Statistical modeling of pulmonary vasculatures with topological priors in CT volumes

Yuki Saeki, Atsushi Saito, Jean Cousty, Yukiko Kenmochi, Akinobu Shimizu

Statistical modeling of pulmonary vasculatures with topological priors in CT volumes *

Yuki Saeki[1], Atsushi Saito[1][0000−0001−9502−8134], Jean Cousty[2][0000−0002−2163−9714], Yukiko Kenmochi[2][0000−0001−9648−326X], and Akinobu Shimizu[1][0000−0002−2719−5923]

[1] Tokyo University of Agriculture and Technology, Tokyo, Japan
[2] LIGM, Univ Gustave Eiffel, CNRS, ESIEE Paris, Marne-la-Vallée, France

Abstract. A statistical appearance model of blood vessels based on variational autoencoder (VAE) is well adapted to image intensity variations. However, images reconstructed with such a statistical model may have topological defects, such as loss of bifurcation and creation of undesired hole. In order to build a 3D anatomical model of blood vessels, we incorporate topological prior into the statistical modeling. Qualitative and quantitative results on 2567 real CT volume patches and on 10000 artificial ones show the efficiency of the proposed framework.

Keywords: Statistical model · Topological data analysis · Variational autoencoder · Vasculature.

1 Introduction

A statistical model, which represents a variation in shape and intensity of an organ with a small number of parameters, plays a key role as priors for medical image analysis [5, 8]. Many statistical shape models have been proposed for parenchymatous organs, such as the liver [1, 3, 14]. However, as far as we know, less study made for building a statistical appearance model that accurately represents variations of blood vessels, whose geometrical and appearance features are useful for the diagnosis of diseases, e.g. pneumonia. This is due to the fact that a simple statistical model is not suitable to describe diverse variations of vessels including difference in not only direction and diameter but also topology such as branching. For example, Principal Component Analysis (PCA), which is a classical statistical analysis method, cannot deal with such complicated distribution. Recently proposed methods, such as manifold learning and deep learning [2], might solve the problem in the modeling of blood vessels. Along this line, the approach employing a Variational AutoEncoder (VAE) [12] and its extension β-Variational AutoEncoder (β-VAE) [9] was previously proposed to build a statistical model of vasculatures of lung CT volumes [13]. Nevertheless, it has been

* Partially supported by MEXT/JSPS KAKENHI Grant Number JP18H03255 and CNRS International Emerging Action Project DiTopAM.

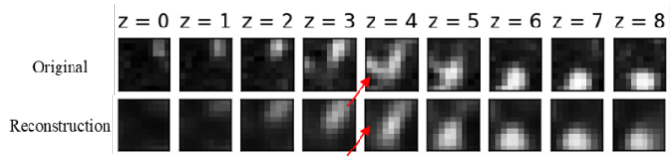

Fig. 1. Cross sections of an original volume patch of $9 \times 9 \times 9$ [voxel] of vasculatures (top) and its failed reconstruction (missing bifurcation marked by the red arrow) by a statistical model based on β-VAE [13] (bottom).

pointed out that the images generated by the statistical model suffer from topological artifacts, such as disconnection of vessels, undesired holes, false cycles, and missing bifurcations (see Fig. 1).

This article proposes a method for building a statistical intensity model of pulmonary vasculatures in CT volume patches that incorporates topological prior of vasculatures in the statistical modeling, which is an incremental contribution over the previous work [13]. We use persistent homology [6] to give topological constraints in neural network [4], which result in topologically correct representation of vasculatures. In particular, in the case of bifurcations and undesired hole artifacts, qualitative and quantitative evaluations show that taking topological prior into account during learning indeed allows for improving the topological soundness of reconstructed images while keeping a satisfying appearance level. We also investigate its combination with a multi-scale approach based on mathematical morphology [15] when the topological correction (hole closing or vessel re-connection) forced by the topological loss is too "thin".

