A Sheaf Theoretic Perspective for Robust Prostate Segmentation

Ainkaran Santhirasekaram[1(✉)], Karen Pinto[2], Mathias Winkler[2], Andrea Rockall[2], and Ben Glocker[1]

[1] Department of Computing, Imperial College London, London, UK
a.santhirasekaram19@ic.ac.uk
[2] Department of Surgery and Cancer, Imperial College London, London, UK

Abstract. Deep learning based methods have become the most popular approach for prostate segmentation in MRI. However, domain variations due to the complex acquisition process result in textural differences as well as imaging artefacts which significantly affects the robustness of deep learning models for prostate segmentation across multiple sites. We tackle this problem by using multiple MRI sequences to learn a set of low dimensional shape components whose combinatorially large learnt composition is capable of accounting for the entire distribution of segmentation outputs. We draw on the language of cellular sheaf theory to model compositionality driven by local and global topological correctness. In our experiments, our method significantly improves the domain generalisability of anatomical and tumour segmentation of the prostate. Code is available at https://github.com/AinkaranSanthi/A-Sheaf-Theoretic-Perspective-for-Robust-Segmentation.git.

1 Introduction

Segmenting the prostate anatomy and detecting tumors is essential for both diagnostic and treatment planning purposes. Hence, the task of developing domain generalisable prostate MRI segmentation models is essential for the safe translation of these models into clinical practice. Deep learning models are susceptible to textural shifts and artefacts which is often seen in MRI due to variations in the complex acquisition protocols across multiple sites [12].

The most common approach to tackle domain shifts is with data augmentation [16,33,35] and adversarial training [11,30]. However, this increases training time and we propose to tackle the problem head on by learning shape only embedding features to build a shape dictionary using vector quantisation [31] which can be sampled to compose the segmentation output. We therefore hypothesise by limiting the search space to a set of shape components, we can improve generalisability of a segmentation model. We also propose to correctly sample

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43901-8_24.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14223, pp. 249–259, 2023.
https://doi.org/10.1007/978-3-031-43901-8_24

and compose shape components with local and global topological constraints by tracking topological features as we compose the shape components in an ordered manner. This is achieved using a branch of algebraic topology called cellular sheaf theory [8,19]. We hypothesise this approach will produce more anatomically meaningful segmentation maps and improve tumour localisation.

The contributions of this paper are summarized as follows: 1. This work considers shape compositionality to enhance the generalisability of deep learning models to segment the prostate on MRI. 2. We use cellular sheaves to aid compositionality for segmentation as well as improve tumour localisation.

2 Preliminaries

2.1 Persistent Homology

Topological data analysis is a field which extracts topological features from complex data structures embedded in a topological space. One can describe a topological space through its connectivity which can be captured in many forms. One such form is the cubical complex. The *cubical complex* $\mathcal{C}$ is naturally equipped to deal with topological spaces represented as volumetric grid structured data such as images [32]. In a 3D image, a cubical complex consists of individual voxels serving as vertices, with information regarding their connections to neighboring voxels captured through edges, squares, and cubes. Matrix reduction algorithms enable us to represent the connectivity of $\mathcal{C}$ in terms of a series of mathematical groups, known as the homology groups. Each homology group encompasses a specific dimension, d of topological features, such as connected components ($d = 0$), holes ($d = 1$), and voids ($d = 2$). The number of topological features present in each group is quantified by the corresponding Betti number. Betti numbers provide useful topological descriptors of the binary label maps as it is a single scale topological descriptor. However, the output, $\mathcal{Y}$ from a segmentation model is continuous. Thus, the Betti number for a cubical complex where vertices are continuous will be a noisy topological descriptor. We therefore use *persistent homology* which tracks changes in the topological features at multiple scales [17]. A cubical complex can be constructed at some threshold, τ over the output defined as: $\mathcal{C}^\tau = \{y \in \mathcal{Y} | \mathcal{Y} \geq \tau\}$. We can now create q cubical complexes over q ordered thresholds. This leads to a sequence of nested cubical complexes shown in Eq. 1 known as a sublevel set filtration. The persistent homology defines d dimensional topological features such as connected components which are born at τ^i and dies at τ^j where $\tau^j > \tau^i$. This creates tuples (τ^i, τ^j) which are stored as points in a persistence diagram (Fig. 2b).

