From Images to Probabilistic Anatomical Shapes: A Deep Variational Bottleneck Approach [*]

Jadie Adams[1,2], **Shireen Elhabian**[1,2]

[1]Scientific Computing and Imaging Institute, University of Utah, UT, USA

[2]School of Computing, University of Utah, UT, USA

Abstract

Statistical shape modeling (SSM) directly from 3D medical images is an underutilized tool for detecting pathology, diagnosing disease, and conducting population-level morphology analysis. Deep learning frameworks have increased the feasibility of adopting SSM in medical practice by reducing the expert-driven manual and computational overhead in traditional SSM workflows. However, translating such frameworks to clinical practice requires calibrated uncertainty measures as neural networks can produce over-confident predictions that cannot be trusted in sensitive clinical decision-making. Existing techniques for predicting shape with aleatoric (data-dependent) uncertainty utilize a principal component analysis (PCA) based shape representation computed in isolation of the model training. This constraint restricts the learning task to solely estimating pre-defined shape descriptors from 3D images and imposes a linear relationship between this shape representation and the output (i.e., shape) space. In this paper, we propose a principled framework based on the variational information bottleneck theory to relax these assumptions while predicting probabilistic shapes of anatomy directly from images without supervised encoding of shape descriptors. Here, the latent representation is learned in the context of the learning task, resulting in a more scalable, flexible model that better captures data non-linearity. Additionally, this model is self-regularized and generalizes better given limited training data. Our experiments demonstrate that the proposed method provides an accuracy improvement and better calibrated aleatoric uncertainty estimates than state-of-the-art methods.

Keywords

Uncertainty Quantification; Statistical Shape Modeling; Bayesian Deep Learning

1 Introduction

Statistical shape modeling (SSM) is an enabling tool in medicine and biology to quantify population-specific anatomical shape variation. SSM can detect pathological morphologies tied to impaired function and answer clinical hypotheses regarding anatomical cohorts (e.g., [3, 4, 11]). Two effective representations of shape in SSM are deformation fields

---

[*]Provisionally accepted for MICCAI 2022.

jadie@sci.utah.edu .

and *landmarks* - the former captures implicit transformations between images/shapes and a pre-defined atlas, while the latter are sets of explicit points consistently defined on the shape surface such that they are in correspondence across the population [20,21]. While the framework we propose is agnostic to the choice of shape representation, we demonstrate the approach using landmarks, both to be consistent with existing relevant methods and because landmarks are more intuitive and interpretable for statistical analyses and visualizing results [20,26]. Existing computational workflows, such as ShapeWorks [8,9], automatically place dense sets of landmarks or *correspondence points* on shapes segmented from 3D medical images. However, generating such a set of correspondence points or *point distribution model* (PDM) requires expert-driven, data-intensive steps: segmenting the anatomy of interest from 3D images, data preprocessing, shape registration, and correspondence optimization along with hyperparameter tuning.

Deep learning has alleviated these burdens by providing end-to-end solutions for mapping unsegmented 3D images (i.e., CT or MRI) to PDM with little preprocessing [1,5,6,23]. While the traditional SSM pipelines take hours, trained networks can discover statistical representations of anatomies directly from new images in seconds, creating the potential to streamline the adoption of SSM in research and practice. DeepSSM [5,6] is one such state-of-the-art framework that provides SSM estimates that perform statistically similar to traditional methods in downstream tasks [7]. DeepSSM relies on a supervised latent representation computed using principal component analysis (PCA) in advance of model training to enable data augmentation and incorporate prior shape knowledge, as has been done in related image-based tasks (e.g., [6,14,18,25,27]). However, PCA supervision imposes a linear relationship between the latent and the output space and restricts the learning task to strictly SSM prediction. Additionally, PCA does not scale in the case of large sets of high-dimensional shape data.

Another caveat of DeepSSM, and deep learning models at large, is that they can produce overconfident estimates that can not be blindly assumed to be accurate, especially in sensitive decision-making settings such as clinical practice. There are two forms of uncertainty in such frameworks; aleatoric (or datadependent) and epistemic (or model-dependent uncertainty, which can be explained away given enough training data) [15]. Learning from medical imaging data poses a challenge because scans can vary widely in quality and often suffer from coarse resolution, artifacts, and noise. These factors signify a strong need to capture uncertainty inherent in the input data; thus here, we chose to focus on aleatoric uncertainty quantification. Progress has been made in inferring probabilistic SSM from images to capture aleatoric uncertainty measures [1,23]. However, existing probabilistic SSM models share the same limitations as DeepSSM in that they rely on a predetermined PCA-based latent representation.