2 Appearance model of 3D blood vessels based on β-VAE

Statistical variations of vessels in chest CT volume patches are modeled in two steps: modeling appearance of vessels using a conventional network, or β-VAE, as a baseline model and incorporating topological prior to the model. In this section, we explain the first step.

Vessel appearance in a volume patch is modeled by the β-VAE (Fig. 2) [13] that consists of four layers, or two fully connected layers with 200 units followed by a rectified linear unit (ReLU), a latent layer, and a fully connected layer with 729 units followed by a sigmoid function to generate an output vector. Note that a volume patch of $9 \times 9 \times 9$ [voxel] is transformed into a vector of size 729 for the input and the output vector is reconstructed to a volume patch.

Let us consider some dataset $\mathbf{X} = \{\mathbf{x}^{(i)}\}_{i=1}^{N}$ of size N. Given a volume patch $\mathbf{x}^{(i)}$ and its reconstruction $\mathbf{y}^{(i)}$, the loss function of β-VAE is defined by

$$\mathcal{L}_{\beta\text{-VAE}} = -\frac{\beta}{2}\sum_{j=1}^{J}\left(1 + \log((\sigma_j^{(i)})^2) - (\mu_j^{(i)})^2 - (\sigma_j^{(i)})^2\right) + \frac{1}{2}\|\mathbf{x}^{(i)} - \mathbf{y}^{(i)}\|_2^2 \quad (1)$$

where $\mu^{(i)}$ and $\sigma^{(i)}$ are the mean and variance of the latent space variables $\mathbf{z}^{(i)}$, both of which are computed using the training data and $\mathbf{x}^{(i)}$. Note that j represents the dimensional index of $\mathbf{z}^{(i)}$, and J is the dimension of the latent space. The first term on the right-hand side is the Kullback-Leibler distance between the predicted distribution in the latent space and the Gaussian distribution $\mathcal{N}(0, I)$, while the second term is the reconstruction error of $\mathbf{x}^{(i)}$. The weight β determines a balance between the two terms.

This statistical model allows us to describe vessels using a small number of parameters and to generate vessels by changing latent parameters. The encoder of the trained β-VAE has the former function, while the decoder has the latter function. Thanks to these two functions, this model enables us to compress a given vessel and also to generate a vessel.

3 Topological loss function based on persistent homology

The second step of the statistical modeling is to incorporate topological prior into the appearance model of Section 2. To this end, computational topology [6] is used to fine-tune the network. We focus on persistent homology [17] and incorporate the topological loss function [4] in the fine-tuning phase.

3.1 Persistent homology

Persistent Homology (PH) aims at describing the topological features of data [6]. Given a function or image defined over a simplicial complex, PH represents the evolution of the topological features of the lower level sets (also called threshold sets) of this function. Thus, the image resulting from the previously presented network must be first converted into a weighted simplicial complex before computing PH. In a complex of dimension d, each topological feature has a dimension between 0 and $d-1$ and, for each dimension k, the number of topological features of dimension k is known as the k-th Betti number, denoted by β_k. In a 3D space, the features of dimension 0 (resp. 1, and 2) correspond to the connected components (resp. to the tunnels, and to the cavities) of the complex. When browsing the successive threshold sets of a weighted complex, at each level, it is possible to track the new topological features that appear, those that disappear at this level, and those that persist from the previous level. Therefore, for each topological feature, it is possible to compute the level at which it appears and

