Cluster Analysis in Latent Space: Identifying Personalized Aortic Valve Prosthesis Shapes using Deep Representations

Jannis Hagenah[1]　　　　　　　　　　　　　　　HAGENAH@ROB.UNI-LUEBECK.DE
Kenneth Kühl[1]
[1] *Institute for Robotics and Cognitive Systems, University of Lübeck, Retzeburger Allee 160, 23562 Lübeck, Germany*

Michael Scharfschwerdt[2]　　　　　　　　　MICHAEL.SCHARFSCHWERDT@UKSH.DE
[2] *Department of Cardiac Surgery, University Hospital Schleswig-Holstein, Ratzeburger Allee 160, 23562 Lübeck, Germany*

Floris Ernst[1]　　　　　　　　　　　　　　　　　ERNST@ROB.UNI-LUEBECK.DE

Abstract

Due to the high inter-patient variability of anatomies, the field of personalized prosthetics gained attention during the last years. One potential application is the aortic valve. Even though its shape is highly patient-specific, state-of-the-art aortic valve prosthesis are not capable of reproducing this individual geometry. An appraoch to reach an economically reasonable personalization would be the identification of typical valve shapes using clustering, such that each patient could be treated with the prosthesis of the type that matches his individual geometry best. However, a cluster analysis directly in image space is not sufficient due to the curse of dimensionality and the high sensitivity to small translations or rotations. In this work, we propose representation learning to perform the cluster analysis in the latent space, while the evaluation of the identified prosthesis shapes is performed in image space using generative modeling. To this end, we set up a data set of 58 porcine aortic valves and provide a proof-of-concept of our method using convolutional autoencoders. Furthermore, we evaluated the learned representation regarding its reconstruction accuracy, compactness and smoothness. To the best of our knowledge, this work presents the first approach to derive prosthesis shapes data-drivenly using clustering in latent space.

Keywords: personalized medicine, representation learning, aortic valve, personalized prosthetics, unsupervised learning

1. Introduction

The geometry of the aortic valve is highly patient-specific. Especially the three valve leaflets differ in their size and shape, while the interdependency of these shapes define the correct function of the valve (De Kerchove et al., 2017). However, state-of-the-art valve prostheses are not capable of remodeling this individual shape. While mechanical prostheses are most commonly designed with only two leaflets or without any leaflet at all, the three leaflets of biological prostheses have the same shape, resulting in a radial symmetry of the valve that is barely found in nature (Pibarot and Dumesnil, 2009). During the last years, multiple studies indicated the high impact of the aortic valve geometry on the whole circulatory system, including the brain circualtion (Al-Atassi et al., 2015), (Blais et al., 2003). Hence, it can be assumed that a personalization of aortic valve prostheses could increase the patient's outcome significantly. Furthermore, it could be shown that a valve that

matches the anatomy of the surrounding tissue has a lower risk of cavitations and reduces the stress inside the leaflet material (Andersen et al., 2006). These effects would lead to a longer lifetime of personalized prostheses and accordingly to a reduced risk of follow-up surgeries.

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 (Hagenah et al., 2018a). Hence, a personalization of the leaflet shapes in a prosthesis is possible in general. However, a completely individual manufaction 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. 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 (Hagenah et al., 2018a) and presents a cost-efficient and hence realistic way of aortic valve prostheses personalization.

In this paper, we present a method to perform a cluster analysis in aortic valve leaflet shapes to identify these valve types. Unfortunately, a clustering in image space where each pixel presents one dimension is not suitable due to an effect called the curse of dimensionality (Keogh and Mueen, 2011). In high dimensional spaces, distance metrics are influenced by the high number of dimemsions and variance in the data can not be incorporated in a sufficient way. Additionally, small translations or rotations of the valve might lead to big distances in image space, even though the shape of the valve stays the same. Hence, we propose to perform the cluster analysis in a latent space description of the valves using representation learning. Utilizing generative modeling, the latent cluster centers 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. To this end, we set up a sufficient data set of pig heart valves and performed a proof-of-concept study using convolutional autoencoders. We analyzed the influence of the number of identified valve types on the capability of reproducing all individual valve shapes that are present in our data set. Furthermore, we performed an analysis of the networks hyperparameters and evaluated the identified valve representation regarding its compactness, accurateness and smoothness.