$$\emptyset = \mathcal{C}^{\tau^1} \subseteq \mathcal{C}^{\tau^2} ... \subseteq \mathcal{C}^{\tau^q} = \mathcal{Y} \tag{1}$$

2.2 Cellular Sheaves

The homology of segmentation maps provides a useful tool for analysing global topology but does not describe how local topology is related to construct global

topological features. Sheaf theory provides a way of composing or 'gluing' local data together to build a global object (new data) that is consistent with the local information [8]. This lends well to modelling compositionality. Formally, a sheaf is a mathematical object which attaches to each open subset or subspace, U in a topological space, Y an algebraic object like a vector space or set (local data) such that it is well-behaved under restriction to smaller open sets [8].

We can consider a topological space, Y such as a segmentation output divided into a finite number of subspaces, $\{\emptyset, Y_1, Y_2 ... Y_n\}$ which are the base spaces for Y or equivalently the patches in a segmentation map. If we sequentially glue base spaces together in a certain order to form increasingly larger subspaces of Y starting with the $\emptyset$, one can construct a filtration of Y such that; $\emptyset \subseteq Y_1 \subseteq Y_1 \cup Y_2 ... \subseteq Y_1 \cup Y_2 ... \cup Y_n \subseteq Y$. We neatly formalise the subspaces and how subspaces are glued together with a poset. A *poset (P)* is a partially ordered set defined by a relation, $\leq$ between elements in P which is reflexive, anti-symmetric, and transitive [19]. In our work, we define a poset by the inclusion relation; $p_i \leq p_j$ implies $p_i \subseteq p_j$ for $p_i, p_j \in P$. Hence, we can map each element in P with a subspace in X which satisfies the inclusion relations in P like in X.

A *cellular sheaf*, $\mathcal{F}$ over a poset is constructed by mapping, each element, $p \in P$ to a vector space $\mathcal{F}(p)$ over a fixed field which preserves the ordering in P by linear transformations, $\rho_{...}$ which are inclusion maps in this case [19]. In our work each element in P maps to the vector space, $\mathbb{R}^2$ which preserves the inclusion relations in P. Specifically, we compute a persistence diagram, $\mathcal{D}$ for the subspace in X associated (homeomorphic) with $p \in P$ whereby (τ^i, τ^j) in the persistence diagram are a set of vectors in the vector space, $\mathbb{R}^2$. A cellular sheaf naturally arises in modelling the connectivity of a segmentation map and provides a mathematically precise justification for using cellular sheaves in our method. We show by approaching the composition of segmentation maps through this lens, one can significantly improved the robustness of segmentation models.

3 Related Work

There have been various deep learning based architectures developed for prostate tumour segmentation [3,15,18]. There is however no work looking at developing models which generalise well to target domains after training on one source domain known as single domain generalisation (SDG). Effective data augmentation techniques, such as CutOut [16], MixUp [34] and BigAug [35] offer a straightforward approach to enhance the generalisability of segmentation models across different domains. Recent methods have utilized adversarial techniques, such as AdvBias [11], which trains the model to generate bias field deformations and enhance its robustness.

RandConv [33] incorporates a randomized convolution layer to learn textural invariant features. Self-supervised strategies such as JiGen [9] can also improve generalisability. The principle of compositionality has been integrated into neural networks for tasks such as image classification [23], generation [2] and more recently, segmentation [26,28] to improve generalisability.

