


S3M: Scalable Statistical Shape Modeling through Unsupervised Correspondences

Lennart Bastian* , Alexander Baumann*, Emily Hoppe, Vincent Bürgin, Ha Young Kim, Mahdi Saleh, Benjamin Busam, and Nassir Navab

Computer Aided Medical Procedures, Technical University Munich, Germany

Abstract. Statistical shape models (SSMs) are an established way to represent the anatomy of a population with various clinically relevant applications. However, they typically require domain expertise, and labor-intensive landmark annotations to construct. We address these shortcomings by proposing an unsupervised method that leverages deep geometric features and functional correspondences to simultaneously learn local and global shape structures across population anatomies. Our pipeline significantly improves unsupervised correspondence estimation for SSMs compared to baseline methods, even on highly irregular surface topologies. We demonstrate this for two different anatomical structures: the thyroid and a multi-chamber heart dataset. Furthermore, our method is robust enough to learn from noisy neural network predictions, potentially enabling scaling SSMs to larger patient populations without manual segmentation annotation. The code is publically available at: <https://github.com/alexanderbaumann99/S3M>


Keywords: Statistical Shape Modeling, Unsupervised Correspondence Estimation, Geometric Deep Learning

1 Introduction

Statistical shape models (SSMs) are a powerful tool to characterize anatomical variations across a population. They have been widely used in medical image analysis and computational anatomy to represent organ structures, with numerous clinically relevant applications such as clustering, classification, and shape regression [21,7,2]. SSMs are generally represented by point-wise correspondences between shapes [13,4], or deformation fields to a pre-defined template [14,28]. Despite the existence of implicit models [2], abstracting shape correspondences in the form of linear point distribution models (PDM) constitutes an appealing and interpretable way to represent shape distributions [3]. Furthermore, many implicit models still rely on correspondence annotations during training [7,2].

Creating SSMs is cumbersome and intricate, as significant manual human annotation is necessary. Domain experts typically first segment images in 3D. The labeled 3D organ surfaces must then be aligned and brought into correspondence, typically achieved through deformable image registration methods

* denotes equal contribution.

: lennart.bastian@tum.de

using manual landmark annotations [6]. This is labor-intensive and error-prone, potentially inducing bias into downstream SSMs and applications [39].

Unsupervised methods have been proposed to estimate correspondences for SSMs [11,24,4]. However, they typically require precisely segmented and smooth surfaces to generate accurate inter-organ correspondences. ShapeWorks has been established to produce high-quality correspondences on several organs such as femurs or left atria [1,7,3]. However, as domain experts carefully curate most medical datasets, the robustness of such methods has not yet been thoroughly evaluated concerning label noise and segmentation inaccuracies. The main obstacles that prevent scaling SSMs to larger patient populations are unsupervised correspondence methods that can handle topological variations in noisy annotations, such as those produced by inexperienced annotators or predictions from deep neural networks. Robust methods to deal with these obstacles are required.

To address these challenges, we propose S3M, which leverages unsupervised deep geometric features while incorporating a global shape structure. Geometric Deep Learning (GDL) provides techniques to process 3D shapes and geometries, which are robust to noise, 3D rotations, and global deformations. We utilize graph neural networks (GNN) and functional mappings to establish dense surface correspondences of samples without supervision. This approach has significant clinical implications as it enables automatically representing anatomical shapes across large patient populations without requiring manual expert landmark annotations. We demonstrate that our proposed method creates objectively superior SSMs from shapes with noisy surface topologies. Moreover, it accurately corresponds regions of complex anatomies with mesh bifurcations such as the heart, which could ease the modeling of inter-organ relations [12].

Our contributions can be summarized as follows:

- We propose a novel unsupervised correspondence method for SSM curation based on geometric deep learning and functional correspondence.
- S3M exhibits superior performance on two challenging anatomies: thyroid and heart. It generates objectively more suitable SSMs from noisy thyroid labels and challenging multi-chamber heart reconstructions.
- To pave the way for unsupervised SSM generation in other medical domains, we open-source the code of our pipeline.

2 Related Work

Point Distribution Models. A population of shapes must be brought into correspondence to construct a PDM. This has been traditionally achieved through pair-wise registration methods [21,19]. However, pairwise approaches can admit bias as they neglect the population during correspondence generation [19]. More recently, group-wise optimization methods such as ShapeWorks [11] have been adopted as they jointly optimize over a cohort, overcoming such biases. They demonstrate superior prediction of clinically relevant anatomical variations [19]. Furthermore, generic models that can perform well across various organs are sought after.

