LEARNING FULL-RANGE AFFINITY FOR DIFFUSION-BASED SALIENCY DETECTION

Keren Fu[A,B] *Irene Y.H. Gu*[A] *Jie Yang*[B*]

[A] Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden
[B] Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, China

ABSTRACT

In this paper we address the issue of enhancing salient object detection through diffusion-based techniques. For reliably diffusing the energy from labeled seeds, we propose a novel graph-based diffusion scheme called affinity learning-based diffusion (ALD), which is based on learning full-range affinity between two arbitrary graph nodes. The method differs from the previous existing work where implicit diffusion was formulated as a ranking problem on a graph. In the proposed method, the affinity learning is achieved in a unified graph-based semi-supervised manner, whose outcome is leveraged for global propagation. By properly selecting an affinity learning model, the proposed ALD outperforms the ranking-based diffusion in terms of accurately detecting salient objects and enhancing the correct salient objects under a range of background scenarios. By utilizing the ALD, we propose an enhanced saliency detector that outperforms 7 recent state-of-the-art saliency models on 3 benchmark datasets.

***Index Terms*—** Saliency detection, graph-based diffusion, affinity learning, semi-supervised learning

1. INTRODUCTION

Many graph-based diffusion processes were reported for salient object detection [1, 2, 3, 4]. Most of them can be considered as computing a saliency map induced by the diffusion result under the following formulation [3]:

$$\mathbf{s} = \mathbf{A}^{*}\mathbf{y} \tag{1}$$

where $\mathbf{A}^{*}$ is a global propagation matrix, $\mathbf{y}$ is a seed vector that makes a preliminary assessment of saliency level of graph nodes, and $\mathbf{s}$ is the diffused result that is used to induce a saliency map. Different diffusion models such as geodesic propagation [2], manifold ranking [1] and quadratic energy model [3], have been employed and reported promising results, however inhomogeneous seed ingredient and complex image contents like textures remain challenging in these techniques, as shown in Fig.1. The resulting erroneous diffusion often requires subsequent post-processing for refinement. For example, [1] proposes a second stage refinement which diffuses on the foreground map obtained by thresholding the result in the first stage. Since such post-processing highly relies on the result from the previous stage diffusion, a natural question arises: *for saliency detection, can we find a better graph-based diffusion model that propagates energy reliably and generates more accurate map?* This paper provides some new insights to this question. The main contributions of this paper are:

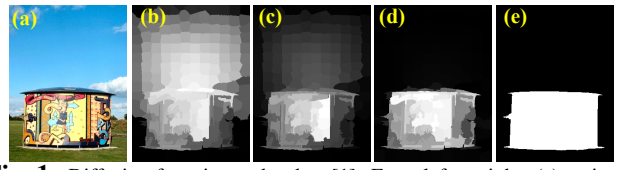

Fig. 1. Diffusion from image borders [1]. From left to right: (a) an input image; (b) manifold ranking [1]; (c) quadratic energy model [3]; (d) the proposed diffusion scheme (ALD); (e) ground truth. All diffusion models are tested under the same parameter settings.

i) We propose a simple yet effective graph-based diffusion scheme: affinity learning-based diffusion (ALD), which is based on learning a full-range affinity matrix. To the best of our knowledge, this work is the first to explicitly address the affinity learning in saliency detection.

ii) Benchmark tests between the ALD and ranking-based diffusion models (e.g. [1, 3]) are conducted. The key findings are given in Table 3, where the best models for learning affinity and ranking can be determined.

iii) Based on the ALD, we propose an enhanced saliency detector, which achieves reliable saliency detection by only one-stage diffusion and outperforms state-of-the-art approaches.

2. AFFINITY LEARNING-BASED DIFFUSION (ALD)

2.1. Problem Formulation

An input image is first represented by a similarity graph $G = (V, E, \mathbf{W})$, where V is a set of vertices (or nodes), E is a set of graph edges, and $\mathbf{W}$ is a symmetrical edge weight matrix with entry w_{ij} ($w_{ij} \geq 0$) encoding similarity between vertices ($w_{ij} = 0$ for non-connected vertices). We first over-segment the image into N SLIC superpixels [5]. Each superpixel, denoted as $v_i, i \in \{1 : N\}$, is a node in V. Superpixels v_i and v_j that satisfy either $\{v_j \in N_i\}$ or $\{\exists k, v_j \in N_k, v_k \in N_i\}$ are connected to form an edge in E, where N_i, N_k denote a set of spatial adjacent superpixels of v_i, v_k, respectively. Furthermore, arbitrary boundary superpixels are connected. An

*This research is partly supported by National Science Foundation, China (No: 61273258, 61105001), Ph.D. Programs Foundation of Ministry of Education of China (No. 20120073110018).

