening operation, m_d fully-connected layers are following with m_n neurons each and *ReLU* activation. Finally, the output layer is attached, featuring *softmax* activation. Fig. 7.27 shows the architecture of the classification model. The training was performed using the *adam* optimizer, binary crossentropy loss and a batch size of 12.

The optimal architecture was identified by evaluating numerous combinations of the hyperparameters m_b, m_c, m_f, m_d and m_n. Thus, as for the typification, a 10-fold Monte-Carlo crossvalidation was performed over all training images (80% training, 20% validation) for each combination of the hyperparameter values given in table 7.6. Within each fold, the autoencoder was trained on the training data using the optimal hyperparameters identified as described in Sec. 7.3.1. The clustering was performed as explained above with a fixed value of $k = 6$ and the optimal cluster centers were assigned to the training and the test images. Then, the CNN was trained on the training data to predict the optimal cluster center based on a pathological image. After training, the CNN was used to estimate the cluster centers for the pathological test data and the classification accuracy was assessed. Additionally, the healthy test images $I_i, i = 1, \dots, N_{test}$ were compared to the synthesized image of the predicted prosthesis $\hat{I}_j^c$ where j is the classification result, respectively, once again using Jaccard similarity, Hausdorff distance, RMSE and ASCD as metrics.

It is important to note that due to the random data splitting during crossvalidation, the training data set is typically imbalanced, i.e. some typical shapes and hence classes occur more often than others. To overcome this class imbalancing issue, oversampling was applied to the minority classes in the training data set so that all classes, i.e. prosthesis shapes, are represented equally [50].

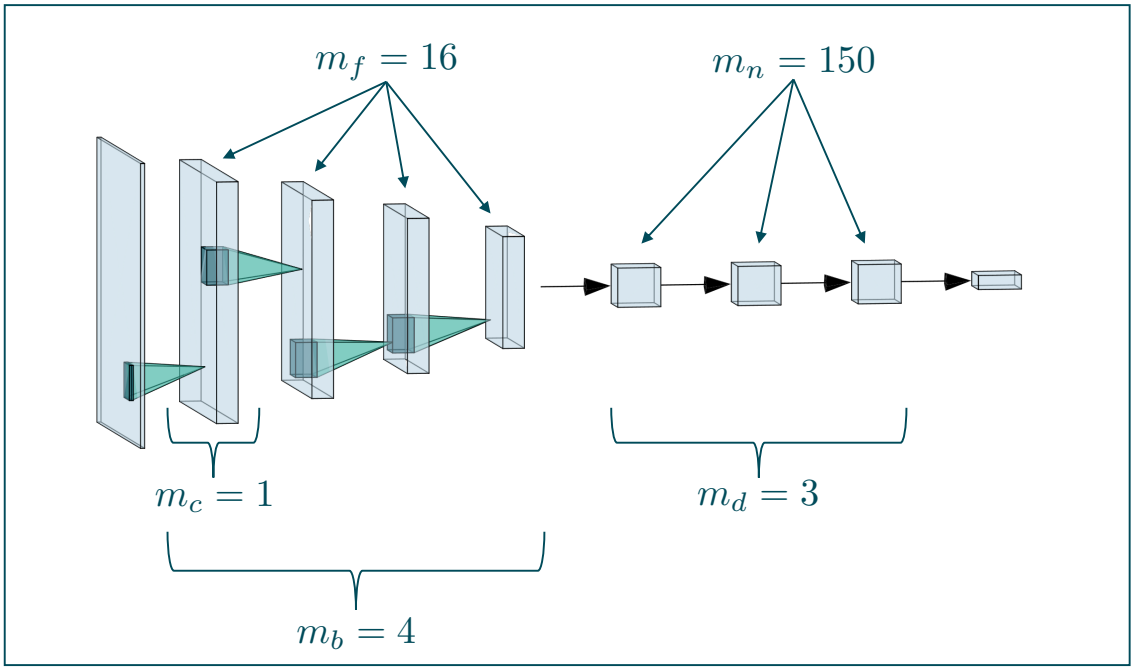

Figure 7.27: Architecture of the classification network, exemplarily shown for $m_b = 4$, $m_c = 1$, $m_f = 16$, $m_d = 3$ and $m_n = 150$.

Results of Type Classification and Discussion Based on the optimal autoencoder architecture, the highest classification accuracy achieved during hyperparameter analysis for the CNN was $77.5 \pm 17.5\%$ with the parameters
$m_b = 4$, $m_c = 1$, $m_f = 16$, $m_d = 3$ and $m_n = 150$. With a value of 67.1%, the classification accuracy was slightly lower on the hold-out test data set. Table 7.7 shows the similarity between the test images and the predicted prosthesis types for all four metrics. For comparison, the latter one is also given for the typification, i.e. with an optimal prosthesis choice. The classification accuracy indicates that the model is able to learn a relationship between the individual pathological root and the desired prosthesis shape. Additionally, the small difference between the classification model and the typification with optimally assigned prosthesis types shows that the classification works robustly and that most of the error relates to the discretization of the prosthesis shapes. The Dice coefficient for the type classification on the test set was 0.62 ± 0.17. Overall, the Jaccard similarity is low. This is most likely due to the relatively thin aortic root wall in the images. Even if a prosthesis approximates a root shape sufficiently, the overlap might still be quite small. Hence, the Hausdorff distance and the ASCD might be more meaningful on this data set. With an ASCD of around 4.5 pixels and a Hausdorff distance of around 8 pixels, the

Table 7.7: Results of the comparison of healthy images and their corresponding prosthesis shapes for typification (optimal prosthesis) and type classification (estimated prosthesis) with $k = 6$, given as ASCD $[pixel]$, Hausdorff distance $[pixel]$, Jaccard similarity and RMSE. For comparison, the results are also given for the continuous approach proposed in [58].

Method	ASCD	Hausdorff	Jaccard	RMSE
Typification on validation set	4.68 ± 4.04	14.48 ± 8.62	0.39 ± 0.11	0.13 ± 0.05
Type Classification on test set	4.48 ± 2.59	8.03 ± 7.60	0.43 ± 0.09	0.12 ± 0.03
Continuous benchmark on test set	24.39 ± 39.07	26.54 ± 38.51	0.17 ± 0.11	0.16 ± 0.01

predicted prosthesis shapes seem to fit the desired root shapes adequately. Example prediction results can be found in Fig. 7.28. As a benchmark, the discrete approach was compared to the continuous, regression-based approach presented in Sec. 6.3.1 utilizing the *Fixed Translation* agent. The proposed method outperforms the regression model by far. This also holds for the best agent, i.e. an agent utilizing $K = 2$ shape primitives, that reaches a Jaccard similarity of 0.41 ± 0.04. It is assumed that this is due to the implicit constraint to realistic and typical shape types. Hence, no out-of-bag-predictions are possible in the discrete approach, making it more robust than the flexible regression approach.

Conclusion and Outlook

Overall, the results indicate that discrete pseudohealthy synthesis for personalized prosthesis shaping in the case of a morphological pathology is possible. Specifically, the type classification seems to work adequately accurately, while most of the errors are introduced by the typification. This might be due to the very small data set, as only 14 training samples are available within each fold. A bigger data set might provide better typification results. Hence, data collection should be an important part of future work. Additionally, the transfer to human data is of high interest regarding clinical application.

Even though discrete pseudohealthy synthesis outperformed a regression approach on this data set, the method is barely capable of dealing with out-of-bag samples, e.g. shapes that are very different from the identified typical ones. For these cases, a robust regression-based shape prediction might be a better choice.