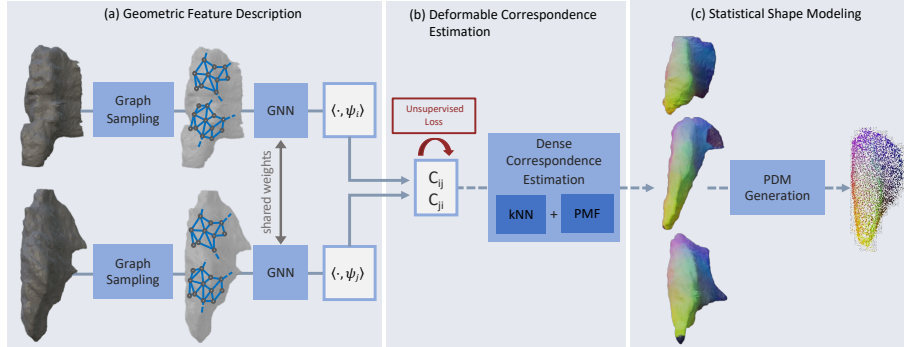


Fig. 1. Our proposed method for unsupervised SSM curation. (a) We use a Siamese GNN as shape descriptor and project the extracted features onto the LBO eigenfunctions ψ to obtain spectral representations. (b) From these, we optimize a functional mapping between pairs of shapes with an unsupervised loss. Gradients are backpropagated to the geometric descriptors. (c) During inference, the dense correspondences are estimated between pairs of shapes based on the learned population parameters, which are then used to construct an SSM.

Graph Neural Networks (GNNs) have been used to enable structural feature extraction through message passing. They constitute a powerful tool to process 3D data and extract geometric features [38,32,33] which can be useful for disease prediction [23,30]. Other medical applications involve brain cortex mesh reconstruction [8] and 3D organ-at-risk (ORA) segmentation [22]. We use GNNs for deformable 3D organ shapes and learn to estimate dense correspondences in the presence of noise and anatomical variations.

Functional Correspondence. Functional maps abstract the notion of point-to-point correspondences by corresponding arbitrary functions, such as texture or curvature, across shapes. They are extensively used to estimate dense correspondences across deformable shapes [29] and can be incorporated into learning frameworks [26,17]. These methods are typically evaluated on synthetic meshes with dense annotations and limited variable surface topology. In contrast, medical shapes exhibit higher variability, requiring robust surface representation for reliable correspondence matching. More recently, unsupervised functional correspondence models have been proposed [31,10]. These methods demonstrate strong performance on synthetic data without manual correspondence annotations. They extract features from the surface geometry using hand-crafted descriptors such as SHOT [36], wave-kernel signatures (WKS) [5] or heat-kernel signatures (HKS) [9]. The extracted features are then typically refined and projected onto the Laplace-Beltrami Operator (LBO) eigenfunctions [29]. μ Match [24] recently leveraged such an unsupervised approach in the medical domain. They employ handcrafted features to extract representations from shapes with a relatively smooth surface topology; however, they fail for shapes with high degrees of surface noise or label inconsistencies. To scale SSM curation to larger

datasets encompassing population variance, our method must be robust to a more variable and complex surface topology.

3 Method

In the following, we propose a method to establish an SSM as a Point Distribution Model (PDM), illustrated in Figure 1. Robust local features from the surface mesh are extracted using GNNs. These features are then projected onto the truncated eigenspace of the Laplace-Beltrami Operator using $m = 20$ eigenfunctions [29]. We perform post-processing with a Product Manifold Filter (PMF) [37] to obtain bijective correspondences for SSM generation. The shape model is subsequently created by aggregating correspondences across a dataset of predicted correspondences using the eigendecomposition of the covariance matrix.

Geometric Feature Description. Handcrafted descriptors [36,5,9] are unable to represent the complex surface topology of medical data adequately. To cope with surface artifacts and irregular morphologies, we use a graph-based descriptor [32]. We first extract a surface mesh from a 3D volumetric grid using marching cubes [27]. Graph nodes are defined as the mesh’s vertices; edges are obtained using a k-nearest neighbor search with $k = 10$. Node features are given by spatial xyz-coordinates. The graph is fed into three topology adaptive layers [18] using graph convolutions with a specific number of hops to define the number of nodes a message is passed to. Increasing the number of hops (we use 1, 2, 3 hops per layer, respectively) increases the receptive field, incorporating features from more distant nodes. Finally, features pass a linear layer before being projected onto the Laplace-Beltrami eigenfunctions.

