Data Set	Metric	B1	B2	B3	DRFI	wCtr*	RC	BSCA	PISA	**LEGS**	**MC**	**MDF**	**Ours**
HKU-IS	maxF	0.510	0.759	0.798	0.776	0.726	0.726	0.723	0.753	0.770	0.798	0.861	0.841
	MAE	0.377	0.289	0.161	0.167	0.140	0.165	0.174	0.127	0.118	0.102	0.076	0.101
PASCAL-S	maxF	0.46	0.681	0.710	0.690	0.655	0.644	0.666	0.660	0.752	0.740	0.764	0.755
	MAE	0.395	0.322	0.212	0.210	0.201	0.227	0.224	0.196	0.170	0.145	0.146	0.162
ECSSD	maxF	0.477	0.748	0.805	0.782	0.716	0.738	0.758	0.764	0.827	0.837	0.847	0.831
	MAE	0.396	0.302	0.180	0.170	0.171	0.186	0.183	0.150	0.137	0.100	0.106	0.129
DUT-OMRON	maxF	0.360	0.534	0.613	0.664	0.630	0.599	0.617	0.630	0.669	0.703	0.694	0.677
	MAE	0.347	0.378	0.154	0.150	0.144	0.189	0.191	0.141	0.133	0.088	0.092	0.138

Table 3: Comparison of quantitative results including maximum F-measure (the larger is the better) and MAE (the smaller is the better). The best three results are shown in red, blue, and green color, respectively. Note that **LEGS**, **MC**, and **MDF** are strongly supervised CNN based approaches, and our method is based on weak supervision cues.

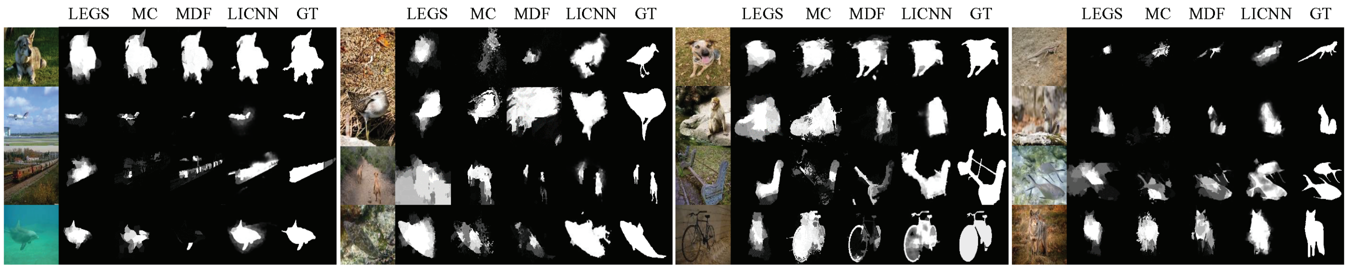

Figure 6: Visual comparison of saliency maps produced by the state-of-the-art CNN based methods, including LEGS, MC, MDF and our LICNN model. The ground truth (GT) is shown in the last column.

is merely based on a pre-trained VGG classifier which only requires the image-level class labels. In contrast, LEGS, MC and MDF strongly rely on the manually labeled segmentation masks of salient objects for model learning. And the three baseline experiments demonstrate the effectiveness of the proposed LI model. Moreover, a visual comparison between LEGS, MC, MDF and LICNN is presented in Fig. 6. LICNN performs much better in difficult cases, e.g., low contrast between objects and background.

ALL these results indicate that the bottom-up salient object detection can be carried out by grouping salient patterns via lateral inhibition. With the help of feedback signals, LICNN can organize different activated patterns effectively and merge them to form different objects. Thus, salient objects can be well perceived by a CNN for classification with only category-level labels.

Conclusion

In this paper, we have proposed a lateral inhibition based attention model, namely LICNN. In contrast to other methods, the LICNN can simultaneously perform object recognition, top-down selective attention, and salient object detection. We reveal that combining lateral inhibition and top-down feedback can construct a competitive environment to ensure that only the most discriminative and salient features are selected. With LICNN, highly discriminative category-specific attention maps are produced, and salient objects are effectively obtained without learning an independent model based on strong segmentation supervision. Both qualitative and quantitative experimental results strongly support the effectiveness of LICNN.

LICNN is an important attempt of modeling visual attention with feedback and lateral inhibition. It provides a new insight to implement brain-inspired concepts, which we believe represents a more promising route in future studies on designing vision algorithms.

Acknowledgments

This work is jointly supported by National Key Research and Development Program of China (2016YFB1001000), National Natural Science Foundation of China (61525306, 61633021, 61572504, 61420106015, 61673362),Youth Innovation Promotion Association CAS (2006121), Beijing Nova Program and Beijing Natural Science Foundation (4162058). This work is also supported by grants from NVIDIA and the NVIDIA DGX-1 AI Supercomputer.

References

Cao, C.; Liu, X.; Yang, Y.; Yu, Y.; Wang, J.; Wang, Z.; Huang, Y.; Wang, L.; Huang, C.; Xu, W.; et al. 2015. Look and think twice: Capturing top-down visual attention with feedback convolutional neural networks. In *Proceedings of the IEEE International Conference on Computer Vision*, 2956–2964.

Casanova, M. F.; Buxhoeveden, D.; and Gomez, J. 2003. Disruption in the inhibitory architecture of the cell minicolumn: implications for autisim. *The Neuroscientist* 9(6):496–507.

Chen, Y., and Fu, H. 2010. Study and application of lateral inhibition models in image's contour enhancement. In *2010*