Introducing Soft Topology Constraints in Deep Learning-based Segmentation using Projected Pooling Loss

Guanghui Fu, Rosana El Jurdi, Lydia Chougar, Didier Dormont, Romain Valabregue, Stéphane Lehéricy, Olivier Colliot

Introducing Soft Topology Constraints in Deep Learning-based Segmentation using Projected Pooling Loss

Guanghui Fu[a], Rosana El Jurdi[a], Lydia Chougar[a,b,c], Didier Dormont[a,c], Romain Valabregue[b,d], Stéphane Lehéricy[b,c,d], Olivier Colliot[a], and the ICEBERG Study Group[b]

[a]Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, F-75013, Paris, France
[b]ICM, Centre de NeuroImagerie de Recherche-CENIR, Paris, France.
[c]AP-HP, Hôpital de la Pitié Salpêtrière, DMU DIAMENT, Dep. of Neuroradiology, F-75013, Paris, France
[d]Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, CNRS, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, F-75013, Paris, France

ABSTRACT

Deep learning methods have achieved impressive results for 3D medical image segmentation. However, when the network is only guided by voxel-level information, it may provide anatomically aberrant segmentations. When performing manual segmentations, experts heavily rely on prior anatomical knowledge. Topology is an important prior information due to its stability across patients. Recently, several losses based on persistent homology were proposed to constrain topology. Persistent homology offers a principled way to control topology. However, it is computationally expensive and complex to implement, in particular in 3D. In this paper, we propose a novel loss function to introduce topological priors in deep learning-based segmentation, which is fast to compute and easy to implement. The loss performs a projected pooling within two steps. We first focus on errors from a global perspective by using 3D MaxPooling to obtain projections of 3D data onto three planes: axial, coronal and sagittal. Then, 2D MaxPooling layers with different kernel sizes are used to extract topological features from the multi-view projections. These two steps are combined using only MaxPooling, thus ensuring the efficiency of the loss function. Our approach was evaluated in several medical image datasets (spleen, heart, hippocampus, red nucleus). It allowed reducing topological errors and, in some cases, improving voxel-level accuracy.

Keywords: Segmentation, Loss function, Anatomical priors, Topology, Connected components, Deep Learning

1. INTRODUCTION

Deep learning methods are the state of the art for medical image segmentation.[1] The standard approach is to supervise at the voxel level, using a voxel-wise loss (typically cross-entropy[2] and Dice loss[3]). However, this may generate anatomically aberrant segmentations.[4] To overcome such problems, one can introduce prior anatomical knowledge in the segmentation procedure. Indeed, medical experts extensively use their anatomical knowledge to perform manual segmentations. In particular, prior knowledge can be introduced through new loss functions. Many loss functions have been designed and we refer the reader to systematic surveys.[5–7] The purpose of such losses is very diverse but some aim at introducing prior information such as for instance shape[8,9] position[10] or size[11,12] of the target region. Topological priors are important because they are usually very robust across patients and because a topologically incorrect segmentation is generally anatomically aberrant. BennTaieb et al.[13] proposed a loss combining topological and geometric priors. However, it is not clear if it is applicable to 3D data. He et al.[14] introduced a cascaded network which preserves the topological relationships between the layers of the retina. The most general approaches to constrain topology[15–17] are based on persistent homology,[18] an algebraic method for characterizing the topology of shapes and functions. Persistent homology offers a principled

Further author information: (Send correspondence to Guanghui Fu)
Guanghui Fu: E-mail: aslanfu123@gmail.com

way to control topology, such as connected components, holes, or voids. However, it is complex to implement and computationally expensive in particular in 3D.

In this paper, we propose a fast loss function to constrain the topology in 3D medical image segmentation. Unlike approaches based on persistent homology, it only provides an approximate control of topology but has the advantage of being computationally efficient and easy to implement.

2. PROPOSED METHOD

Our approach has two components: we first project the 3D volume into 2D space according to three views; we then characterize topological features using pooling layers with different kernel sizes (Figure 1). The topological loss corresponds to the difference in topological features between the ground truth and the prediction.