Deformable Correspondence Estimation. PDMs require correspondences between samples to model the statistical distribution of the organ. Inspired by methods for geometric shape correspondence, we propose to estimate a functional mapping T to correspond high-level semantics from two input shapes, X_i and X_j . The LBO extends the Laplace operator to Riemannian manifolds, capturing intrinsic characteristics of the shape independent of its position and orientation in Euclidean space. It can be efficiently computed on a surface mesh using, for example, the cotangent weight discretization scheme [15]. This results in a matrix representation of the LBO from which one can then calculate the associated eigenfunctions $\psi_i \in \mathbf{R}^{n \times m}$ for a shape $X_i \in \mathbf{R}^{n \times 3}$. Given a feature vector D_i extracted from a surface mesh of shape X_i and a neural network T_ϕ , we can approximate a functional mapping between shapes by solving the following optimization problem:

$$\min_{\phi} \sum_{(X_i, X_j)} \mathcal{L}(C_{ij}, C_{ji}) \quad \text{where} \quad C_{ij} = \arg \min_C \|CA_{T_\phi(D_i)} - A_{T_\phi(D_j)}\| \quad (1)$$

Here, $A_{T_\phi(D_i)} \in \mathbf{R}^{m \times m}$ denotes the transformed descriptor, written in the basis of the LBO eigenfunctions of shape X_i and $C_{ij} \in \mathbf{R}^{m \times m}$ represents the optimal functional mapping from the descriptor space of X_i to the one of X_j . Inspired

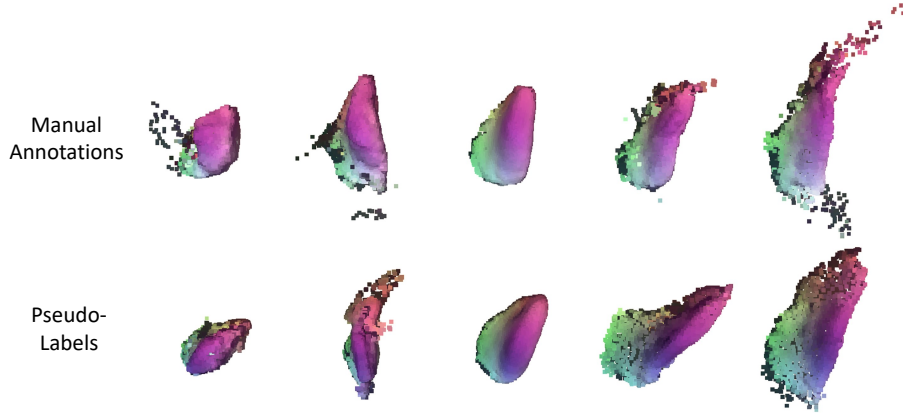


Fig. 2. PDM results from the proposed method S3M on the thyroid dataset [25]. The top row depicts the PDM generated from manual annotations and the bottom row from network pseudo-labels. From left to right, we depict $-3\sqrt{\lambda_1}$, $-3\sqrt{\lambda_2}$, the mean shape, $+3\sqrt{\lambda_2}$, $+3\sqrt{\lambda_1}$, with λ_i the eigenmode corresponding to the i -th largest eigenvalue. Similar colors indicate corresponding regions predicted by the model.

by existing works on shape correspondence [10,31,34], our loss function enforces four separate characteristics on the learned functional mapping, including bijectivity, orthogonality, and isometric properties. We refer to the supplementary materials for the complete definition. Notably, none of these losses uses ground truth correspondences, making the entire process unsupervised.

Training and Inference. During training, two shapes are sampled from the dataset, and the pipeline is optimized with Equation 1. We increase model robustness by augmenting with rotations and small surface deformations. The point cloud is sub-sampled in each training iteration using farthest point sampling with random initialization. During inference, our model predicts pairwise correspondences. To accumulate these over an entire dataset of N shapes, we choose a template shape $X_T = \arg \min_{X_i} \sum_{j=0}^N \mathbb{1}_{i \neq j} \mathcal{L}(C_{ij}, C_{ji})$ as the instance with the lowest average loss to all other shapes in the dataset. As in [29], we extract point-to-point correspondences between the template X_T and another shape X_i by matching the transformed LBO eigenfunctions of X_i , namely $C_{iT}\psi_i$, with the LBO eigenfunctions of the template ψ_T using nearest neighbors. As PDMs require bijective correspondences, we subsequently post-process the results with PMF [37].