1.1. Contribution of this work

The contribution of this work is two-fold. First, our resulting model presents the first data-driven typisation of aortic valve shapes regarding personalized medicine. Second, to the best of our knowledge, the proposed method is the first approach to identify prosthesis shapes using clustering in a latent space description in general. It is transferable to comparable problems and could be the basis for a new subfield of personalized prosthetics.

2. Material and Methods

This section is divided into three parts. First, the generation of the porcine dataset and the preprocessing pipeline is described (adapted from (Hagenah et al., 2018b)). Second, since a clustering directly in the image space is not feasible, an autoencoder is used to learn a compact representation
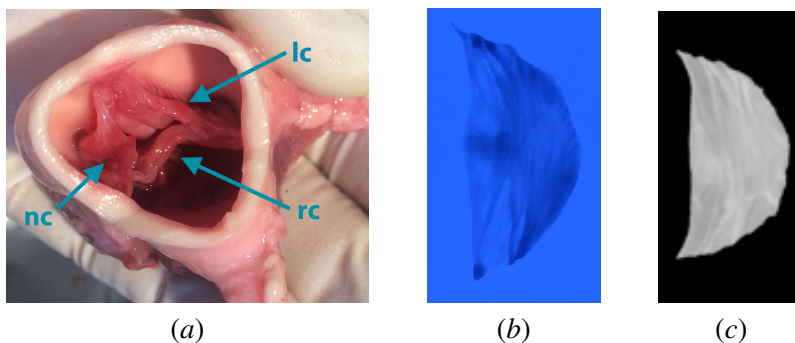
Figure 1: Data set generation. (a) Extracted aortic root from above with the right-coronary (rc), left-coronary (lc) and non-coronary (nc) leaflet. (b) Raw photo of one leaflet with blue backlight illumination. (c) Processed leaflet image.

of the valve geometry. In the third subsection, the clustering method as well as the evaluation of the identified prosthesis shapes in image space is presented.

2.1. Data Set Generation and Preprocessing

Due to their very thin structure and their fast movement during the cardiac cycle, detailed imaging of the leaflets is a tough task. Hence, we bypass this problem by cutting out the aortic valve leaflets of fresh porcine aortic roots, assuming that the shape of the aortic valve is given by the shape of its three leaflets. The pig heart is a common animal model in the context of aortic valve research due to its quite similar shape compared to humans (Crick et al., 1998). The hearts were bought at a slaughterhouse, so this study is not related to any ethical issue. To extract the leaflets, the aortic root was cut out of the pig heart in a first step (see Fig. 1(a)). Then, the root was cut vertically in between the left-coronary and the right-coronary leaflet. Hence, the root could be opened and the three sinuses were separated using vertical incisions in between the leaflets. Finally, the leaflets were cut off of the aortic root wall. To acquire the individual shape of the extracted leaflets, they were spread on an illuminated plate and a photograph was taken. Special attention was paid to preserve the leaflets natural shape while spreading them on the plate. The illumination was monochromatic with a wavelength of 470 *nm*. Light of this color is strongly absorbed in the collagen fibers of the leaflets, resulting in a good contrast of the very thin structures (see Fig. 1(b)). Details on the setup can be found in (Hagenah et al., 2018b). This procedure was performed for 56 porcine hearts, resulting in images of 168 leaflets.