Here, we present a novel formulation for inferring statistical representations of anatomies from unsegmented images that mitigates these limitations by learning a latent representation in tandem via a variational information bottleneck [2]. In information bottleneck (IB) theory, a stochastic encoding $\mathbf{z}$ captures the minimal sufficient statistics required of input $\mathbf{x}$ to predict the output $\mathbf{y}$ [22]. The encoding $\mathbf{z}$ and model parameters Θ are estimated by maximizing the IB objective:

$$\text{argmax}_{\Theta}\, I(\mathbf{y}, \mathbf{z}; \Theta) - \beta I(\mathbf{x}, \mathbf{z}; \Theta) \tag{1}$$

where I is the mutual information and β, is a Lagrangian multiplier that controls the trade-off between predictive accuracy (encouraging $\mathbf{z}$ to be maximally expressive about $\mathbf{y}$) and model complexity (encouraging $\mathbf{z}$ to be maximally compressive about $\mathbf{x}$). By leveraging this theory to learn a latent representation in our proposed model, we achieve better accuracy and uncertainty calibration than existing state-of-the-art techniques. The advantages of this formulation include the following.

– The latent representation is learned in the context of the learning task rather than in isolation, resulting in better accuracy and a more scalable framework.

– The proposed model better captures the true shape distribution by allowing for a non-linear relationship between the latent encoding and output space.

– This formulation is self-regularized and thus generalizes better under limited training data without ad-hoc regularization methods.

– Because the latent space is unsupervised, the proposed model is general enough to accommodate other learning tasks. For example, it could be used to predict SSM plus additional clinically relevant quantities of interest.

– The proposed model outperforms existing state-of-the-art methods in both predictive accuracy and uncertainty calibration.

2 Methods

Given an unsegmented image $\mathbf{x}_n \in \mathbb{R}^{H \times W \times D}$, the task is to predict a statistical representation of an anatomy-of-interest in the form of a PDM $\mathbf{y}_n \in \mathbb{R}^{3M}$ which is a set of M 3D correspondence points, and the associated point-wise estimates of aleatoric uncertainty $\mathbf{a}_n \in \mathbb{R}^{3M}$. This task is solved by training a model with parameters Θ using a population of N–unsegmented 3D images $\mathcal{X} = \{\mathbf{x}_n\}_{n=1}^{N}$ and their corresponding PDMs denoted $\mathcal{Y} = \{\mathbf{y}_n\}_{n=1}^{N}$ The proposed method and state-of-the-art models used in comparison utilize a bottleneck and latent encoding $\mathcal{Z} = \{\mathbf{z}_n\}_{n=1}^{N}$ where $\mathbf{z}_n \in \mathbb{R}^{L}$ and $L \ll 3M$.

2.1 VIB-DeepSSM Formulation

In the proposed method, denoted VIB-DeepSSM, the latent encoding is learned by minimizing the IB objective (Eq. 1). Given that directly calculating mutual information is ill-posed in this context, variational inference provides a way to approximate the problem given an empirical data distribution. In "Deep Variational Information Bottleneck" [2], the IB model is parameterized via a neural network by minimizing the derived theoretical lower bound on the IB objective:

$$VIB = \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}_{\epsilon \sim p(\epsilon)}[-\log q_{\phi}(\mathbf{y}_n \mid \mathbf{z}_{n,\epsilon})] + \beta \text{KL}[p_{\theta}(\mathbf{z}_n \mid \mathbf{x}_n) \parallel r(\mathbf{z})] \tag{2}$$