Statistical Shape Modeling. We use the PDM [13] as the underlying method of the shape model. It takes input points of the form $X \in \mathbf{R}^{N \times d}$, where N , d are the number of shape samples and coordinates per shape, respectively. It returns a multi-variate normal distribution. In our case, $d = 3n$ given each shape has n points. We calculate the mean shape \bar{X} and the empirical covariance matrix $S = \text{cov}(X)$ over the N samples [13]. Since S has rank $N - 1$, it has $N - 1$ real

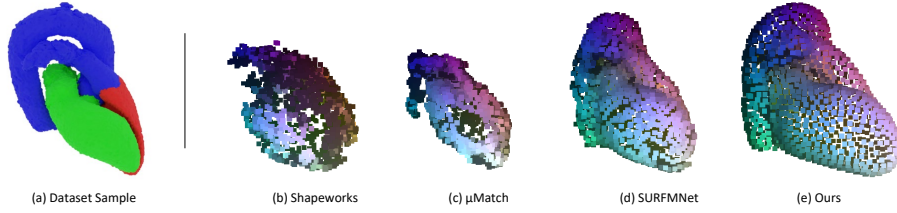


Fig. 3. Qualitative Analysis of Whole-Heart SSMs. A sample from the heart dataset (a). Composition 1 (right ventricle) is denoted in red. Composition 2 (both ventricles and atria) combines red and green regions. Composition 3 additionally incorporates the vessels, denoted in blue. The predicted SSM mean shapes for composition 3 are portrayed for ShapeWorks (b), μ Match (c), SURFMNet (d), and S3M (e).

eigenvectors v_j with eigenvalues λ_j . If we consider the sum

$$s = \bar{X} + \sum_{j=1}^{N-1} \alpha_j \lambda_j v_j, \quad \alpha_j \sim \mathcal{N}(0, 1) \quad (2)$$

then $s \sim \mathcal{N}(\bar{X}, S)$, which is the desired distribution of the model. For the above, the points must be in correspondence across the samples. We thus use the correspondences generated in section 3 to construct the PDM.

4 Experiments

All experiments are carried out using two publicly available datasets: thyroid ultrasound scans and heart MRI acquisitions. Our model is implemented in PyTorch 1.12 using CUDA 11.6. Training takes between 2.5 – 3 hours on an Nvidia A40 GPU and inference about 0.71 seconds for a pair of shapes. We use publicly available implementations for all baseline methods.

Thyroid Dataset (SegThy) [25]. The dataset comprises 3D freehand US scans of healthy thyroids from 32 volunteers aged 24-39. For each volunteer, three physicians acquired three scans each. Ultrasound scans generally exhibit noise induced by physical properties such as phase aberrations and attenuation. This leads to label inconsistencies or topological irregularities that pose a challenge for shape modeling (see Figure 2). US sweeps were compounded to a 3D grid of resolution $0.12 \times 0.12 \times 0.12 \text{mm}^3$. A single scan from each of the 16 volunteers was manually annotated by experts (ground truth) and used to train QuickNAT [25]. The remaining scans were pseudo-labeled through QuickNAT segmentation predictions exhibiting moderate degrees of noise and inaccuracies (dice score of 0.94 [25]). We divide the dataset into manual and pseudo-label predictions and evaluate them separately. The pseudo-label experiment evaluates the model’s performance under topological noise and inaccuracies and is limited to 100 scans

due to ShapeWorks memory constraints [20]. We extract a surface mesh for each scan using marching cubes and subsample the meshes to 5000 vertices.

Heart Dataset [35]. The data constitutes 30 MRI scans from a single cardiac phase of the heart. Each image has a voxel resolution of $1.25 \times 1.25 \times 2.7 \text{ mm}^3$. Segmentation is carried out using an automated method, with subsequent manual corrections by domain experts. Labels are provided for: the right/left ventricle, right/left atrium, aorta, and pulmonary artery. We evaluate the capability of the models to reconstruct complex organs using three hierarchical compositions of the heart chambers. *Composition 1* consists of the right ventricle, *Composition 2* of the left and right atrium and ventricle, and *Composition 3* of the whole heart, including the aorta and pulmonary artery.

Table 1. SSM quality metrics for the Thyroid dataset [25]

Metrics	Ground-Truth Segmentation		Network Pseudo-Labels	
	Generality [mm] ↓	Specificity [mm] ↓	Generality [mm] ↓	Specificity [mm] ↓
SURFMNet	2.20 ± 0.20	3.20 ± 0.29	-	-
μ Match	1.92 ± 0.07	2.81 ± 0.18	1.90 ± 0.08	2.84 ± 0.10
Shapeworks	1.94 ± 0.27	1.81 ± 0.06	1.60 ± 0.04	1.75 ± 0.07
S3M (Ours)	1.25 ± 0.11	1.59 ± 0.06	0.95 ± 0.07	1.84 ± 0.08

