Experimental Results

May 27, 2014

1 Positive Results

We apply mmodes algorithm to a segmentation task. Using CRFs basing on hierarchical merging trees (Fig. 1) has been used in image segmentation tasks [3, 2]. There are

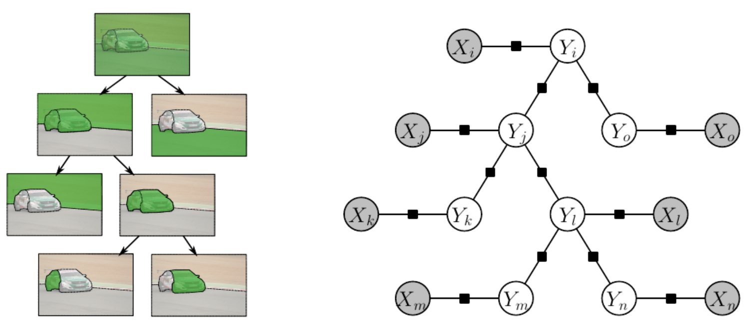

Figure 1: An illutration of the hierarchical tree. Picture from [3]. Need to reproduce.

two classes of segmentation tasks. The first class is to partition an image into segments. In the second class, one is not only expected to partition the image, but also to assign semantic labels (e.g. sky, cows or cars) to these segments. We focus on the first class and use a 3-labeled model in which each node has three labels, -1 (undersegment), 0 (segment) or 1 (oversegment). A labeling y of the tree corresponds to a segmentation under the following set of restrictions

- The root could only be labeled $\{-1, 0\}$;
- A leaf node could only be labeled $\{0, 1\}$;
- For any parent-child pair (p, c), $y_c \geq y_p \geq y_c - 1$;
- $(y_p, y_c) \neq (0, 0)$.

With these constraints enforced, a CRF can be trained and used for predictions. We use our method to predict M predictions and measure the best accuracy achieved by the top M predictions. In practice, we do not know which of the top M predictions is the best to use. Choosing a better prediction among the M is another interesting research problem and is out of the scope of this paper [4]. We compare our method to M-diverse [1], M-non-maximum-suppression, and sampling. There are different scoreing functions on this task. We show that our method outperforms other methods in VOI. Since with bigger δ, we may have insufficient modes to provide M predictions. Therefore, we compensate the method by adding additional predictions aquired by non-maximum-suppression.

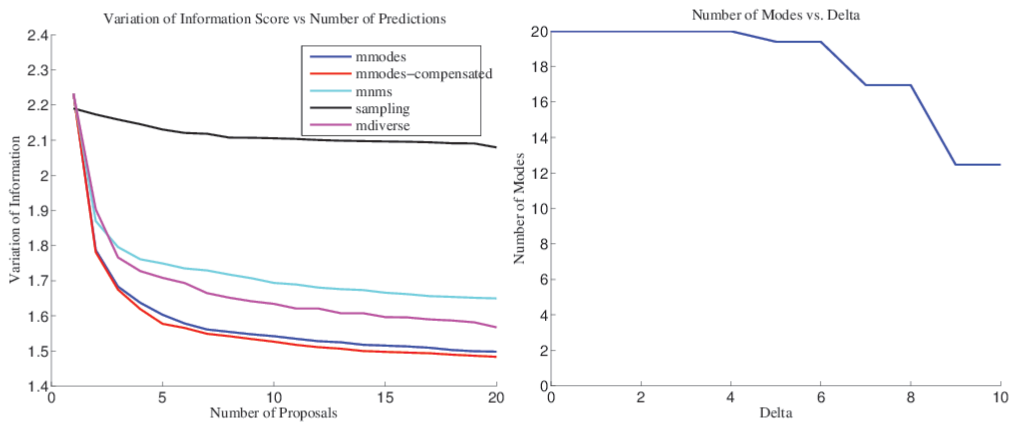

Figure 2: Left: MModes outperforms other methos in VOI. Right: Average number of modes decreases as δ increases.

In Fig. 3, we illutrate how mmodes outperforms other methods in some images.

2 Negative Results

For Berkeley dataset, our method is not winning in another score function RI (Fig. 4). The fact that RI is boundary based score while VOI is region based score might be related to this performance discrepancy.

We also test on NYUDepth dataset. However in this dataset, there seems to be a very small number of modes in general. Thus MModes does not help much. To my surprise, this “uni-modal” behavior is not changing even when the graph size is large (1200 nodes). See Fig. 5.

References

[1] D. Batra, P. Yadollahpour, A. Guzman-Rivera, and G. Shakhnarovich. Diverse m-best solutions in markov random fields. *Computer Vision–ECCV 2012*, pages 1–16, 2012.