Table 1. Four commonly used manifold regularization models. MR_{sn}: regularization with symmetric normalized weight matrix; MR_{un}: regularization with un-normalized weight matrix; MR_{sp}: simple regularization; MR_{ln}: regularization with left-normalized weight matrix.

Model	Learned Affinity	Regularization formulation
MR_{sn}	$\mathbf{A} = (\mathbf{I} - \alpha \mathbf{D}_{\mathbf{W}}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}_{\mathbf{W}}^{-\frac{1}{2}})^{-1}$	$E(\mathbf{A}_{:,k}) = \sum_{i,j=1}^{N} \frac{1}{2} w_{ij} (\frac{a_{ik}}{\sqrt{d_i}} - \frac{a_{jk}}{\sqrt{d_j}})^2 + \mu \sum_{i=1}^{N} (a_{ik} - b_{ki})^2$
MR_{un}	$\mathbf{A} = (\mathbf{D}_{\mathbf{W}} - \alpha \mathbf{W})^{-1}$	$E(\mathbf{A}_{:,k}) = \sum_{i,j=1}^{N} \frac{1}{2} w_{ij} (a_{ik} - a_{jk})^2 + \mu \sum_{i=1}^{N} d_i (a_{ik} - \frac{b_{ki}}{d_i})^2$
MR_{sp}	$\mathbf{A} = (\mathbf{D}_{\mathbf{W}} - \mathbf{W} + \mu \mathbf{I})^{-1}$	$E(\mathbf{A}_{:,k}) = \sum_{i,j=1}^{N} \frac{1}{2} w_{ij} (a_{ik} - a_{jk})^2 + \mu \sum_{i=1}^{N} (a_{ik} - b_{ki})^2$
MR_{ln}	$\mathbf{A} = (\mathbf{I} - \alpha \mathbf{D}_{\mathbf{W}}^{-1} \mathbf{W})^{-1}$	$E(\mathbf{A}_{:,k}) = \sum_{i,j=1}^{N} \frac{1}{2} w_{ij} (a_{ik} - a_{jk})^2 + \mu \sum_{i=1}^{N} d_i (a_{ik} - b_{ki})^2$

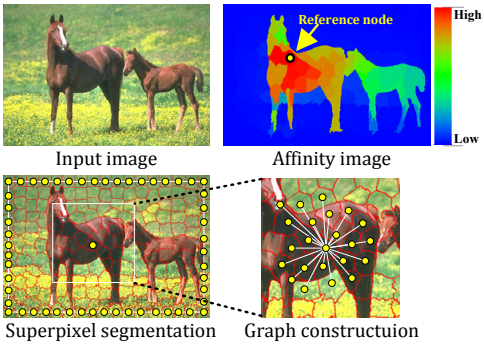

Fig. 2. Graph structure and affinity learning. Upper left: An input image. Upper right: Learned affinity with respect to a reference node. Bottom: Superpixel segmentation and graph construction. Yellow circle dots are vertices and white lines are graph edges.

illustration of graph structure is given in Fig.2.

Given such a graph $G = (V, E, \mathbf{W})$, the key idea of affinity learning-based diffusion (ALD) is to learn a full-range affinity matrix $\mathbf{A}$, with which the diffusion is formulated as:

$$\mathbf{s} = \mathbf{D}_{\mathbf{A}}^{-1} \mathbf{A} \mathbf{y} \tag{2}$$

where $\mathbf{y}$ is a seed vector, $\mathbf{A}$ is a symmetric full-range affinity matrix with entry a_{ij} ($a_{ij} \geq 0$) that encodes pairwise affinity between ith and jth nodes, $\mathbf{D}_{\mathbf{A}}$ is the diagonal degree matrix of $\mathbf{A}$ (i.e., with ith diagonal entry be $\sum_j a_{ij}$), and $\mathbf{s}$ is the diffused result. A desired full-range affinity matrix $\mathbf{A}$ satisfies that, if two superpixels (ith and jth) locate in a same visually coherent region, the affinity value a_{ij} between them is large. Otherwise, a_{ij} is small or $a_{ij} \to 0$ (see the "affinity image" in Fig.2). With learned $\mathbf{A}$, $\mathbf{D}_{\mathbf{A}}^{-1}\mathbf{A}$ serves as a filtering kernel (i.e. with each row summed to 1) and performs the diffusion on $\mathbf{y}$. This filtering kernel has a conservation property [6, 7]. If one conducts a linear transformation $\mathbf{y} \leftarrow k\mathbf{y} + c$ (k, c are constants) that does not change the ranking order of seed elements, the resulting $\mathbf{s}$ will maintain the same transformation $\mathbf{s} \leftarrow k\mathbf{s} + c$, yielding to the same $\mathbf{s}$ after linearly normalizing $\mathbf{s}$ into [0,1].

2.2. Determine A by Semi-Supervised Learning

To learn the long range affinity between superpixels in an image, inspired by [8] which proposes to learn the full affinity for spectral segmentation, we employ graph-based semi-supervised learning (GSSL) [9, 10]. The GSSL is used to obtain a relevance score between each node and a reference node [8]. We use such relevance scores as our affinity values. Note that there are three categories in the GSSL [10]: mincut, harmonic function, and manifold regularization. The first two categories are only applicable to multi-class semi-supervised learning, whereas our affinity learning is a one-class problem. For the manifold regularization, four common models in machine learning [9, 10] can be used as summarized in Table 1. In the following, we take the first model (MR_{sn}) in Table 1 as an example and show how the affinity values are learned. Recall that the graph has N nodes, i.e. $|V| = N$, and using MR_{sn}, the affinity vector $\mathbf{A}_{:,k}$ (kth column of $\mathbf{A}$) containing the relevance scores of all nodes with respect to a reference node v_k is computed as:

$$\mathbf{A}_{:,k} = (\mathbf{I} - \alpha \mathbf{D}_{\mathbf{W}}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}_{\mathbf{W}}^{-\frac{1}{2}})^{-1} \mathbf{b}_k \tag{3}$$

where $\mathbf{b}_k$ is a label vector, whose element corresponding to the reference node v_k (i.e., the kth node) is 1 whereas the remaining unlabeled elements are 0, $\mathbf{D}_{\mathbf{W}}$ is the diagonal degree matrix of $\mathbf{W}$, and $\mathbf{I}$ is the identity matrix. One can prove (3) is the closed-form solution to minimizing the following regularization function:

$$E(\mathbf{A}_{:,k}) = \sum_{i,j=1}^{N} \frac{1}{2} w_{ij} (\frac{a_{ik}}{\sqrt{d_i}} - \frac{a_{jk}}{\sqrt{d_j}})^2 + \mu \sum_{i=1}^{N} (a_{ik} - b_{ki})^2 \tag{4}$$

where a_{ik}, b_{ki} denote the ith element of $\mathbf{A}_{:,k}$, $\mathbf{b}_k$, respectively, and d_i is the ith diagonal entry of $\mathbf{D}_{\mathbf{W}}$. The first summation term in (4) is the *pairwise smoothness* term. It forces similar neighbor nodes to take similar relevance scores. The second summation term in (4) is the *data fitness* term, which requests the relevance scores fitting to the original labels. Here we have $\alpha = \frac{1}{1+\mu}$. By stacking all affinity vectors in a row, one obtains the learned affinity matrix $\mathbf{A}$:

$$\begin{aligned} \mathbf{A} &= [\mathbf{A}_{:,1}, \mathbf{A}_{:,2}, ..., \mathbf{A}_{:,N}] \\ &= (\mathbf{I} - \alpha \mathbf{D}_{\mathbf{W}}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}_{\mathbf{W}}^{-\frac{1}{2}})^{-1} [\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_N] \\ &= (\mathbf{I} - \alpha \mathbf{D}_{\mathbf{W}}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}_{\mathbf{W}}^{-\frac{1}{2}})^{-1} \end{aligned} \tag{5}$$

It is easy to see that $\mathbf{A}$ computed is symmetric, and all entries of $\mathbf{A}$ are greater than 0, making $\mathbf{A}$ a valid full-range affinity matrix applicable to (2). Likewise, one may also apply three other GSSL models in Table 1 to the affinity learning task. For all models in Table 1, we have $\alpha = \frac{1}{1+\mu}$ if any. It is worth noting that in Table 1, $\mathbf{A}$ from the fourth model MR_{ln} is not symmetric, therefore $\mathbf{A} \leftarrow (\mathbf{A} + \mathbf{A}^T)/2$ is computed.