SSM Evaluation. We compare Shapeworks [11], μ Match [24] and SURFMNet [31], with S3M. A four-fold cross-validation is employed. SURFMNet, μ Match, and S3M are trained on the training folds and correspondences are predicted on the training and validation set. Since the particle-based optimization of ShapeWorks does not generalize to unseen data, it uses all folds for correspondence estimation. The SSM is built using correspondences from the training set, and evaluated with respect to two standardized metrics: generality and specificity [16]. For generality, we measure how well the SSM can represent unseen instances from the fourth fold through the Chamfer distance between the original shape and its SSM reconstruction. Specificity indicates how well random samples from the SSM represent the training data. We sample from the PDM 1000 times and calculate each sample’s minimum Chamfer distance to the training shapes. Generalization and specificity are reported in mm. Numbers in bold indicate statistically significant results by a one-sided t-test ($p < 0.05$).

5 Results & Discussion

Experiment 1: Thyroid SSM. Table 1 depicts the performance for all methods on the thyroid shapes. SURFMNet results for the thyroid pseudo-labels are omitted, as the method did not converge. Consistent trends can be observed across all methods for both sets of thyroid labels. The two existing functional

map-based methods were outperformed by Shapeworks, while the proposed S3M exceeded the latter’s scores. The descriptor is the most significant difference between the three learned functional map methods. Hand-crafted shape descriptors like SHOT and a simple fully-connected residual network architecture do not adequately represent thyroids’ noisy and heterogeneous shapes.

Our proposed method significantly outperformed Shapeworks in all metrics except the specificity of pseudo-labeled thyroids, where the results are not statistically significant. This was despite the advantage of optimizing correspondences across all shapes in training and validation. S3M can better cope with topological noise and generalizes to unseen samples, demonstrating potential in scaling SSM generation to larger datasets. Furthermore, it does not suffer from increasing memory requirements with the number of samples.

Table 2. Whole Heart Statistical Shape Modeling

Metrics	Composition 1		Composition 2		Composition 3	
	Generality [mm] ↓	Specificity [mm] ↓	Generality [mm] ↓	Specificity [mm] ↓	Generality [mm] ↓	Specificity [mm] ↓
SURFMNet	1.10 ± 0.39	1.79 ± 0.87	1.37 ± 0.16	2.00 ± 0.30	1.71 ± 0.17	2.54 ± 0.36
μ Match	3.39 ± 0.26	4.65 ± 0.16	2.82 ± 0.14	4.81 ± 0.41	3.30 ± 0.07	5.80 ± 0.08
ShapeWorks	0.89 ± 0.08	1.40 ± 0.04	2.57 ± 0.17	3.60 ± 0.06	3.07 ± 0.19	4.99 ± 0.22
S3M (Ours)	0.85 ± 0.07	1.30 ± 0.01	1.30 ± 0.13	1.72 ± 0.05	1.63 ± 0.17	2.14 ± 0.07

Experiment 2: Whole Heart SSM. Table 2 depicts the results of the different models on the three heart chamber compositions as previously defined. For the single-organ right atrium (composition 1), our proposed method fares comparably to ShapeWorks. For the more complex compositions 2 and 3, we observe larger increases in generalization and specificity for Shapeworks. μ Match fails to create a convincing SSM for any heart composition. Interestingly, SURFMNet can represent the more complex compositions 2 and 3 better than ShapeWorks, showing the strength of functional maps at representing complex high-level structures. S3M still exceeded the performance of SURFMNet, possibly due to the graph descriptor being better able to represent the surface topology.

From the qualitative results in Figure 3, it becomes more apparent that ShapeWorks does not generate an adequate SSM for the more complex compositions. This further supports our proposed method’s ability to learn correspondences for intricate and complex surface topologies, even consisting of meshes with bifurcations. The flexibility of our surface representation enables unsupervised correspondence estimation from multiple hierarchical sub-shapes, which is invaluable in multi-organ modeling such as for the heart [6,12].

Experiment 3: Thyroid Pseudo-label Generalization

To further highlight the proposed methods’ robustness to network-generated segmentation labels, we additionally measure the reconstruction ability of SSMs created from pseudo-labels on manually annotated thyroid labels under the Chamfer distance (Ours: $1.05 \pm 0.10\text{mm}$, Shapeworks: $1.84 \pm 0.40\text{mm}$). Notably, the proposed PDM on pseudo-labels generalizes better than the SSM built on

few manual labels ($1.25 \pm 0.11\text{mm}$; see Table 1), suggesting that more data can improve the SSM even if the labels are inaccurate. This is further supported by differences in shape (suppl. fig. 1); the SSM’s mean shape generated from pseudo-labels approximates the mean shapes on GT labels (and thus, the true organ shape) more closely.

