

# Learning Complementary Maps for Light Field Salient Object Detection — Supplementary Material —

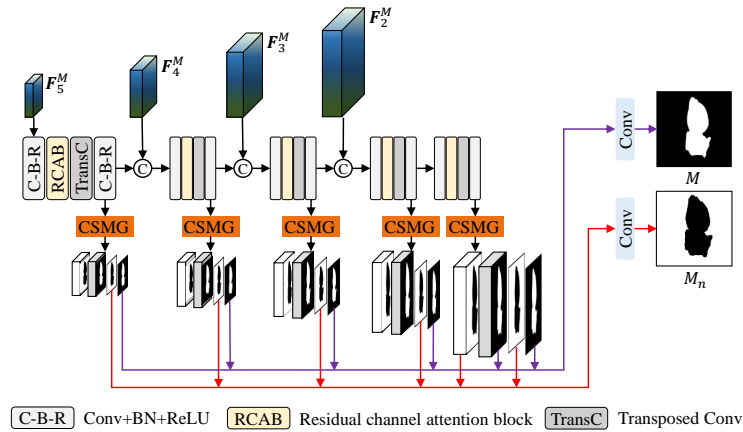
Zeyu Xiao, Jiateng Shou, and Zhiwei Xiong<sup>(✉)</sup>

MoE Key Laboratory of Brain-Inspired Intelligent Perception and Cognition,  
University of Science and Technology of China  
{zeyuxiao, shoujt}@mail.ustc.edu.cn    zwxiong@ustc.edu.cn

This supplementary document is organized as follows:  
Section 1 provides the detailed structure of the decoder.  
Section 2 provides more visual comparisons on benchmark datasets.  
Section 3 provides more ablation studies.

## 1 Detailed Structure of the Decoder

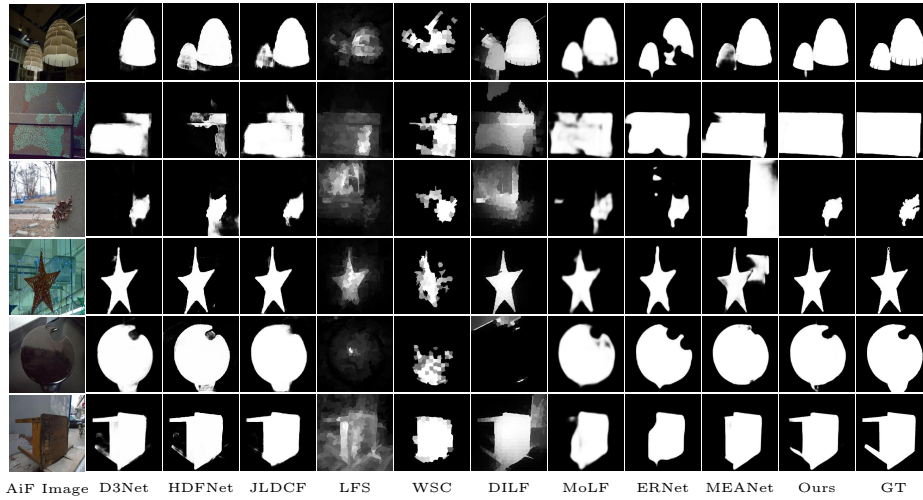
Here we show the detailed structure of the decoder in Figure 1. As can be seen in the figure, the hierarchical generated salient and non-salient maps, as well as their feature maps, are fed to the decoder to generate final saliency and non-saliency predictions.



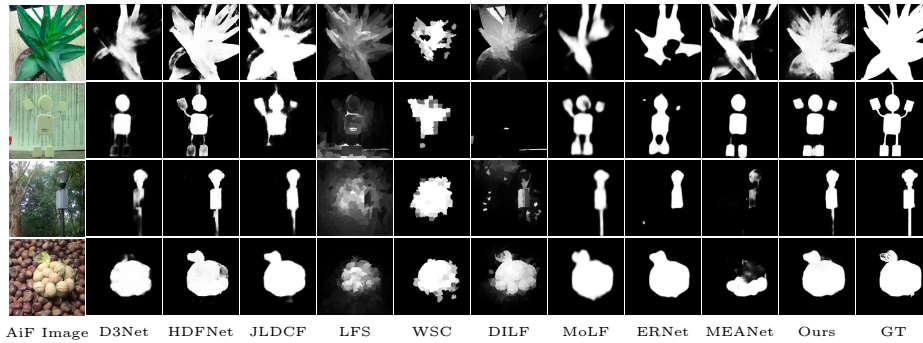
**Fig. 1:** The structure of the decoder.