All of these images were preprocessed by applying the following pipeline. At first, the images were transformed to grayscale images and inverted. In these images, the background pixels were set to 0 using 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 rotating the leaflet so that the commissure points are vertically aligned. Finally, the images were downsampled to a size of 128×64 pixels with a resolution of 0.34 $\frac{mm}{pixel}$. The result is exemplarily shown in Fig. 1(c).

2.2. Representation Learning

We trained an artificial neural network to find a compact yet meaningful representation of a leaflet image in an unsupervised manner. To this end, we used a convolutional autoencoder AE, which is known to be capable of encoding images with good representations in the bottleneck layer, i.e. the latent space (Hinton and Salakhutdinov, 2006). The autoencoder is divided into two different subnetworks: the encoder network $enc : I \in \mathscr{R}^{128 \times 64} \rightarrow z \in \mathscr{R}^{n_z}$ and the decoder network $dec : z \in \mathscr{R}^{n_z} \rightarrow I \in \mathscr{R}^{128 \times 64}$ such that

$$AE(I) = dec(enc(I)) = I_{reco}, \qquad (1)$$

where I is an image of one leaflet and n_z is the dimensionality of the latent space, i.e. the number of neurons in the bottleneck layer.

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, we used 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) for the encoder. In all layers, ReLU-Activation was applied. The decoder architecture was identical but mirrored. The autoencoder was implemented in *Keras* using the *tensorflow*-backend (Chollet et al., 2015). It was trained with mean-squared-error loss using the ADAM-optimizer. An illustration of the proposed architecture can be found in the appendix in Fig. 4.

To evaluate the quality of the representation, we analyzed the reconstruction accuracy of the autoencoder by propagating the leaflet images through the encoder and generate an image of the resulting latent representation using the decoder. These predicted images were compared to their ground truth using two metrics: the Jaccard-Coefficient d_J and the Hausdorff-Metric d_H (Yeghiazaryan and Voiculescu, 2015). The Jaccard-Coeffienct measures the overlap of two leaflets and hence evaluates the overall shape similarity:

$$d_J(A,B) = \frac{A \cap B}{A \cup B}, \qquad (2)$$

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

$$d_H(X,Y) = max \left(\sup_{x \in X} \inf_{y \in Y} d(x,y), \sup_{y \in Y} \inf_{x \in X} d(y,x) \right), \qquad (3)$$

where d is the euclidian distance and X and Y are sets of contour pixels, respectively. Since the reconstruction accuracy is highly dependant of the network architecture, we performed this analysis additionally 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). Once again, the decoder architecture was the mirrored encoder architecture. The number of latent dimensions n_z was set to 20.

Furthermore, we evaluated the reconstuction accuracy of the autoencoder 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. 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, that will barely match one existing valve. Hence, we compared our proposed architecture to two alternative methods. First, we used the same autoencoder but introduced data augmentation (random shearing and zooming by up to 5%, translations in any direction with a maximum of 2.2*mm*). Second, we evaluated a variational autoencoder (VAE) of the same encoder/decoder architecture as VAEs are optimized to identify a smooth latent space implicitly (Kingma and Welling, 2013). 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.
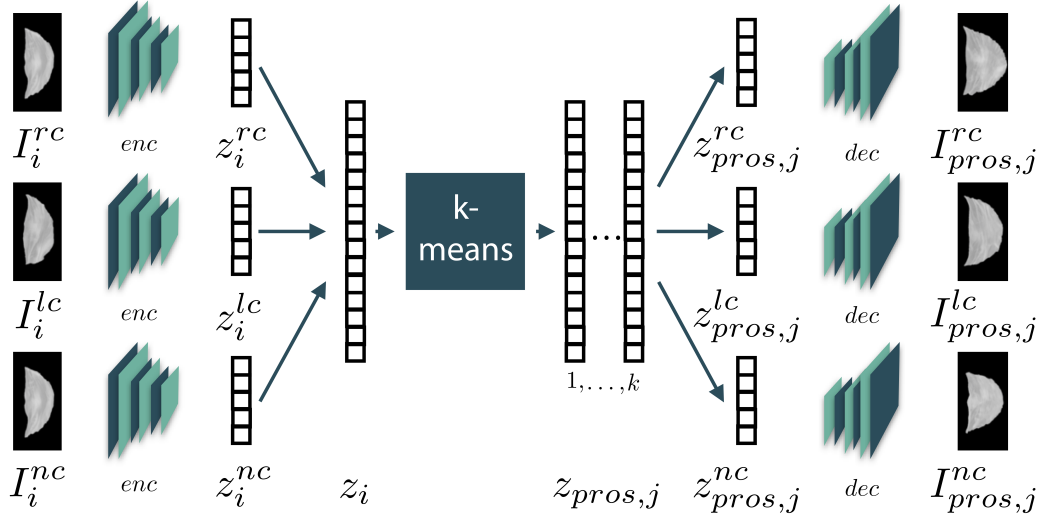