6 Conclusion

We present an unsupervised approach for learning correspondences between shapes that exhibit noisy and irregular surface topologies. Our method leverages the strengths of geometric feature extractors to learn the intricacies of organ surfaces, as well as high-level functional bases of the Laplace-Beltrami operator to capture more extensive organ semantics. S3M outperforms existing methods on both manual labels, and label predictions from a network, demonstrating the potential to scale existing SSM pipelines to datasets that encompass more substantial population variance without additional annotation burden. Finally, we show that our model has the potential to learn correspondences between complex multi-organ shape hierarchies such as chambers of the heart, which would ease the manual burden of SSM curation for structures that currently still require meticulous manual landmark annotations.

Acknowledgements This work originated as a TUM Data Innovation Lab project through the Munich Data Science Institute. Vincent Bürgin is supported by the DAAD program Konrad Zuse Schools of Excellence in Artificial Intelligence, sponsored by the Federal Ministry of Education and Research. The authors declare no additional conflicts of interest.

References

1. Adams, J., Bhalodia, R., Elhabian, S.: Uncertain-deepssm: From images to probabilistic shape models. In: ShapeMI 2020, Held in Conjunction with MICCAI 2020. pp. 57–72. Springer (2020)
2. Adams, J., Elhabian, S.: From images to probabilistic anatomical shapes: A deep variational bottleneck approach. In: MICCAI 2022. pp. 474–484. Springer (2022)
3. Adams, J., Khan, N., Morris, A., Elhabian, S.: Spatiotemporal cardiac statistical shape modeling: A data-driven approach. In: STACOM MICCAI 2022. pp. 143–156. Springer (2023)
4. Agrawal, P., Whitaker, R.T., Elhabian, S.Y.: Learning deep features for shape correspondence with domain invariance. arXiv preprint arXiv:2102.10493 (2021)
5. Aubry, M., Schlickewei, U., Cremers, D.: The wave kernel signature: A quantum mechanical approach to shape analysis. In: ICCV Workshops 2011
6. Banerjee, A., Zacur, E., Choudhury, R.P., Grau, V.: Automated 3d whole-heart mesh reconstruction from 2d cine mr slices using statistical shape model. In: 2022 IEEE EMBS. pp. 1702–1706. IEEE (2022)
7. Bhalodia, R., Elhabian, S., Adams, J., Tao, W., Kavan, L., Whitaker, R.: Deepssm: A blueprint for image-to-shape deep learning models. arXiv preprint arXiv:2110.07152 (2021)

8. Bongratz, F., Rickmann, A.M., Pölsterl, S., Wachinger, C.: Vox2cortex: fast explicit reconstruction of cortical surfaces from 3d mri scans with geometric deep neural networks. In: CVPR 2022. pp. 20773–20783 (2022)
9. Bronstein, M.M., Kokkinos, I.: Scale-invariant heat kernel signatures for non-rigid shape recognition. In: CVPR 2010. pp. 1704–1711. IEEE (2010)
10. Cao, D., Bernard, F.: Unsupervised deep multi-shape matching. In: Computer Vision–ECCV 2022. pp. 55–71. Springer (2022)
11. Cates, J., Elhabian, S., Whitaker, R.: ShapeWorks. In: Statistical Shape and Deformation Analysis, pp. 257–298. Elsevier (2017)
12. Cerrolaza, J.J., Picazo, M.L., Humbert, L., Sato, Y., Rueckert, D., Ballester, M.Á.G., Linguraru, M.G.: Computational anatomy for multi-organ analysis in medical imaging: A review. *Medical image analysis* **56**, 44–67 (2019)
13. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Training Models of Shape from Sets of Examples. In: BMVC92 (1992)
14. Cootes, T.F., Twining, C.J., Babalola, K.O., Taylor, C.J.: Diffeomorphic statistical shape models. *Image and Vision Computing* **26**(3), 326–332 (2008)
15. Crane, K.: Discrete differential geometry: An applied introduction. *Notices of the AMS, Communication* pp. 1153–1159 (2018)
16. Davies, R.H.: Learning Shape: Optimal Models for Analysing Natural Variability. The University of Manchester (United Kingdom) (2002)
17. Donati, N., Sharma, A., Ovsjanikov, M.: Deep geometric functional maps: Robust feature learning for shape correspondence. In: CVPR 2020. pp. 8592–8601 (2020)
18. Du, J., Zhang, S., Wu, G., Moura, J.M., Kar, S.: Topology adaptive graph convolutional networks. *arXiv preprint arXiv:1710.10370* (2017)
19. Goparaju, A., Iyer, K., Bone, A., Hu, N., Henninger, H.B., et al.: Benchmarking off-the-shelf statistical shape modeling tools in clinical applications. *Medical Image Analysis* **76**, 102271 (2022)
20. Gutiérrez-Becker, B., Wachinger, C.: Learning a conditional generative model for anatomical shape analysis. In: IPMI 2019. pp. 505–516. Springer (2019)
21. Heimann, T., Meinzer, H.P.: Statistical shape models for 3D medical image segmentation: A review **13**(4), 543–563 (2009)
22. Henderson, E.G., Green, A.F., van Herk, M., Vasquez Osorio, E.M.: Automatic identification of segmentation errors for radiotherapy using geometric learning. In: MICCAI 2022. pp. 319–329. Springer (2022)
23. Kazi, A., Shekarforoush, S., Arvind Krishna, S., Burwinkel, H., Vivar, G., Kortüm, K., Ahmadi, S.A., Albarqouni, S., Navab, N.: Inceptiongcn: receptive field aware graph convolutional network for disease prediction. In: IPMI 2019. pp. 73–85. Springer (2019)
24. Klatzow, J., Dalmasso, G., Martínez-Abadías, N., Sharpe, J., Uhlmann, V.: μ Match: 3D Shape Correspondence for Biological Image Data **4** (2022)
25. Krönke, M., Eilers, C., Dimova, D., Köhler, M., Buschner, G., et al.: Tracked 3d ultrasound and deep neural network-based thyroid segmentation reduce interobserver variability in thyroid volumetry. *Plos one* **17**(7) (2022)
26. Litany, O., Remez, T., Rodola, E., Bronstein, A., Bronstein, M.: Deep functional maps: Structured prediction for dense shape correspondence. In: ICCV. pp. 5659–5667 (2017)
27. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3d surface construction algorithm. *ACM siggraph computer graphics* **21**(4), 163–169 (1987)
28. Lüdke, D., Amiranashvili, T., Ambellan, F., Ezhov, I., Menze, B.H., Zachow, S.: Landmark-free statistical shape modeling via neural flow deformations. In: MICCAI 2022. pp. 453–463. Springer (2022)