As mentioned above, the classification method presented in this study assumes that the model can be trained on paired training data. With an increasing number of long-term screening studies with large cohorts, like for example the SHIP study [75], it is likely that paired data is available for a wide range of applications. However, extending the framework to be capable of dealing with unpaired data is an interesting research question

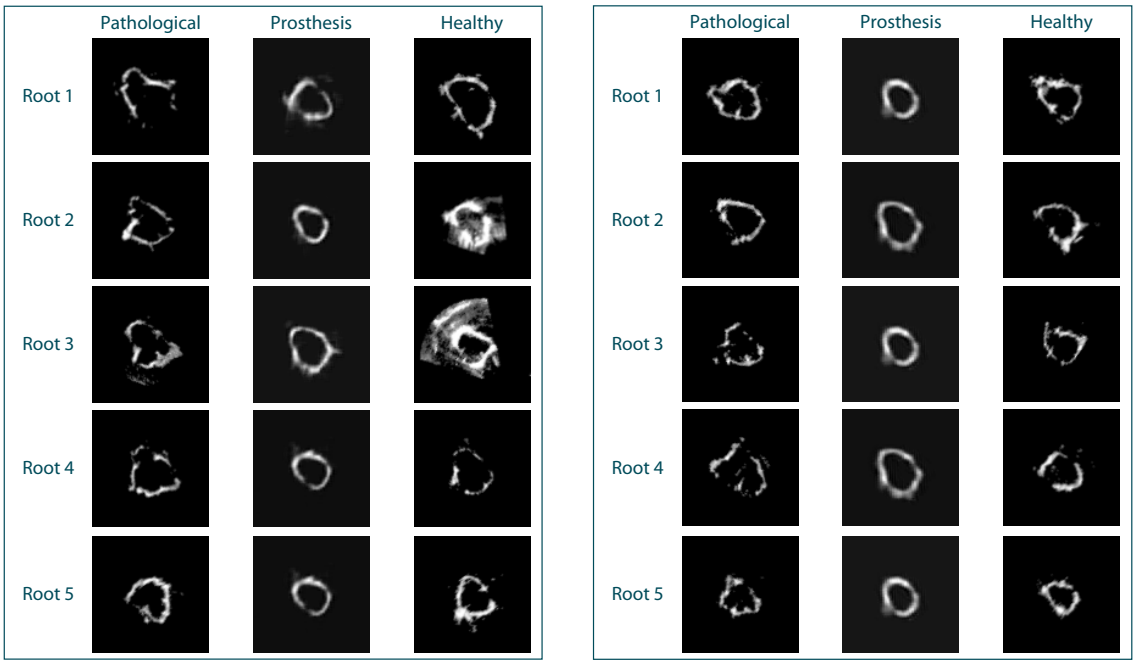

Figure 7.28: Qualitative results of type classification for the test images of two random folds (a) and (b). For each valve, the pathologically dilated shape, the predicted prosthesis shape as well as the healthy ground truth is shown.

and should be addressed in future work. Additionally, it might be possible to generalize from this study on the given small, paired data set to larger, unpaired ones using transfer learning.

In this proof-of-concept study, the reliable error quantification was possible due to the available ground truth segmentation on the 2D image slices. The presented method is easily extendable to also work on 3D volumes by only adding another dimension to all convolutional layers in the autoencoder as well as in the classification network. One might also think about utilizing geometric deep learning to process pointclouds instead of volumetric data. The framework for typification and type classification is capable of working with such kind of data, highlighting the flexibility of the proposed approach.

In contrast to the continuous approach presented in Sec. 6.3.1, the discrete approach in its current form does not provide candidate prediction. To integrate the surgeon into the decision making pipeline, this could be included easily by not only classifying the optimal shape but for example the best three ones. This is a classical concept in computer

vision typically used for assessing error rates in image classification tasks [115]. Another interesting option would be to introduce ordinal classification to the discrete approach. Thus, miclassifications would be penalized differently, depending on the similarity or dissimilarity between the misclassified prosthesis shape and the desired one. Hence, a wrongly classified but sufficiently similar prosthesis would be penalized less than a highly dissimilar prosthesis. This would force the classifier to avoiding strong misclassifications while small errors are still acceptable, resembling the nature of a discretized prosthesis shaping approach.

In this work, the novel approach for personalized prosthesis shaping, called discrete pseudohealthy synthesis, was adapted to the case of a morphological pathology. Thus, a framework for fully automatic aortic root shape typification as well as a type classification approach was developed to estimate the optimal prosthesis type for a given pathological morphology. Furthermore, a proof-of-concept study on personalized aortic root prosthesis shaping was presented, including a vast hyperparameter analysis. The results indicate that approximating the variance of natural aortic root shapes using a specific set of prosthesis types is possible and that the pathologically dilated aortic root shape carries enough information to classify the optimal prosthesis type only based on this dilated shape. As regulatory challenges and manufacturing costs are way lower for a finite set of typical prostheses instead of fully personalized ones, this study presents an important step towards clinical application of personalized prosthetics in morphological pathology scenarios.

7.4 Conclusion

The aim of this chapter was to answer the research question

Q3: Is it possible to identify optimal sets of general prostheses?

Therefore, a framework for automatic typification was developed, proposing to identify typical shapes by performing clustering in a latent space description. Based on this typification, classification models were presented to predict the individually optimal prosthesis shape type. The whole framework was evaluated for the case of structural as well as morphological pathologies.

In both cases, it was possible to identify typical shapes completely data-driven. All typical shapes, i.e. prostheses candidates, appeared realistic and all shapes within one set were complementary. Even though the accuracy of the classification models vary between the pathological cases, the comparison of the predicted shapes with their respective ground truths reveals a good approximation of the individual anatomy by the set of shape types. Hence, the research question *Q3* can be answered with yes: It is not only possible to identify optimal sets of typical shapes for both kinds of pathologies, but it is also possible to predict the individually optimal one given surrogate information on the patient's anatomy by introducing the concept of discrete pseudohealthy synthesis.

In the case of a morphological pathology, the discrete approach even outperformed the continuous pseudohealthy synthesis method presented in Sec. 6.3.1 regarding the Dice coefficient (0.62 ± 0.17 vs. 0.58 ± 0.08). While these area-based metrics are still higher for the continuous approach for a structural pathology, the discrete method provided a lower ASCD (2.58 ± 1.42 vs. 2.84 ± 5.01), indicating a better overall contour line fitting. These results are surprising as a discretization of the possible shapes should lead to a coarser approximation. However, they indicate that the classification problem is easier to solve than the regression problem, which might be due to its nature: The classification method is not only constrained to predict realistic shapes but also to predict shapes that are typical. Hence, the chances to predict very poorly fitting shapes is comparably low. As these advantages become less relevant with more observations, it is possible that the regression-based method might outperform the discrete approach on a larger data set. Such a comparison should be part of future work. However, the continuous approach presented in Sec. 6.3.1 still provides relevant advantages like the possibility for candidate prediction.

The results of this chapter show that approximating the variety of anatomical shapes using a set of prostheses is a promising alternative to full personalization. By introducing

a novel concept for discrete pseudohealthy synthesis, it is possible to identify an optimal set of typical prostheses shapes completely data-driven and classify the optimal type given a patient's surrogate information. Due to the promising results and the lower costs as well as regulatory challenges, this discrete method presents an interesting approach to translate personalized cardiovascular prostheses shaping into clinical application.

8 Discussion

In the previous chapters, all research questions could be answered with yes. This chapter aims at framing these results in a general context. Therefore, basic methodological choices are discussed and limitations as well as the significance of the results are described. Finally, the possibilities for further research that open up on the basis of this work are presented.

Throughout this thesis, representation learning was achieved utilizing various kinds of autoencoder models. Another classical method for describing the variety of anatomical shapes is the usage of Statistical Shape Models (SSM). Besides their limited capabilities of dealing with complex manifolds [65], comparative studies revealed that SSMs outperform autoencoders in terms of generalization only in scenarios with very small data sets [140]. If the number of observations is greater than 20, it was shown that autoencoders outperform all assessed SSM variants. As the clinical application of the methods developed in this thesis requires a sufficiently large data set to cover the full variety of shapes, it can be assumed that autoencoders are the superior concept for representing anatomical shapes. Hence, this thesis proposes the usage of autoencoder variants throughout all developed methods.