Figure 2: Sketch of the proposed method. The trained encoder model is used to receive the latent vector of each leaflet. These vectors are stitched together to represent a valve. In this space, the clustering is performed. Afterwards, the latent representation of the single leaflets of an identified prosthesis are extracted and the corresponding images are generated using the decoder.

2.3. Clustering

Our aim is to perform a cluster analysis in latent space. To this end, we assume that the geometry of the whole valve can be described by the shape of its three leaflets. Thus, we propagated the three leaflets of one valve through the trained encoder network to get their latent representations $z_i^{rc} \in \mathscr{R}^{n_z}$, $z_i^{lc} \in \mathscr{R}^{n_z}$ and $z_i^{nc} \in \mathscr{R}^{n_z}$, 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 stitching the latent vectors together, we retrieve a latent representation $z_i \in \mathscr{R}^{3n_z}$ of the whole valve. In this representation, a k-means-clustering was performed, aiming on identifying k cluster centers

$z_{pros,j}, j = 1 \ldots k$ that could serve as prosthesis types (Lloyd, 1982). Even though a clustering in the latent space is convenient, the evaluation of the identified prosthesis types in latent space is not intuitive. Hence, we transformed the prosthesis back to image space using the decoder. For this purpose, the latent prosthesis representation was split up into the three parts $z_{pros,j}^{rc} \in \mathscr{R}^{n_z}$, $z_{pros,j}^{lc} \in \mathscr{R}^{n_z}$ and $z_{pros,j}^{nc} \in \mathscr{R}^{n_z}$ corresponding to the three leaflets. These vectors were propagated through the decoder, resulting in images of the three leaflets of the prosthesis, i.e. $I_{pros,j}^{rc} \in \mathscr{R}^{128 \times 64}$, $I_{pros,j}^{lc} \in \mathscr{R}^{128 \times 64}$ and $I_{pros,j}^{nc} \in \mathscr{R}^{128 \times 64}$. These images could be compared to the images of all valves that correspond to this cluster using the metrics described above. Hence, we can calculate the mean Jaccard coefficient $\overline{d_{J,k}}$ and the mean Hausdorff metric $\overline{d_{H,k}}$ over all valves in dependency of the number of identified prosthesis types k as

$$\overline{d_{J,k}} = \frac{1}{56} \sum_{j=1}^{k} \sum_{I_c \in C_j} \frac{1}{3} \left(d_J \left(I_c^{rc}, I_{pros,j}^{rc} \right) + d_J \left(I_c^{lc}, I_{pros,j}^{lc} \right) + d_J \left(I_c^{nc}, I_{pros,j}^{nc} \right) \right) \qquad (4)$$

$$\overline{d_{H,k}} = \frac{1}{56} \sum_{j=1}^{k} \sum_{I_c \in C_j} \frac{1}{3} \left(d_H \left(I_c^{rc}, I_{pros,j}^{rc} \right) + d_H \left(I_c^{lc}, I_{pros,j}^{lc} \right) + d_H \left(I_c^{nc}, I_{pros,j}^{nc} \right) \right), \qquad (5)$$

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 our data set. To analyze the amount of valve prosthesis types needed to ensure a good shape approximation, we performed the clustering 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. For the clustering study, n_z was set to 20.