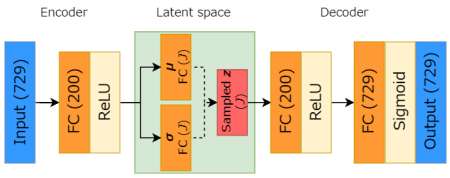

Fig. 2. Network architecture of β-VAE [13] for statistical modeling of vessels

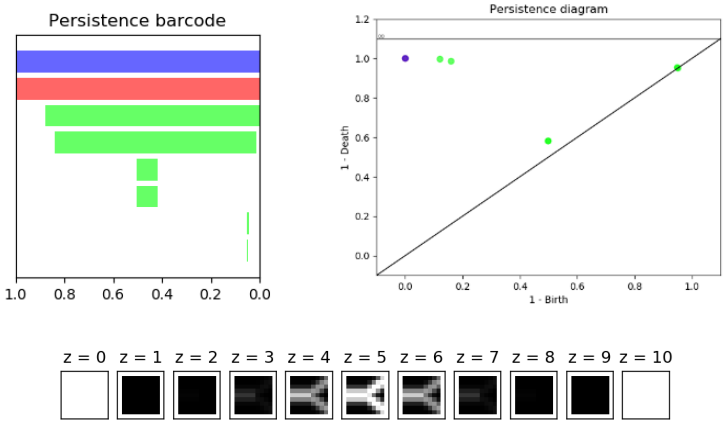

Fig. 3. Persistence barcode (top, left) and diagram (top, right) of a 3D bifurcated vessel image with a frame padded (bottom)

the one at which it disappears. These levels are called the birth and death times of the topological feature, respectively. PH then produces a *Persistence Barcode (PB)* such as the one shown in Fig. 3, where a bar is displayed for each topological feature and where the starting and ending points of the bar correspond to the birth and death times of the considered feature. In the figure, blue (resp. green and red) bars represent the features of dimension 0 (resp. 1 and 2) composing the Betti number β_0 (resp. β_1 and β_2). A *Persistence Diagram (PD)* can also be drawn where each feature is represented as a point whose coordinates correspond to its birth and death times. The length of a bar in a PB indicates the lifetime of each topological feature. When the lifetime is long, the corresponding point of a PD is plotted close to the upper left corner and considered as a stable feature. In this example, blue and red points are at stable positions. When the lifetime is short, the corresponding points are plotted close to the diagonal line and such components are considered as noise.

3.2 Topological loss functions and priors

Based on the idea presented in [4], we assume that topological information of 3D blood vessels is known a priori, and incorporate this knowledge into deep learning based modeling.

Let us consider the l-th longest bar of the k-th Betti number of PB, whose birth and death times are denoted by $b_{k,l}$ and $d_{k,l}$, respectively. Suppose that we have the topological prior such that the PB contains β_k^* long bars for the k-th Betti number. In other words, those corresponding points of PD are located far from the diagonal line and close to the upper left corner. With such topological

prior, the topological loss function is defined by

$$\mathcal{L}_{\text{topo}}(\beta_0^\star, \beta_1^\star, \beta_2^\star) = \sum_{k\in\{0,1,2\}} \left\{ \lambda_k^* \sum_{l=1}^{\beta_k^*} \left(1 - |b_{k,l} - d_{k,l}|^2\right) + \lambda_k \sum_{l=\beta_k^*+1}^{\infty} |b_{k,l} - d_{k,l}|^2 \right\} \tag{2}$$

where λ_k^* and λ_k are weights for the first and second terms with respect to k-th Betti number, respectively. Note that we assume that the first β_k^* long bars are the "correct" ones. Given topological prior β_k^* for each k, this loss function is minimized during the learning process. Note that minimizing this loss implies increasing the lengths of the β_k^* long bars of each k-th Betti number on PB (the first term of (2)) and decreasing the lengths of the rest (the second term of (2)). This will encourage the β_k^* points of PD in the first term to move to the upper left point while the rest moves to the diagonal line.

