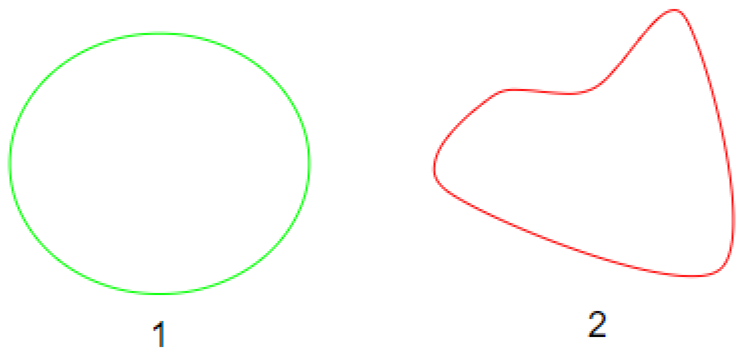
Segmentation Losses Classification Part - 2

COMBINED REGION - BOUNDARY LOSS FUNCTION

ACTIVE CONTOUR WITHOUT EDGES (LENGTH AND REGION)

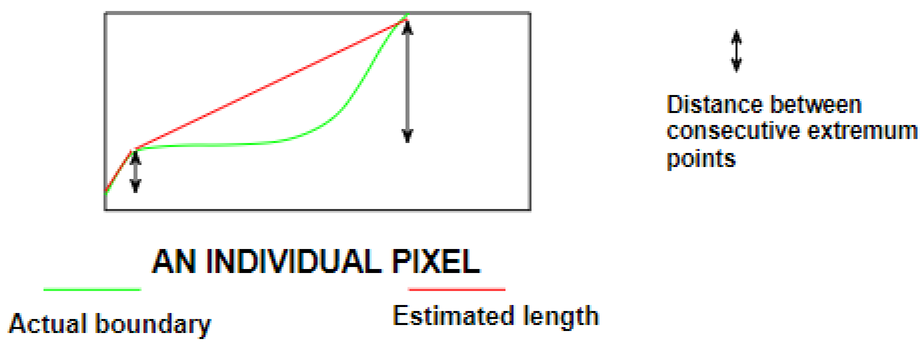
- Active contour without edges are different from active contour or curve evolution models
  - Curve evolution is an energy minimization method, where we have internal (shape of the curve) and external energy (image gradients) differential equations
  - Internal energy has first order and second order derivative terms, where the first order takes care of stretchiness of the curve and the second order takes care of the bendiness of the curve.
    - Stretchiness describes whether the points on the curve are uniformly placed or not. First order derivative terms have $f(x + \delta x) - f(x)$ in the numerator, if the points are uniformly distributed then the numerator would be less, hence the first order derivative would be less
    - Similarly if the second order derivative is less, then the curve has less twists and turns (bendiness).



Curve 1 has lower internal energy than curve 2

- External energy has image edges or gradients, so it helps the curve finally evolve to fit the boundary of object.
- The problem here is, in the process of minimization of energy of the curve, it may be stuck in local minima and give sub-optimal results.