[2] V. S. Lempitsky, A. Vedaldi, and A. Zisserman. Pylon model for semantic segmentation. In *NIPS*, volume 24, pages 1485–1493, 2011.

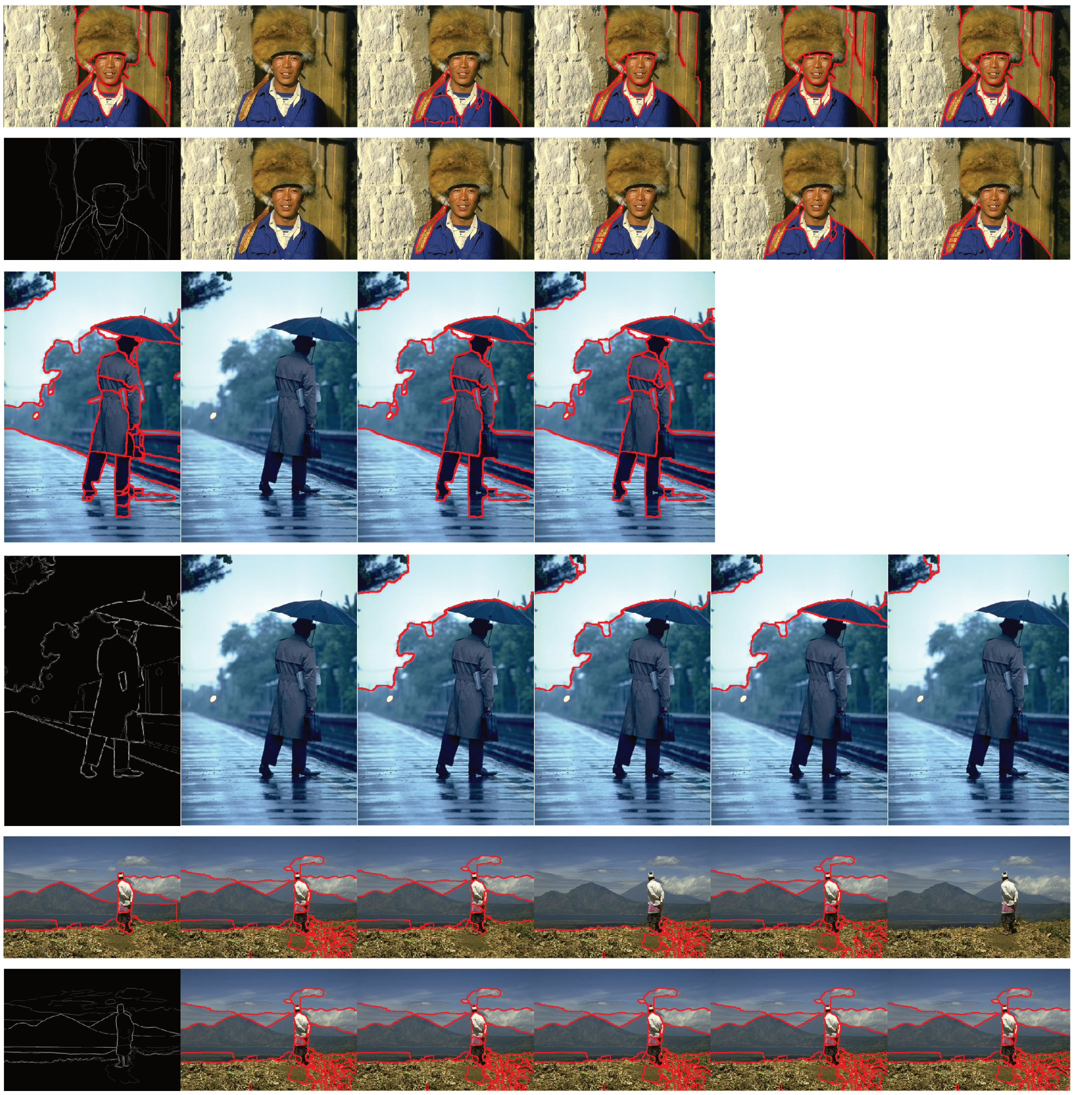

Figure 3: First column: Ground truth. Second to sixth column: the first to the fifth predictions of mmodes (top) and mdiverse (bottom).