This article is aimed at modeling 3D blood vessels in volume patches. Toward this end, we consider various topological loss functions. First, we propose to set

- $(\beta_0^*, \beta_1^*, \beta_2^*) = (1, 0, 0)$ in order to favor one connected component

existing in a volume patch, denoted by $\mathcal{L}_{\text{topo}}(1, 0, 0)$. As this topological prior is not sufficient to distinguish a single tube and a bifurcation structure, we also consider the padded volume patches with a one-voxel border of value 1 (see Fig.3 (bottom)), to which we propose to set

- $(\beta_0^*, \beta_1^*, \beta_2^*) = (1, 1, 1)$ for a single tube structure, and
- $(\beta_0^*, \beta_1^*, \beta_2^*) = (1, 2, 1)$ for a single bifurcation structure (Y-shaped)

as the topological priors, denoted by $\mathcal{L}^{\square}_{\text{topo}}(1, 1, 1)$ and $\mathcal{L}^{\square}_{\text{topo}}(1, 2, 1)$. A similar idea of using padded images for 2D image segmentation can be seen in [10].

In some cases, the topological correction (hole closing or vessel re-connection) forced by the topological loss is too "thin" compared to the width of the vessel (see Fig. 4–line 4). It is then desirable to also consider a topological loss applied to an erosion of the generated image (insertion of an erosion layer before computing the topological loss), leading to multi-scale topological loss functions, where the scale is given by the size of the structuring element used in the erosion. The effect of this multi-scale loss function can be seen in Fig. 4–line 5 where the topological correction is thick enough.

4 Incorporating topological priors into statistical model

The topological priors proposed in the previous section are incorporated into the deep learning generator by using the following loss function

$$\mathcal{L} = \mathcal{L}_{\beta\text{-VAE}} + \mathcal{L}_{\text{topo}} \tag{3}$$

instead of simply using $\mathcal{L}_{\beta\text{-VAE}}$ of (1). Note that the effect of homology based loss on the total loss can be controlled by setting values of weights λ_k^* and λ_k

inserted in $\mathcal{L}_{\mathrm{topo}}$ (see Eq. (2)). In this article, the function $\mathcal{L}_{\mathrm{topo}}$ can be either the topological function that defines the number of connected components $\mathcal{L}_{\mathrm{topo}}(1,0,0)$, the structure of non-bifurcation $\mathcal{L}^{\square}_{\mathrm{topo}}(1,1,1)$, the structure of bifurcation $\mathcal{L}^{\square}_{\mathrm{topo}}(1,2,1)$, their multi-scale version, or their linear combination. It should be noted that such a combination allows us to represent more complex shapes while it provides more parameter tuning and more computation time.

5 Experiments

The network is pre-trained with only the β-VAE loss (1) and fine-tuned by adding the topological loss (2). The idea behind this strategy is that the network is first optimized coarsely with the β-VAE loss and then refined with the topological loss. This is because the topological loss is likely to provide many local optimums everywhere in the search space. The method is evaluated in the following two contexts: blood vessels containing (1) hole artifacts, created in image acquisition process such as random noise, and (2) bifurcations. The topological priors are chosen depending on the contexts. Volume patches of size $9 \times 9 \times 9$ are treated and intensities of every patch are normalized to $[0, 1]$. Both in the pre-training and the fine-tuning, the number of epochs is decided such that the validation loss becomes minimum.

5.1 Blood vessels containing hole artifacts

Datasets Artificial images of blood vessels containing hole artifacts are generated from spatial lines with a Gaussian intensity profile whose center is positioned on the lines and whose parameter values are tuned (no hole image). For each image, a hole is then generated by reversing the intensity values of a ball, whose center is on the centerline of a blood vessel (hole images) and which is located inside the blood vessel. Among the generated images, 6000, 2000 and 2000 volume patches are used for training, validation and testing.

Chosen topological loss function and parameter setting The chosen topological prior is of one connected component $\mathcal{L}_{\mathrm{topo}}(1,0,0)$. The latent dimensions J of β-VAE was set to 6, and the parameters of the loss function were set to $\beta = 0.1$ and $\lambda_k^* = \lambda_k = 60000$. Adam [11] was used with $\alpha = 10^{-5}, \beta_1 = 0.9, \beta_2 = 0.999$ as the optimization method, where the batch size was set to 128. The number of epochs was set to 3072 (resp. 371) for the pre-training (resp. the fine-training).