3. Results

The results of this study are divided into two parts: the analysis of the learned representation and the identification of valve types using clustering in the latent space.

3.1. Representation Analysis

The results of the experiment regarding the reconstruction accuracy of our proposed architecture as well as of alternative architectures is given in Table 1. Compared to a shallower and a deeper architecture, the proposed autoencoder achieves the highest Jaccard-Coefficient and the smallest value of the Hausdorff-Metric, given as the mean over all leaflets.

Fig. 3(a) 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 only lead to a slight increase of the reconstruction accuracy. The introduction of data augmentation as well as the use of a variational autoencoder provide comparable results as the proposed architecture (see Table 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 the appendix in Fig. 5.

Table 1: Reconstruction accuracy of the proposed autoencoder (marked in bold), compared to a shallower and a deeper architecture regarding the Jaccard-Coefficient d_J and the Hausdorff-Metric 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.

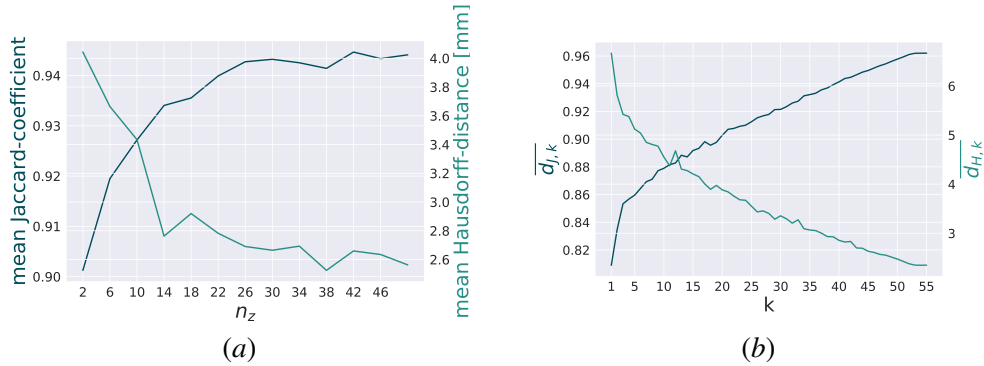| Architecture	d_J	d_H [mm]
Shallower architecture	0.9272	3.17
Proposed architecture	**0.9471**	**2.60**
Deeper architecture	0.9283	3.18
Proposed AE with Data Augmentation	0.9424	2.70
VAE of same architecture	0.9423	2.74

(a) (b)

Figure 3: Resulting reconstruction accuracies. (a) Reconstruction accuracy of the autoencoder in dependency of the number of latent dimensions n_z. (b) Capability of reproducing all individual valve shapes given in the data set in dependency of the number of prosthesis type.

3.2. 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 our data set is shown in Fig. 3(b). 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 the appendix in Fig. 6 . To critically question the whole approach of our proposed autoencoder pipeline, we also performed the clustering and evaluation routine regarding the image description using principal components. To this end, we performed Principal Component Analysis (PCA) on the vectorized images and used the 20 most relevant components as a representation. While the mean Jaccard coefficient was significantly smaller compared to the autoencoder approach, outliers inn the range of up to 20mm were observed in the mean Hausdorff metric. Detailed results of this comparison can be found in the appendix in Fig. 7.

4. 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, we achieve a good trade-off between compactness of the representation and reconstruction accuracy with 20 latent dimensions. Our 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 (Chen et al., 2016). Another possible explanation is the higher number of trainable parameters on the relatively small data set.

