Topology Conservation for Image Segmentation

Thomas Aussaguès (student, IMT Atlantique) & **Pierre-Henri Conze**
(Associate Prof., IMT Atlantique, LaTIM UMR 1101)
name.surname@imt-atlantique.net

I-Introduction

Image segmentation is a widely used process across various fields such as blood vessel segmentation for retinal images or road segmentation from aerial images. When considering such applications, preserving the connectivity of the different vessels or roads is a critical task.
Previously used metrics do not take into account this characteristics and lead to topological errors, as shown by Figure 1. In this work, we benchmarked a new proposed metric, **clDice** [1], which ensures topology preservation.

Index terms : Deep learning, Image segmentation, U-Net architecture, Topology preservation

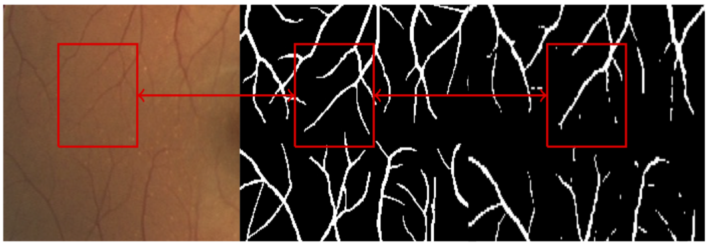

Fig. 1: Original retinal image [2], ground truth and predicted masks using **dice** metric. From a topological perspective, the prediction is not accurate as shown by the red rectangles which highlight connectivity errors.

II-Methodology

We used a convolutionnal encoder-decoder network[a] architecture (U-Net) proposed in [3] with 4 encoding and decoding layers as shown by figure 2. The network was trained using $\mathcal{L}_\alpha$ loss function involving a weighted combination between **dice** and **clDice** scores.

$$\mathcal{L}_\alpha = (1-\alpha)\textbf{dice} + \alpha(1-\textbf{clDice}) \quad (1)$$

where $\alpha \in [0, 0.5]$.

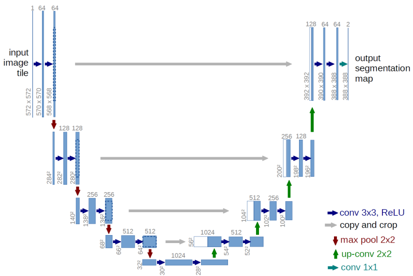

Fig. 2: U-Net architecture proposed in [3]

The first term of equation 1 refers to the **dice** metric which is a similarity score between two sets (or masks in image segmentation) X and Y given by :

$$\textbf{dice} = \frac{2|X \cap Y|}{|X|+|Y|} \quad (2)$$

The **clDice** score is computed using a skeletonization algorithm as proposed by Suprosanna Shit *et al.* More detailed definitions and theoretical material about **clDice** are provided in [1].

[a] All `python` codes can be found on the following github repo ThomasAussagues/unet_cl_dice

III-Experiments

During the experiments, we applied the proposed framework on both **DRIVE** (1) and **ROADS** (2) [4] datasets. Using affine transformations, training examples were augmented before feeding the network[a].
We used the following experimental setups for the training :

1. 200 epochs with a learning rate of $1 \cdot 10^{-3}$ and $\alpha = 0.5$,
2. 200 epochs with a learning rate of $1 \cdot 10^{-4}$ and $\alpha = 0.5$.

We compared various scores using two networks : one trained using $\mathcal{L}_\alpha$ loss function and a second one with **dice** cost function.

[a] Only for DRIVE which contains 20 training examples

IV-Results & discussion

Qualitative The tested framework led to significant topological improvements as shown in Figure 3 where one can notice the higher connectivity of the predicted masks using $\mathcal{L}_\alpha$ loss function.

Quantitative We evaluated the obtained masks using the following metrics : **dice**, **clDice** and **accuracy**. The table 1 presents the obtained scores for each loss function.

Dataset\Loss function	dice ↑			$\mathcal{L}_\alpha$ ↑		
DRIVE	**0.767**	0.778	**0.951**	0.749	**0.790**	0.950
ROADS	0.749	0.222	0.222	0.222	0.222	0.222

Table 1: Obtained scores (**dice**, **clDice** and **accuracy**)

The results show enhanced scores with $\mathcal{L}_\alpha$ loss function for **clDice** measure which accounts for topological correctness. Therefore, quantitative results confirm qualitative observations.

The tested framework achieves higher topological accuracy on both retina blood vessels and roads masks. Nevertheless, using $\mathcal{L}_\alpha$ loss function requires a large dataset. Moreover, we introduced a new parameter α which must be tuned according to the dataset. Furtherwork will extend **clDice** metric to multi-class segmentation problems.

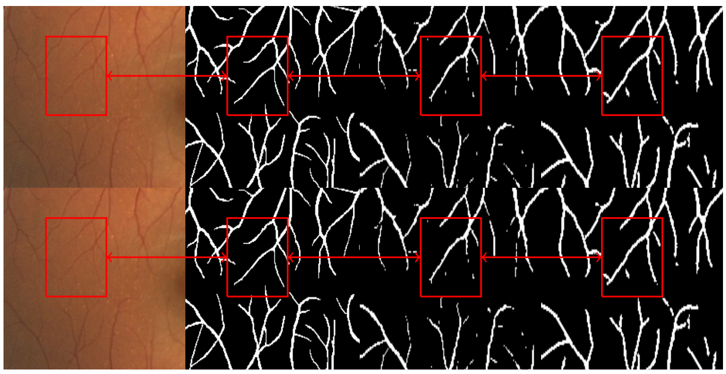

Fig. 3: Obtained results for both **DRIVE** and **ROADS** datasets. From left to right : original image, ground truth mask, predicted masks with **dice** and $\mathcal{L}_\alpha$ loss functions.

References

[1] Suprosanna Shit et al. "clDice - a Topology-Preserving Loss Function for Tubular Structure Segmentation." In: *CoRR* abs/2003.07311 (2020). URL: https://arxiv.org/abs/2003.07311.

[2] Joes Staal et al. "Ridge-Based Vessel Segmentation in Color Images of the Retina". In: *IEEE transactions on medical imaging* 23 (Apr. 2004), pp. 501–9. DOI: 10.1109/TMI.2004.825627.

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation.* 2015. arXiv: 1505.04597 [cs.CV].

[4] Volodymyr Mnih. "Machine Learning for Aerial Image Labeling". PhD thesis. University of Toronto, 2013.