It should be noted that most methods developed in this thesis were evaluated using crossvalidation. However, in contrast to suggestions from the literature [116], experiments for analyzing hyperparameters were also performed using a crossvalidation on the full data set without a hold-out test set, with an exception in Chap. 7, where a hold-out evaluation was performed in the morphological pathology case. Exploring the hyperparameters on the full data set was done intentionally to maximize the number of data samples in the data subsets for training and evaluating the model. As the size of all data sets used in the scope of this thesis is relatively small, this approach was necessary to ensure reliable answers to the research questions. As stated in the introduction, the goal of this work was methodological development. Therefore, the presented results fulfill the role of proof-of-concept studies rather than accuracy benchmarks. As the used data sets are sufficiently similar but still not clinically realistic, the final hyperparameters have to be fine-tuned on clinical data anyway and the exact quantitative results are not in focus.

Hence, the results of this work are meaningful and significant even though no hold-out test set was used during hyperparameter optmization.

This is related to the fact that throughout the thesis, different autoencoder architectures were identified as optimal dealing with the same data. This shows that a final judgement of the set of hyperparameters cannot be done on these limited data sets. However, it also highlights the robustness of the developed frameworks as adequate results could be achieved for different kinds of autoencoders. Once again, the final set of hyperparameters should be identified using a clinically realistic data set.

All results presented in this work were achieved on idealized ex-vivo porcine data sets. During image acquisition, there was neither movement of the organ nor surrounding tissue harming the image analysis. In addition, the manual dilation of the aortic roots to simulate a morphological pathology is not likely to mimic a real pathological deformation exactly. However, all approaches developed in this thesis have in common that no deformation or image translation is modeled explicitly. Every approach presented in this thesis relies on data-driven learning to predict the desired healthy shape based on the available surrogate information. Therefore, the results can be interpreted as proof-of-concepts that the developed methods are capable of performing accurately for structural as well as morphological pathologies. Hence, it can be assumed that they also provide an adequate solution on clinical data of realistic pathologies.

In the case of a structural pathology, it was assumed that the shape of a valve prosthesis can be completely described using the shapes of its three leaflets. Obviously, this assumption is greatly simplifying as the stent geometry has a strong influence on the valve's shape. However, the stent's shape is highly constrained by the commissure lines of the three leaflets and, hence, is implicitly learned by the presented models. Furthermore, if the commissure lines can be assessed in 3D, the optimal stent shape can be predicted from the surrogate information in the same way the leaflet shapes are predicted. Hence, the methods developed in this thesis deliver a promising approach for also solving the stent shaping problem. One possibility of assessing the 3D commissure line is the imaging of silicone molds of aortic roots [139].

The images of the leaflets in the data sets collected in the scope of this thesis were acquired in a planar, unpressurized state. As several sources of stress and strain apply to them during the cardiac cycle, they might appear with different dimensions in the collected data sets than in clinical data. However, in the scope of personalized prosthesis development, this is actually desired. If the fabrication material closely resembles the biomechanical properties of the native leaflet tissue, the artificial leaflets deformation un-

der stress and strain will be very close to the native one. Thus, the unpressurized state of the collected data is beneficial for developing personalized aortic valve prostheses.

The field of medical imaging keeps on advancing and, hence, it can be assumed that imaging currently unassessible structures like the aortic valve's leaflets might be possible in the near future. However, the presented approaches for predicting the optimal prosthesis shape in the case of a structural pathology are not limited to the predicting the shape. If the latent representation is capable of encoding more detailed features, it should be possible to predict these features as well. In the case of personalized aortic valve prostheses, it might be possible to encode not only the leaflet's shape but also its micro-structure, i.e. the distribution and thickness of the prominent collagen fibers on the leaflet. As they have a strong influence on the biomechanical behaviour of the valve, reproducing them might be an important long-term goal. However, as the fiber's size is in the micrometer range, it is unlikely that it is possible to assess them using in-vivo medical imaging in the near future. Similarly, the prediction could be performed for even smaller structures on a microscopic scale. The only constraint is that ex-vivo data collection is necessary and that the representation learning model should be capable of encoding the information adequately. The methodological framework itself stays the same. This highlights the great flexibility of the developed approaches and the relevance of the them, even if the field of medical imaging advances.

In this thesis, it could be shown that personalized shaping of cardiovascular prostheses is possible using pseudohealthy synthesis. Thus, proof-of-concept studies provided acceptable qualitative and quantitative results on ex-vivo porcine data sets. However, the desired target accuracy remains unclear. Since the problem of identifying the individually optimal prostheses shape had not been solved, the field of cardiovascular surgery lacks of personalized prosthetics. Therefore, the influence of a personalized prosthesis on the blood flow and on the patient's outcome in general has never been scientifically investigated. This work presents the basis for such investigations and therefore potentially for an improved quality of life for the patient. Based on the results of this work, comparisions between native aortic roots or valves, state-of-the-art implants as well as personalized ones can - in cooporation with a prothesis manufacturer - easily be examined in a left heart simulator setup [14]. Therefore, state-of-the-art prostheses as well as personalized prostheses should be assessed and compared to the measurement of the native anatomies. Such an experiment would allow for quantifying the potential superiority of personalized cardiovascular prosthetics compared to classical ones. In the scope of this experiment, it might also be interesting to compare the continuous, regression-based approaches presented in Chap. 6 to the discrete ones proposed in Chap. 7.

9 Conclusion

In this thesis, it was investigated whether personalized shaping of cardiovascular prostheses can be achieved using pseudohealthy synthesis. Thus, the problem of estimating the desired healthy shape only based on available surrogate information was formulated as a pseudohealthy synthesis problem for the first time. Two different pathological cases that lead to different prediction methods were identified and examined: structural pathology with the example of personalized aortic valve prostheses and morphological pathology with the example of personalized aortic root prostheses. For both pathological cases, acquisition setups and experimental workflows to acquire ex-vivo porcine data sets were identified and developed. The first data sets for training and evaluating pseudohealthy synthesis methods in the scope of personalized cardiovascular prostheses were collected for both pathological cases, resulting in five data sets in total.

Al three research questions could be answered with yes.

Q1: It could be shown that the surrogate data carries a sufficient amount of information to predict geometric key features of the desired prosthesis shape for both pathological cases.

Q2: It could be shown that it is possible to synthesize an image of the full desired healthy shape only based on the surrogate information for both pathological cases.

Q3: It could be shown that it is possible to identify a set of typical prostheses shapes and that a classification of the individually optimal prosthesis type is possible.

Therefore, it can be concluded that personalized cardiovascular prosthesis shaping is possible and that pseudohealthy synthesis presents a promising approach to solve it. As classical methods for pseudohealthy synthesis are not sufficient for this application, novel concepts, methods and frameworks for pseudohealthy synthesis were developed, specifically tailored to personalized prosthesis shaping. All of them were validated in proof-of-concept studies and showed a remarkable performance. They are the first methods to predict the shape of cardiovascular prostheses and hence present the basis for personalizing them. The novel approach of interpreting image deformation as a translation in latent space and solving this using deformation agents presents a completely new

concept for learning-based image deformation. The application possibilities are manifold and range from image registration to time-critical simulations.

To increase the clinical acceptance, different approaches for including the surgeon into the decision making pipeline were proposed and evaluated. Furthermore, the novel concept of discrete pseudohealthy synthesis addresses the high costs and regulatory challenges of individually tailoring prostheses for each and every patient by providing a set of general prosthesis shapes that can be manufactured in serial production. By proposing and proving a framework for identifying this set of prostheses in a completely data-driven way as well as predicting the individually optimal shape type for a patient, the possibility of translating the personalization concept to clinical application increased significantly. Additionally, from an academic perspective, this framework presents the first discrete approach for pseudohealthy synthesis and might be applied in a wide range of applications.

The methodological advancements presented in this thesis lay the basis for manufacturing personalized cardiovascular prostheses and benchmarking them against state-of-the-art ones. Since all pseudohealthy shape synthesis methods developed in the scope of this thesis work model-free, their application is not limited to the cardiovascular domain. In general, each kind of prosthesis can be personalized using the presented methods as long as sufficient data sets are available. Therefore, this thesis opens up the new subfield of machine learning based personalized medicine, aiming at prostheses that mimic the patient's original, healthy organ shape as close as possible.