We could also show 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 contor lines or detailed shape differences, which also explains the overall high values of the Jaccard-Coefficient. In the qualitative smoothness identification, slight differences between the proposed autoencoder and the 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-Coefficient for $k = 1$ is already at 0.8. This means, that the whole potential of personalization lies within a range of a Jaccard-Coefficient increase of 0.2. Just by introducing three different prosthesis types, we can close about one quarter of this personalization gap. 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 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 manufaction of 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 herat valves, deformations of the leaflet shapes cannot be entirely avoided. However, we assume that these errors would be normally distributed, hence, no systematic bias towards specific shapes would 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.

To the best of our knowledge, 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 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, our method could present an important step towards personalized prosthetics.

Acknowledgments

References

Talal Al-Atassi, Hadi Daood Toeg, Reza Jafar, Benjamin Sohmer, Michel Labrosse, and Munir 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.

Tina S Andersen, Peter Johansen, Bekka O Christensen, Peter K Paulsen, Hans Nygaard, and J Michael Hasenkam. Intraoperative and postoperative evaluation of cavitation in mechanical heart valve patients. *The Annals of Thoracic Surgery*, 81(1):34–41, 2006.

Claudia Blais, Jean G Dumesnil, Richard Baillot, Serge Simard, Daniel Doyle, and Philippe Pibarot. Impact of valve prosthesis-patient mismatch on short-term mortality after aortic valve replacement. *Circulation*, 108(8):983–988, 2003.

Xi Chen, Diederik P Kingma, Tim Salimans, Yan Duan, Prafulla Dhariwal, John Schulman, Ilya Sutskever, and Pieter Abbeel. Variational lossy autoencoder. *arXiv preprint arXiv:1611.02731*, 2016.

François Chollet et al. Keras. *GitHub*, https://github.com/fchollet/keras, 2015.

Simon J Crick, Mary N Sheppard, Siew Yen Ho, Lior Gebstein, and Robert H Anderson. Anatomy of the pig heart: comparisons with normal human cardiac structure. *The Journal of Anatomy*, 193 (1):105–119, 1998.

Laurent De Kerchove, Mona Momeni, Gaby Aphram, Christine Watremez, Xavier Bollen, Ramadan Jashari, Munir Boodhwani, Parla Astarci, Philippe Noirhomme, and Gebrine El Khoury. Free margin length and coaptation surface area in normal tricuspid aortic valve: an anatomical study. *European Journal of Cardio-Thoracic Surgery*, 2017.

Jannis Hagenah, Tizian Evers, Michael Scharfschwerdt, Achim Schweikard, and Floris Ernst. An svr-based data-driven leaflet modeling approach for personalized aortic valve prosthesis development. *Computing in Cardiology 2018*, 2018a.

Jannis Hagenah, Michael Scharfschwerdt, and Floris Ernst. Towards personalised aortic valve prostheses: A compact description of the individual valve geometry. *Computing in Cardiology 2018*, 2018b.

Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.

Eamonn Keogh and Abdullah Mueen. Curse of dimensionality. In *Encyclopedia of machine learning*, pages 257–258. Springer, 2011.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Stuart Lloyd. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28(2): 129–137, 1982.

Philippe Pibarot and Jean G Dumesnil. Prosthetic heart valves. *Circulation*, 119(7):1034–1048, 2009.