Evaluation strategies The performance of the statistical model for intensity distributions was evaluated by the following two indices, generalization, which measures the ability to represent unknown data correctly, and specificity, which measures the ability to eliminate an unnatural shape [16]; they are made by comparison with test images and their reconstructed images. Note that the smaller

Table 1. Evaluations of the three methods (Pre, Fine and Fine with erosion) on their reconstructions of a volume containing a hole artifact (see Fig. 4) with three measures based on generalization, specificity and topological loss function.

	Generalization	Specificity	Topology
Pre	**0.00251**	**0.0134**	0.430
Fine	0.00478	0.0143	0.0976
Fine + Erosion	0.00728	0.0156	**0.00139**

index leads to better performance since they measure the distances between the image generated by the model and the original images.

For evaluating the topological performance of the model, the topological loss function was calculated with $\lambda_k^* = \lambda_k = 1$ (the smaller the index, the better the performance). The null hypothesis that the values achieved by each method (Pre, Fine and Fine (Erosion)) have the same distribution was tested: Wilcoxon's signed quartiles test was used for generalization and topological loss function, and Mann-Whitney's U test was used for specificity.

Results The method without topological prior is denoted as Pre while the proposed one with topological prior without/with multiscale is denoted as Fine and Fine (Erosion), respectively. Table 1 shows that Fine is inferior to Pre in the evaluation of intensity values, but it improves the performance in the evaluation of topology. However, the 4th row of Fig. 4 shows that the result still contains an artifact. With the multiscale version of this topological loss, this is improved qualitatively (the 5th row of Fig. 4) and quantitatively with topological evaluation (Table 1). Null hypothesis (H0: there is no difference in performance between any pair of the three methods, Pre, Fine and Fine (Erosion)) was rejected ($p < 0.01$) for all evaluation measures, except for the specificity between Fine and Fine (Erosion).

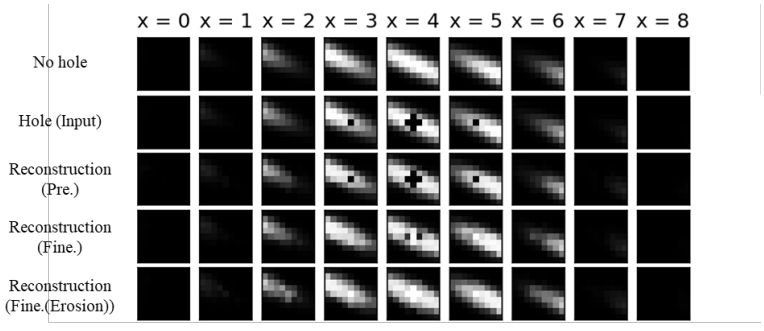

Fig. 4. Modeling of a volume containing a hole artifact (Input): resulting reconstruction without/with topological prior (Pre/Fine), and with the multiscale version (Erosion).

Table 2. Evaluations of the two methods (Pre and Fine) on their reconstructions of a volume containing a bifurcation (see Fig. 5) with three measures based on generalization, specificity and bottleneck distance with the original image via the PDs.

	Generalization	Specificity	Bottleneck distance
Pre	**0.00704**	**0.0283**	0.288
Fine	0.0102	0.0410	**0.236**

5.2 Blood vessels with bifurcations

Datasets CT images taken at Tokushima University Hospital (TUAT Ethics Committee Application No. 30-29) were used in the experiments. From 47 cases, volume patches of size $9\times9\times9$ containing Y-shaped blood vessels with a thickness of approximately from 1 to 4 mm, located between the hilar and peripheral regions, were extracted using a method based on Hessian filter [7]. Among them, 1533, 517 and 517 volume patches are used for training, validation and testing. The training images were augmented by flipping the original volume patches with probability 0.5 before inputting to the network.