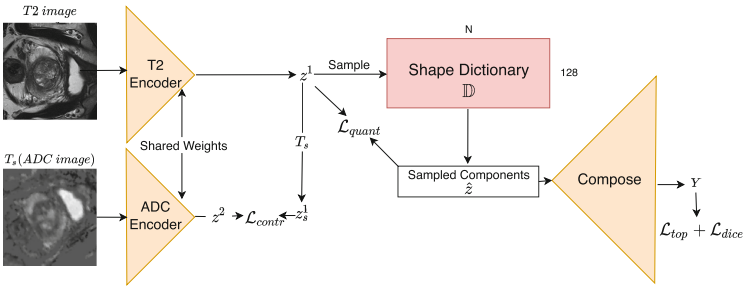

Fig. 1. Three part overview: 1. Use ADC maps to help learn shape equivariant features from T2 weighted images. 2. Construct a shape dictionary, $\mathbb{D}$ from the shape equivariant features. 3. Model the composition of the sampled shaped components to form the segmentation output using a cellular sheaf.

The utilization of persistent homology in deep learning-based segmentation is restricted to either generating topologically accurate segmentations in the output space [21] or as a subsequent processing step [14]. The novel approach of topological auto-encoders [27] marks the first instance of incorporating persistent homology to maintain the topological structure of the data manifold within the latent representation. Cellular sheaves were used to provide a topological insight into the poor performance of graph neural networks in the heterophilic setting [7]. Recently, cellular sheaves were used as a method of detecting patch based merging relations in binary images [20]. Finally, [4] recently proposed using sheaf theory to construct a shape space which allows one to precisely define how to glue shapes together in this shape space.

4 Methods

4.1 Shape Equivariant Learning

Given spatial, T_s and textural, T_i transformations of the input space, $\mathcal{X}$, the goal is to learn an encoder, Φ_e to map $\mathcal{X}$ to lower dimensional embedding features, $\mathcal{E}$ which are shape equivariant and texture invariant as shown in Eq. 2.

$$\Phi_e(T_s(T_i(\mathcal{X}))) = T_s(\mathcal{E}) \tag{2}$$

We assume T2 and ADC MRI images share the same spatial information and only have textural differences. We exploit this idea in Fig. 1, where firstly an ADC image under spatial transformation, T_s is mapped with an encoder, Φ_e to z^2 and the $T2$ image is mapped with the same encoder to z^1. Shape equivariance and texture invariance is enforced with the contrastive loss, $\mathcal{L}_{contr} = \|T_s(z^1) - z^2\|_2^2$. Specifically, we apply transformations from the dihedral group (D4) which consists of 90° rotations in the z plane and 180° rotations in the y plane. Note, a contrastive only learns equivariance as opposed to constraining the convolutional kernels to be equivariant. z^1 containing 128 channels is spatially quantised before

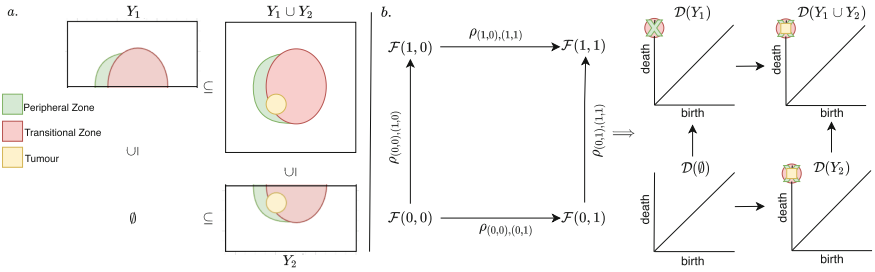

Fig. 2. Figure 2a shows 2 patches (Y_1, Y_2) glued together in an ordered manner defined by the poset, P, via inclusions maps to form the prostate segmentation output $Y = Y_1 \cup Y_2$. Figure 2b shows a cellular sheaf $\mathcal{F}$ over the poset P.

passing into the composer. In the test phase, the ADC image is not required and only the T2 image is used as input. T2 segmentations are used as the label.