This loss balances a Kullback-Leibler divergence (KL) and a data fidelity (or reconstruction) term formulated as negative log-likelihood (NLL) with the Lagrangian multiplier $\beta \in [0, 1]$ to learn model parameters $\Theta = \{\varphi, \theta\}$. The KL term encourages maximal compression of $\mathbf{x}$, while the reconstruction term encourages maximal expression of $\mathbf{y}$. We define a stochastic encoder of the form $p_\theta(\mathbf{z} \mid \mathbf{x}) = \mathcal{N}\big(\mathbf{z} \mid f_e^\mu(\mathbf{x}; \theta), f_e^\Sigma(\mathbf{x}; \theta)\big)$ where f_e is a convolutional network parameterized by θ that maps input image to a Gaussian latent distribution. Here, $r(\mathbf{z})$ is a variational approximation to the marginal distribution of $\mathbf{z}$, which we select to be a fixed L-dimensional spherical unit Gaussian, $r(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{0}, \mathbb{I})$. The decoder, $q_\varphi(\mathbf{y}|\mathbf{z})$, serves as a variational approximation to the intractable $p(\mathbf{y}|\mathbf{z})$. We define it as $q_\varphi(\mathbf{y}|\mathbf{z}) = f_d(\mathbf{z}; \varphi)$, where f_d is an MLP parameterized by the weights φ (with non-linear activation functions) that maps the latent representation to correspondence points (Fig. 1). To enable gradient calculation, posterior samples $\mathbf{z}_{n,\in}$ are acquired using the reparameterization trick [17]: $\mathbf{z}_{n,\epsilon} = f_e^\mu(\mathbf{x}_n; \theta) + \epsilon f_e^\Sigma(\mathbf{x}_n; \theta)$ where $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbb{I})$. In our experiments, the expectation is estimated using 30 samples. The uncertainty is quantified as the entropy of this distribution $p(\mathbf{y}|\mathbf{z})$.

2.2 Baseline Models in Comparison

We compare the proposed model against the state-of-the-art deterministic formulation that does not predict uncertainty, PCA-DeepSSM [6], as well as two state-of-the-art stochastic formulations that do predict uncertainty: PPCA-DeepSSM [1] and PPCA-Offset-DeepSSM [23] (Fig. 1). In these frameworks, the latent representation is supervised via PCA scores computed from the training correspondence points $\mathbf{y}_n$. The latent space dimension, L, is chosen such that 95% of variability is preserved. We denote the PCA scores $\mathcal{Z}^{pca} = \{\mathbf{z}_n^{pca}\}_{n=1}^N$ where $\mathbf{z}_n^{pca} \in \mathbb{R}^L$ Thus $\mathbf{y}_n \approx \mathbf{U}\mathbf{z}_n^{pca} + \boldsymbol{\mu}$ where $\mathbf{U} \in \mathbb{R}^{3M \times L}$ is the matrix of eigenvectors and $\boldsymbol{\mu} \in \mathbb{R}^{3M}$ is the mean of the training correspondence points. In the proposed model, VIB-DeepSSM, the latent space is not supervised, yet its dimensionality should be predefined. We found that PCA provides a reasonable, reproducible way to estimate latent dimension size in the unsupervised case as well, thus in VIB-DeepSSM, we use the same L-value, though this is not a requirement.

PCA-DeepSSM [6] is a deterministic framework that uses PCA scores as a supervised latent space. The decoder is a single fixed, fully connected layer without activation with the PCA basis as weights and mean shape as bias. The loss function is mean square error (MSE) between ground-truth and predicted PCA scores.

PPCA-DeepSSM [1] is equivalent to Uncertain-DeepSSM [1], except dropout is removed as we are not analyzing epistemic uncertainty quantification. Here the latent space is supervised and the encoder is stochastic, where PCA variance serves as a regularization term and is learned implicitly via the loss (NLL of the true PCA scores given the predicted distribution). The decoder is fixed PCA reconstruction, as it is in PCA-DeepSSM [6], but here the predicted correspondence points are acquired by averaging over posterior samples.

PPCA-Offset-DeepSSM [23] is based on the approach in Tóthová *et al.* [23], but here the architecture is altered to predict 3D PDM instead of 2D shape vertices. In this formulation, the encoder also predicts a distribution of PCA scores; however, in addition, an offset term $f_e^s(\mathbf{x}) \in \mathbb{R}^{3M}$ is predicted and added to the point prediction (Fig 1). The offset term helps

