clDice score for topology preservation

Astarag Mohapatra

June 2020

1 What problems are addressed here?

1. In segmentation problems related to hemodynamics (the study of dynamics of blood flow), Alzheimer's disease prediction or stroke modeling the most important job is to conserve the topology
2. The item of interest here are tubular or network like structures which are sub-optimal if we consider the performance metrics like Jaccard index or dice score.
3. Here we are more concerned if there is any miss in the segmentation mask rather than over or under segmentation. Over or under segmentation means how well the segmentation mask fits within the boundary of the image.
   - Here we are not concerned how well the mask fits the object boundary or if it is marginally thicker or thinner compared to object width, rather we are concerned about misses (false negative) or ghosts (false positive) in the mask.
4. Misses in segmentation mask can derive results of stroke in case of brain vascular vessels or insufficient road layouts in case of road network structures.
5. The metrics like dice score or Jaccard are not properly optimized for varied radius like in the case of vessels which have large, medium and small radii. Here we have a wide radius range and these performance don't paint a accurate picture with varying radius
6. So we need to develop a loss function which preserves topology primarily and a performance metric

2 Performance metric to address topology conservation

1. Here we use clDice or centerline-in mask dice coefficient.

2. We have two binary masks: the ground truth mask (V_L) and the predicted segmentation mask (V_P).The skeletons extracted from ground truth and predicted mask are S_L and S_P respectively.

3. Here we have two parameter

   - $T_{prec}(S_P, V_L) = \frac{|S_P \cap V_L|}{|S_P|}$,and it tells what fraction of S_P lies within V_L
   - $T_{sens}(S_L, V_P) = \frac{|S_L \cap V_P|}{|S_L|}$, it tells what fraction of S_L lies within V_P

4. $T_{prec}(S_P, V_L)$ is susceptible or sensitive to false positives (the ghosts) and $T_{sens}(S_L, V_P)$ is susceptible or sensitive to false negatives (the misses)

   - We get ghosts when we have some extra segmentation mask at some parts compared to the ground truth
   - And we have misses when the segmentation mask misses some part compared to the ground truth
   - So susceptible or sensitive means,if we have more false positive or ghosting then $T_{prec}(S_P, V_L)$ is less because the numerator $|S_P \cap V_L|$ is less

5. $clDice(V_P, V_L) = 2*(\mathrm{T}_{prec}(S_P, V_L)*T_{sens}(S_L, V_P)/(T_{prec}(S_P, V_L)+T_{sens}(S_L, V_P))))$

3 Loss function based on topology conservation

1. We need to have a have an end-to-end differentiable function so we adopt a loss function based on soft-skeletonization.

2. For skeletonization, previously euclidean distance is used but it doesn't have an end-to-end differentiable function.

3. So here we use iterative min- and max- pooling which is a proxy for erosion and dilation in morphological operation

4. The algorithm for soft-skeleton is

   - Given: I (mask to skeletonized), k (iterations)
   - $I^{'} \leftarrow maxpool(minpool(I)$
   - $S \leftarrow ReLU(I - I^{'})$
   - for $i \leftarrow 0$ to k do
     - $I \leftarrow minpool(I)$
     - $I^{'} \leftarrow maxpool(minpool(I))$
     - $S \leftarrow S + (1 - S) \circ ReLU(I - I^{'})$
     - The output is S

5. As we increase k (number of iterations) we get proper skeletons but the computational time increases. As k increases, we are able to skeletonize bigger radius so k is chosen between 5-10.

4 How the performance metric and loss function helped to preserve the topology

1. Betti numbers helps us to quantify differences between two different topology and we use 3 betti numbers for this purpose
   - b_0 represents number of connected components
   - b_1 represents number of circular holes and it is 2D
   - b_2 represents number of cavities and it is 3D
2. Morphological operation is topology preserving so we can compare the topology of ground truth and predicted mask after skeletonization.
3. Here we define the topology of both foreground and back ground so, P_f and P_b represent predicted mask in foreground and background and similarly L_f and L_b for ground truth.
   - Note: As we have unbalanced data sets so the background instance is more pronounced than the foreground, so we use clDice score and the loss function on foreground object only.
4. Here we define two type of topology **Top 1** and **Top 2**
   - **Top 1**, if $S_P \subset V_L$ is satisfied then predicted skeleton is inside the ground truth mask. It is no ghost condition
   - **Top 2**, if $S_L \subset V_P$ is satisfied then ground truth skeleton is inside the predicted mask. It is no misses condition
5. So now let us proof that minimizing the topology differences will increase the clDice score
   - Let the clDice score for foreground be 1 (highest possible score). If clDice score is 1 then $T_{prec}\, and\, T_{sens}$ are 1. So if $T_{prec} = 1 \Rightarrow$ then $S_P \subset V_L$ is S_P so the predicted skeleton perfectly fits inside the ground truth mask.
   - If there is any ghosting in the skeleton then T_{prec} decreases as $S_P \subset V_L$ is less because some part of predicted skeleton is outside the ground truth mask. This also decreases the clDice score.
   - So the topology difference is implicitly dependent on clDice score.
6. The final loss function is

$$L_c = \alpha(1 - softDice) + (1 - \alpha)(1 - softclDice)$$

   - $\alpha \in$ [0,0.5], as α increases the performance also increases which shows the effectiveness of dice score.

5 Datasets, Architecture and Experimentation

1. Datasets used here are
   - In 2D dataset, we used DRIVE retina dataset and the Massachusetts Roads dataset.
   - In 3D, a synthetic and a real brain vessel dataset.
2. Architecure, we used 2D and 3D U-Net and FCN (Fully convolutional network).
3. Evaluation metrics used here are
   - For overlap-based, dice coefficient, accuracy and clDice
   - For topology based metric, length of the vascular network, ratio of detected bifurcation points with respect to ground truth and Euler characteristics
     - Bifurcation ratio is similar to graph theory where we calculate the ratio of number of nodes etc.
     - Euler characteristics is the ratio of $X = V - E + F$ where V is the number of vertices, E is the number of edges and F is the number of faces. We take the ratio X for both predicted and true mask
4. Increasing α in the range of [0,0.5] we get accurate topology results. It gave good results at $\alpha = 0.5$ most of the time.
5. Despite training on foreground class only, the performance was not affected and the networks converged properly.
6. Using the above loss function gave better results than loss function based on dice score.

6 TL;DR

1. Here we are looking at segmentation of tubular and vascular structures like blood vessels and road networks
2. The main aim is to have proper network of segmentation rather than concentrating if the mask is marginally thicker or thinner compared to object boundary.
3. We are concerned to have the proper network of segmentation and not to have any misses (false negative) or ghosts (false positive).
4. The previously established metrics and loss function don't conserve topology and also are not good at evaluating variable radius segmentation problems, so here a topology based loss function and performance metric is proposed

5. In short we derive the skeletons of the ground truth and predicted mask and compare the topology of them by an end-to-end trainable loss function and clDice score