

Saliency Detection via Background Seeds by Object Proposals

Muwei Jian^{*,†}, Runxia Zhao[†], Junyu Dong[†], and Kin-Man Lam^{††}

^{*}School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan, 250014, China

E-mail: jianmuwei@163.com Tel: +86-531-88596228

[†]Department of Computer Science and Technology, Ocean University of China, Qingdao, China

^{††}Centre for Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong

E-mail: enkmlam@polyu.edu.hk Tel: +852-27662607

Abstract—In the recent research of saliency detection, many graph-based algorithms are applied, which use the border of an image as a background query. This frequently leads to undesired errors and retrieval outputs when the boundaries of the salient objects concerned touch, or connect with, the image's border. In this paper, we propose a novel bottom-up saliency-detection algorithm to tackle and overcome the above issue. First, we utilize object proposals to collect the background seeds reliably. Then, the Extended Random Walk algorithm is adopted to propagate the prior background labels to the rest of the pixels in an image. Finally, we refine the saliency map by taking both the textural and structural information into consideration simultaneously. Experiments on publicly available data sets show that our proposed approach achieves competitive results against the state-of-the-art methods.

I. INTRODUCTION

The aim of saliency detection is to identify the most attractive object or the most important regions in an image. As one of the most fundamental problems in computer vision, saliency detection has been applied to numerous computer-vision tasks, such as image segmentation [1], object recognition [2, 40, 41], image compression [3], traffic congestion analysis [35, 42] and content-based image retrieval [4, 24].

Generally, the development of saliency detection can be categorized into top-down (supervised) [5, 6] and bottom-up [7-17, 40, 43] (unsupervised) models. The former saliency-detection framework is often viewed as a task-driven or application-oriented process, which is influenced by the prior knowledge constructed from the training images. In contrast, the bottom-up models are mainly involuntary, biologically plausible, and visual cue-driven.

As one of the earliest bottom-up models in the computer-vision domain, Itti et al. [7] pointed out that the human visual system (HVS) is sensitive to meaningful regions with higher contrast, thus they designed a series of different low-level features, including intensity, color and direction for saliency detection. Cheng et al. [8] integrated a histogram-based contrast (HC) framework into spatial information to generate a region-based contrast (RC) saliency map for detecting salient regions in an image. In [9], Murray et al. explored an efficient model of color appearance, which mimics human-vision perception, to obtain a saliency map. Based on

integrating features nonlinearly using region covariances, Erdem et al. [10] proposed a method to employ covariance matrices of image features for saliency detection. In [11], Shi et al. proposed a computational model, such that the final saliency values are inferred by combining all the saliency cues in different scales, using a hierarchical inference algorithm. Later, graph-based frameworks were proposed to achieve the goal of saliency detection [12, 15-17]. For example, Graph-Based Visual Saliency (GBVS) employed a bottom-up method for saliency detection, by normalizing and integrating feature maps to highlight the salient regions in an image [15]. Then, based on a discriminative propagation mechanism, a discriminative similarity metric was proposed to separate similar regions for saliency detection in [12]. Furthermore, more models have been proposed to improve saliency-detection performance [13, 14, 37-39]. By constructing a background-based map using color and space contrast, a Hierarchical Cellular Automata (HCA) model is devised to detect salient objects in [13]. In [14], Huang et al. formulated the saliency-detection problem as a multiple instance learning (MIL) issue. In this method, the superpixels contained in each proposal are as the bags of instances of MIL. In [37], Yan et al. proposed an effective sparsity pursuit-based method for visual saliency detection, in which the learned overcomplete sparse bases were employed to compute saliency values in an image. Later, a decomposition model was developed to increase the precision of the saliency-detection task, by analyzing high-dimensional tensor-like visual data [38]. Jian et al. [39] designed a saliency-detection method by integrating Quaternionic Distance Based Weber Descriptor (QDWD) with pattern distinctness and local contrast for underwater salient-object detection.