Chosen topological loss function and parameter setting The network was pre-trained with only the β-VAE loss (1) and fine-tuned by adding topological prior (2) of bifurcation case $\mathcal{L}^{\square}_{\mathrm{topo}}(1,2,1)$. The latent dimensions J of β-VAE was set to 22, which corresponds to the number of axes of PCA whose cumulative contribution exceeds 0.8 using the training and validation data. The parameters of the loss function were experimentally set to $\beta = 0.1$ and $\lambda_k^* = \lambda_k = 50$. Adam [11] was used with $\alpha = 10^{-3}, \beta_1 = 0.9, \beta_2 = 0.999$ as the optimization method, where the batch size was set to 128. To reduce the computational cost, the topological loss was calculated using $\frac{1}{10}$ of the batch size. The number of epochs was set to 346 (resp. 36) for the pre-training (resp. the fine-training).

Evaluation strategies Similarly to the previous experiment, generalization and specificity were calculated. For evaluating the topological performance of

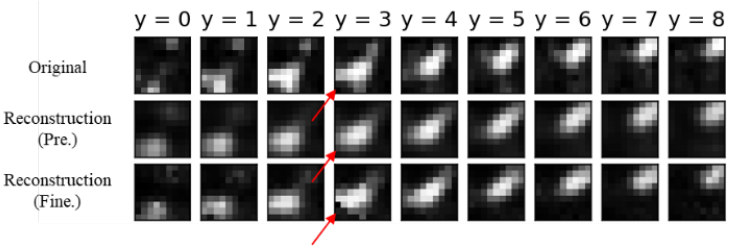

Fig. 5. Modeling of a volume containing a bifurcation (Original): resulting reconstruction without/with topological prior (Pre/Fine)

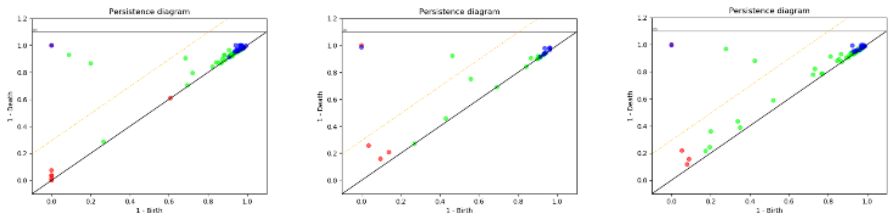

Fig. 6. PDs of the original (left) and reconstructed images of Fig. 5: Pre. (middle) and Fine (right).

the model, the Bottleneck distance [6] between the PDs of the original and reconstructed images was used rather than the topological loss function, which is simpler to be calculated, as we consider in this experiment that the reconstruction must be close to the original. Note that the smaller the distance, the better the performance. Wilcoxon's signed quartiles test was used for Bottleneck distance as well as generalization while Mann-Whitney's U test was used for specificity.

Results The method without topological prior is denoted as Pre and the proposed method with topological prior is denoted as Fine. Table 2 shows that the Fine is inferior to the Pre in the evaluation of intensity values, but it improves the performance in the evaluation of topology. Null hypothesis (H0: there is no difference in performance between Pre and Fine) was rejected ($p < 0.01$) for all evaluation measures.

The reconstructed images are shown in Fig. 5 with the red arrows which indicate the typical difference between the two methods. The arrows confirm that the proposed model (Fine) improves the representation of bifurcation compared to the conventional model (Pre).

The PDs corresponding to the original and reconstructed images in Fig. 5 are shown in Fig. 6. In the original (left) and the Fine (right), two green points are positioned close to the left corner point $(0, 1)$ and apart from other noise components close to the diagonal line. On the other hand, in the Pre (middle) there is only one stable green point that is not a noise component. This indicates that the Pre cannot represent bifurcations of blood vessels, while the Fine, which incorporates the topological prior, can construct a model that generates blood vessels with bifurcations closer to the ones of the original.

