An End-to-end Tree based approach for Instance Segmentation

KV Manohar[1] Yusuke Niitani[2]

[1]kvmanohar22@gmail.com [2]niitani@preferred.jp

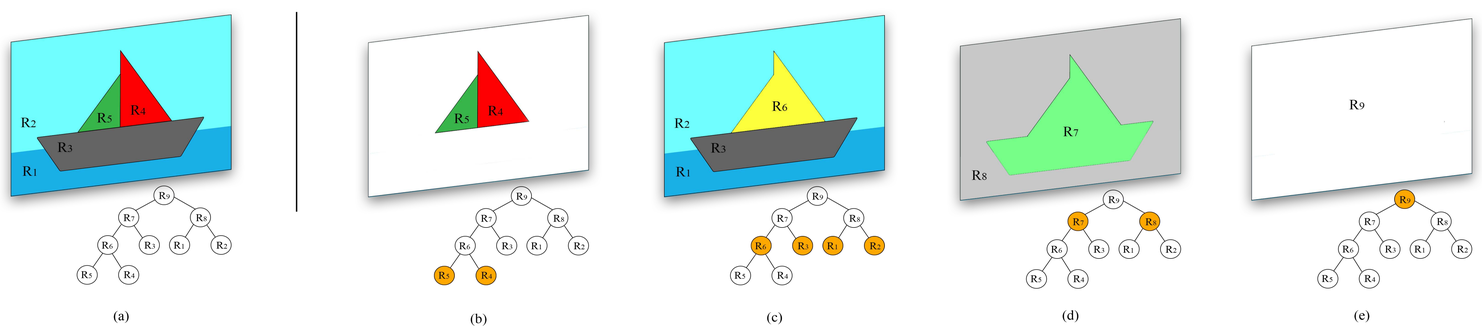
[1]Department of Mathematics, IIT Kharagpur [2]Preferred Networks Inc., Japan

1. Problem Statement

Estimate energies of regions in an hierarchical region tree using Convolutional Tree-LSTM leveraging the tree-structured network topology.

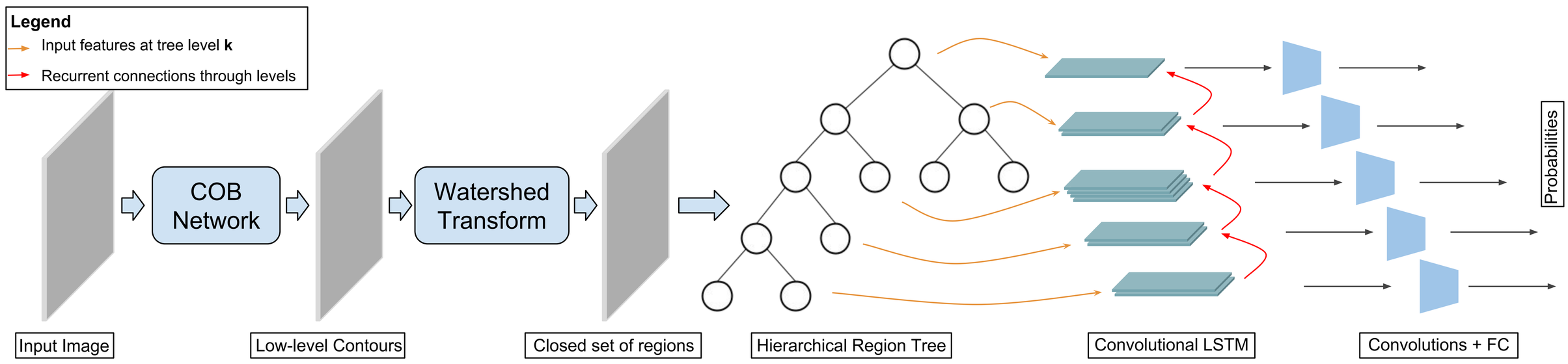
1. Estimate low level contours and orientations [1]
2. Obtain closed set of regions using Watershed Transform
3. Construct Ultrametric Contour Map (UCM)
4. Construct Hierarchical Region Tree [2]
5. Estimate energies of regions using Conv-LSTM
6. Top-down traversal yields optimal tree-cut

2. Processing Hierarchical Region Tree



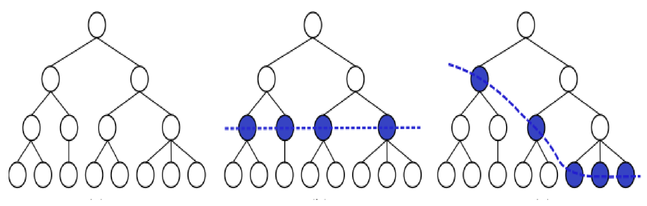
(a) Representation of initial closed set of regions as an Hierarchical Region Tree.
(b)-(e) Process each region at each level starting from the bottom all the way to the top of the tree.

3. Model Architecture



4. Tree-cuts

Left: Hierarchical region tree. **Middle**: Simple horizontal tree cut **Right:** Non-linear cut. Image credits [3]



6. Conclusions

- We proposed an unique approach for bottom-up instance segmentation.
- Our method produces comparative results with good trade-off between accuracy and processing time.
- We would like to further investigate an end-to-end network predicting contours in tandem with the estimation of energies.

