Medical Image Analysis
Homework IV

1) Shape-aware loss functions

Shape is an important feature of an object which fundamentally differs from other objects. This feature also helps people to recognize an object and people can distinguish the object from others. When it comes to image segmentation, the shape feature has been used in active shape models in order to restrict the solution space of segmentations to a class of learned shapes. Shape-aware loss functions help deep networks recognize the object and segment more precisely considering a shape of an object.

The one shape-aware loss function I'd like to consider is from the "Shape-Aware Deep Convolutional Neural Network for Vertebrae Segmentation" paper by S. M. M. R. Al Arif et al. In this paper, authors use a loss function which has a term that is aware of segmentation shapes.

$$\hat{\boldsymbol{W}} = \arg\min_{\boldsymbol{W}} \sum_{n=1}^{N} L_t \left(\{x^{(n)}, y^{(n)}\}; \boldsymbol{W} \right)$$

$$L_s(\{x, y\}; \boldsymbol{W}) = -\sum_{i \epsilon \hat{\Omega}_p} \sum_{j=1}^{M} y_i^j E_i \log P(y_i^j = 1 | x_i; \boldsymbol{W}); \quad E_i = D(\hat{C}, C_{GT}),$$

$$P(y_i^j = 1 | x_i; \boldsymbol{W}) = \frac{\exp(a_j(x_i))}{\sum_{k=1}^{M} \exp(a_k(x_i))},$$

L_s is the shape-aware log loss per image. N is the number of training samples and $\{x^{(n)}, y^{(n)}\}$ represents the n-th example in the training set with corresponding manual segmentation. $a_j(x_i)$ is the output of the activation layer for the pixel x_i, Ω_p is the pixel space and M is the total number of segmentation class labels. P is the class probability.

$\hat{C}$ is the curve surrounding the predicted regions and C_{GT} is the ground truth curve. The function, D(·), computes the average point to curve Euclidean distance between the predicted shape and the ground truth shape. $\hat{\Omega}_p$ contains the set of pixels where the prediction mask does not match the ground truth mask.

In comparison to the cross entropy loss function, shape-aware loss function restricts its set to pixels where the prediction mask does not match the ground truth mask. Moreover, shape-aware loss function considers the distance between predicted shape and the ground truth shape. Other parts are pretty similar when it comes to their behavior and purpose.

2) Comparison of Wasserstein loss and minimax loss functions

$$-[\mathbb{E}(\log{(d(x))}) + \mathbb{E}(1 - \log{(d(g(z)))})] \qquad \mathbb{E}(c(x)) - \mathbb{E}(c(g(z)))$$

Minimax loss function | Wassertein loss function

Minimax loss function generates a value between 0 and 1 whereas W-Loss (Wassertein loss) functions outputs a real number. Additionally, the forms of the cost functions are very similar, but W-Loss doesn't use any logarithmic function within it, and that's because it's a measure of how far the prediction of the critic for the real is from its prediction on the fake. The critic is allowed to improve without degrading its feedback back to the generator. And this is because it doesn't have a vanishing gradient problem, and this will mitigate against mode collapse, because the generator will always get useful feedback back.

So, in summary, W-Loss looks very similar to BCE Loss, but it isn't as complex and a much simpler way to minimize the loss. Under the hood what it does is, approximates the Earth Mover's Distance, so it prevents mode collapse in vanishing gradient problems.