Shape Equivariant Learning for Robust MRI Segmentation

Author name(s) withheld EMAIL(S) WITHHELD
Address withheld

Editors: Under Review for MIDL 2023

Abstract

The reliability of deep learning based segmentation models is essential to the safe translation of these models into clinical practise. Unfortunately, these models are sensitive to distributional shifts. This is particularly notable in MRI, where there is a large variation of acquisition protocols across different domains leading to varying textural profiles. We hypothesise that the constrained anatomical variability across subjects can be leveraged to discretize the latent space to a dictionary of shape components. We achieve this by using multiple MRI sequences to learn texture invariant and shape equivariant features which are used to construct a shape dictionary using vector quantisation. This dictionary is then sampled to compose the segmentation output. Our method achieves SOTA performance in the task of single domain generalisation (SDG) for prostate zonal segmentation.

Keywords: Shape Equivariance, Robustness, Segmentation, MRI

1. Introduction

Magnetic resonance imaging involves a complex acquisition process which differs across subjects and domains. This can lead to varying textural profiles and artefacts. Deep learning based segmentations are however not robust to textural shifts and unencountered artefacts at test time. Domain generalisability for deep learning has been traditionally tackled through augmentation based strategies such as CutOut (DeVries and Taylor, 2017) and BigAug (Zhang et al., 2020). AdvBias (Chen et al., 2020) is an adversarial technique for MRI data which learns to generate bias field deformations to improve model robustness for segmentation. RandConv (Xu et al., 2020) which is perhaps the most related work, attempts to learn textural invariant features by using a randomised convolutional input layer. Here, we propose an alternative method to learn shape equivariant features based on the principle that in MRI, T2 weighted images and ADC maps calculated from diffusion weighted imaging contain the exact same spatial information and only differ in their textural profiles. There is anatomical consistency across subjects meaning there is reduced spatial variation in the segmentation outputs. Therefore, we propose to constrain the latent space to a dictionary of shape components which is sampled to construct the segmentation output. We hypothesise this will improve the generalisability of any segmentation model which maps the input space, $\mathcal{X}$ to a lower dimensional embedding space, $\mathcal{E}$ using an encoder, Φ_e before mapping to the segmentation output, $\mathcal{Y}$ with a decoder, Φ_d. This is achieved using vector quantisation (Van Den Oord et al., 2017) of the shape equivariant features to create a discrete shape space. We assume the dictionary is complete and sufficient to capture the entire distribution of segmentation outputs after composition of the discrete shape space using the decoder. We evaluate the capability of our method to improve domain generalisability in the task of prostate zonal segmentation with two labels (transitional and peripheral zone) when training on a single domain.

© 2023 CC-BY 4.0, Anonymous.

2. Method

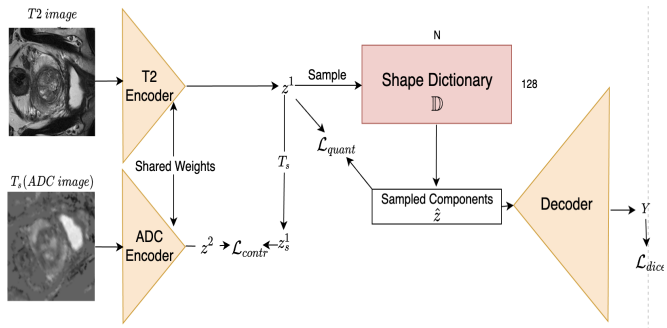

Figure 1: Overview of our method demonstrating using the ADC map to learn shape equivariant features which is quantised to construct a shape dictionary, $\mathbb{D}$.

We start with the image input which is the T2 weighted image, $x \in \mathbb{R}^{1\times256\times256\times24}$ and apply an intensity transformation, T_i which is equivalent to acquiring the ADC map. We also apply a spatial transformation, T_s to the ADC map which involves rotations and flipping (horizontal and vertical). The T2 image and spatially transformed ADC map are passed through an encoder to produce their respective embeddings, z^1 and z^2 as shown in Figure 1. Shape equivariance and texture invariance is enforced by satisfying the equation: $\Phi_e(T_s(T_i(x))) = T_s(z^1)$. Therefore, we minimise the contrastive loss: $\mathcal{L}_{contr} = \|T_s(z^1) - z^2\|_2^2$.