4.2 Shape Component Learning

We posit that there is limited shape variation in the low dimensional embedding space across subjects which can be fully captured in N discrete shapes. N discrete shapes form a shape dictionary, $\mathbb{D}$ shown in Fig. 1 which is learnt with vector quantisation [31]. Given we enforce a texture invariant continuous embedding space and hence only contains shape information, quantisation converts this continuous embedding space to discrete shape features, $\hat{z}$. The quantisation process involves minimising the Euclidean distance between the embedding space, z^1 divided into m features, z_i^1 and its nearest shape component, $e_k \in \mathbb{D}$ shown in Eq. 3 where $k = argmin_j \|z_i^1 - e_j\|_2$ and $m = 3048$. Next, sampling $\mathbb{D}$ such that z_i^1 is replaced by e_k produces the spatially quantized embedding space $\hat{z}$. Straight-through gradient approximation is applied for backpropagation through the sampling process to update z^1 and $\mathbb{D}$ [31]. Gradient updates are applied to only the appropriate operands using stop gradients (sg) during optimization.

$$\mathcal{L}_{Quant} = \frac{1}{m} \sum_{i=0}^{i=m-1} \|sg(z_i^1) - e_k\|_2 + \beta \|z_i^1 - sg(e_k)\|_2 \tag{3}$$

4.3 Cellular Sheaves for Shape Composition

Shapes in $\mathbb{D}$ sampled with a uniform prior can lead to anatomically implausible segmentations after composition which we tackle through the language of cellular sheaves to model the connectivity of patches in an image which provides a connectivity-based loss function.

Composition: The quantised embedding space, $\hat{z}$, is split into c groups. The composition of each group in $\hat{z}$ to form each class segmentation, Y_c in the output,

Y involves two steps. Initially, a decoder with grouped convolutions equal to the number of classes followed by the softmax function maps, $\hat{z} \in \mathbb{R}^{128\times16\times16\times12}$ to $C \in \mathbb{R}^{p\times c\times256\times256\times24}$ where p is the number of patches for each class c. The second step of the composition uses a cellular sheaf to model the composition of Y_c by gluing the patches together in an ordered manner defined by a poset while tracking its topology using persistent homology. This in turn enforces $\mathbb{D}$ to be sampled in a topological preserving manner as input into the decoder/composer to improve both the local and global topological correctness of each class segmentation output, Y_c after composition.

Illustration: We illustrate our methodology of using cellular sheaves with a simple example in Fig. 2. Here, we show Y as perfectly matching the ground truth label (not one-hot encoded) divided into 2 patches. Y is a topological space with the subspaces, $\mathcal{V} = \{\emptyset, Y_1, Y_2, Y_1 \cup Y_2\}$. A 4-element poset, $P = \{(0,0),(1,0),(0,1),(1,1)\}$ is constructed where given $(x_1,x_2),(y_1,y_2) \in P$ then $(x_1,x_2) \subseteq (y_1,y_2)$ only if $x_1 \leq y_1 \wedge x_2 \leq y_2$. Each element in P is associated with a subspace in $\mathcal{V}$ such that the inclusion relationship is satisfied. Therefore, in Fig. 2a, P defines that Y_1 and Y_2 associated with $(1,0)$ and $(0,1)$ respectively are glued together to form $Y = Y_1 \cup Y_2$ which maps with $(1,1)$. A cellular sheaf $\mathcal{F}$ over P is created by assigning a vector space to $p \in \mathcal{P}$ by deriving a persistence diagram, $\mathcal{D}$ for each element in $\mathcal{V}$ associated with p as shown in Fig. 2b. The arrows in Fig. 2b are inclusion maps defined as $\rho_{.,.}$. Persistence diagrams are computed from the sequence of nested cubical complexes of each subspace in $\mathcal{V}$. The persistence diagrams in Fig. 2b are formed by overlapping the persistence diagrams for each class segmentation. Note, persistence diagrams contain infinite points in the form (τ,τ) (diagonal line in persistence diagrams) which always allows a bijection between two persistence diagrams. The main advantage of our approach is that in addition to ensuring correct local topology (patch level) and global topology (image level), we also force our network to produce topologically accurate patches correctly merged together in a topology preserving manner which matches the ground truth. For example in Fig. 2b, Y_2 contains 3 connected components glued onto Y_1 containing 2 connected components to form Y, which also has 3 connected components. This means an extra connected component is added by Y_2 due to tumour which therefore improves patch-wise tumour localisation. It also indicates the other 2 connected components in Y_2 are merged into the 2 connected components in Y_1 to form 2 larger connected components (peripheral and transitional zone) in Y. Hence, the same 2 vectors present in both $\mathcal{F}(1,0)$ and $\mathcal{F}(0,1)$ representing the peripheral and transitional zone are also in $\mathcal{F}(1,1)$. This is also known as a local section in $\mathcal{F}$.