## 2 More Experimental Results

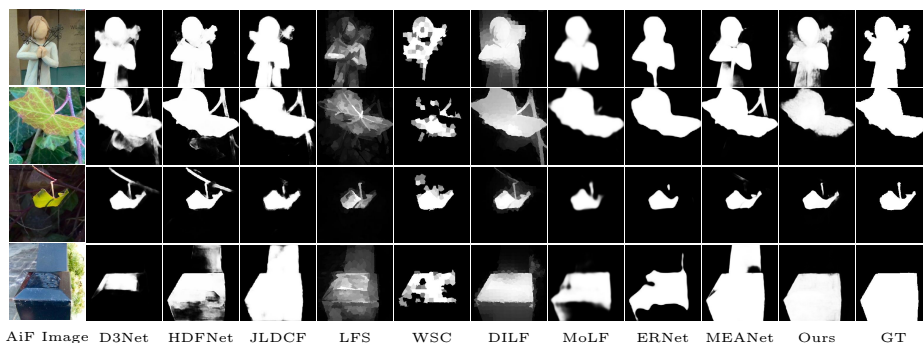
In this section, we show more visual comparisons on three benchmark datasets of different methods (*e.g.*, MEANet [3], MoLF [8], DILF [7], WSC [4], LFS [5], D3Net [1], HDFNet [6], and JLDCF [2]). As can be seen in Figure 2, Figure 3, and Figure 4, the proposed method achieves saliency maps with clear and accurate object boundaries, segmenting salient objects more accurately.



**Fig. 2:** Visual comparisons of different models on the DUTLF-FS testset. Please zoom in for better visualization.



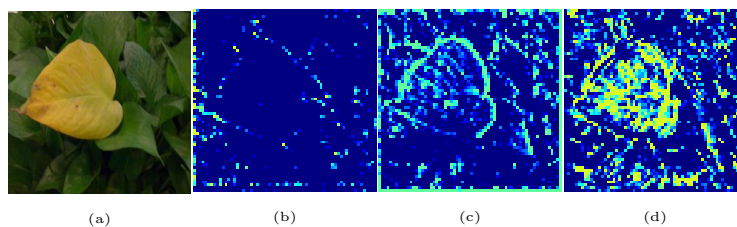
**Fig. 3:** Visual comparisons of different models on the HFUT testset. Please zoom in for better visualization.



**Fig. 4:** Visual comparisons of different models on the LFSOD testset. Please zoom in for better visualization.

### 3 More Ablation Results

Here, we visualize the intermediate feature maps in the CMA module. As seen in Figure 5, edge features are enhanced when the coarse feature map is aggregated and rich texture details are fully exploited when the fine feature map is aggregated. After feeding the feature map (a) to the CMA module, pixels are activated in a coarse-to-fine manner and respond progressively to regions and edges of the salient object.



**Fig. 5:** We visualize the intermediate feature maps in the CMA module. (a) The AiF image, (b) the feature map before aggregation (the input feature map of RRF), (c) the coarse aggregated feature map (the output feature map of RRF), and (d) the fine aggregated feature map (the output feature map of CMA).

### References

1. Fan, D.P., Lin, Z., Zhang, Z., Zhu, M., Cheng, M.M.: Rethinking rgb-d salient object detection: Models, data sets, and large-scale benchmarks. *IEEE Trans. Neural Networks Learn. Syst.* **32**(5), 2075–2089 (2020)

2. Fu, K., Fan, D.P., Ji, G.P., Zhao, Q.: JI-dcf: Joint learning and densely-cooperative fusion framework for rgb-d salient object detection. In: CVPR (2020)
3. Jiang, Y., Zhang, W., Fu, K., Zhao, Q.: Meanet: Multi-modal edge-aware network for light field salient object detection. *Neurocomputing* **491**, 78–90 (2022)
4. Li, N., Sun, B., Yu, J.: A weighted sparse coding framework for saliency detection. In: CVPR (2015)
5. Li, N., Ye, J., Ji, Y., Ling, H., Yu, J.: Saliency detection on light field. In: CVPR (2014)
6. Pang, Y., Zhang, L., Zhao, X., Lu, H.: Hierarchical dynamic filtering network for rgb-d salient object detection. In: ECCV (2020)
7. Zhang, J., Wang, M., Gao, J., Wang, Y., Zhang, X., Wu, X.: Saliency detection with a deeper investigation of light field. In: IJCAI (2015)
8. Zhang, M., Li, J., Wei, J., Piao, Y., Lu, H.: Memory-oriented decoder for light field salient object detection. *NeurIPS* (2019)