SALIENCY DETECTION VIA LOCAL SINGLE GAUSSIAN MODEL

Nan Xu, Yanqing Guo*, Xiangwei Kong
School of Information and Communication Engineering, Dalian University of Technology
Email: guoyq@dlut.edu.cn

Motivation

- **Goal:** Improve the saliency detection performance

① Highlight the salient objects uniformly ②Suppress the background noises effectively

- **Approach: Local Single Gaussian Model (LSGM) based saliency detection algorithm**

-LSGM is used to construct saliency maps. The LSGM is a dynamic model and each LSGM indicates the geometric structure information surrounding a superpixel node.

-Two LSGMs are constructed based on the background and foreground dictionaries, respectively, and then are integrated into together to get a LSGM-based saliency map.

Proposed Method

- **Learn an initial saliency map**

We utilize the bottom-up saliency map provided in [PR 2015] as a rough initial map which is denoted as S_b.

- **Dictionary construction**

$$SN(i)=\begin{cases}0 & S_b(i)<\lambda_1\\ 1 & S_b(i)>\lambda_2\end{cases} \quad i=1,2,\dots,t.$$

We define the superpixel nodes corresponding to zero elements and non-zero elements of *SN* as background nodes and foreground nodes, respectively.

- **Local Single Gaussian Model construction**

For each node r_i , we select the *K* nearest neighbors from the original dictionary as the local dictionary

$$P(r_i)=\frac{1}{\sqrt{2\pi|\Sigma_i|}}\exp[-\frac{1}{2}(f_i-\mu_i)^T\Sigma_i^{-1}(f_i-\mu_i)].$$

When background dictionary is used: $S_t^b(r_i)=\frac{1}{2\sigma^2}\exp(-\frac{P^b(r_i)}{2\sigma^2}).$

When foreground dictionary is used: $S_t^f(r_i)=\frac{1}{2\sigma^2}(1-\exp(-\frac{P^f(r_i)}{2\sigma^2})).$

Then we integrate the two saliency maps together: $S_t(r_i)=S_t^b(r_i)\times S_t^f(r_i).$

- **The final saliency map**

$S=\omega\times S_b+(1-\omega)\times S_t.$ (*w* is a balance factor)

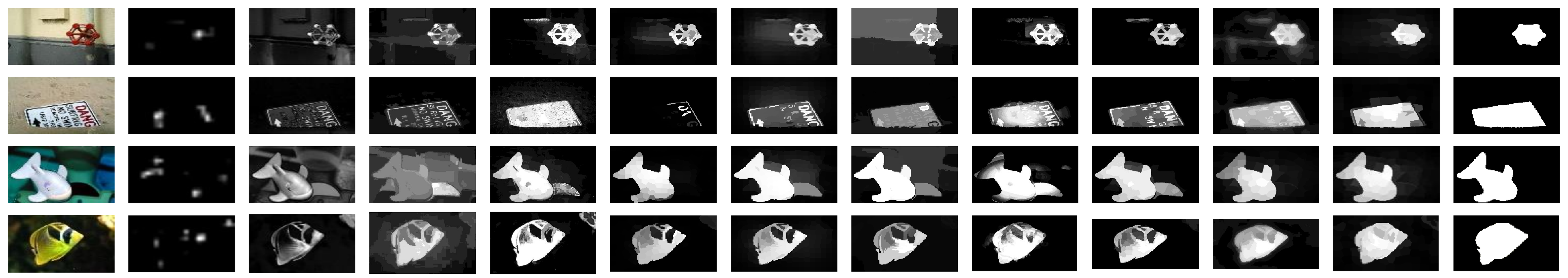
Experiments

- **Datasets**

ECSSD dataset, SOD dataset, MSRA5000 dataset, THUS dataset,

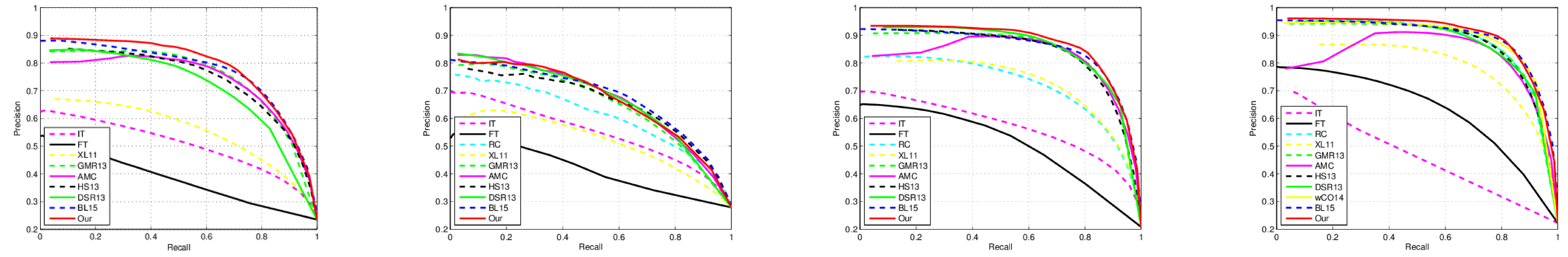
- **Results**

Saliency maps. (a) input images, (b) IT, (c) FT, (d) RC, (e) XL11, (f) GMR13, (g) AMC, (h) DSR13, (i) wCO14, (j) Tong15, (k) BL15, (l) Our method and (m) ground truth.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k) (l) (m)

P-R curves. From left to right: Experimental results on the ECSSD, SOD, MSRA5000 and THUS dataset.



AUC values on the ECSSD, SOD, MSRA5000 and THUS datasets.

	IT	FT	RC	XL11	GMR13	AMC	DSR13	wCO14	Tong15	BL15	**Ours**
ECSSD	.7920	.6296	-	.8135	.8827	.9079	.8619	-	.9117	.9142	**.9162**
SOD	.7838	.5974	.7924	.7597	.7899	.8391	.8210	-	.8366	.8466	**.8473**
MSRA	.8504	.7363	.8951	.9098	.9261	.9476	.9382	-	.9544	.9534	**.9588**
THUS	.6169	.7849	.9357	.9243	.9283	.9464	.9369	.9437	.9615	.9622	**.9638**

- **Analysis**

Our method can detect saliency accurately, suppress the background noises effectively, and uniformly highlight the salient objects.