Implementation: In practise during training, a cellular sheaf is built for each class in the output Y and label $\hat{Y}$, denoted Y^c and $\hat{Y}^c$ respectively. Y^c and $\hat{Y}^c$ are divided into $2 \times 2 \times 1$ patches. Starting with the top left patch, we sequentially glue on each patch in a zigzag manner until the entire image is formed which is formally defined by posets, P^c for Y^c and $\hat{P}^c$ for $\hat{Y}^c$, each containing 7 elements this time. Each element in the poset is associated with a patch i.e. Y_i^c or a sequences of patches glued together. For example, P^c is

Table 1. Mean dice, Hausdorff distance (HD) Betti error±standard deviation using our method, several SDG methods and the nnUNet for zonal segmentation.

	$RUNMC \rightarrow BMC$			$Internal \rightarrow RUNMC$		
	Dice	HD	Betti Error	Dice	HD	Betti Error
Baseline	0.51±0.13	0.40±0.11	2.98±0.91	0.67±0.17	0.35±0.09	2.01±0.72
nnUNet [16]	0.57±0.15	0.32±0.10	1.90±0.82	0.72±0.15	0.30±0.12	1.10±0.44
AdvBias [11]	0.56±0.13	0.33±0.15	1.92±0.16	0.73±0.19	0.29±0.13	1.09±0.22
RandConv [33]	0.59±0.15	0.29±0.08	1.54±0.24	0.73±.017	0.23±0.11	0.99±0.20
BigAug [35]	0.63±0.15	0.25±0.12	1.39±0.49	0.75±0.18	0.21±0.07	0.86±0.38
Jigen [9]	0.54±0.25	0.38±0.17	2.72±1.17	0.68±0.13	0.33±0.15	1.89±0.93
vMFNet [26]	0.61±0.15	0.28±0.14	1.48±0.39	0.72±0.16	0.24±0.09	0.99±0.28
Ours	**0.65±0.10**	**0.20±0.10**	**0.93±0.27**	**0.77±0.14**	**0.18±0.07**	**0.69±0.20**

Table 2. The average dice score, Betti error, Hausdorff distance(HD), sensitivity, specificity and positive predictive value (PPV) ± standard deviations using our method, several SDG methods and the nnUNet for tumour segmentation.

	Dice	HD	Betti Error	Sensitivity	Specificity	PPV
Baseline	0.38±0.17	1.03±0.32	5.11±2.90	0.37±0.11	0.60±0.20	0.29±0.11
nnUNet [16]	0.45±0.15	0.81±0.22	4.43±2.32	0.46±0.11	0.70±0.14	0.37±0.14
AdvBias [11]	0.42±0.10	0.90±0.19	4.41±2.16	0.45±0.13	0.66±0.21	0.35±0.17
RandConv [33]	0.43±0.18	0.80±0.27	4.19±2.01	0.47±0.16	0.65±0.17	0.35±0.19
BigAug [35]	0.47±0.12	0.68±0.19	4.03±1.89	0.48±0.18	0.73±0.22	0.40±0.19
Jigen [9]	0.42±0.11	0.88±0.21	4.51±2.43	0.42±0.18	0.65±0.13	0.33±0.09
vMFNet [26]	0.46±0.12	0.80 ±0.20	3.33±1.18	0.47±0.15	0.66±0.21	0.38±0.11
Ours	**0.51±0.13**	**0.57±0.16**	**2.99±0.97**	**0.50±0.18**	**0.79±0.20**	**0.45±0.18**

bijective with $\{Y_1^c, Y_2^c, Y_3^c, Y_4^c, Y_1^c \cup Y_2^c, Y_1^c \cup Y_2^c \cup Y_3^c, Y^c\}$. We construct cellular sheaves, $\mathcal{F}$ over P^c and $\hat{P}^c$ and minimise the distance between these cellular sheaves.