29. Ovsjanikov, M., Ben-Chen, M., Solomon, J., Butscher, A., Guibas, L.: Functional Maps: A Flexible Representation of Maps Between Shapes. *ACM ToG* **31**(4) (2012)
30. Parisot, S., Ktena, S.I., Ferrante, E., Lee, M., Moreno, R.G., Glocker, B., Rueckert, D.: Spectral graph convolutions for population-based disease prediction. In: *MICCAI 2017*. pp. 177–185. Springer (2017)
31. Roufosse, J.M., Sharma, A., Ovsjanikov, M.: Unsupervised deep learning for structured shape matching. In: *ICCV 2019*
32. Saleh, M., Dehghani, S., Busam, B., Navab, N., Tombari, F.: Graphite: Graph-induced feature extraction for point cloud registration. In: *2020 3DV*. pp. 241–251. IEEE (2020)
33. Saleh, M., Wu, S.C., Cosmo, L., Navab, N., Busam, B., Tombari, F.: Bending graphs: Hierarchical shape matching using gated optimal transport. In: *CVPR 2022*. pp. 11757–11767 (2022)
34. Sharma, A., Ovsjanikov, M.: Weakly supervised deep functional maps for shape matching. *NeurIPS* **33**, 19264–19275 (2020)
35. Tobon-Gomez, C., Geers, A.J., Peters, J., Weese, J., Pinto, K., Karim, R., Ammar, M., Daoudi, A., Margeta, J., Sandoval, Z., et al.: Benchmark for algorithms segmenting the left atrium from 3d ct and mri datasets. *IEEE transactions on medical imaging* **34**(7), 1460–1473 (2015)
36. Tombari, F., Salti, S., Stefano, L.D.: Unique signatures of histograms for local surface description. In: *ECCV 2010*
37. Vestner, M., Litman, R., Rodola, E., Bronstein, A., Cremers, D.: Product manifold filter: Non-rigid shape correspondence via kernel density estimation in the product space. In: *CVPR 2017*
38. Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M.: Dynamic graph cnn for learning on point clouds. *Acm ToG* **38**(5), 1–12 (2019)
39. Zhang, L., Tanno, R., Xu, M.C., Jin, C., Jacob, J., Cicarrelli, O., Barkhof, F., Alexander, D.: Disentangling human error from ground truth in segmentation of medical images. *NeurIPS* **33**, 15750–15762 (2020)

S3M: Supplementary Materials

1 Correspondence Loss

The functional loss consists of four components.