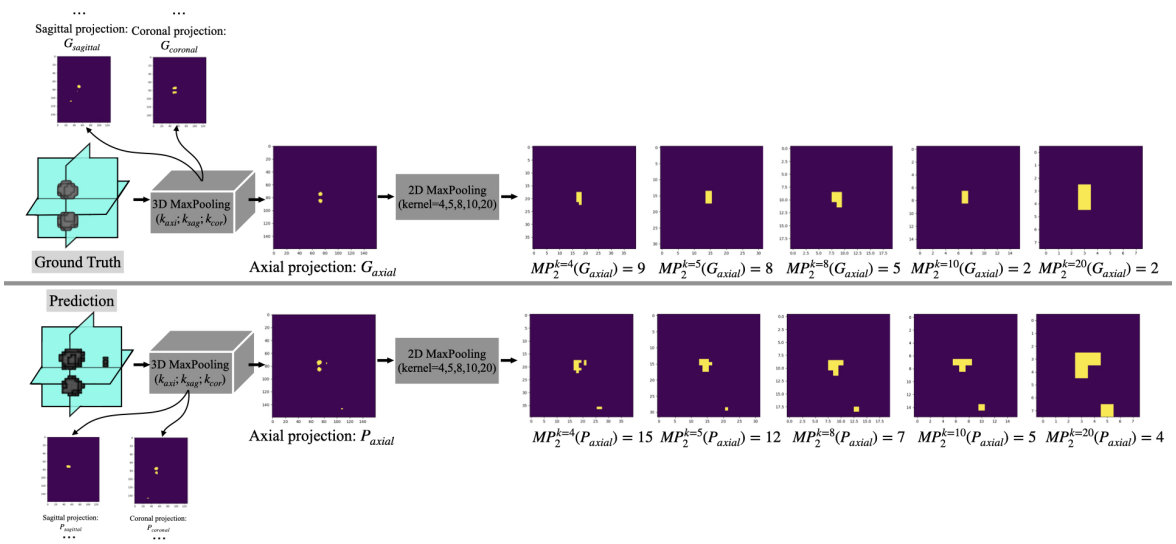

Figure 1. Overview of the approach. The method contains two steps: 3D MaxPooling for projection and 2D MaxPooling with different kernel sizes for topological features extraction. Due to space limitation, we do not show the 2D MaxPooling step for sagittal and coronal planes but the specific steps remain the same as for axial plane.

2.1 Volume projection by 3D-MaxPooling

We propose to use 3D MaxPooling to project the volume onto three planes (axial, coronal, and sagittal) to create a 2D representation of the 3D object. The projection can be easily obtained by using a MaxPooling operation. The volume projections onto the axial, sagittal and coronal planes are denoted respectively as P_{axial}, $P_{sagittal}$ and $P_{coronal}$ and computed as follows:

$$P_{axial} = {MP_3}^{k_{axi}}(V_{w,h,s}); P_{sagittal} = {MP_3}^{k_{sag}}(V_{w,h,s}); P_{coronal} = {MP_3}^{k_{cor}}(V_{w,h,s}) \quad (1)$$

where $V_{w,h,s}$ denotes the 3D volume, w, h and s are the width, height and number of slices of the volume respectively, MP_3^k denotes the 3D-MaxPooling with kernel k and $k_{axi} = (1, 1, s)$, $k_{sag} = (w, 1, 1)$ and $k_{cor} = (1, h, 1)$ are the kernels for projecting onto the axial, sagittal and coronal planes respectively. Note that P_{axial} is of dimension (w, h), $P_{sagittal}$ is of dimension (h, s) and $P_{coronal}$ is of dimension (w, s).

The above procedure concentrates the target objects along each axis. The idea is to be able to magnify the effect of small errors on the overall loss. Indeed, a small segmentation error (such as for example a small extra connected component) has little effect on the overall loss. For example, the proportion of a voxel mispredicted in the volume $V_{w,h,s}$ is $\frac{1}{w \times h \times s}$, while if we project it onto the axial plane, the error proportion is $\frac{1}{w \times h}$. Thus, projecting the prediction into axial planes augments the weight of this error by s times.