The overview of the proposed ALD scheme is in Table 2. It is worth noting that all models in Table 1 can be used to perform regularization directly on the seed vector $\mathbf{y}$, where the diffusion is formulated as **a ranking problem** (for details

Table 2. Affinity Learning-based Diffusion (ALD).

1. Construct a local neighbor graph $G = (V, E, \mathbf{W})$;
2. Learn the full-range affinity matrix $\mathbf{A}$ from $\mathbf{W}$;
3. Set diagonal of $\mathbf{A}$ to zeros, and use self-normalized matrix $\mathbf{D}_{\mathbf{A}}^{-1}\mathbf{A}$ to diffuse from the seed vector $\mathbf{y}$;

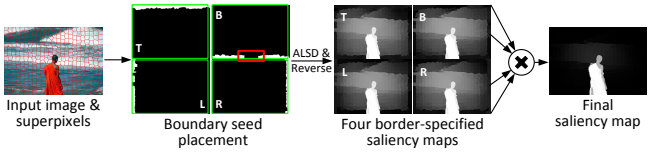

Fig. 3. Processing pipeline of the proposed method. "T,B,L,R" are short for "Top, Bottom, Left, Right", respectively. The proposed method is able to achieve good result with less refinement.

see [1, 3]). In Section 4, we compare between using the four models in Table 2 for both ALD and ranking-based diffusion.

3. SALIENCY DETECTION BY USING ALD

By choosing the most effective affinity learning model in Table 1, we may achieve more accurate diffusion. By utilizing the ALD, we propose an enhanced saliency detector inspired by [1]. The block diagram of our method is shown in Fig.3. We firstly use four image borders as seeds but remove boundary nodes that are likely to belong to boundary-cropping objects. Next, we use the proposed ALD to perform diffusion on the seeds separately, and reverse each map (similar to [1]) to suppress the background and highlight potential objects. Finally, four border-specified maps are superpixel-wisely multiplied to obtain the final saliency map. For the affinity learning model we use, see Section 4.2. Details of the proposed scheme are as following:

Graph edge weight matrix W: The weights among nodes are constructed by incorporating both color and image edge cues among superpixels [11]:

$$w_{ij} = \sqrt{\underbrace{\exp(-\lambda_c||\mathbf{c}_i - \mathbf{c}_j||)}_{\text{color term}} \cdot \underbrace{\exp(-\lambda_e \max_{i' \in \bar{ij}} f_{i'})}_{\text{edge term}}} \quad (6)$$

where $\mathbf{c}_i$/$\mathbf{c}_j$ are mean CIELab colors of v_i/v_j, $\bar{ij}$ is the straight line on the image plane connecting centroids of two superpixels, i' refers to a pixel on $\bar{ij}$, and $f_{i'}$ indicates the edge magnitude at i', and λ_c/λ_e are parameters controlling the damping rate of the two terms. We derive $f_{i'}$ from an edge detector [12]. Using (6), large color difference and strong edges between two superpixels lead to small similarity weights [11].

Border seed placement: To alleviate the possible over-suppression of a salient object that touches any border (see Fig.3), border superpixels in the background seed vector $\mathbf{y}$ are filtered according to a boundary connectivity prior [13]. We first define a soft region $\mathcal{A}_k$ associated with each boundary node v_k using the geodesic method in [13], and its corresponding prior is defined as:

$$\mathcal{P}_k = 1 - \exp\{-\frac{Ratio^2(v_k)}{2\sigma_{bc}^2}\} \quad (7)$$

where $Ratio(v_k) = Len(\mathcal{A}_k)/\sqrt{Area(\mathcal{A}_k)}$, $Len(\mathcal{A}_k)$ is the cropped length of $\mathcal{A}_k$ at image borders, and $Area(\mathcal{A}_k)$ is the area of $\mathcal{A}_k$. The prior value is close to 0 if the boundary node belongs to an object that adjoins image boundary, and is close to 1 if it belongs to the background. $\sigma_{bc} = 1$ is set according to [13]. The kth element of $\mathbf{y}$ is then defined accordingly as:

$$y_k = \begin{cases} \mathcal{P}_k & \text{if } \mathcal{P}_k > T_{\text{bc}} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where T_{bc} is a predefined threshold used to clip the weak priors. The function of T_{bc} is shown in Fig.3 by the red rectangle.