address the linearity restriction, but it is unregularized and thus prone to overshoot. An ad hoc regularization scheme is utilized in this formulation in the form of KL divergence between an assumed prior distribution of $\mathbf{z}$ (selected to be $p(\mathbf{z}) = \mathrm{N}(0, I)$) and the observed one in the training set. The loss is the PDM NLL plus this KL term (Fig 1). Note that in this loss, the sum of the $N \mathbf{z}$ predicted distributions (i.e., the aggregate posterior) is regularized by the KL term, and in the VIB loss (Eq. 2), each of the $N \mathbf{z}$ predicted distributions are individually regularized. Hence, the VIB loss effectively minimizes the KL divergence between the aggregate posterior and the latent prior while minimizing the mutual information between $\mathbf{x}$ and $\mathbf{z}$ [13]. This behavior is dictated by the information bottleneck; learn the $\mathbf{z}$ descriptor that is minimally informative of the input image $\mathbf{x}$ but maximally predictive of the shape $\mathbf{y}$.

3 Results

To best analyze accuracy and uncertainty calibration across model variants, we select a large, highly variable dataset for experiments. In the original DeepSSM [6] formulation, a data augmentation scheme is used to create additional training samples. However, this technique relies on PCA, which we intend to remove as a requirement; thus here training is done without augmentation.

3.1 Left Atrium Dataset

In experiments, we use a set of 1001 anonymized LGE MRI images of the left atrium (LA) from unique atrial fibrillation patients. The images were segmented by domain experts with spatial resolution $0.65 \times 0.65 \times 2.5$ mm^3 and the endocardium wall was used to cut off pulmonary veins. Significant morphological variations are expected to be in the left atrium appendage or LAA (varies immensely in size across patients), the pulmonary veins (vary in number and size across patients), and in the mitral valve (for which there are no defining image features for segmentation). This dataset is appropriate for aleatoric uncertainty analysis because, in addition to the large shape variation, the input images vary widely in intensity and quality, and LA boundaries are blurred and have low contrast with the surrounding structures (Fig. 2). To enable calibration analysis, we define a specific testing set of images with high uncertainty using an outlier degree computed on input images that combines within- and off-subspace distances [1,19]. By thresholding on outlier degree at 3 (Fig. 2), we define an *outlier test* set of size 78. We randomly split the remaining samples (90%, 10%, 10%) to get a training set of 739, a validation set of 92, and an *inlier test* set of 92.

Acquiring target correspondence points for the training images requires generating a PDM via the traditional SSM pipeline (i.e., segmentation, preprocessing, and optimization). We employ the open-source *ShapeWorks* software [8] to align and crop the images and binary segmentations and optimize a PDM with 1024 landmarks per shape. The resulting input images are size (166, 120, 125). Given the optimized training PDM, we generate target PDMs for the validation and test sets for analysis by optimizing landmarks positions on these samples with the training landmarks fixed so that the statistics of the validation and

test samples are not reflected in the training PDM. The first two modes of variations in the training PDM are shown in Fig. 3.

3.2 Training Scheme

VIB-DeepSSM and the baseline methods are trained on 100%, 80%, 40%, 20%, and 10% of the training data to analyze robustness in low-sample-size scenarios. We select subsets of the training data using stratified random sampling by clustering the training PDMs and selecting data subsets so that each cluster is equally represented. All variants are implemented in PyTorch and training is run on GTX1080Ti GPUs with Adam optimization [16], a fixed learning rate of $5e^{-5}$, a batch size of 6, Xavier weight initialization [10], and Parametric ReLU [12] activation. Models are trained until the validation PDM MSE has not decreased in 50 epochs. A loss burn-in scheme is used for stochastic variants to allow the model to learn to predict deterministic shapes before estimating variance. This is done using a linear combination of the given stochastic loss and deterministic loss (PDM MSE), where the weight changes over a given number of epochs from entirely deterministic to entirely stochastic. We set the number of burn-in epochs to be proportional to the training data size (i.e., 100 epochs for 100%, 80 epochs for 80%, etc.). To increase stability when training stochastic encoders, we predict the log of the latent variance, ensuring positivity and removing the potential for division by zero. We empirically set hyperparameters using the validation MSE and find for PPCA-Offset-DeepSSM [23], $\lambda = 100$ (tested range: $[1e^{-4}, 1e^{4}]$) and for VIB-DeepSSM, $\beta = 0.01$ (tested range: $[1e^{-8},1]$).