We firstly plot persistence diagrams, $\mathcal{D}$ from the set of vectors (τ^i, τ^j) in $\mathcal{F}(P_i^c)$ and $\mathcal{F}(\hat{P}_i^c)$. Next, we minimise the total p^{th} Wasserstein distance (topological loss) between the persistence diagrams $\mathcal{D}(\mathcal{F}(P_i^c))$ and $\mathcal{D}(\mathcal{F}(\hat{P}_i^c))$ shown in Eq. 4 where $\eta : \mathcal{D}(\mathcal{F}(P_i^c)) \rightarrow \mathcal{D}(\mathcal{F}(\hat{P}_i^c))$ is a bijection between the persistence diagrams [27] and $p = 2$. This loss function is proven to be stable to noise [29] and differentiable [10]. We add a dice loss between Y and $\hat{Y}$. The total loss to train our entire framework is: $\mathcal{L}_{total} = \mathcal{L}_{dice}(Y, \hat{Y}) + \mathcal{L}_{contr} + \mathcal{L}_{quant} + \mathcal{L}_{top}$.

$$\mathcal{L}_{top} = \sum_{c=1}^{c=4} \sum_{i=1}^{i=7} \left(inf_{\eta : \mathcal{D}(\mathcal{F}(P_i^c)) \rightarrow \mathcal{D}(\mathcal{F}(\hat{P}_i^c))} \sum_{x \in \mathcal{D}(\mathcal{F}(P_i^c))} \|x - \eta(x)\|_\infty^p \right)^{\frac{1}{p}} \quad (4)$$

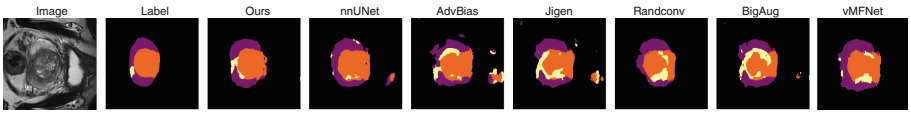

Fig. 3. Tumour and zonal Segmentation of a prostate slice from the RUNMC datasets when training on the internal dataset.

5 Experiments

5.1 Experimental Setup

Datasets: We firstly assess single domain generalisability of prostate zonal segmentation on 2 external datasets which has 2 labels; the transitional and peripheral zone. The training set contains the Prostate dataset from the Medical Segmentation Decathlon which has 32 T2 weighted and ADC images acquired from the Radboud University Nijmegen Medical Centre (RUNMC) [1]. The test set consists of 30 T2 weighted scans acquired from the Boston Medical Centre (BMC) as part of the NCI-ISBI13 Challenge [5,6,13]. The second experiment is a 3 class segmentation problem assessing single domain generalisability of prostate tumour and zonal segmentation. We compose an internal dataset for training which contains 168 T2 weighted and ADC images obtained from a single site across 6 years. The test set has 68 T2 weighted images from the PROSTATEx-2 challenge [24,25] acquired from RUNMC. All experiments were validated on 10 samples. See supplementary material for more details on the datasets.

Pre-processing: All images are resampled to $0.5 \times 0.5 \times 3\,\text{mm}$, centre cropped to $256 \times 256 \times 24$ and normalised between 0 and 1.

Model: In order to address the anisotropic characteristics of Prostate MRI images, we have chosen a hybrid 2D/3D UNet as our baseline model. We use the same encoder and decoder architecture as the baseline model in our method. See supplementary material for further details.

Comparison: We compare our method with the nnUNet [22] and several approaches to tackle SDG segmentation namely, RandConv [33], AdvBias [11], Jigen [9] and BigAug [35] applied to the baseline model. We also compare to a compositionality driven segmentation method called the vMFNet [26].