2.2 Topology feature extraction by 2D-MaxPooling

In a second step, we characterize the topology of the projected view by using two-dimensional MaxPooling with different kernel sizes. The output can be roughly treated as connected components. Some small areas such as noise in predictions are very common, and we don't want to ignore them. We achieve this by using a set of 2D-MaxPooling with different kernel sizes. The topological characterization of the prediction and the ground truth is then defined as follows:

$$\begin{aligned} P^k_{topo} &= MP^k_2(P_{axial}) + MP^k_2(P_{sagittal}) + MP^k_2(P_{coronal}) \\ G^k_{topo} &= MP^k_2(G_{axial}) + MP^k_2(G_{sagittal}) + MP^k_2(G_{coronal}) \end{aligned} \tag{2}$$

where MP^k_2 denotes 3D-MaxPooling with square kernel of size k, G_{plane} (resp. P_{plane}) is the projection of the ground truth (resp. of the prediction) onto a given *plane*. After 2D MaxPooling (MP^k_2) under different kernels k, the correct part of the prediction will occupy the same position as the label.

We then compute the absolute difference between P^k_{topo} and P^k_{topo} for each kernel size. The topology loss L_{topo} is obtained by averaging over all kernels and over the three projection planes:

$$L_{topo} = \frac{|\sum_{k \in K} G^k_{topo} - P^k_{topo}|}{3 \times |K|}. \tag{3}$$

where K is the set of kernels and $|K|$ is the number of kernels. regarded as geometric appreciation.

2.3 Final loss

The final loss function is a combination of the voxel-wise Dice loss and the proposed topological loss: $\mathbf{L}_{total} = L_{topo} + \lambda L_{dice}$ where λ is the trade-off parameter. In our experiments, we used a fixed value of $\lambda = 1$.

3. EXPERIMENTS AND RESULTS

3.1 Implementation

Our loss function is compatible with any convolutional neural network. In our experiments, we used the standard 3D-UNet.[19] In some experiments, we used early stopping to avoid over-fitting. The early stopping was done using only the validation set in order to obtain an unbiased performance on the test set. In all experiments, the hyperparameters were identical. Because our approach needs to be pre-trained, each training involved two steps, both using Adam:[20] (1) pre-train for 150 epochs with learning rate 10^{-3}; (2) fine tuning for the same number of epochs with learning rate 10^{-4}. Finally, we used the following sequence of kernels sizes: 4, 5, 8, 10, 20. We use the open-source Python library TorchIO[21] * for medical dataset preprocessing that include: reshaping images to the same size for each task, min-max normalization. Our code is developed based on the PyTorch framework.[22]

Table 1. Characteristics of the imaging datasets.

Dataset	Task	Train+val	Test	Image Size
Public	Spleen	25+7	9	160,160,128
	Heart	12+4	4	160,160,128
	Hippocampus	166+42	52	64, 64, 48
Local	Red Nucleus	51+13	16	160,160,128

*https://torchio.readthedocs.io/

3.2 Datasets and evaluation framework

Our method was assessed using three public datasets from the Medical Segmentation Decathlon (MSD)[23] and one local dataset. The three tasks from MSD were segmentation of spleen, heart and hippocampus (see[23] for details). For the local dataset, the task is to segment the red nucleus from Quantitative Susceptibility Mapping (QSM) brain images. The cohort included 18 healthy subjects, 46 patients with early Parkinson's disease, and 16 patients with prodromal parkinsonism (isolated rapid eye movement sleep behavior disorders) from the ICEBERG study (NCT02305147). The institutional ethical committee approved the study (CPP Paris VI/RCB: 2014-A00725-42). All participants gave written informed consent. QSM images were manually segmented by a trained neuroradiologist (L.C.).

Each dataset was split into training, validation and test sets. The splits were done at the participant level to avoid any data leakage.[24] The dataset distribution can be seen in the Table 1. For the local dataset, we also explored the performance with a reduced training set size by randomly sampling 30% of the training set (15 MRIs).