Among the numerous saliency-detection models, the graph-based approach is an effective solution to achieve high-quality salient maps. Yang et al. [16] proposed a saliency-detection method via graph-based manifold ranking, in which salient maps can be obtained by estimating the similarity of the image elements (i.e. pixels or regions) with foreground and background cues. By the aid of a graph-based foreground representation of an image, a spectral graph-based salient object detection model with a predetermined connectivity rule was proposed in [17]. Therefore, graph-based algorithms have

been developed as an important research branch in saliency detection. However, the backgrounds and foregrounds may not be adequately labeled with the abovementioned measures, which may lead to unexpected results during propagation.

To overcome and solve the abovementioned issues, we select potential region proposals of input images to compute the background seeds. Different from traditional methods, our proposed method can coarsely locate the object and obtain accurate background seeds, according to the object-level prior knowledge. Then, the background seeds can be propagated with the Extended Random Walk (ERW) algorithm more reliably than the previous approaches. Finally, the texture and global structure of the input image are also considered in order to produce smooth and desirable saliency maps.

The rest of this paper is organized as follows. In Section II, we introduce our proposed framework for saliency detection in detail. Section III illustrates the experiment results on publicly available data sets and presents comparisons with the state-of-the-art methods. Finally, conclusions and discussions are given in Section IV.

II. THE PROPOSED FRAMEWORK FOR SALIENCY DETECTION

In this section, we describe our saliency-detection framework. Firstly, we compute the background seeds based on selected object proposals. Then, a coarse map is obtained by propagating the saliency information to the rest of the image through the ERW algorithm. At last, we refine the generated map by diffusing saliency values among superpixels.

A. Selecting object proposals based on background seeds

Object proposals are the candidate regions of an object in a whole image. Proposals can be considered as rough instructions where an object occurs in an image. Therefore, they can be used to facilitate salient-object detection. Given an image I , we firstly segment it into a fixed number of superpixels by the Simple Linear Iterative Clustering (SLIC) method [18]. Then, we use the Edge Boxes [19] method to generate a series of object proposals. During the generation process, thousands of object proposals can be produced. However, most of the proposals actually cannot locate the object well, and do not even contain any objects, as shown in Fig. 1(c). Therefore, in this paper, we firstly focus on how to choose the reasonable object proposals as background seeds for saliency detection.

In the beginning stage, we aim to select some proposals, in terms of the objectness score [20, 21]. A proposal with a higher score is more likely to contain an object, as shown in Fig. 1(b). Therefore, in our algorithm, the N proposals with the highest scores are chosen to form a target set A . Then, we delete those bounding boxes, which contain mostly background regions, as shown in Fig. 1(c). Inspired by [14, 22], we use the boundary connectivity, based on the superpixels, to further refine the proposals. As in [23], we use $BndCon(sp_i)$ to quantify how close the i^{th} superpixel sp_i of

each bounding box is connected to the image boundaries, which is defined as follows:

$$BndCon(sp_i) = \frac{Len_{bnd}(sp_i)}{\sqrt{Area(sp_i)}}, \quad (1)$$

where $Len_{bnd}(sp_i) = |\{(sp_i \in bnd) \cap (sp_i \in R)\}|$ is the number of superpixels in the current proposal R , adjacent to the boundary, bnd , and $Area(sp_i)$ represents the number of superpixels in the corresponding proposal R . The larger the value of $BndCon(sp_i)$ is, the greater the number of superpixels are on the boundary. Therefore, a large value for $BndCon(sp_i)$ means that most parts of the proposal may belong to the image background, and this kind of proposal should be removed. Let $A^- = \{b_1, \dots, b_i, \dots, b_n\}$ be the set of target proposals, where b_i is the i^{th} proposal and n is the number of proposals available. Suppose that sp_i is removed, then the set of proposals is updated as $A^- = \{b_1, \dots, b_{i-1}, b_{i+1}, \dots, b_n\}$. The set now contains $n-1$ proposals and the proposals are relabeled, so the final set $A^- = \{b_1, \dots, b_{i-1}, b_i, b_{i+1}, \dots, b_{n-1}\}$.

