Region Mutual Information Loss for Semantic Segmentation

Abraham Jose

1 Summary

The author proposes a new loss function for training a segmentation model which is the Region Mutual Information(RMI) loss which models the dependencies among the pixels(and its immediate neighbours spatially) more simply and effectively. The loss encodes the prediction and the ground truth image to a multi-dimensional(multi-variate) distribution of the relationship between pixels and by maximizing the mutual information(MI) of the distributions we can achieve higher order consistency while training the model. In order to reduce the computation while maximizing the MI, the lower bound of the MI is maximized instead. RMI outperforms every other techniques like Conditional Random Field(CRF) and pixel affinity method on inference time and the final accuracy achieved.

2 Good points

The proposed loss function is a novel loss function considering the fact that the pixel-wise loss function ignores the dependencies between the pixels and their relationship. Compared to other pixel relationship based models like Conditional Random Field(CRF) and pixel affinity methods which requires lot of computational resource and updated architecture which makes it less feasible. Also, the proposed loss function is trying to increase the consistency between the prediction and the ground truth by maximizing the MI by maximizing the the mower bound of MI.

3 Weak points

The loss model can have many redundant computation as the loss relies on higher dimensional feature set from the neighboring pixels. The same data/information is redundantly present in the higher dimensional data in the neighboring pixels as well. So minimizing the lower bound of the mutual information can result in certain redundant computation which can be optimized.

4 Questions

Which other techniques can be used to explore the relationship between pixels? For example, losses based on Hausdorff distance can be used to get the relationship of the distance information between each set of the segments/regions.
How CRF(Conditional Random Field) can be used to calculate the loss in semantic segmentation problem?.

5 Ideas

None.