Bibliography

[1] T. Al-Atassi, H. D. Toeg, R. Jafar, B. Sohmer, M. Labrosse, and M. Boodhwani. Impact of Aortic Annular Geometry on Aortic Valve Insufficiency: Insights from a Preclinical, Ex Vivo, Porcine Model. *The Journal of Thoracic and Cardiovascular Surgery*, 150(3):656–664, 2015.

[2] S. Andermatt, A. Horváth, S. Pezold, and P. Cattin. Pathology Segmentation Using Distributional Differences to Images of Healthy Origin. In A. Crimi, S. Bakas, Kuijf H, Keyvan F, Reyes M, and van Walsum T, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018, Lecture Notes in Computer Science*, volume 11383. Springer, Cham, 2019.

[3] M. Arsalan and T. Walther. Durability of Prostheses for Transcatheter Aortic Valve Implantation. *Nature Reviews Cardiology*, 13(6):360–367, 6 2016.

[4] I. Assent. Clustering High Dimensional Data. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(4):340–350, 7 2012.

[5] P. Astudillo, P. Mortier, J. Bosmans, O. De Backer, P. de Jaegere, M. De Beule, and J. Dambre. Enabling Automated Device Size Selection for Transcatheter Aortic Valve Implantation. *Journal of Interventional Cardiology, Article ID 3591314*, 2019, 11 2019.

[6] J. Barrios-Muriel, F. Romero-Sánchez, F. J. Alonso-Sánchez, and D. R. Salgado. Advances in Orthotic and Prosthetic Manufacturing: A Technology Review. *Materials, Article Number 295*, 13(2), 5 2020.

[7] M. Bashir, A. Harky, S. Froghi, B. Adams, M. Garner, P. Gupta, A. Oo, and R. Uppal. A Systematic Review and Meta-Analysis of Mechanical vs Biological Composite Aortic Root Replacement, Early and 1-year Results. *General Thoracic and Cardiovascular Surgery*, 67(1):70–76, 1 2019.

[8] C. F. Baumgartner, L. M. Koch, K. C. Tezcan, J. X. Ang, and E. Konukoglu. Visual Feature Attribution Using Wasserstein GANs. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8309–8319. IEEE, 6 2018.

[9] C. Baur, B. Wiestler, S. Albarqouni, and N. Navab. Deep Autoencoding Models for Unsupervised Anomaly Segmentation in Brain MR Images. In Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, and van Walsum T, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science*, volume 11383. Springer, Cham, 2019.

[10] T. Bechsgaard, T. Lindskow, T. Lading, J. Hasenkam, D. Røpcke, H. Nygaard, P. Johansen, and S. L. Nielsen. Biomechanical Characterization of the Native Porcine Aortic Root. *Journal of Biomechanics*, 74:156–162, 6 2018.

[11] A. Beckmann, R. Meyer, J. Lewandowski, A. Markewitz, and J. Gummert. German Heart Surgery Report 2019: The Annual Updated Registry of the German Society for Thoracic and Cardiovascular Surgery. *The Thoracic and Cardiovascular Surgeon*, 68(04):263–276, 6 2020.

[12] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle. Greedy Layer-Wise Training of Deep Networks, 2007.

[13] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl. Algorithms for Hyper-Parameter Optimization. In Shawe-Taylor J, Zemel R, Bartlett P, Pereira F, and Weinberger K Q, editors, *Advances in Neural Information Processing Systems*, volume 24, pages 2546–2554. Curran Associates, Inc., 2011.

[14] L. P. Bezerra and C. N. Pai. A Left Heart Ventricle Simulator Manufactured by 3D Printing. *IFAC Proceedings Volumes*, 47(3):11599–11604, 2014.

[15] M. Bode, M. Mueller, H. Zernetsch, and B. Glasmacher. Electrospun Vascular Grafts with Anti-Kinking Properties. *Current Directions in Biomedical Engineering*, 1(1):524–528, 9 2015.

[16] M. Bongert, M. Geller, W. Pennekamp, and V. Nicolas. Simulation of Personalised Haemodynamics by Various Mounting Positions of a Prosthetic Valve using Computational Fluid Dynamics. *Biomedical Engineering / Biomedizinische Technik*, 64(2):147–156, 4 2019.

[17] B. E. Boser, I. M. Guyon, and V. N. Vapnik. A Training Algorithm for Optimal Margin Classifiers. In *Proceedings of the fifth Annual Workshop on Computational Learning Theory - COLT '92*, pages 144–152, New York, New York, USA, 1992. ACM Press.

[18] G. M. Bosi, C. Capelli, M. H. Cheang, N. Delahunty, M. Mullen, A. M. Taylor, and S. Schievano. A Validated Computational Framework to Predict Outcomes in TAVI. *Scientific Reports, Article Number 9906*, 10(1), 5 2020.

[19] C. Bowles, C. Qin, C. Ledig, R. Guerrero, R. Gunn, A. Hammers, E. Sakka, D. A. Dickie, M. V. Hernández, N. Royle, J. Wardlaw, H. Rhodius-Meester, B. Tijms, A. W. Lemstra, W. van Der Flier, F. Barkhof, P. Scheltens, and D. Rueckert. Pseudo-Healthy Image Synthesis for White Matter Lesion Segmentation. In Tsaftaris S, Gooya A, Frangi A, and Prince J, editors, *Simulation and Synthesis in Medical Imaging. SASHIMI 2016. Lecture Notes in Computer Science*, volume 9968, pages 87–96. Springer, Cham, 2016.

[20] L. Breiman. Random Forests. *Machine Learning*, 45(1):5–32, 2001.

[21] M. M. Bronstein, J. Bruna, Y. Lecun, A. Szlam, and P. Vandergheynst. Geometric Deep Learning: Going beyond Euclidean Data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

[22] J. L. Bruse, S. Schievano, M. A. Zuluaga, A. Khushnood, K. McLeod, H. N. Ntsinjana, T.-Y. Hsia, M. Sermesant, X. Pennec, and A. M. Taylor. Detecting Clinically Meaningful Shape Clusters in Medical Image Data: Metrics Analysis for Hierarchical Clustering Applied to Healthy and Pathological Aortic Arches. *IEEE Transactions on Biomedical Engineering*, 64(10):2373–2383, 10 2017.

[23] J. G. Byrne, T. Gubjartsson, A. N. Karavas, T. Mihaljevic, B. J. Phillips, S. F. Aranki, J. D. Rawn, and L. H. Cohn. Biological vs. Mechanical Aortic Root Replacement. *European Journal of Cardio-Thoracic Surgery*, 23(3):305–310, 3 2003.

[24] J. Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, 11 1986.

[25] C.-C. Chang and C.-J. Lin. LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[26] S. Chechani, R. Suresh, and K. A. Patwardhan. Aortic Root Segmentation in 4D Transesophageal Echocardiography. In K. Mori and N. Petrick, editors, *Proc. SPIE 10575, Medical Imaging 2018: Computer-Aided Diagnosis*, volume 105750W. SPIE, 2 2018.

[27] X. Chen, D. P. Kingma, T. Salimans, Y. Duan, P. Dhariwal, J. Schulman, I. Sutskever, and P. Abbeel. Variational Lossy Autoencoder. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017.

[28] X. Chen and E. Konukoglu. Unsupervised Detection of Lesions in Brain MRI using Constrained Adversarial Auto-Encoders. In *Proceedings of the 1st International Conference on Medical Imaging with Deep Learning*. OpenReview.net, 6 2018.

[29] F. Chollet. keras, 2015.

[30] S. J. Crick, M. N. Sheppard, S. Y. Ho, L. Gebstein, and R. H. Anderson. Anatomy of the Pig Heart: Comparisons with Normal Human Cardiac Structure. *Journal of Anatomy*, 193(1):105–119, 7 1998.

[31] T. E. David, C. David, A. Woo, and C. Manlhiot. The Ross Procedure: Outcomes at 20 Years. *The Journal of Thoracic and Cardiovascular Surgery*, 147(1):85–94, 1 2014.

[32] T. E. David and C. M. Feindel. An Aortic Valve-Sparing Operation for Patients with Aortic Incompetence and Aneurysm of the Ascending Aorta. *The Journal of Thoracic and Cardiovascular Surgery*, 103(4):617–622, 1992.