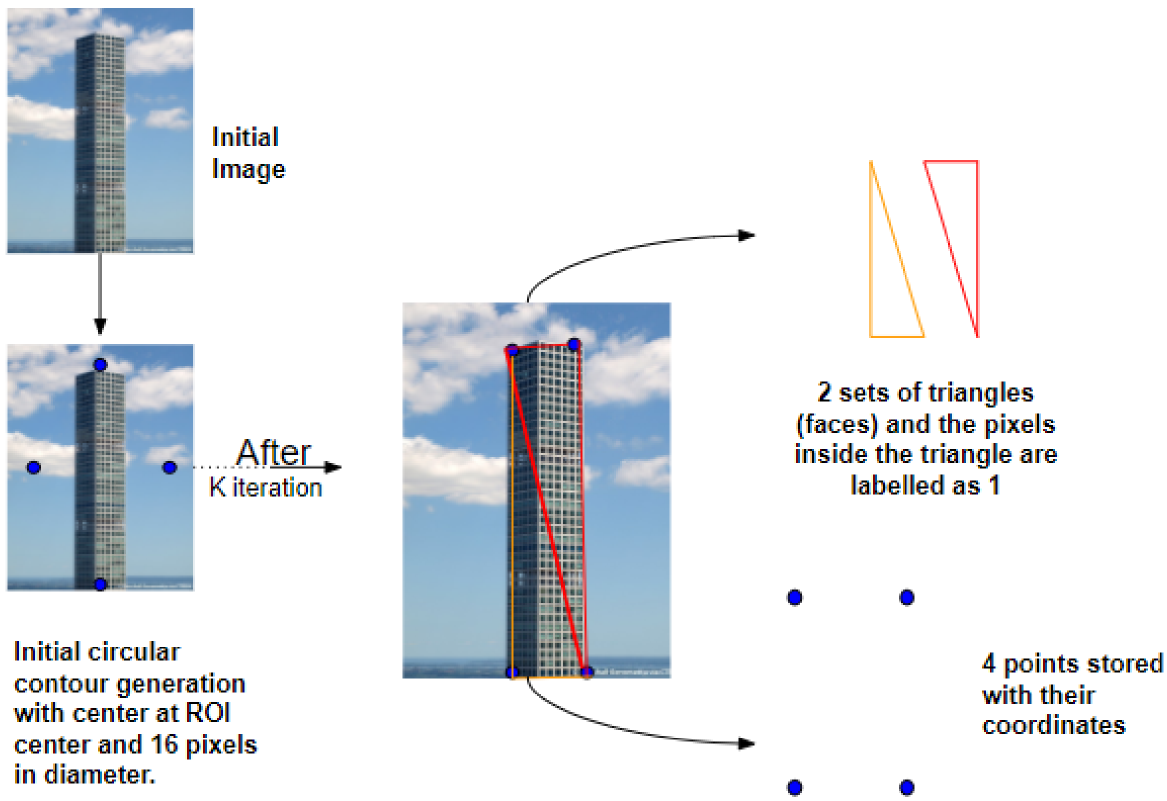
- So here we choose the length and regional area of our object of interest. As we are taking the length and area of the image, the model will be robust to noise.
  - Also the model is **dataset-specific**, the authors trained it on heart images so it is very accurate at segmenting heart images
- The methodology is
  - For length, the estimated length of boundary can be given as hypotenuse length of the horizontal and vertical projection. It is like moving along the curve and finding the length of y-projection between two consecutive extremum points



- For region, the conventional method assigning the pixel either to foreground or to the background method is used. The foreground and background of ground truth is labelled as logical 1 and 0 respectively.
- The final loss function is combination of length and regional loss function.

DIFFERENTIABLE RENDERING

- In active contour models we have a curve initialization, but curve are not so good at fitting straight edges like skyscrapers. This method applies well in images where we have straight edges.
- Every shape (2D or 3D) can be made of triangles, like a rectangle can be made of two right angled triangles. So the methodology is as follows



- The methodology is
  - We initialise a circular contour centred at the ROI's centre and of diameter 16 pixels and number of points are chosen so that we get proper boundary edge meanwhile conserving computation. It requires no user-intervention.
  - Then it is passed through an encoder and decoder layer, which outputs a displacement field J. J, helps in updating the initialised points towards the boundary of ROI
    - It gives the displacement vector along x and y direction of each point.
  - In each iteration the points are displaced according to the displacement field to cover the boundary of the region
    - $p_j^t = p_j^{t-1} + J(p_j^{t-1})$, here p_j means the vertices of polygon and t^{th} iteration
    - J guides the points to the object boundary.

        - After K-iteration, with the points we follow Delaunay triangulation to generate triangular faces
            - Delaunay triangulation is forming set of triangles from discrete points in a plane so that no point is present inside the circumcircle of a triangle.
            - The pixels inside the triangle (or the faces) are labelled as 1 (foreground) and others are labelled as 0. This is how you get predicted segment binary map.
    - The loss function is MSE distance loss between predicted contour and ground truth contour. This loss function is combined to balloon term, which helps the initial contour to expand or "balloon" if it is initialised inside the object, and a curvature reduction term, which takes the average between two points to reduce curvature.

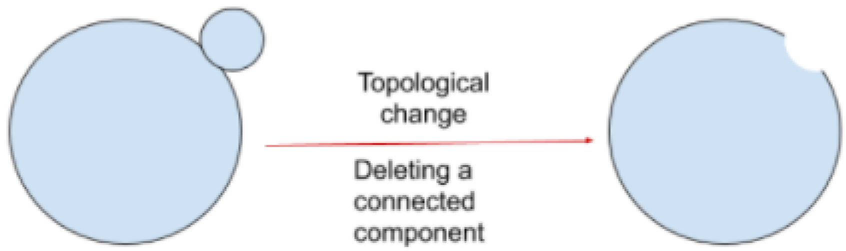
TOPOLOGY PRESERVING LOSS FUNCTIONS

- We expressed our challenges for per-pixel accuracy in challenges to segmentation. Topology preserving loss function can help ameliorate this problem.

cl-DICE SCORE (Centre-Line in mask)

- Blood vessels are very thin and tubular structures, have pixel level changes. So to conserve pixel-level accuracy we use a modification of Dice score, called as centre-line in mask dice coefficient. Another application is in road-networks detection
- The main intuition if you can ensure that your predicted segment does not have any ghosts or misses compared to ground truth, you can ensure topological correctness
    - Another key point is, here we are talking about combinatorial topology. The topology that is concerned with vertices, edges and faces.
        - Here Euler number was used as evaluation metric, which is equal to $V - E + F$, where V, E, and F are the number of vertices, edges, and faces of an object

- If we can ensure there are no misses in the foreground after the topological change (after each iteration), then we can conserve topology.