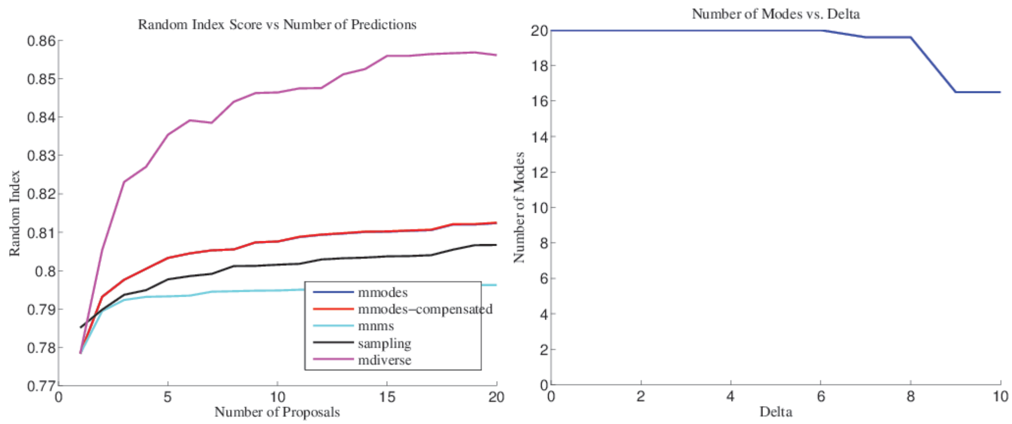

Figure 4: Left: MModes loses to MDiverse in RI. Right: Average number of modes decreases as δ increases.

[3] S. Nowozin, P. V. Gehler, and C. H. Lampert. On parameter learning in CRF-based approaches to object class image segmentation. In *ECCV*, pages 98–111. Springer, 2010.

[4] P. Yadollahpour, D. Batra, and G. Shakhnarovich. Discriminative re-ranking of diverse segmentations. *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.

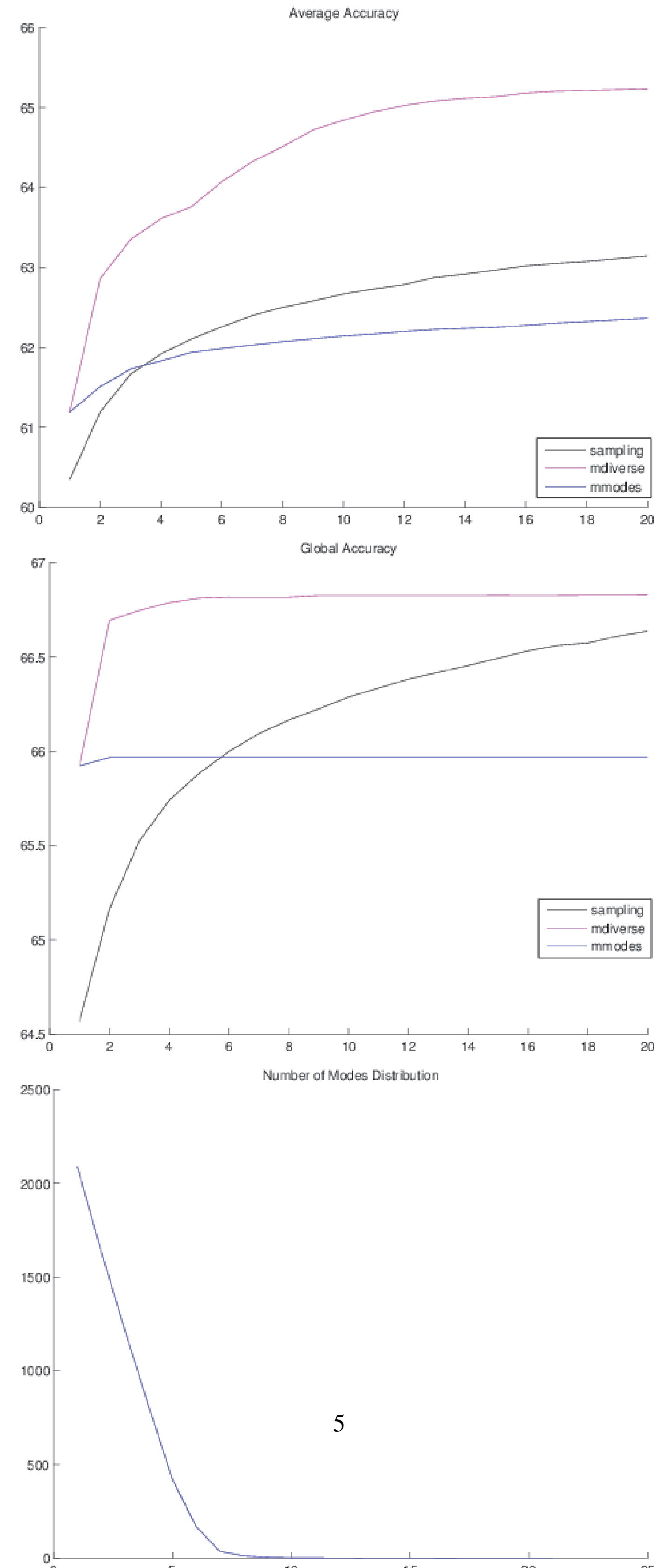

Figure 5: Left: average accuracy; Middle: global accuracy; Right: number of modes.