4. RESULTS AND COMPARISONS

Two types of experiments were conducted and the results are reported in this paper: 1) to compare the performance of using the four models in Table 1 for the ALD as well as for ranking-based diffusion. 2) to compare the proposed saliency detector to 7 state-of-the-art saliency detectors.

Three benchmark datasets: MSRA-1000 [14] (1000 images), SOD [15] (300 images), and ECSSD [16] (1000 images with texture background) are used. Evaluations are conducted in terms of two metrics: (1) precision-recall (PR) curve [14, 17, 1, 18, 19]. (2) weighted F-measure (F_β^w) [20].

4.1. Diffusion through using Different Models

Setups: It is important to compare different diffusion models under the **same settings**, i.e. μ, $\mathbf{W}$, and $\mathbf{y}$. For fair comparison, we follow the original graph configuration in [1, 3] on both graph topology (the same as introduced in Section 2) and edge weight computation. In the setting, 200 superpixels are used and w_{ij} is computed solely by the color term in (6) ($\lambda_c = 10$ according to [1]). The regularization parameter μ of all models in Table 1 is fixed as 0.01, leading to $\alpha = 0.99$. Two types of seed vectors were used. The first type is regarding superpixels on each image border of an image as *background seeds*. The second type is deriving an initial *foreground seed*. For the latter, we use the most effective foreground seed in [3]. This seed, called "backgroundness" is obtained by computing the color similarity between a superpixel and all boundary superpixels as follows:

$$\mathbf{y}_i = 1 - \mathcal{N}\{\sum_{j:v_j \in B} Sim(v_i, v_j)\}$$

where B is the boundary superpixel set, $\mathcal{N}$ is the normalization function, and the color similarity function $Sim(v_i, v_j)$ is consistent to the one for computing w_{ij}. We use notation "MR$_{sn}$" to refer to directly using a model for ranking-based diffusion, and use "MR$_{sn}$_ALD" to refer to using the same model for affinity learning-based diffusion (ALD).

Results: Evaluations by F_β^w on three datasets are shown in Table 3. One can see that the ALD generally outperforms using the models for ranking saliency directly. The best performing diffusion models are MR$_{un}$_ALD, MR$_{sn}$_ALD, which significantly improves the diffusion accuracy over

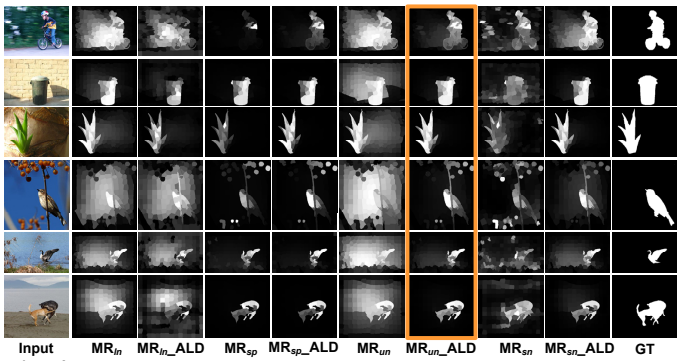

Fig. 4. Visual comparison of diffusion via different models from the background seeds. The model which performs generally the best is highlighted in the brown dash box. GT represents ground truth.

Table 3. Evaluation by F_β^w. The bracket with "B"/"F" after a dataset indicates diffusion from background/foreground seeds as mentioned in Section 4.1. The bracket after a F_β^w score indicates ranking position on a certain dataset. The best score on a dataset is highlighted in bold.