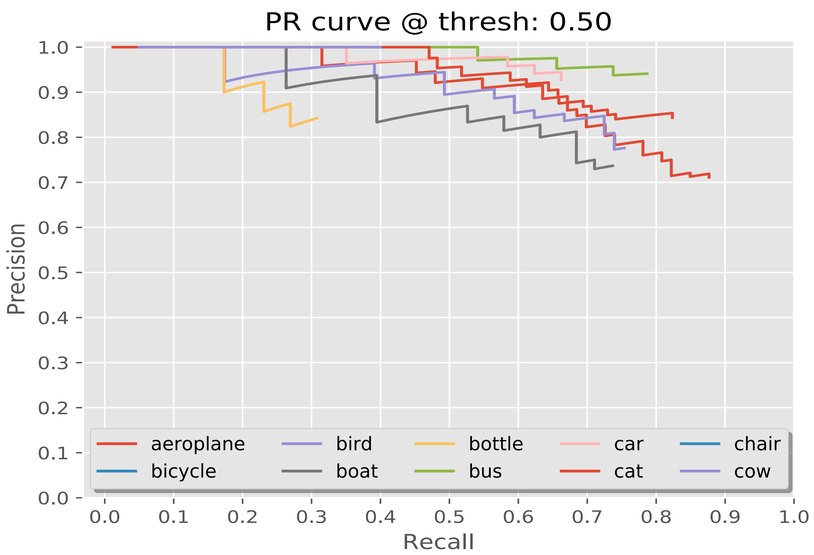
7. References

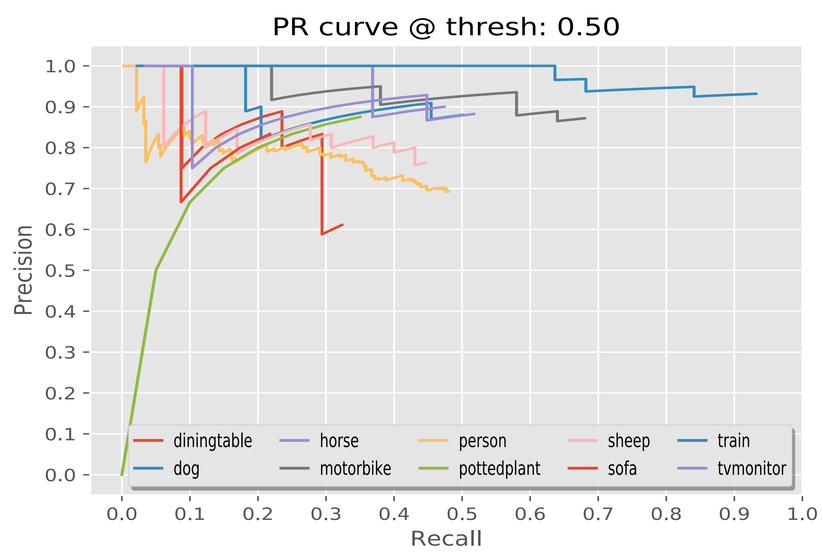
[1] K.K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool. Convolutional oriented boundaries. In *European Conference on Computer Vision (ECCV)*, 2016.

[2] P. Arbeláez, J. Pont-Tuset, J. Barron, F. Marques, and J. Malik. Multiscale combinatorial grouping. In *Computer Vision and Pattern Recognition*, 2014.

[3] T. Pham, T.-T. Do, N. Sünderhauf, and I. Reid. SceneCut: Joint Geometric and Object Segmentation for Indoor Scenes. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2018.

5. Experiments





Comparison of Jaccard Index (IoU)

Method	N	std	Plane	Cycle	Bird	Boat	Bottle	Bus	Car	Cat	Chair	Cow	Table	Dog	Global
MCG	100	0	70.2	38.8	73.6	67.7	55.3	68.5	50.6	82.4	54.4	78.1	67.7	77.7	63.7
	51	32	68	15.2	64.7	58	26.3	73.3	50.9	73.2	11.1	41.5	26.2	58.4	47.8
	36	21	68.5	16.3	63.9	56.1	24.5	72.6	50.5	73.8	11	42	27.6	58.9	47.9
Ours	23	12	69.6	14.9	67.6	57.3	34.5	72	57.2	77.3	10.2	48.8	29.2	64.8	**50.7**
	14	6	68.2	15.5	65.4	56.3	30.8	73	53.8	77	11.7	46.8	34.3	67	**50.6**

Run-time comparison (in sec)

Method	Segmentation	Generation	Total
MCG [2]	24.4 $\pm$ 3.6	9.9 $\pm$ 3.5	34.3 $\pm$ 6.2
SCG [2]	3.2 $\pm$ 0.4	1.5 $\pm$ 0.5	4.7 $\pm$ 0.7
SceneCut [3]	0.79	3.76	4.55
Ours	0.79	0.06	0.85