3.3 Accuracy Analysis

Prediction accuracy is evaluated by calculating the relative root mean square error (RRMSE) between the true and predicted correspondence points. This is evaluated sample-wise, $\textbf{RRMSE} = \sqrt{\left(\|\mathbf{y}_n - \hat{\mathbf{y}}_n\|^2\right)/3M}$ and point-wise where for point i, $\textbf{Point-RRMSE} = \sqrt{\left(\|\mathbf{y}_n^i - \hat{\mathbf{y}}_n^i\|^2\right)/3}$. We also analyze prediction accuracy using the **surface-to-surface** distance between the ground truth mesh of the anatomy and one constructed from the predicted PDM. To construct a mesh from predicted points, we find the closest example in the training set and apply the warp between the points to its mesh using [24]. The accuracy using different training sizes is summarized in Fig. 4. The proposed method has a similar or better accuracy on both testing inliers and outliers than the baseline methods given any training size. Accuracy is most notably improved when training data is limited. Note that for PPCA-Offset-DeepSSM [23], accuracy does not significantly improve with increased training data. This suggests that the model is over-regularized, however decreasing the λ parameter results in unstable training. This suggests VIB-DeepSSM provides a more principled way to capture data non-linearity and self-regularization.

3.4 Uncertainty Calibration Analysis

To assess uncertainty calibration, we consider the correlation between (1) the input image outlier degree and entropy of $p(\mathbf{y}|\mathbf{z})$, (2) RRMSE and entropy of $p(\mathbf{y}|\mathbf{z})$, and (3) point-RRMSE and predicted point-wise standard deviation (volume of the 3D Gaussian for each landmark/point). We quantify correlation using the Pearson correlation coefficient, where a higher value suggests the model more effectively identifies input with high

uncertainty and areas of low prediction confidence. Our models predicted uncertainty is better calibrated both in relation to outlier degree and error (Fig. 5). The heat maps generated via interpolation in Fig. 5 show that point uncertainty is correctly predicted to be higher in these areas where the error is highest (LAA, pulmonary veins, and mitral valve).

4 Conclusion

We presented a novel principled framework based on the IB theory for predicting probabilistic SSM from unsegmented 3D images and demonstrated that it provides better calibrated aleatoric uncertainty quantification than existing state-of-the-art techniques without sacrificing accuracy. By learning a latent representation in the context of the task, we provide a more scalable, flexible framework that better captures data non-linearity. Additionally, the proposed method is self-regularized and generalizes better under limited training data. These contributions increase the feasibility of using SSM in research and medicine by both bypassing the time and cost-prohibitive steps of traditional SSM and providing the necessary safeguard against model over-confidence. This has the potential to improve medical standards and increase patient accessibility to statistic-based diagnosis. In future work, we plan to analyze the effectiveness of this model for predicting SSM with clinically relevant quantities, and to explore techniques for learning the optimal latent dimension size in tandem with the task.

Acknowledgements

References

1. Adams J, Bhalodia R, Elhabian S: Uncertain-deepssm: From images to probabilistic shape models. Shape in Medical Imaging : International Workshop, ShapeMI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Proceedings 12474, 57–72 (2020)

2. Alemi A, Fischer I, Dillon J, Murphy K: Deep variational information bottleneck. In: ICLR (2017), https://arxiv.org/abs/1612.00410

3. Atkins PR, Elhabian SY, Agrawal P, Harris MD, Whitaker RT, Weiss JA, Peters CL, Anderson AE: Quantitative comparison of cortical bone thickness using correspondence-based shape modeling in patients with cam femoroacetabular impingement. Journal of Orthopaedic Research 35(8), 1743–1753 (2017) [PubMed: 27787917]

4. Bhalodia R, Dvoracek LA, Ayyash AM, Kavan L, Whitaker R, Goldstein JA: Quantifying the severity of metopic craniosynostosis: A pilot study application of machine learning in craniofacial surgery. Journal of Craniofacial Surgery (2020)

5. Bhalodia R, Elhabian S, Adams J, Tao W, Kavan L, Whitaker R: Deepssm: A blueprint for image-to-shape deep learning models (2021)

6. Bhalodia R, Elhabian SY, Kavan L, Whitaker RT: Deepssm: A deep learning framework for statistical shape modeling from raw images. CoRR abs/1810.00111 (2018), http://arxiv.org/abs/1810.00111