Methods	MSRA(B)	ECSSD(B)	SOD(B)	MSRA(F)	ECSSD(F)	SOD(F)
MR$_{ln}$	0.709(4)	0.466(3)	0.398(3)	0.666(5)	0.423(2)	0.371(3)
MR$_{ln}$_ALD	0.609(7)	0.374(6)	0.312(8)	0.557(7)	0.363(6)	0.322(7)
MR$_{sp}$	0.651(6)	0.352(7)	0.330(6)	0.679(4)	0.372(5)	0.333(5)
MR$_{sp}$_ALD	0.717(3)	0.434(5)	0.373(5)	**0.694(1)**	0.403(4)	0.356(4)
MR$_{un}$	0.660(5)	0.440(4)	0.381(4)	0.617(6)	0.340(7)	0.323(6)
MR$_{un}$_ALD	**0.763(1)**	0.471(2)	**0.408(1)**	0.681(3)	0.420(3)	**0.377(1)**
MR$_{sn}$	0.509(8)	0.331(8)	0.325(7)	0.349(8)	0.272(8)	0.245(8)
MR$_{sn}$_ALD	0.747(2)	**0.479(1)**	**0.408(1)**	0.682(2)	**0.427(1)**	0.376(2)

MR$_{un}$ and MR$_{ln}$. The best model for ranking-based diffusion is MR$_{ln}$. This reveals that MR$_{un}$ used by [1] is a less good choice. The worst diffusion model is MR$_{sn}$.

When comparing the performance between diffusion on the background and foreground seeds in Table 3, one can observe that the diffusion from image borders [1] is more effective than from the foreground seed [3].

Visual comparison among different diffusion models are shown in Fig.4, which reveals that choosing an appropriate diffusion model is crucial to obtaining an accurate saliency map, and our method consistently improves diffusion quality over MR$_{sp}$, MR$_{un}$, MR$_{sn}$. As a result, the background is suppressed effectively and entire objects are emphasized.

4.2. Comparison with 7 State-of-the-art Methods

Setups: Section 4.1 allows one to choose the best diffusion model for the ALD. We choose MR$_{un}$_ALD since it performs favorably on diffusion from both the foreground and background seeds. MR$_{sn}$_ALD can be chosen as well since it achieves reasonably good performance similar to MR$_{un}$_ALD. In the proposed method $N = 400$ superpixels are used for a richer representation of image contents. $\lambda_c = \lambda_e = 10$ and $T_{\text{bc}} = 0.4$ are empirically set. The edge maps are produced by the structured random forest edge detector [12], which works on multi-scales and achieves state-of-the-art performance with fast speed. Other edge detectors can be used if they produce reasonably good edge maps.

We compare with 7 recent state-of-the-art methods including: GS (Geodesic Saliency) [21], HS (Hierarchical Saliency) [16], PCA [22], DRFI (Discriminative Regional Feature Integration) [23], GC (Global Cue) [18], GMR (Graph-based Manifold Ranking) [1], wCtrO (background weighted Contrast with Optimization) [13]. The implementation of [3] is not publicly available. Other methods exist, but most of them are inferior to the aforementioned ones.

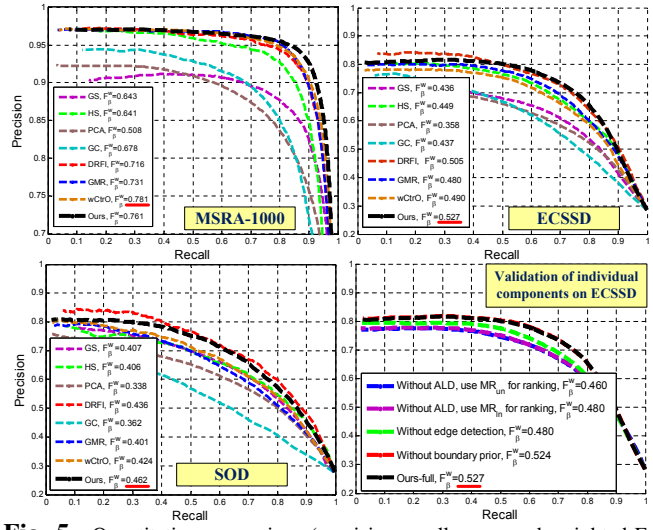

Fig. 5. Quantitative comparison (precision-recall curve and weighted F-measure F_β^w). The best F_β^w is underlined in red.

Results: Quantitative evaluation of different methods on the three datasets are shown in Fig.5. The proposed method by simply improving the diffusion technique has achieved very competitive performance to the best performing state-of-the-arts, some of which have adopted multi-stage refinement/optimization (e.g. GMR, wCtrO) or supervised learning (e.g. DRFI), whereas our method only uses one-stage diffusion. On F_β^w, our method achieves the best results on ECSSD, SOD, and the second best on MSRA-1000.