[33] R. De Paulis, R. Scaffa, A. Salica, L. Weltert, and I. Chirichilli. Biological Solutions to Aortic Root Replacement: Valve-Sparing versus Bioprosthetic Conduit. *Journal of Visualized Surgery*, 4(5):855–861, 5 2018.

[34] R. Demirci. *Generative Adversarial Networks for Estimating the Healthy Anatomy from Ultrasound Images of the Pathological Aortic Root [M.Sc. Thesis]*. Universität zu Lübeck, 2020.

[35] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255. IEEE, 6 2009.

[36] T. C. Doehring, M. Kahlen, and I. Vesely. Mesostructures of the Aortic Valve. *Journal of Heart and Valve Disease*, 14(5):679–686, 2005.

[37] C. Doersch. Tutorial on Variational Autoencoders. *arXiv preprint, https://arxiv.org/abs/1606.05908*, 6 2016.

[38] P. M. Dohmen. Tissue Engineered Aortic Valve. *HSR Proceedings in Intensive Care and Cardiovascular Anesthesia*, 4:89–93, 2012.

[39] J. Dominik and P. Zacek. *Heart Valve Surgery*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[40] A. Erasmi, H.-H. Sievers, M. Scharfschwerdt, T. Eckel, and M. Misfeld. In Vitro Hydrodynamics, Cusp-Bending Deformation, and Root Distensibility for Different Types of Aortic Valve-Sparing Operations: Remodeling, Sinus Prosthesis, and Reimplantation. *The Journal of Thoracic and Cardiovascular Surgery*, 130(4):1044–1049, 10 2005.

[41] F. Ernst and A. Schweikard. *Fundamentals of Machine Learning - Support Vector Machines Made Easy*, volume 1. UVK Verlag, Tübingen, 2020.

[42] C. D. Etz, F. F. Girrbach, K. von Aspern, R. Battellini, P. Dohmen, A. Hoyer, M. Luehr, M. Misfeld, M. A. Borger, and F. W. Mohr. Longevity After Aortic Root Replacement: Is the Mechanically Valved Conduit Really the Gold Standard for Quinquagenarians? *Circulation, Article Number 18168*, 128(21):– undefined, 9 2013.

[43] B. Fan, N. Tomii, H. Tsukihara, E. Maeda, H. Yamauchi, K. Nawata, A. Hatano, S. Takagi, I. Sakuma, and M. Ono. Attention-Guided Decoder in Dilated Residual Network for Accurate Aortic Valve Segmentation in 3D CT Scans. In H. e. a. Liao, editor, *Machine Learning and Medical Engineering for Cardiovascular Health and Intravascular Imaging and Computer Assisted Stenting. MLMECH 2019, CVII-STENT 2019. Lecture Notes in Computer Science*, volume 11794. Springer, Cham, 2019.

[44] M. Galiñanes, A. Meduoye, I. Ferreira, and A. Sosnowski. Totally Biological Composite Aortic Stentless Valved Conduit for Aortic Root Replacement: 10-Year Experience. *Journal of Cardiothoracic Surgery*, 6(86), 12 2011.

[45] A. Gholami, S. Subramanian, V. Shenoy, N. Himthani, X. Yue, S. Zhao, P. Jin, G. Biros, and K. Keutzer. A Novel Domain Adaptation Framework for Medical Image Segmentation. In Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, and van Walsum T, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science*, volume 11384, pages 289–298. Springer, Cham, 2019.

[46] P. Gifani, H. Behnam, A. Shalbaf, and Z. A. Sani. Automatic Detection of End-Diastole and End-Systole from Echocardiography Images using Manifold Learning. *Physiological Measurement*, 31(9), 5 2010.

[47] L. N. Girardi. Composite Root Replacement with a Mechanical Conduit. *Operative Techniques in Thoracic and Cardiovascular Surgery*, 13(3):148–160, 9 2008.

[48] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, Cambridge, 2016.

[49] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative Adversarial Nets. In Ghahramani Z, Welling M, Cortes C, Lawrence N, and Weinberger K Q, editors, *Advances in Neural Information Processing Systems*, volume 27, pages 2672–2680. Curran Associates, Inc., 2014.

[50] A. Gosain and S. Sardana. Handling Class Imbalance Problem using Oversampling Techniques: A Review. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 79–85. IEEE, 9 2017.

[51] K. J. Grande, R. P. Cochran, P. G. Reinhall, and K. S. Kunzelman. Stress Variations in the Human Aortic Root and Valve: The Role of Anatomic Asymmetry. *Annals of Biomedical Engineering*, 26(4):534–545, 7 1998.

[52] S. Grbic, T. Mansi, R. Ionasec, I. Voigt, H. Houle, M. John, M. Schoebinger, N. Navab, and D. Comaniciu. Image-Based Computational Models for TAVI Planning: From CT Images to Implant Deployment. In Mori K, Sakuma I, Sato Y, Barillot C, and Navab N, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2013. MICCAI 2013. Lecture Notes in Computer Science*, volume 8150. 2013.

[53] J. Hagenah. *Individuelle Prädiktion der Prothesengröße bei der Klappenerhaltenden Aortenwurzelrekonstruktion [M.Sc. thesis]*. Universität zu Lübeck, 2016.

[54] J. Hagenah and F. Ernst. Discrete Pseudohealthy Synthesis: Aortic Root Shape Typification and Type Classification with Pathological Prior. *Proceedings of The 4th International Conference on Medical Imaging with Deep Learning, Proceedings of Machine Learning Research*, 143:252–267, 2021.

[55] J. Hagenah, T. Evers, M. Scharfschwerdt, A. Schweikard, and F. Ernst. A Support Vector Regression-Based Data-Driven Leaflet Modeling Approach for Personalized Aortic Valve Prosthesis Development. In *2018 Computing in Cardiology Conference (CinC)*, pages 1–4, 12 2018.

[56] J. Hagenah, M. Heinrich, and F. Ernst. Deep Transfer Learning for Aortic Root Dilation Identification in 3D Ultrasound Images. *Current Directions in Biomedical Engineering*, 4(1):71–74, 9 2018.

[57] J. Hagenah, K. Kuehl, M. Scharfschwerdt, and F. Ernst. Cluster Analysis in Latent Space: Identifying Personalized Aortic Valve Prosthesis Shapes using Deep Representations. In M. J. Cardoso, A. Feragen, B. Glocker, E. Konukoglu, I. Oguz, G. Unal, and T. Vercauteren, editors, *Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning, Proceedings of Machine Learning Research*, volume 102, pages 236–249. PMLR, 2019.

[58] J. Hagenah, M. Mehdi, and F. Ernst. Generating Healthy Aortic Root Geometries From Ultrasound Images of the Individual Pathological Morphology Using Deep Convolutional Autoencoders. In *2019 Computing in Cardiology (CinC)*, pages 1–4. IEEE, 2019.

[59] J. Hagenah, M. Mehdi, and F. Ernst. Fully Data-Driven Pseudohealthy Synthesis for Planning Valve-Sparing Aortic Root Reconstruction using Conditional Variational Autoencoders. *Current Directions in Biomedical Engineering*, 6(3):284–287, 9 2020.

[60] J. Hagenah, M. Scharfschwerdt, and F. Ernst. Towards Personalized Aortic Valve Prostheses - A Sparse Representation of the Individual Leaflet Shape. In *2018 Computing in Cardiology Conference (CinC)*, pages 1–4, 12 2018.

[61] J. Hagenah, M. Scharfschwerdt, A. Schweikard, and C. Metzner. Combining Deformation Modeling and Machine Learning for Personalized Prosthesis Size Prediction in Valve-Sparing Aortic Root Reconstruction. In M. Pop and G. Wright, editors, *Functional Imaging and Modelling of the Heart. FIMH 2017. Lecture Notes in Computer Science*, volume 10263. Springer, Cham, 2017.

[62] J. Hagenah, E. Werrmann, M. Scharfschwerdt, F. Ernst, and C. Metzner. Prediction of Individual Prosthesis Size for Valve-Sparing Aortic Root Reconstruction Based on Geometric Features. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3273–3276. IEEE, 2016.