We report performance at the voxel level and the topological level. We used the following seven classical voxel-level evaluation metrics: Dice score, Precision, Recall, Volume error rate (VER), Absolute volume error rate (AVER), Pearson's r between predicted and ground-truth volume, 95 percentile Hausdorff distance (95HD). Topological errors may have only a small impact on voxel-level metrics if they are of limited size. However, topological errors make the predictions anatomically inconsistent. We thus used additional metrics for evaluation of the topology, computing the 3D connected component error as well as the average 2D connected component error. Topological metrics were the 3D connected component error (CCE) and the average 2D CCE.

3.3 Results

Results are displayed in Table 2. The proposed loss performs comparably with respect to the Dice loss baseline in terms of voxel-level metrics. However, the proposed loss substantially decreases the topological errors both in 2D and in 3D for the spleen, heart and red nucleus datasets. For the hippocampus dataset, results were comparable. Furthermore, in the case of the red nucleus, when training on a smaller subset, the Dice score and other voxel-level metrics were also improved. Figure 2 shows an example of result on the red nucleus task. One can observe that only a few iterations of the topology loss are needed to improve topological correctness and Dice score.

Table 2. Results with our loss and with the standard Dice loss (trained with either 150 or 150+150 epochs).

Task	**Setting**		**Voxel level**							Topology level	
	Loss	Epoch	Dice	95HD	Precision	Recall	MVER	MAVER	Pearson's r	2D CCE	3D CCE
Spleen	Dice	150	0.887	9.811	0.849	0.938	0.118	0.164	0.890	0.524	2
	Dice	150+150	0.894	21.26	0.853	0.948	0.124	0.163	0.897	0.506	1.667
	Ours	150+150	0.895	8.405	0.902	0.898	-0.001	0.098	0.894	0.289	0.222
Heart	Dice	150	0.872	6.923	0.821	0.930	0.136	0.136	0.873	0.543	2.75
	Dice	150+150	0.872	5.722	0.844	0.902	0.072	0.074	0.872	0.403	0.75
	Ours	150+150	0.864	7.104	0.861	0.870	-0.086	0.087	0.865	0.373	0.25
Hippocampus	Dice	150	0.861	1.368	0.864	0.861	0.001	0.053	0.877	0.2	0.019
	Dice	150+150	0.861	1.339	0.869	0.858	-0.006	0.054	0.879	0.192	0.019
	Ours	150+150	0.853	1.375	0.877	0.833	-0.048	0.067	0.872	0.203	0
Red Nucleus (30%: 15 QSM)	Dice	150	0.725	7.646	0.595	0.948	0.645	0.645	0.748	0.601	1.025
	Dice	150+150	0.738	5.493	0.610	0.954	0.616	0.616	0.76	0.568	1.125
	Ours	150+150	0.808	3.612	0.783	0.849	0.107	0.184	0.812	0.349	0.375
Red Nucleus (full: 51 QSM)	Dice	150	0.896	1.005	0.862	0.938	0.095	0.129	0.898	0.232	0.662
	Dice	150+150	0.900	1.005	0.872	0.933	0.087	0.115	0.902	0.211	0.525
	Ours	150+150	0.905	1.005	0.896	0.918	0.03	0.09	0.906	0.144	0.05

4. DISCUSSION AND CONCLUSION

We proposed a new loss function that integrates soft topological priors for 3D medical image segmentation. We demonstrated that it allows reducing topological errors, without the need for subsequent postprocessing. In some