Varduhi Yeghiazaryan and Irina Voiculescu. An overview of current evaluation methods used in medical image segmentation. Technical report, Tech. Rep. CS-RR-15-08, Department of Computer Science, University of Oxford, Oxford, UK, 2015.
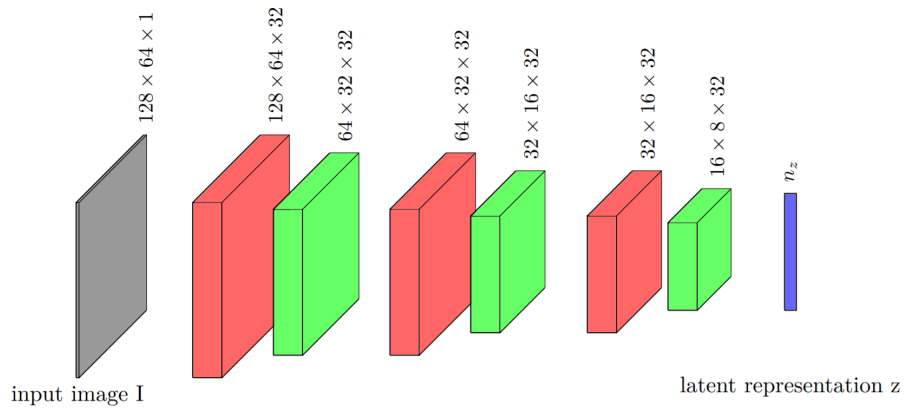
Appendix A. Encoder Architecture

Figure 4: The architecture of the proposed encoder network, consisting of 3 convolutional layers (red), 3 maximum pooling layers (red) and a fully-connected layer (blue). The decoder consists of a similar, but mirrored architecture.
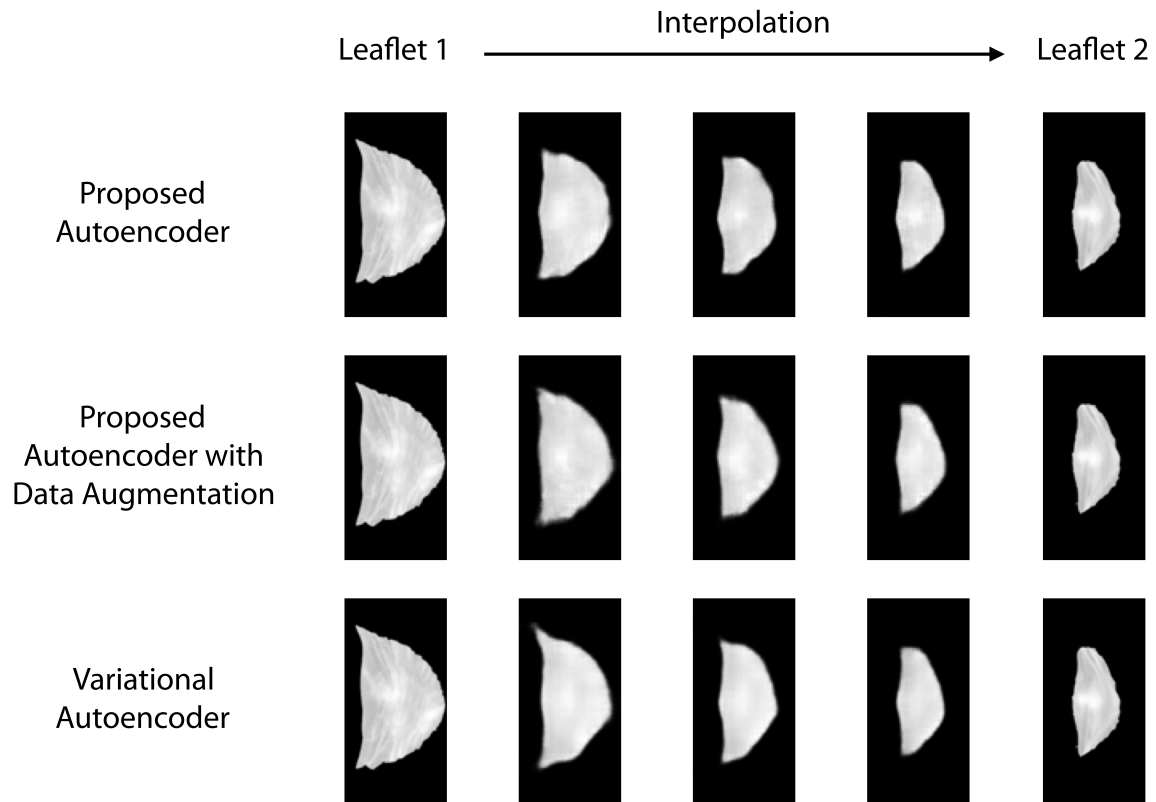
Appendix B. Qualitative Analysis of Smoothness