To show the effectiveness of individual modules, evaluation was conducted on ECSSD by ablating each component from our full implementation in Fig.3. The results are shown in the bottom-right sub-figure in Fig.5. One can see that the diffusion technique contributes most to the performance gain. Changing the diffusion to ranking-based methods MR$_{ln}$/MR$_{un}$ largely degrades the performance. Besides, incorporating edge cue leads to moderate improvement, whereas the introduced boundary connectivity prior marginally boosts the performance.

5. CONCLUSION

A novel diffusion scheme based on affinity learning is proposed for saliency detection. Our results from the proposed scheme and comparisons with 7 state-of-the-art methods on three benchmark datasets have shown that the proposed diffusion scheme consistently outperforms existing ranking-based diffusion methods by even using a simple one-stage diffusion. This shows that diffusion models are crucial to high quality saliency detection, and the proposed diffusion method is effective. Future work will be conducted on applying the proposed diffusion to other computer vision applications.

6. REFERENCES

[1] C. Yang and L. Zhang et al, "Saliency detection via graph-based manifold ranking," in *CVPR*, 2013.

[2] K. Fu, C. Gong, I. Gu, and J. Yang, "Geodesic saliency propagation for image salient region detection," in *ICIP*, 2013.

[3] S. Lu and V. Mahadevan et al, "Learning optimal seeds for diffusion-based salient object detection," in *CVPR*, 2014.

[4] K. Fu, I. Gu, C. Gong, and J. Yang, "Robust manifold-preserving diffusion-based saliency detection by adaptive weight construction," *Neurocomputing*, vol. 175, pp. 336–347, 2016.

[5] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Ssstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 34, no. 11, pp. 2274–2282, 2012.

[6] R. Coifman and S. Lafon, "Diffusion maps," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 5–30, 2006.

[7] X. Yang, S. Köknar-Tezel, and L. Latecki, "Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval," in *CVPR*, 2009.

[8] T.H Kim et al, "Learning full pairwise affinities for spectral segmentation," *TPAMI*, vol. 35, no. 7, pp. 1690–1703, 2013.

[9] D. Zhou et al, "Learning with local and global consistency," in *NIPS*, 2003.

[10] X. Zhu and A. Goldberg, "Introduction to semi-supervised learning," *Synthesis lectures on artificial intelligence and machine learning*, vol. 3, no. 1, pp. 1–130, 2009.

[11] K. Fu, C. Gong, I. Gu, and J. Yang, "Normalized cut-based saliency detection by adaptive multi-level region merging," *TIP*, vol. 24, no. 12, pp. 5671–5683, 2015.

[12] P. Dollár and C.L. Zitnick, "Structured forests for fast edge detection," in *ICCV*, 2013.

[13] W. Zhu, S. Liang, and Y. Wei, "Saliency optimization from robust background detection," in *CVPR*, 2014.

[14] R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, "Frequency-tuned salient region detection," in *CVPR*, 2009.

[15] V. Movahedi and J. Elder, "Design and perceptual validation of performance measures for salient object segmentation," in *IEEE Computer Society Workshop on Perceptual Organization in Computer Vision*, 2010.

[16] Q. Yan et al, "Hierarchical saliency detection," in *CVPR*, 2013.

[17] M. Cheng, G. Zhang, N. Mitra, X. Huang, and S. Hu, "Global contrast based salient region detection," in *CVPR*, 2011.

[18] M. Cheng and J. Warrell et al, "Efficient salient region detection with soft image abstraction," in *ICCV*, 2013.

[19] F. Perazzi and P. Krahenbul et al, "Saliency filters: Contrast based filtering for salient region detection," in *CVPR*, 2012.

[20] R. Margolin, L. Zelnik-Manor, and A. Tal, "How to evaluate foreground maps," in *CVPR*, 2014.

[21] Y. Wei, F. Wen, W. Zhu, and J. Sun, "Geodesic saliency using background priors," in *ECCV*, 2012.

[22] R. Margolin et al, "What makes a patch distinct," in *CVPR*, 2013.

[23] H. Jiang et al, "Salient object detection: A discriminative regional feature integration approach," in *CVPR*, 2013.