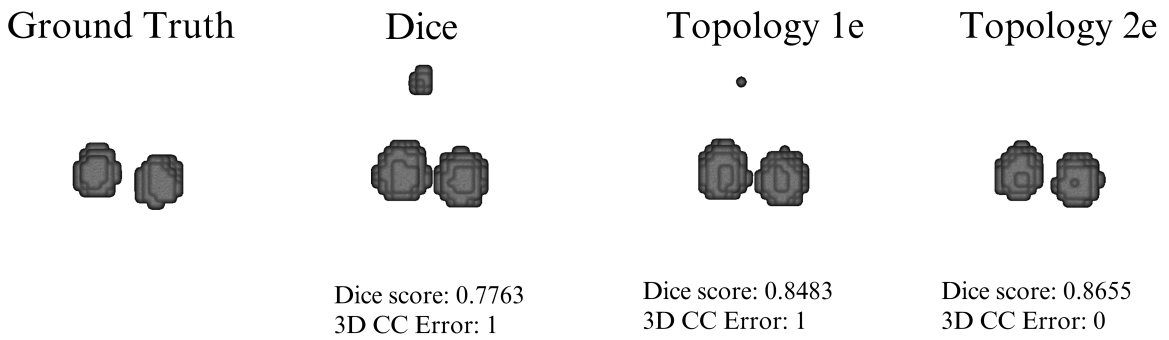

Figure 2. An illustrative example of red nucleus segmentation. Dice is the result of the Dice loss. Topological ne denotes correspond to n epochs using the proposed loss.

cases, it also allowed to improve voxel-level metrics. A limitation of our approach is that, unlike those based on persistent homology,[15–17] it does not strictly guarantee topological correctness. However, it is computationally efficient and easy to implement. Another limitation is that the sequence of kernel sizes needs to be chosen. However, we obtained suitable results for various anatomical structures without changing the kernel size.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the French government under management of Agence Nationale de la Recherche as part of the "Investissements d'avenir" program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute) and reference ANR-10-IAIHU-06 (Agence Nationale de la Recherche-10-IA Institut Hospitalo-Universitaire-6). The ICEBERG study is supported by the European Research Council (ERC) under grant agreement No. 678304, the European Union's Horizon 2020 research and innovation program under grant agreement No. 826421 (TVB-Cloud), Agence Nationale de la Recherche (ANR) under grant agreements ANR-10-IAIHU-06 (IHU ICM), ANR-11-INBS-0006, and ANR-19-JPW2-000 (JPND E-DADS), association France Parkinson (PRECISE-PD project), the Fondation d'Entreprise EDF, Biogen Inc., Fondation Thérèse and René Planiol, Fondation Saint Michel. It received unrestricted support for Research on Parkinson's disease from Energipole (M. Mallart), M. Villain and the Société Française de Médecine Esthétique (M. Legrand). Guanghui Fu is supported by the Chinese Government Scholarship provided by China Scholarship Council (CSC). Lydia Chougar is supported by a Poste d'accueil Inria/AP-HP.

REFERENCES

[1] Hesamian, M. H., Jia, W., He, X., and Kennedy, P., "Deep learning techniques for medical image segmentation: Achievements and challenges," *Journal of digital imaging* **32**(4), 582–596 (2019).

[2] Ronneberger, O., Fischer, P., and Brox, T., "U-net: Convolutional networks for biomedical image segmentation," in [*International Conference on Medical image computing and computer-assisted intervention*], 234–241, Springer (2015).

[3] Milletari, F., Navab, N., and Ahmadi, S.-A., "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in [*2016 fourth international conference on 3D vision (3DV)*], 565–571, IEEE (2016).

[4] Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P.-A., Cetin, I., Lekadir, K., Camara, O., Ballester, M. A. G., et al., "Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved?," *IEEE transactions on medical imaging* **37**(11), 2514–2525 (2018).

[5] Ma, J., Chen, J., Ng, M., Huang, R., Li, Y., Li, C., Yang, X., and Martel, A. L., "Loss odyssey in medical image segmentation," *Medical Image Analysis* **71**, 102035 (2021).

[6] El Jurdi, R., Petitjean, C., Honeine, P., Cheplygina, V., and Abdallah, F., "High-level prior-based loss functions for medical image segmentation: A survey," *Computer Vision and Image Understanding* **210**, 103248 (2021).

[7] Liu, L., Wolterink, J. M., Brune, C., and Veldhuis, R. N., "Anatomy-aided deep learning for medical image segmentation: a review," *Physics in Medicine & Biology* **66**(11), 11TR01 (2021).