In order to obtain the final background seeds, we calculate the union of the selective proposals A^- based on the objectness score. The target score $O(sp_i)$ of the superpixel sp_i is defined as follows:

$$O(sp_i) = \sum_{j=0}^n o(b_j) \cdot \delta(sp_i \in b_j), \quad (2)$$

where $o(b_j)$ represents the target score of the proposal b_j . If sp_i belongs to b_j , $\delta(sp_i \in b_j)$ is 1, otherwise $\delta(sp_i \in b_j)$ is 0. Then, a score map is computed based on $O(sp_i)$ of all the superpixels in A^- . Finally, we set a threshold φ to divide the score map, in order to obtain the background seeds, as follows:

$$\Phi(sp_i) = \begin{cases} 0, & \text{if } O(sp_i) \leq \varphi, \\ O(sp_i), & \text{otherwise,} \end{cases} \quad (3)$$

where φ is set directly proportional to the average score of selective proposals.

Fig. 1(e) shows the final results of the proposed method, in which the background region is masked in black; while Fig. 1(d) illustrates the results produced by [33], in which the background seeds from image-border superpixels are collected by using boundary information and clustering. In comparison with our proposed scheme, taking Fig. 1(d) as an example, a part of the tiger's tail is mistakenly chosen as the background seed. Furthermore, the method of Fig. 1(d) only coarsely finds the border of an image as background seeds. As a result, this usually leads to unexpected mistakes, especially when the salient object is close to or is connected with the border of an image. Compared to the method [33] with results shown in Fig. 1(d), our background seed-selection algorithm achieves better performance, in terms of quality. Our method can generate more precise and compact background seeds, while maintaining the completeness of salient objects in the foreground. In the subsequent stage for saliency detection, we take the black regions, shown in Fig. 1(e), as the background seeds in the ERW algorithm.

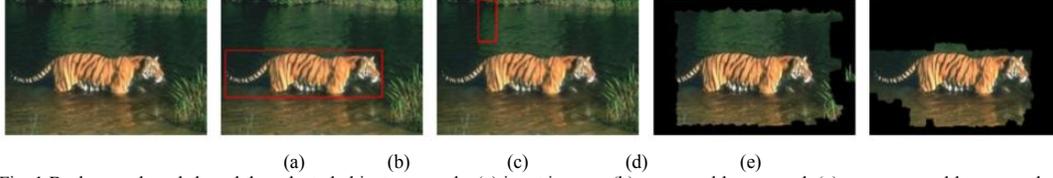


Fig. 1 Background seeds based on the selected object proposals: (a) input images, (b) a reasonable proposal, (c) an unreasonable proposal, (d) background seeds generated by [33], and (e) our proposed background seeds for saliency detection.

B. Saliency map formulation based on ERW ranking

The background seeds generated by our proposed scheme, which is based on selecting optimal object proposals, can provide accurate initial salient positions for the saliency-estimation process. In the proposed framework, the background seeds are propagated based on a constructed undirected graph. Classic propagation models, in general, can be categorized into the following major approaches: conditional random field models, Markov models, and manifold ranking models. In order to produce more accurate saliency maps, we experimentally employ the ERW algorithm [24] for propagation in our saliency-detection model. Compared with other traditional propagation methods, ERW can efficiently propagate saliency information to the rest of the image regions during saliency detection with high accuracy.