7. Bhalodia R, Goparaju A, Sodergren T, Whitaker RT, Morris A, Kholmovski E, Marrouche N, Cates J, Elhabian SY: Deep learning for end-to-end atrial fibrillation recurrence estimation. In: Computing in Cardiology, CinC 2018, Maastricht, The Netherlands, September 23–26, 2018 (2018)

8. Cates J, Elhabian S, Whitaker R: Shapeworks: Particle-based shape correspondence and visualization software. In: Statistical Shape and Deformation Analysis, pp. 257–298. Elsevier (2017)

9. Cates J, Fletcher PT, Styner M, Shenton M, Whitaker R: Shape modeling and analysis with entropy-based particle systems. In: IPMI. pp. 333–345. Springer (2007)

10. Glorot X, Bengio Y: Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research, vol. 9, pp. 249–256. PMLR (13–15 May 2010)

11. Harris MD, Datar M, Whitaker RT, Jurrus ER, Peters CL, Anderson AE: Statistical shape modeling of cam femoroacetabular impingement. Journal of Orthopaedic Research 31(10), 1620–1626 (2013). https://doi.org/10.1002/jor.22389, 10.1002/jor.22389 [PubMed: 23832798]

12. He K, Zhang X, Ren S, Sun J: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. CoRR abs/1502.01852 (2015), http://arxiv.org/abs/1502.01852

13. Hoffman MD, Johnson MJ: Elbo surgery: yet another way to carve up the variational evidence lower bound. In: Workshop in Advances in Approximate Bayesian Inference, NIPS. vol. 1 (2016)

14. Huang W, Bridge CP, Noble JA, Zisserman A: Temporal heartnet: Towards human-level automatic analysis of fetal cardiac screening video. In: MICCAI 2017. pp. 341–349. Springer International Publishing (2017)

15. Kendall A, Gal Y: What uncertainties do we need in bayesian deep learning for computer vision? CoRR abs/1703.04977 (2017), http://arxiv.org/abs/1703.04977

16. Kingma D, Ba J: Adam: A method for stochastic optimization. International Conference on Learning Representations (12 2014)

17. Kingma DP, Welling M: Auto-encoding variational bayes. CoRR abs/1312.6114 (2014)

18. Milletari F, Rothberg A, Jia J, Sofka M: Integrating statistical prior knowledge into convolutional neural networks. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins DL, Duchesne S (eds.) Medical Image Computing and Computer Assisted Intervention. pp. 161–168. Springer International Publishing, Cham (2017)

19. Moghaddam B, Pentland A: Probabilistic visual learning for object representation. IEEE Transactions on pattern analysis and machine intelligence 19(7), 696–710 (1997)

20. Sarkalkan N, Weinans H, Zadpoor AA: Statistical shape and appearance models of bones. Bone 60, 129–140 (2014) [PubMed: 24334169]

21. Thompson D: On Growth and Form. Cambridge University Press (1917)

22. Tishby N, Pereira FC, Bialek W: The information bottleneck method (2000)

23. Tóthová K, Parisot S, Lee MCH, Puyol-Antón E, Koch LM, King AP, Konukoglu E, Pollefeys M: Uncertainty quantification in cnn-based surface prediction using shape priors. CoRR abs/1807.11272 (2018), http://arxiv.org/abs/1807.11272

24. Wang Y, Jacobson A, Barbi J, Kavan L: Linear subspace design for real-time shape deformation. ACM Transactions on Graphics (TOG) 34(4), 1–11 (2015)

25. Xie J, Dai G, Zhu F, Wong EK, Fang Y: Deepshape: deep-learned shape descriptor for 3d shape retrieval. IEEE transactions on pattern analysis and machine intelligence 39(7), 1335–1345 (2017) [PubMed: 27482929]