[63] P. E. Hammer, C. A. Pacak, R. D. Howe, and P. J. del Nido. Collagen Bundle Orientation Explains Aortic Valve Leaflet Coaptation. In Ourselin S., Rueckert D., and Smith N., editors, *Functional Imaging and Modeling of the Heart. FIMH 2013. Lecture Notes in Computer Science*, volume 7945. Springer, Cham, 2013.

[64] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer New York, New York, NY, 2009.

[65] T. Heimann and H.-P. Meinzer. Statistical Shape Models for 3D Medical Image Segmentation: A Review. *Medical Image Analysis*, 13(4):543–563, 8 2009.

[66] G. E. Hinton. Reducing the Dimensionality of Data with Neural Networks. *Science*, 313(5786):504–507, 7 2006.

[67] G. A. Holzapfel and R. W. Ogden. Modelling the Layer-Specific Three-Dimensional Residual Stresses in Arteries, with an Application to the Human Aorta. *Journal of The Royal Society Interface*, 7(46):787–799, 5 2010.

[68] J. L. Honge, J. Funder, E. Hansen, P. M. Dohmen, W. Konertz, and J. M. Hasenkam. Recellularization of Aortic Valves in Pigs. *European Journal of Cardio-Thoracic Surgery*, 39(6):829–834, 6 2011.

[69] D. Horstkotte and F. Loogen. The Natural History of Aortic Valve Stenosis. *European Heart Journal*, 9(suppl E):57–64, 4 1988.

[70] C.-W. Hsu and C.-J. Lin. A Comparison of Methods for Multiclass Support Vector Machines. *IEEE Transactions on Neural Networks*, 13(2):415–425, 3 2002.

[71] H. Huang, Z. Li, Z. Sun, and T. Tan. IntroVAE: Introspective Variational Autoencoders for Photographic Image Synthesis. In Bengio S, Wallach H, Larochelle H, Grauman K, Cesa-Bianchi N, and Garnett R, editors, *Advances in Neural Information Processing Systems*, volume 31, pages 52–63. Curran Associates, Inc., 2018.

[72] M. Ilse, J. M. Tomczak, C. Louizos, M. Welling, and M. W. Nl. DIVA: Domain Invariant Variational Autoencoders. In T. Arbel, , I. Ben Ayed, M. de Bruijne, M. Descoteaux, H. Lombaert, and C. Pal, editors, *Proceedings of the Third Conference on Medical Imaging with Deep Learning, Proceedings of Machine Learning Research*, volume 121, pages 322–348. PMLR, 2020.

[73] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-Image Translation with Conditional Adversarial Networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5967–5976. IEEE, 7 2017.

[74] K. K. Jain. *Textbook of Personalized Medicine*. Springer New York, New York, NY, 2009.

[75] U. John, E. Hensel, J. Lüdemann, M. Piek, S. Sauer, C. Adam, G. Born, D. Alte, E. Greiser, U. Haertel, H. W. Hense, J. Haerting, S. Willich, and C. Kessler. Study of Health in Pomerania (SHIP): A health examination survey in an east German region: Objectives and design. *Sozial- und Präventivmedizin*, 46(3):186–194, 2001.

[76] S. Kaule, S. Siewert, A. Dierke, J. Keiler, A. Wree, N. Grabow, N. Degen, A. Öner, K.-P. Schmitz, and M. Stiehm. 3D-Printing of the Aortic Root for In Vitro Hydrodynamic Assessment of Transcatheter Aortic Valve Prostheses. *Transactions on Additive Manufacturing Meets Medicine*, 1(1), 2019.

[77] E. Keogh and A. Mueen. Curse of Dimensionality. In *Encyclopedia of Machine Learning and Data Mining*. Springer US, Boston, MA, 2017.

[78] P. Khandelwal and P. Yushkevich. Domain Generalizer: A Few-Shot Meta Learning Framework for Domain Generalization in Medical Imaging. In Albarqouni S et al., editor, *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning. DART 2020, DCL 2020. Lecture Notes in Computer Science*, volume 12444, pages 73–84. Springer, Cham, 2020.

[79] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. In Bengio Y and LeCun Y, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 12 2015.

[80] D. P. Kingma, D. J. Rezende, S. Mohamed, and M. Welling. Semi-Supervised Learning with Deep Generative Models. In Ghahramani Z, Welling M, Cortes C, Lawrence N, and Weinberger K Q, editors, *Advances in Neural Information Processing Systems*, volume 4, pages 3581–3589. Curran Associates, Inc., 2014.

[81] D. P. Kingma and M. Welling. Auto-Encoding Variational Bayes. In Bengio Y and LeCun Y, editors, *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014.

[82] W. M. Kouw and M. Loog. A Review of Domain Adaptation without Target Labels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(3):766–785, 3 2021.

[83] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In Pereira F, Burges C J G, Bottou L, and Weinberger K Q, editors, *Advances in Neural Information Processing Systems*, volume 25, pages 1097–1105. Curran Associates, Inc., 2012.

[84] M. Kruse, M. Greuel, F. Kreimendahl, T. Schneiders, B. Bauer, T. Gries, and S. Jockenhoevel. Electro-Spun PLA-PEG-Yarns for Tissue Engineering Applications. *Biomedizinische Technik*, 63(3):231–243, 6 2018.

[85] M. R. Labrosse, C. J. Beller, M. Boodhwani, C. Hudson, and B. Sohmer. Subject-Specific Finite-Element Modeling of Normal Aortic Valve Biomechanics from 3D+t TEE Images. *Medical Image Analysis*, 20(1):162–172, 2 2015.

[86] Y. LeCun, Y. Bengio, and G. Hinton. Deep Learning. *Nature*, 521(7553):436–444, 2015.

[87] Y. LeCun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. E. Hubbard, and L. D. Jackel. Handwritten Digit Recognition with a Back-Propagation Network. In Touretzky D, editor, *Advances in Neural Information Processing Systems*, volume 2, pages 396–404. Morgan-Kaufmann, 1989.

[88] S. A. LeMaire, S. Y. Green, K. Sharma, C. K. Cheung, A. Sameri, P. I. Tsai, G. Adams, and J. S. Coselli. Aortic Root Replacement With Stentless Porcine Xenografts: Early and Late Outcomes in 132 Patients. *The Annals of Thoracic Surgery*, 87(2):503–512, 2 2009.

[89] J. A. Leopold. Cellular Mechanisms of Aortic Valve Calcification. *Circulation: Cardiovascular Interventions*, 5(4):605–614, 8 2012.

[90] C. Li, F. Wang, G. Douglas, Z. Zhang, R. Guidoin, and L. Wang. Comprehensive Mechanical Characterization of PLA Fabric Combined with PCL to Form a Composite Structure Vascular Graft. *Journal of the Mechanical Behavior of Biomedical Materials*, 69:39–49, 5 2017.

[91] J. Li, A. M. Adry, J. Peebles, and L. Schmidt. On the Limitations of First-Order Approximation in GAN Dynamics. In J. Dy and A. Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80, pages 3005–3013. PMLR, 2018.

[92] R. Li, W. Cao, Q. Jiao, S. Wu, and H.-S. Wong. Simplified Unsupervised Image Translation for Semantic Segmentation Adaptation. *Pattern Recognition*, 105, 9 2020.

[93] J. Lin, Y. Pang, Y. Xia, Z. Chen, and J. Luo. TuiGAN: Learning Versatile Image-to-Image Translation with Two Unpaired Images. In Vedaldi A, Bischof H, Brox T, and Frahm JM, editors, *Computer Vision - ECCV 2020. ECCV 2020. Lecture Notes in Computer Science*, volume 12349, pages 18–35. 2020.

[94] M.-Y. Liu, X. Huang, A. Mallya, T. Karras, T. Aila, J. Lehtinen, and J. Kautz. Few-Shot Unsupervised Image-to-Image Translation. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 10550–10559. IEEE, 10 2019.

[95] S. Lloyd. Least Squares Quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 3 1982.

[96] M. Loukas, E. Bilinsky, S. Bilinsky, C. Blaak, R. S. Tubbs, and R. H. Anderson. The Anatomy of the Aortic Root. *Clinical Anatomy*, 27(5):748–756, 7 2014.