We quantise, $z^1 \in \mathbb{R}^{128\times16\times16\times12}$ using vector quantisation by dividing z^1 into $16\times16\times12$ components and replacing each component in z^1 denoted z_i^1 with its nearest component, $e_k \in \mathbb{D}$ where $k = argmin_j \|z_i^1 - e_j\|_2$. This produces the discrete shape latent space, $\hat{z}$ which is inputted into the decoder to construct the segmentation output. The quantisation loss minimises the euclidean distance between z_i^1 and its nearest component, $e_k \in \mathbb{D}$ formally defined as: $\mathcal{L}_{Quant} = \frac{1}{m}\sum_{i=0}^{i=m-1} \|sg(z_i^1) - e_k\|_2 + \beta\|z_i^1 - sg(e_k)\|_2$. Stop gradients (sg) are applied to the correct operand. We compute the dice loss between the output, $\hat{y}$ and the T2 segmentation label, y. The total loss for training our framework is $\mathcal{L}_{total} = \mathcal{L}_{dice}(\hat{y}, y) + \mathcal{L}_{contr} + \mathcal{L}_{quant}$. Note, only T2 weighted images are required as input during inference.

3. Experiments and Results

Dataset: The training set comprises the Prostate dataset obtained from the Medical Segmentation Decathlon (Antonelli et al., 2022), consisting of 32 T2-weighted and ADC images captured at the Radboud University Nijmegen Medical Centre (RUNMC). We use the 30 T2 weighted images in the NCI-ISBI13 Challenge (Bloch et al., 2015) which was acquired from Boston Medical Centre (BMC) for our test set. All images are centre cropped to $256 \times 256 \times 24$ and normalised between 0 and 1.

Baseline Model and Comparison: We use a hybrid 2D/3D UNet as our baseline model in order to deal with the anisotropic Prostate MRI images. The encoder and decoder is made up of 5 levels consisting of 2D pre-activation residual blocks in the top 4 levels and a 3D pre-activation residual block in the bottleneck level. We use the same encoder and decoder architecture for our method. We compare our method to the following SDG methods applied to the baseline model: CutOut (DeVries and Taylor, 2017), BigAug (Zhang et al., 2020), AdvBias (Chen et al., 2020) and RandConv (Xu et al., 2020). All models were trained for up to 500 epochs using Adam optimisation with a learning rate of 0.001.

	Baseline	**CutOut**	**BigAug**	**AdvBias**	**RandConv**	**Ours**
Dice	0.51±0.13	0.53±0.17	0.63±0.15	0.56±0.13	0.59±0.15	**0.64±0.11**
HD	0.40±0.11	0.37±0.19	0.25±0.12	0.33±0.15	0.29±0.08	**0.23±0.10**

Table 1: Dice score and Hausdorff distance(HD) ± standard deviation for different SDG methods compared to our approach

Results and Discussion: In Table 1, we show that our method outperforms other SDG methods in terms of the Dice score and Hausdorff distance. We therefore show that one can improve the domain generalisability of a segmentation model in an anatomical segmentation task by constraining the latent space to a finite set of shape components.

Acknowledgments

Acknowledgments withheld.

References

Michela Antonelli, Annika Reinke, Spyridon Bakas, Keyvan Farahani, Annette Kopp-Schneider, Bennett A Landman, Geert Litjens, Bjoern Menze, Olaf Ronneberger, Ronald M Summers, et al. The medical segmentation decathlon. *Nature communications*, 13(1):4128, 2022.

Nicholas Bloch, Anant Madabhushi, Henkjan Huisman, John Freymann, Justin Kirby, Michael Grauer, Andinet Enquobahrie, Carl Jaffe, Larry Clarke, and Keyvan Farahani. Nci-isbi 2013 challenge: automated segmentation of prostate structures. *The Cancer Imaging Archive*, 370:6, 2015.

Chen Chen, Chen Qin, Huaqi Qiu, Cheng Ouyang, Shuo Wang, Liang Chen, Giacomo Tarroni, Wenjia Bai, and Daniel Rueckert. Realistic adversarial data augmentation for mr image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 667–677. Springer, 2020.

Terrance DeVries and Graham W Taylor. Improved regularization of convolutional neural networks with cutout. arxiv 2017. *arXiv preprint arXiv:1708.04552*, 2017.

Aaron Van Den Oord, Oriol Vinyals, et al. Neural discrete representation learning. *Advances in neural information processing systems*, 30, 2017.

Zhenlin Xu, Deyi Liu, Junlin Yang, Colin Raffel, and Marc Niethammer. Robust and generalizable visual representation learning via random convolutions. *arXiv preprint arXiv:2007.13003*, 2020.

Ling Zhang, Xiaosong Wang, Dong Yang, Thomas Sanford, Stephanie Harmon, Baris Turkbey, Bradford J Wood, Holger Roth, Andriy Myronenko, Daguang Xu, et al. Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation. *IEEE transactions on medical imaging*, 39(7):2531–2540, 2020.