Training: In all our experiments, the models were trained using Adam optimization with a learning rate of 0.0001 and weight decay of 0.05. Training was run for up to 500 epochs on three NVIDIA RTX 2080 GPUs. The performance of the models was evaluated using the Dice score, Betti error [21] and Hausdorff distance. We evaluate tumour localisation by determining a true positive if the tumour segmentation overlaps by a minimum of one pixel with the ground truth.

In our ablation studies, the minimum number of shape components required in $\mathbb{D}$ for the zonal and zonal + tumour segmentation experiments was 64 and 192 respectively before segmentation performance dropped. See supplementary material for ablation experiments analysing each component of our framework.

5.2 Results

In the task of anatomical segmentation, the first two columns of Table 1 show the results for the domain shift from RUNMC in the decathlon dataset to BMC. Here, we demonstrate that our method improves segmentation performance in all evaluation metrics compared to the baseline, nn-UNet and the other SDG methods. Similar findings are noted for the domain shift from the internal dataset to the RUNMC data in the ProstateX2 dataset (second two columns of Table 1).

In Table 2, we note our method significantly improves tumour segmentation and localisation performance. We visualise our findings with an example in Fig. 3, where there is improved localisation of the tumour and the correct number of tumour components enforced by our topological loss. This significantly reduces the false positive rate highlighted in Table 2. Also, note the more anatomically plausible zonal segmentations. However, our method is restricted by the number of low dimensional shape components in the shape dictionary used to compose the high dimensional segmentation output. Therefore, our approach can fail to segment the finer details of prostate tumours due to its high shape variability which leads to coarser but better localised tumour segmentations.

6 Conclusion

In conclusion, we propose shape compositionality as a way to improve the generalisability of segmentation models for prostate MRI. We devise a method to learn texture invariant and shape equivariant features used to create a dictionary of shape components. We use cellular sheaf theory to help model the composition of sampled shape components from this dictionary in order to produce more anatomically meaningful segmentations and improve tumour localisation.

Acknowledgements. This work was supported and funded by Cancer Research UK (CRUK) (C309/A28804).

References

1. Antonelli, M., et al.: The medical segmentation decathlon. Nat. Commun. **13**(1), 4128 (2022)
2. Arad Hudson, D., Zitnick, L.: Compositional transformers for scene generation. Adv. Neural. Inf. Process. Syst. **34**, 9506–9520 (2021)
3. Arif, M., et al.: Clinically significant prostate cancer detection and segmentation in low-risk patients using a convolutional neural network on multi-parametric MRI. Eur. Radiol. **30**, 6582–6592 (2020)
4. Arya, S., Curry, J., Mukherjee, S.: A sheaf-theoretic construction of shape space. arXiv preprint arXiv:2204.09020 (2022)
5. Bloch, N., et al.: Cancer imaging archive Wiki. http://doi.org/10.7937/K9/TCIA.2015.zF0vlOPv (2015)
6. Bloch, N., et al.: NCI-ISBI 2013 challenge: automated segmentation of prostate structures. Cancer Imaging Arch. **370**, 6 (2015)