[8] Huang, H., Zheng, H., Lin, L., Cai, M., Hu, H., Zhang, Q., Chen, Q., Iwamoto, Y., Han, X., Chen, Y.-W., et al., "Medical image segmentation with deep atlas prior," *IEEE Transactions on Medical Imaging* **40**(12), 3519–3530 (2021).

[9] Shit, S., Paetzold, J. C., Sekuboyina, A., Zhylka, A., Ezhov, I., Unger, A., Pluim, J. P. W., Tetteh, G., and Menze, B. H., "clDice - a Topology-Preserving Loss Function for Tubular Structure Segmentation," in [*Medical Imaging Meets NeurIPS 2019 Workshop*], (2019).

[10] Kervadec, H., Dolz, J., Wang, S., Granger, E., and Ayed, I. B., "Bounding boxes for weakly supervised segmentation: Global constraints get close to full supervision," in [*Medical Imaging with Deep Learning*], 365–381, PMLR (2020).

[11] Chen, X., Williams, B. M., Vallabhaneni, S. R., Czanner, G., Williams, R., and Zheng, Y., "Learning active contour models for medical image segmentation," in [*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*], 11632–11640 (2019).

[12] Jurdi, R. E., Petitjean, C., Honeine, P., Cheplygina, V., and Abdallah, F., "A surprisingly effective perimeter-based loss for medical image segmentation," in [*Medical Imaging with Deep Learning*], 158–167, PMLR (2021).

[13] BenTaieb, A. and Hamarneh, G., "Topology aware fully convolutional networks for histology gland segmentation," in [*International conference on medical image computing and computer-assisted intervention*], 460–468, Springer (2016).

[14] He, Y., Carass, A., Yun, Y., Zhao, C., Jedynak, B. M., Solomon, S. D., Saidha, S., Calabresi, P. A., and Prince, J. L., "Towards topological correct segmentation of macular oct from cascaded fcns," in [*Proc. International Workshop on Ophthalmic Medical Image Analysis*], 202–209, Springer (2017).

[15] Hu, X., Li, F., Samaras, D., and Chen, C., "Topology-preserving deep image segmentation," *Advances in Neural Information Processing Systems* **32** (2019).

[16] Clough, J., Byrne, N., Oksuz, I., Zimmer, V., Schnabel, J., and King, A., "A topological loss function for deep-learning based image segmentation using persistent homology," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **In Press** (2020).

[17] Byrne, N., Clough, J. R., Montana, G., and King, A. P., "A persistent homology-based topological loss function for multi-class cnn segmentation of cardiac MRI," in [*Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*], Puyol Anton, E., Pop, M., Sermesant, M., Campello, V., Lalande, A., Lekadir, K., Suinesiaputra, A., Camara, O., and Young, A., eds., 3–13, Springer International Publishing, Cham (2021).

[18] Edelsbrunner, H., Harer, J., et al., "Persistent homology - a survey," *Contemporary mathematics* **453**, 257–282 (2008).

[19] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O., "3D U-net: Learning dense volumetric segmentation from sparse annotation," in [*International conference on medical image computing and computer-assisted intervention*], 424–432, Springer (2016).

[20] Kingma, D. P. and Ba, J., "Adam: A method for stochastic optimization," *arXiv preprint* **1412.6980** (2014).

[21] Pérez-García, F., Sparks, R., and Ourselin, S., "Torchio: a python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning," *Computer Methods and Programs in Biomedicine* **208**, 106236 (2021).

[22] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al., "Pytorch: an imperative style, high-performance deep learning library," *Advances in neural information processing systems* **32** (2019).

[23] Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Landman, B. A., Litjens, G., Menze, B., Ronneberger, O., Summers, R. M., van Ginneken, B., et al., "The medical segmentation decathlon," *arXiv preprint* **2106.05735** (2021).

[24] Thibeau-Sutre, E., Diaz, M., Hassanaly, R., Routier, A., Dormont, D., Colliot, O., and Burgos, N., "Clinicadl: an open-source deep learning software for reproducible neuroimaging processing," *Computer Methods and Programs in Biomedicine* **220**, 106818 (2022).