$$\min_{\phi} \sum_{(X_i, X_j)} \mathcal{L}(C_{ij}, C_{ji}) \quad \text{where} \quad C_{ij} = \arg \min_C \|CA_{T_\phi(D_i)} - A_{T_\phi(D_j)}\| \quad (1)$$

$$\mathcal{L} = \mathcal{L}_{bij} + \mathcal{L}_{orth} + \mathcal{L}_{iso} + \mathcal{L}_{point} \quad (2)$$

We first require the functional maps to be approximately bijective by regularizing their composition with the identity operator:

$$\mathcal{L}_{bij}(C_{ij}, C_{ji}) = \|C_{ij}C_{ji} - I\|_F + \|C_{ji}C_{ij} - I\|_F \quad (3)$$

Furthermore, we enforce their orthogonality to maintain their local area preservation characteristics with \mathcal{L}_{orth} [2], and a further regularization to ensure commutation with the LBO to maintain isometric properties on the surfaces with \mathcal{L}_{iso} :

$$\mathcal{L}_{orth}(C_{ij}, C_{ji}) = \|C_{ij}^T C_{ji} - I\|_F + \|C_{ji}^T C_{ij} - I\|_F \quad (4)$$

$$\mathcal{L}_{iso}(C_{ij}, C_{ji}) = \|C_{ij}\Lambda_i - \Lambda_j C_{ij}\|_F + \|C_{ji}\Lambda_j - \Lambda_i C_{ji}\|_F \quad (5)$$

where Λ_i is the diagonal matrix of the LBO eigenvalues of shape X_i and $\|\cdot\|_F$ denotes the frobenius norm. Finally, in addition to regularization through these “soft” correspondences defined by the functional map, we further regularize “hard” point-to-point correspondences based on pointwise descriptors matched using the functional maps in the truncated LBO bases of each shape with \mathcal{L}_{point} [1,3].

References

1. Nogneng, D., Ovsjanikov, M.: Informative descriptor preservation via commutativity for shape matching. In: Computer Graphics Forum. vol. 36, pp. 259–267. Wiley Online Library (2017)
2. Ovsjanikov, M., Ben-Chen, M., Solomon, J., Butscher, A., Guibas, L.: Functional Maps: A Flexible Representation of Maps Between Shapes. ACM ToG **31**(4) (2012)
3. Roufousse, J.M., Sharma, A., Ovsjanikov, M.: Unsupervised deep learning for structured shape matching. In: ICCV 2019

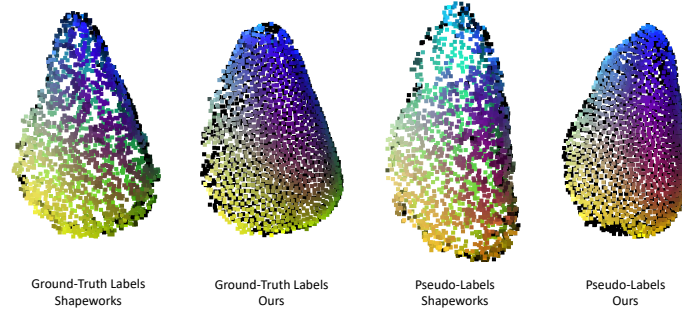


Fig. 1. Qualitative analysis of thyroid SSM mean shapes. Mean shapes on manually labeled shapes are depicted on the left, and on pseudo-labels from the QuickNat are on the right.

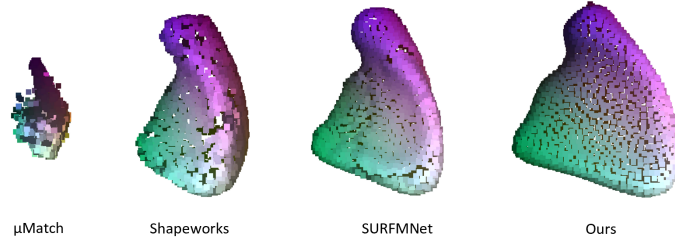


Fig. 2. Qualitative analysis of SSM mean shapes on Composition 1 (right ventricle) of heart dataset.

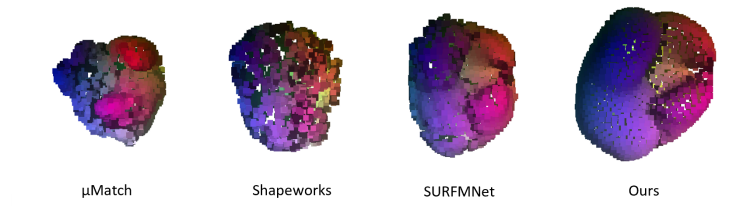


Fig. 3. Qualitative analysis of SSM mean shapes on Composition 2 of heart dataset.