7. Bodnar, C., Di Giovanni, F., Chamberlain, B.P., Liò, P., Bronstein, M.M.: Neural sheaf diffusion: a topological perspective on heterophily and oversmoothing in GNNs. arXiv preprint arXiv:2202.04579 (2022)
8. Bredon, G.E.: Sheaf Theory, vol. 170. Springer, Heidelberg (2012)
9. Carlucci, F.M., D'Innocente, A., Bucci, S., Caputo, B., Tommasi, T.: Domain generalization by solving jigsaw puzzles. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2229–2238 (2019)
10. Carriere, M., Chazal, F., Glisse, M., Ike, Y., Kannan, H., Umeda, Y.: Optimizing persistent homology based functions. In: International Conference on Machine Learning, pp. 1294–1303. PMLR (2021)
11. Chen, C., et al.: Realistic adversarial data augmentation for MR image segmentation. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 667–677. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59710-8_65
12. Chen, Y.: Towards to robust and generalized medical image segmentation framework. arXiv preprint arXiv:2108.03823 (2021)
13. Clark, K., et al.: The cancer imaging archive (TCIA): maintaining and operating a public information repository. J. Digit. Imaging **26**, 1045–1057 (2013)
14. 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. arXiv preprint arXiv:1910.01877 (2019)
15. De Vente, C., Vos, P., Hosseinzadeh, M., Pluim, J., Veta, M.: Deep learning regression for prostate cancer detection and grading in bi-parametric MRI. IEEE Trans. Biomed. Eng. **68**(2), 374–383 (2020)
16. DeVries, T., Taylor, G.W.: Improved regularization of convolutional neural networks with cutout. arxiv 2017. arXiv preprint arXiv:1708.04552 (2017)
17. Edelsbrunner, H., Harer, J., et al.: Persistent homology-a survey. Contemp. Math. **453**(26), 257–282 (2008)
18. Gunashekar, D.D., et al.: Explainable AI for CNN-based prostate tumor segmentation in multi-parametric MRI correlated to whole mount histopathology. Radiat. Oncol. **17**(1), 1–10 (2022)
19. Hu, C.S.: A brief note for sheaf structures on posets. arXiv preprint arXiv:2010.09651 (2020)
20. Hu, C.S., Chung, Y.M.: A sheaf and topology approach to detecting local merging relations in digital images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4396–4405 (2021)
21. Hu, X., Li, F., Samaras, D., Chen, C.: Topology-preserving deep image segmentation. In: Advances in Neural Information Processing Systems, vol. 32 (2019)
22. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods **18**(2), 203–211 (2021)
23. Kortylewski, A., He, J., Liu, Q., Yuille, A.L.: Compositional convolutional neural networks: a deep architecture with innate robustness to partial occlusion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8940–8949 (2020)
24. Litjens, G., Debats, O., Barentsz, J., Karssemeijer, N., Huisman, H.: Cancer imaging archive Wiki. URL https://doi.org/10.7937/K9TCIA (2017)
25. Litjens, G., et al.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)

26. Liu, X., Thermos, S., Sanchez, P., O'Neil, A.Q., Tsaftaris, S.A.: vMFNet: compositionality meets domain-generalised segmentation. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) Medical Image Computing and Computer Assisted Intervention-MICCAI 2022: 25th International Conference, Singapore, 18–22 September 2022, Proceedings, Part VII, pp. 704–714. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16449-1_67
27. Moor, M., Horn, M., Rieck, B., Borgwardt, K.: Topological autoencoders. In: International Conference on Machine Learning, pp. 7045–7054. PMLR (2020)
28. Santhirasekaram, A., Kori, A., Winkler, M., Rockall, A., Glocker, B.: Vector quantisation for robust segmentation. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 663–672. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16440-8_63
29. Skraba, P., Turner, K.: Wasserstein stability for persistence diagrams. arXiv preprint arXiv:2006.16824 (2020)
30. Tramer, F., Boneh, D.: Adversarial training and robustness for multiple perturbations. In: Advances in Neural Information Processing Systems, vol. 32 (2019)
31. Van Den Oord, A., Vinyals, O., et al.: Neural discrete representation learning. In: Advances in Neural Information Processing Systems, vol. 30 (2017)
32. Wagner, H., Chen, C., Vuçini, E.: Efficient computation of persistent homology for cubical data. In: Peikert, R., Hauser, H., Carr, H., Fuchs, R. (eds.) Topological Methods in Data Analysis and Visualization II: Theory, Algorithms, and Applications, pp. 91–106. Springer, Cham (2011). https://doi.org/10.1007/978-3-642-23175-9_7
33. Xu, Z., Liu, D., Yang, J., Raffel, C., Niethammer, M.: Robust and generalizable visual representation learning via random convolutions. arXiv preprint arXiv:2007.13003 (2020)
34. Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: beyond empirical risk minimization. arXiv preprint arXiv:1710.09412 (2017)
35. Zhang, L., et al.: Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation. IEEE Trans. Med. Imaging **39**(7), 2531–2540 (2020)