26. Zachow S: Computational planning in facial surgery. Facial Plastic Surgery 31(05), 446–462 (2015) [PubMed: 26579861]

27. Zheng Y, Liu D, Georgescu B, Nguyen H, Comaniciu D: 3d deep learning for efficient and robust landmark detection in volumetric data. In: MICCAI 2015. pp. 565–572. Springer International Publishing (2015)
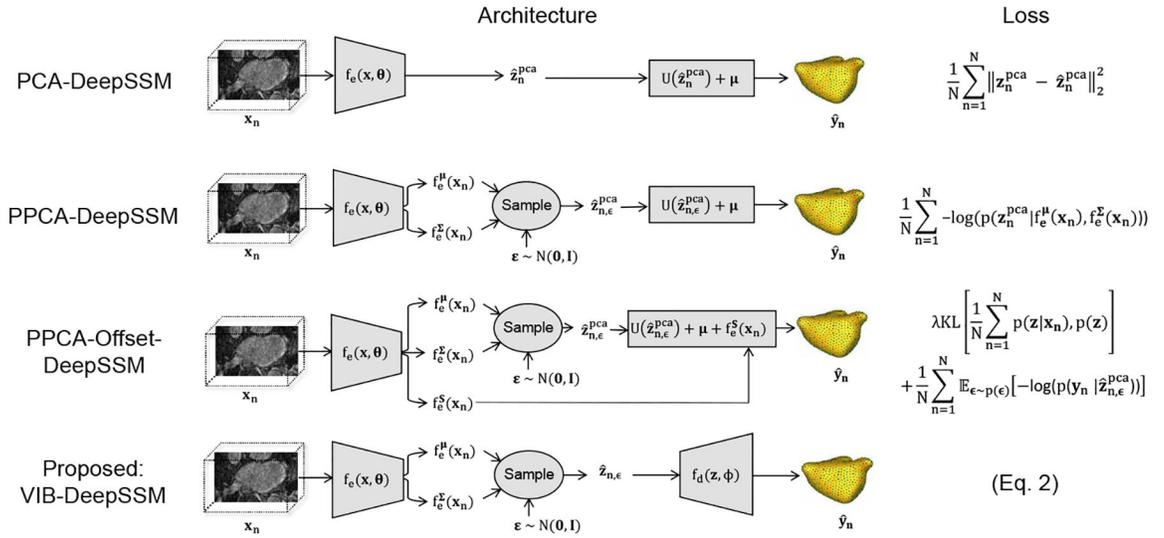
Fig. 1: Model Variants:

Architecture and loss of the proposed method and baseline models in comparison are shown. In each case, the encoder $f_e^\mu(\mathbf{x}; \theta)$ has the same architecture as in DeepSSM [6]; five convolutional layers with batch normalization followed by two fully connected layers. Only the output size of the last layer of the encoder is variant-dependent. In VIB-DeepSSM, the decoder $f_d(\mathbf{z}; \varphi)$ is comprised of three fully connected layers with non-linear activations.
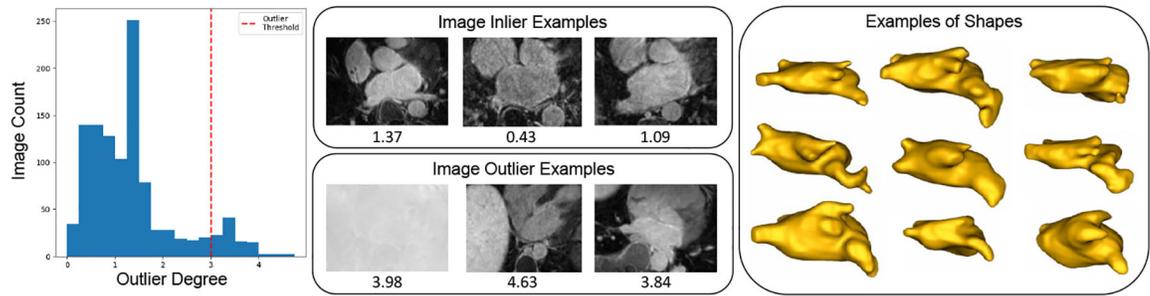
Fig. 2:

The distribution of image outlier degrees is displayed with examples of image slices and the corresponding outlier degree. Outlier images tend to be over-exposed with low contrast or contain artifacts. Examples of ground truth meshes are displayed from the top view to demonstrate the significant shape variability.
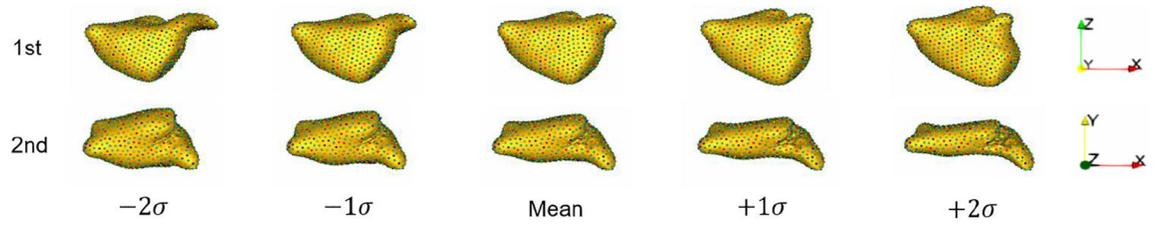
Fig. 3:
Modes of variation from the training PDM generated using ShapeWorks. The primary mode shown from anterior view captures the LAA elongation. The secondary mode shown from top view captures the LA sphericity.
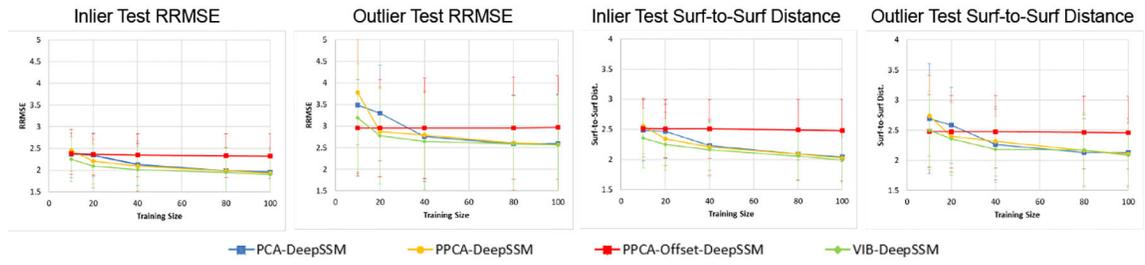
Fig. 4: Accuracy:

The mean accuracy is shown for each model variant with error bars representing standard deviation. Overall, VIB-DeepSSM performs best.
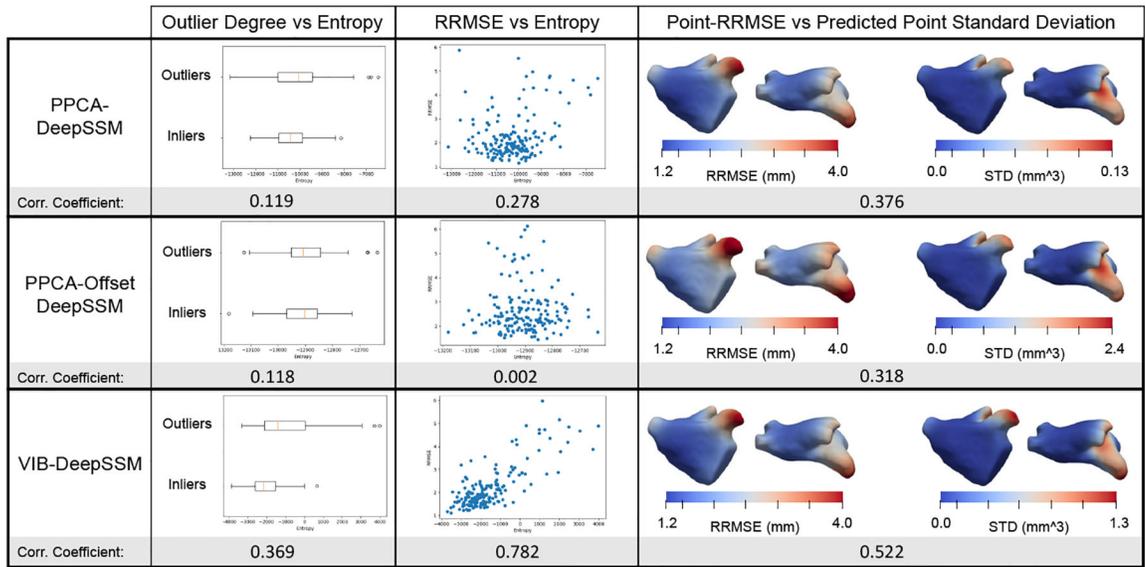
Fig. 5: Uncertainty Calibration

Plots and correlation coefficients show the uncertainty calibration of each model variant.

Heat maps show average quantities on a representative mesh.