Figure 5: 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.
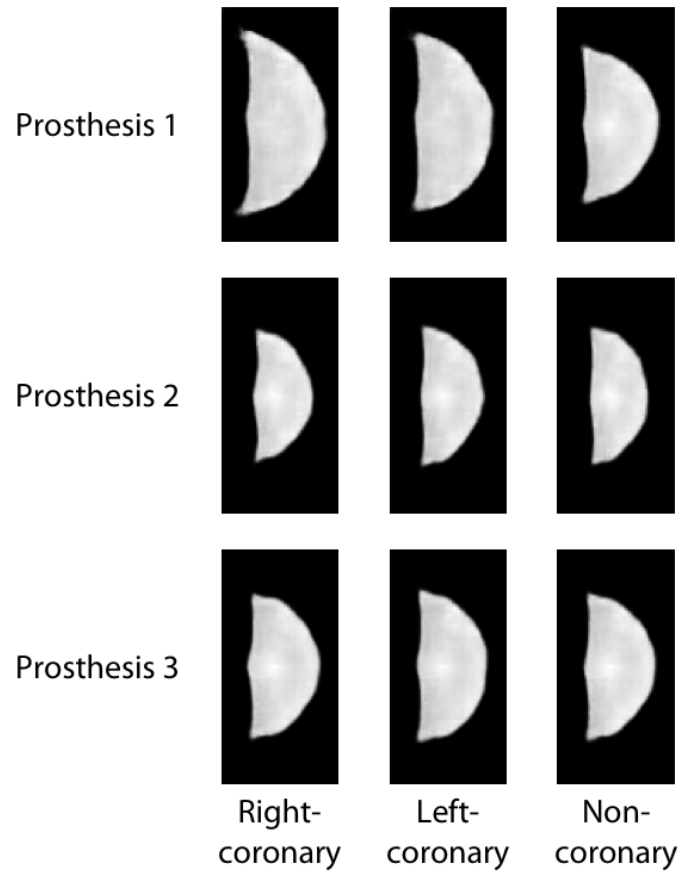
Appendix C. Resulting prosthesis shapes

Figure 6: Resulting prosthesis shapes, exemplarily shown for $k = 3$.

Appendix D. Comparison to PCA approach
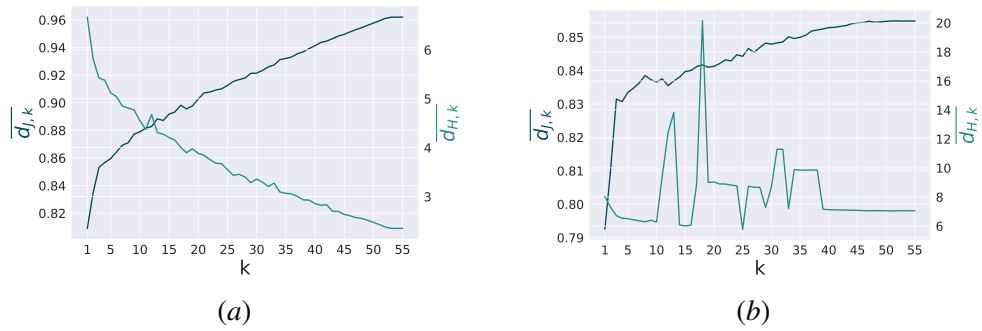

(a) (b)

Figure 7: Capability of reproducing all individual valve shapes given in the data set in dependency of the number of prosthesis type for the proposed autoencoder (a, copied from Fig. 3(*b*)) and when Principal Component Analysis (PCA) is used for deriving the latent space representation (b). For comparison, the number of principal components was set to $n_z = 20$.