[97] G. B. Luciani, F. Santini, and A. Mazzucco. Autografts, Homografts, and Xenografts: Overview on Stentless Aortic Valve Surgery. *Journal of Cardiovascular Medicine*, 8(2):91–96, 2 2007.

[98] J.-M. Marin, K. Mengersen, and C. P. Robert. Bayesian Modelling and Inference on Mixtures of Distributions. In Dey D K and Ra C R, editors, *Handbook of Statistics*, volume 25, chapter Bayesian Thinking, pages 459–507. Elsevier, 2005.

[99] G. Marom, R. Haj-Ali, M. Rosenfeld, H. J. Schäfers, and E. Raanani. Aortic Root Numeric Model: Annulus Diameter Prediction of Effective Height and Coaptation in Post-Aortic Valve Repair. *The Journal of Thoracic and Cardiovascular Surgery*, 145(2):406–411, 2 2013.

[100] S. Marsland. *Machine Learning: An Algorithmic Perspective*. Chapman and Hall/CRC, Boca Raton, 2009.

[101] J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber. Stacked Convolutional Auto-Encoders for Hierarchical Feature Extraction. In Honkela T, Duch W, Girolami M, and Kaski, editors, *Artificial Neural Networks and Machine Learning - ICANN 2011. ICANN 2011. Lecture Notes in Computer Science*, volume 6791. Springer, Berlin, Heidelberg, 2011.

[102] M. Misfeld and H.-H. Sievers. Heart Valve Macro- and Microstructure. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1484):1421–1436, 8 2007.

[103] A. W. Moore. Very Fast EM-Based Mixture Model Clustering Using Multiresolution kd-Trees. In Kearns M, Solla S, and Cohn D, editors, *Advances in Neural Information Processing Systems*, volume 11, pages 543–549. MIT Press, 1999.

[104] A. Myronenko, X. Song, and M. A. Carreira-Perpinán. Non-Rigid Point Set Registration: Coherent Point Drift. In Sch\"{o}lkopf B, Platt J, and Hoffman T, editors, *Advances in Neural Information Processing Systems*, volume 19, pages 1009–1016. MIT Press, 2007.

[105] E. Nalisnick, A. Matsukawa, Y. W. Teh, D. Gorur, and B. Lakshminarayanan. Do Deep Generative Models Know What They Don't Know? In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 10 2018.

[106] M. A. Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2018.

[107] Y. Pang, J. Lin, T. Qin, and Z. Chen. Image-to-Image Translation: Methods and Applications. *arXiv preprint arXiv:2101.08629*, 1 2021.

[108] T. Park, M.-Y. Liu, T.-C. Wang, and J.-Y. Zhu. Semantic Image Synthesis With Spatially-Adaptive Normalization. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2332–2341. IEEE, 6 2019.

[109] S. Pashneh-Tala, S. MacNeil, and F. Claeyssens. The Tissue-Engineered Vascular Graft - Past, Present, and Future. *Tissue Engineering Part B: Reviews*, 22(1):68–100, 2 2016.

[110] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros. Context Encoders: Feature Learning by Inpainting. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2536–2544. IEEE, 6 2016.

[111] M. J. D. Powell. Radial Basis Functions for Multivariable Approximation: A Review. In *Proceedings of the IMA Conference on Algorithms for the Approximation of Functions and Data*, volume 61, pages 143–167. Oxford University Press (to appear). Interpolation by Radial Functions, 1985.

[112] D. Richardt, A. Karluss, C. Schmidtke, H.-H. Sievers, and M. Scharfschwerdt. A New Sinus Prosthesis for Aortic Valve-Sparing Surgery Maintaining the Shape of the Root at Systemic Pressure. *The Annals of Thoracic Surgery*, 89(3):943–946, 3 2010.

[113] K. A. Rocco, M. W. Maxfield, C. A. Best, E. W. Dean, and C. K. Breuer. In Vivo Applications of Electrospun Tissue-Engineered Vascular Grafts: A Review. *Tissue Engineering Part B: Reviews*, 20(6):628–640, 12 2014.

[114] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning Representations by Back-Propagating Errors. *Nature*, 323(6088):533–536, 10 1986.

[115] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3):211–252, 12 2015.

[116] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Pearson Education Inc., New Jersey, 2 edition, 2003.

[117] M. Russo, M. Taramasso, A. Guidotti, A. Pozzoli, F. Nietilspach, L. von Segesser, and F. Maisano. The Evolution of Surgical Valves. *Cardiovascular Medicine*, 20(12):285–292, 12 2017.

[118] M. S. Sacks, D. B. Smith, and E. D. Hiester. The Aortic Valve Microstructure: Effects of Transvalvular Pressure. *Journal of Biomedical Materials Research*, 41(1):131–141, 1998.

[119] M. S. Sacks and A. P. Yoganathan. Heart Valve Function: A Biomechanical Perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1484):1369–1391, 8 2007.

[120] A. Saeyeldin, M. A. Zafar, C. A. Velasquez, K. Ip, A. Gryaznov, A. J. Brownstein, Y. Li, J. A. Rizzo, Y. Erben, B. A. Ziganshin, and J. A. Elefteriades. Natural History of Aortic Root Aneurysms in Marfan Syndrome. *Annals of Cardiothoracic Surgery*, 6(6):625–632, 11 2017.

[121] Y. Sahasakul, W. D. Edwards, J. M. Naessens, and A. Tajik. Age-Related Changes in Aortic and Mitral Valve Thickness: Implications for Two-Dimensional Echocardiography Based on an Autopsy Study of 200 Normal Human Hearts. *The American Journal of Cardiology*, 62(7):424–430, 9 1988.

[122] S. Saito, L. Hu, C. Ma, H. Ibayashi, L. Luo, and H. Li. 3D Hair Synthesis using Volumetric Variational Autoencoders. *ACM Transactions on Graphics*, 37(6), 1 2019.

[123] R. Saraei. *Construction and evaluation of an optical measuring system for automatical determination of the thickness of biological materials for heart valve implants [B.Sc. Thesis]*. Universität zu Lübeck, 2013.

[124] M. A. Sarsam and M. Yacoub. Remodeling of the Aortic Valve Anulus. *The Journal of Thoracic and Cardiovascular Surgery*, 105(3):435–438, 3 1993.

[125] S. Schaal, J. Peters, J. Nakanishi, and A. Ijspeert. Learning Movement Primitives. In Dario P Chatila R, editor, *Robotics Research. The Eleventh International Symposium. Springer Tracts in Advanced Robotics*, volume 15, pages 561–572. Springer, Berlin, Heidelberg, 2005.

[126] M. Scharfschwerdt, H.-H. Sievers, A. Hussein, E. G. Kraatz, and M. Misfeld. Impact of Progressive Sinotubular Junction Dilatation on Valve Competence of the 3F Aortic and Sorin Solo Stentless Bioprosthetic Heart Valves. *European Journal of Cardio-Thoracic Surgery*, 37(3):631–634, 2010.

[127] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs. Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery. In Niethammer M. et al., editor, *Information Processing in Medical Imaging. IPMI 2017. Lecture Notes in Computer Science*, volume 10265. Springer, Cham, 2017.

[128] S. Schleidgen, C. Klingler, T. Bertram, W. H. Rogowski, and G. Marckmann. What is Personalized Medicine: Sharpening a Vague Term Based on a Systematic Literature Review. *BMC Medical Ethics*, 14, 55(1), 12 2013.

[129] B. Schölkopf and A. J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, 2002.

[130] D. Sharifikia, M. Salem Yafia, G. Fradet, and H. Mohammadi. Design and Fabrication of a 3D Scaffold for the Aortic Root Tissue Engineering Application. *Journal of Medical and Biological Engineering*, 38(2):211–221, 4 2018.

[131] H. H. Sievers, M. Scharfschwerdt, and A. Hof. Biologische Herzklappenprothese. *Patent DE102014223522, Germany*, 2016.

[132] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 9 2015.

[133] A. J. Smola and B. Schölkopf. A Tutorial on Support Vector Regression. *Statistics and Computing*, 14(3):199–222, 2004.