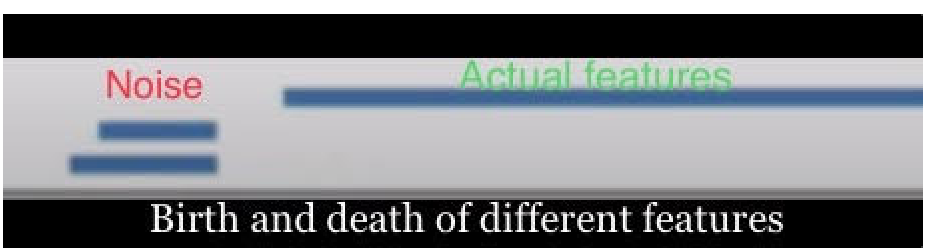


- We have predicted (V_P) and ground truth (V_L) binary masks and then we generate skeletons of the two masks, named S_P and S_L.
  - Skeletonization is an iterative min-max pooling layer so that we can include our skeletonization layer in our CNN architecture.
  - Now we will go into the mathematical formulation
    - $T_{prec}(S_P, V_L) = \frac{|S_P \cap V_L|}{|S_P|}$,and it tells what fraction of S_P lies within V_L
      - Precision takes FP or ghosts into account, the places where the model thinks the segmented object to be present but it is not in the ground truth. So T_{prec} penalizes FP
    - $T_{sens}(S_L, V_P) = \frac{|S_L \cap V_P|}{|S_L|}$, it tells what fraction of S_L lies within V_P
      - Sensitivity or recall takes FN or misses into account, the places where the model didn't predict anything, but we have some part of object according to the ground truth. So T_{sens} penalizes FN
    - **The IoU score between T_{prec} and T_{sens} is called as cl-Dice (centre-line) score** and the associated loss function is $1 - clDicescore$. This loss function is combined dice score function as final loss function expression.

TOPOLOGY PRESERVING BY USING BETTI NUMBER

- In segmentation, our main idea is to generate predicted segment similarly shaped to the ground truth. Geometrical equivalence which means to ensure we have same length and angle is a computational overload given that out predicted segment changes in each iteration. So we look for other ways of equivalence.
- Betti number is used to distinguish between two topological structures, like number of connected components ($b0$), 2D holes or connecting handles ($b1$), 3D holes or cavities ($b2$) and so on. Betti number helps in establishing topological equivalence.
    - Two structures are topologically equivalent, if they can be deformed into one another by bending, twisting, stretching and shrinking.
    - Like a square and circle have same topology, they both divide the space into two parts (inner and outer) and a square can be deformed into a circle and vice versa. They both have $b0 = 1$ and $b1 = 1$, so they belong to same homology class.
    - Topological equivalence is less computationally demanding compared to geometric equivalence.
    - But directly comparing betti number is an discrete operation, so we need a continuous process which can be end-to-end trainable.

- Our main intuition here is, if we vary the threshold of the image then some features will die and some features will be born
    - Features means connected components and handles.
    - The features that have longer survival time or we can say the features that "persists" for longer time are true features else they are noise. This is known as persistence homology



    - We compare this birth and death time of the features in predicted and ground truth
- METHOD
    - We vary the threshold of the predicted and ground truth from 0.0 to 1.0
    - The features that stay for longer duration are plotted as persistent dots on persistence diagram where we have x-axis as (1-birth_time) and y-axis as (1-death_time). This is done for both ground truth and predicted segment
    - Then we take MSE loss between the persistence dots of ground truth and predicted segment. Also we take BCE loss in the final loss function. The topological loss is more accurate on pixels compared to BCE loss.
        - The Betti number derived loss function is computationally expensive and performs well on critical pixels (the pixels which are hard to segment). So at initial part of training, we use BCE loss function which helps in regional classification of foreground and at later stage of training, we use Betti number derived loss function.
        - It is like saving your expensive and players who can perform better in critical games, for later part of the tournament.