6 Conclusion

In this article, we tackled the problem of modeling blood vessels with topological correctness based on a statistical appearance model, which has been a challenge for conventional methods. In order to incorporate topological priors

of bifurcations and undesired hole existence into statistical modeling, we integrated topological loss into the deep generative model (β-VAE). More precisely, the network is pre-trained with only the β-VAE loss and fine-tuned by adding the topological The topological loss is indeed used for refining the network as it may provide local optimums everywhere in the search space. In the experiments, we used a dataset of vascular patches with bifurcation structures to build a model using the proposed method. The experimental results showed that the proposed method improved the performance of the topology evaluation values while the performance of the intensity evaluation value decreased. Qualitatively, we also identified improvements of hole closing and representation of bifurcation structures in the reconstructed volumes. Future work would be to apply the constructed statistical appearance model of blood vessels to associated tasks such as vessel segmentation and super-resolution.

References

1. Bailleul, J., Ruan, S., Bloyet, D., Romaniuk, B.: Segmentation of anatomical structures from 3D brain MRI using automatically-built statistical shape models. In: ICIP 2004. vol. 4, pp. 2741–2744 (2004)
2. Bengio, Y., Courville, A., Vincent, P.: Representation Learning: A Review and New Perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence **35**(8), 1798–1828 (2013)
3. Blackall, J.M., King, A.P., Penney, G.P., Adam, A., Hawkes, D.J.: A statistical model of respiratory motion and deformation of the liver. In: MICCAI 2001. pp. 1338–1340 (2001)
4. Clough, J.R., Byrne, N., Oksuz, I., Zimmer, V.A., Schnabel, J.A., King, A.P.: A Topological Loss Function for Deep-Learning based Image Segmentation using Persistent Homology. IEEE Transactions on Pattern Analysis and Machine Intelligence (2020)
5. Cremers, D., Rousson, M., Deriche, R.: A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape. International Journal of Computer Vision **72**(2), 195–215 (2007)
6. Edelsbrunner, H., Harer, J.: Computational Topology - an Introduction. American Mathematical Society (2010)
7. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: MICCAI'98, vol. 1496, pp. 130–137 (1998)
8. Heimann, T., Meinzer, H.P.: Statistical shape models for 3D medical image segmentation: A review. Medical Image Analysis **13**(4), 543–563 (2009)
9. Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., Mohamed, S., Lerchner, A.: β-VAE: Learning basic visual concepts with a constrained variational framework. In: 2nd ICLR (2017)
10. Hu, X., Li, F., Samaras, D., Chen, C.: Topology-Preserving Deep Image Segmentation. In: NeurIPS 2019 (2019)
11. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: 3rd ICLR (2015)
12. Kingma, D.P., Welling, M.: Auto-encoding variational bayes. In: 2nd ICLR (2014)
13. Saeki, Y., Saito, A., Ueno, J., Harada, M., Shimizu, A.: Statistical intensity model of lung vessels in a ct volume using β-vae. In: CARS (2019)

14. Saito, A., Shimizu, A., Watanabe, H., Yamamoto, S., Nawano, S., Kobatake, H.: Statistical shape model of a liver for autopsy imaging. International Journal of Computer Assisted Radiology and Surgery **9**(2), 269–281 (2014)
15. Serra, J.: Image Analysis and Mathematical Morphology. Academic Press, USA (1983)
16. Styner, M.A., Rajamani, K.T., Nolte, L.P., Zsemlye, G., Székely, G., Taylor, C.J., Davies, R.H.: Evaluation of 3d correspondence methods for model building. In: Information Processing in Medical Imaging. vol. 2732, pp. 63–75 (2003)
17. Zomorodian, A., Carlsson, G.: Computing Persistent Homology. Discrete & Computational Geometry **33**(2), 249–274 (2005)