[134] K. Sohn, X. Yan, and H. Lee. Learning Structured Output Representation using Deep Conditional Generative Models. In Cortes C, Lawrence N, Lee D, Sugiyama M, and Garnett R, editors, *Advances in Neural Information Processing Systems*, volume 28, pages 3483–3491. Curran Associates, Inc., 2015.

[135] N. Srivastava, H. Goh, and R. Salakhutdinov. Geometric Capsule Autoencoders for 3D Point Clouds. *arXiv 1912.03310*, 2019.

[136] L. Sun, J. Wang, Y. Huang, X. Ding, H. Greenspan, and J. Paisley. An Adversarial Learning Approach to Medical Image Synthesis for Lesion Detection. *IEEE Journal of Biomedical and Health Informatics*, 24(8):2303–2314, 2020.

[137] S. Sun, H. Shi, and Y. Wu. A Survey of Multi-Source Domain Adaptation. *Information Fusion*, 24:84–92, 7 2015.

[138] J. ten Kate, G. Smit, and P. Breedveld. 3D-Printed Upper Limb Prostheses: A Review. *Disability and Rehabilitation: Assistive Technology*, 12(3):300–314, 4 2017.

[139] M. J. Thubrikar, M. R. Labrosse, K. J. Zehr, F. Robicsek, G. G. Gong, and B. L. Fowler. Aortic Root Dilatation May Alter the Dimensions of the Valve Leaflets. *European Journal of Cardio-Thoracic Surgery*, 28(6):850–855, 12 2005.

[140] H. Uzunova, P. Kaftan, M. Wilms, N. D. Forkert, H. Handels, and J. Ehrhardt. Quantitative Comparison of Generative Shape Models for Medical Images. In Tolxdorff T., Deserno T., Handels H., Maier A., Maier-Hein K., and Palm C., editors, *Bildverarbeitung für die Medizin 2020. Informatik aktuell*, Wiesbaden, 2020. Springer Vieweg.

[141] H. Uzunova, S. Schultz, H. Handels, and J. Ehrhardt. Unsupervised Pathology Detection in Medical Images using Conditional Variational Autoencoders. *International Journal of Computer Assisted Radiology and Surgery*, 14(3):451–461, 3 2019.

[142] L. van der Maaten and G. Hinton. Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9(86):2579–2605, 2008.

[143] E. Vorontsov, P. Molchanov, C. Beckham, W. Byeon, S. De Mello, V. Jampani, M.-Y. Liu, S. Kadoury, and J. Kautz. Towards Semi-Supervised Segmentation via Image-to-Image Translation. *arXiv 1904.01636*, 2019.

[144] Z. Wan, Y. Zhang, and H. He. Variational Autoencoder Based Synthetic Data Generation for Imbalanced Learning. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–7. IEEE, 11 2017.

[145] T. Xia, A. Chartsias, and S. A. Tsaftaris. Adversarial Pseudo Healthy Synthesis Needs Pathology Factorization. In M. J. Cardoso, A. Feragen, B. Glocker, E. Konukoglu, I. Oguz, G. Unal, and T. Vercauteren, editors, *Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning*, volume 102, pages 512–526. PMLR, 2019.

[146] T. Xia, A. Chartsias, and S. A. Tsaftaris. Pseudo-healthy synthesis with pathology disentanglement and adversarial learning. *Medical Image Analysis*, 64, 8 2020.

[147] M. H. Yacoub, P. J. Kilner, E. J. Birks, and M. Misfeld. The Aortic Outflow and Root: A Tale of Dynamism and Crosstalk. *The Annals of Thoracic Surgery*, 68(3):37–43, 9 1999.

[148] C. A. Yankah, Y.-G. Weng, and R. Hetzer. *Aortic Root Surgery - The Biological Solution*. Steinkopff, Heidelberg, 2010.

[149] Y. Zhan, M. Dewan, and X. S. Zhou. Cross Modality Deformable Segmentation Using Hierarchical Clustering and Learning. In H. D. Yang GZ., Rueckert D., Noble A., and Taylor C, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2009. MICCAI 2009. Lecture Notes in Computer Science*, volume 5762. Springer, Cham, 2009.

[150] Zhou and Chellappa. Computation of optical flow using a neural network. In *IEEE International Conference on Neural Networks*, pages 71–78. IEEE, 1988.

[151] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2242–2251. IEEE, 10 2017.

[152] O. C. Zienkiewicz, R. L. Taylor, P. Nithiarasu, and J. Z. Zhu. *The Finite Element Method*, volume 3. McGraw-Hill, London, 1977.

[153] K. H. Zou, S. K. Warfield, A. Bharatha, C. M. Tempany, M. R. Kaus, S. J. Haker, W. M. Wells, F. A. Jolesz, and R. Kikinis. Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index. *Academic Radiology*, 11(2):178–189, 2 2004.

List of Figures

List of Tables

Danksagung

Diese Arbeit würde nicht existieren ohne die langjährige Unterstützung zahlreicher Menschen, die mich auf dem Weg hierher begleitet haben. Zunächst möchte ich mich bei Floris Ernst für die persönliche, wertschätzende und immer unterstützende Betreuung bedanken. Danke für all die Freiräume und Chancen, das stets offene Ohr und deine unglaubliche Flexibiltät in jeder Lebenslage. Weiterer Dank gebührt Michael Scharfschwerdt für die ausführlichen fachlichen und praktischen Erläuterungen, die mich schon früh an die Thematik dieser Arbeit herangeführt haben. Vielen Dank auch für die uneingeschränkt mögliche Labornutzung und die stets unkomplizierte Zusammenarbeit. Danke an Prof. Schweikard, der mir zu Beginn meiner Promotionszeit jede thematische Freiheit gelassen und damit die Möglichkeit für diese Arbeit gegeben hat. Weiterhin danke ich allen Studierenden, die mich bei dieser Arbeit unterstützt haben, insbesondere Rabia Demirci, Ingvild Detjens, Tizian Evers, Meike Fünderich, Marius Krusen, Kenneth Kühl, Kristian Küssner, Mohamad Mehdi, Pascal Merks und Erik Werrmann.

Ein riesengrosses Dankeschön gilt allen aktuellen und ehemaligen Kolleginnen und Kollegen, vom ROB und darüber hinaus. Zu Beginn meiner Promotionszeit sagte ein Kollege bei seiner Verteidigung: Ohne euch hätte ich das niemals geschafft. Heute kann ich das nur wiederholen. Ohne all die Diskussionen und Ratschläge, die Pausen am Kickertisch und das Freude- bzw. Frustbier, wenn ein Paper angenommen oder abgelehnt wurde, hätte ich das niemals geschafft. Ganz besonders hervorheben möchte ich Maximilian Wattenberg für seine unglaubliche Unterstützung auf allen Ebenen sowie Maria Henke, Svenja Ipsen, Conni Rieckhoff und Patrick Stüber für die wertvollen Anmerkungen zu dieser Arbeit.

Ich möchte mich von ganzem Herzen bei meiner Familie und meinen Freunden bedanken, für all die Unterstützung, Geduld und die Nachsicht, wenn mal wieder eine Deadline anstand und ich nichts von mir habe hören lassen. Und natürlich für all die wundervollen Momente jenseits des Schreibtisches. Mama und Papa, vielen Dank für eure unumstößliche Unterstützung auf meinem Lebensweg, eure wertvollen Ratschläge und euren steten Glauben an mich. Samuel, vielen Dank für dein Strahlen, das mir Energie geschenkt hat, wenn die Akkus schon und die Seiten noch leer waren.

Aber der größte Dank geht an Antonia. Für dein Verständnis, wenn die Arbeitstage lang und die Schreibnächte noch länger waren. Dafür, dass du mir den Rücken freigehalten und dich so rührend um Samuel gekümmert hast. Und für die aufmunternden Worte, die du trotz all der Belastung immer wieder finden konntest. Ohne dich hätte ich es nicht geschafft. Danke für alles, alles ist gut!

Erklärung an Eides Statt

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbständig sowie ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen fÃ¼r Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Ich versichere an Eides Statt, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

(Lübeck, 19.05.2021) (Unterschrift)