In the following sub-section, we will briefly introduce ERW, which is employed in our saliency-detection method. As a kind of random walk method [26][27][28], ERW can be constructed via a graph $G = (V, E)$, where the node set $V = \{sp_1, sp_2, \dots, sp_m\}$ contains a set of superpixels and m represents the number of superpixels. E represents all the connections of any two nodes in G , quantified by a weight matrix $W = [w_{ij}]_{m \times m}$, where the weight between two adjacent nodes is defined as:

$$w_{ij} = \exp\left(-\frac{\|c_{sp_i} - c_{sp_j}\|^2}{\sigma^2}\right), \quad (4)$$

where c_{sp_i} denotes the mean color feature of the node sp_i , and σ is a controlling constant factor. The degree matrix $D = \text{diag}\{d_1, \dots, d_m\}$ is generated by $d_i = \sum_j w_{ij}$. We further define the Laplacian matrix L as $L = D - W$. Then, we use ERW to propagate the saliency information by minimizing the following energy function:

$$\begin{aligned} & \arg \min_f \frac{1}{2} \sum_{i,j} w_{ij} (f_i - f_j)^2 + \\ & \alpha \sum_{i=1}^m (d_i f_i - \sum_j w_{ij} f_j)^2 + \beta \sum_{i=1}^m (f_i - y_i)^2, \text{ s.t. } f_i = 1, \end{aligned} \quad (5)$$

where f_i denotes the label of node sp_i and $f_i = 1$ means sp_i is a background seed, otherwise $f_i = 0$; and α, β are trade-off parameters. The first term is the traditional random walk formulation, while the second term is an additional Laplacian term in order to propagate the influence to more distant nodes. In the third term, y_i denotes the output of an external classifier. The minimum solution can be computed by setting

the derivative of (5) to zero. Denote B as the set of background seeds, and u as the set of unlabeled nodes. Then, the solution can be written as follows:

$$\mathbf{f}_u = \mathbf{P}_{uu}^{-1}(-\mathbf{P}_{ub}\mathbf{f}_B + \beta\mathbf{y}_u), \quad (6)$$

where $\mathbf{P} = \mathbf{L} + \alpha\mathbf{L}^2 + \beta\mathbf{I}$ and \mathbf{I} is an identity matrix. \mathbf{f}_B represents the label of background seeds with $\mathbf{f}_B = 1$. Through equation (6), we can compute the labels \mathbf{f}_u for other nodes accordingly, and we use S_{erw} to represent the label results.

C. Refinement based on texture and structure information

In our method, the saliency map S_{erw} of an image is estimated based on superpixels and ERW propagation. However, the saliency map may contain some unexpected noises and inaccurate predictions of saliency. For example, two superpixels, which resemble and share similar features, may be estimated with different saliency values. To obtain a better pixel-level result, we refine the saliency map S_{erw} by taking both texture and structure information into consideration, by diffusing the saliency values among superpixels. To accomplish this objective, we firstly apply a mid-level clustering algorithm, proposed in [29], to divide the image into larger subregions. The proposed idea is based on the observation that the superpixels in the same subregion may have the same structure information. Assume that there are K subregions with M superpixels in each subregion. Then, we can further refresh the saliency values in the superpixel i by the other superpixels in the subregion k , which is expressed as follows:

$$S_{ki} = aS_{erw_{ki}} + b \sum_{j=1, j \neq i}^M \frac{q_{i,j}}{\sum_{j=1}^M q_{i,j}} S_{erw_{kj}}, \quad (7)$$

where S_{ki} is the i -th superpixel in the k -th subregion, and a and b are the weights for the two terms in (7). The first term S_{erw} is the saliency map generated using ERW propagation in Section 2.2. The second term is the weighted average of the saliency values of the other superpixels in the subregion, according to the texture and structure information. In the same subregion, $q_{i,j}$ is defined as follows:

$$q_{i,j} = \exp\left(-\frac{\|t_i - t_j\|}{2\sigma_t^2}\right) + \gamma, \quad (8)$$

where t_i represents the texture feature of the i^{th} superpixel. If the two superpixels, with indices i and j , are not in the same subregion, $\gamma = 1$, otherwise $\gamma = 0$. In other words, $q_{i,j}$ is only dependent on the texture information.

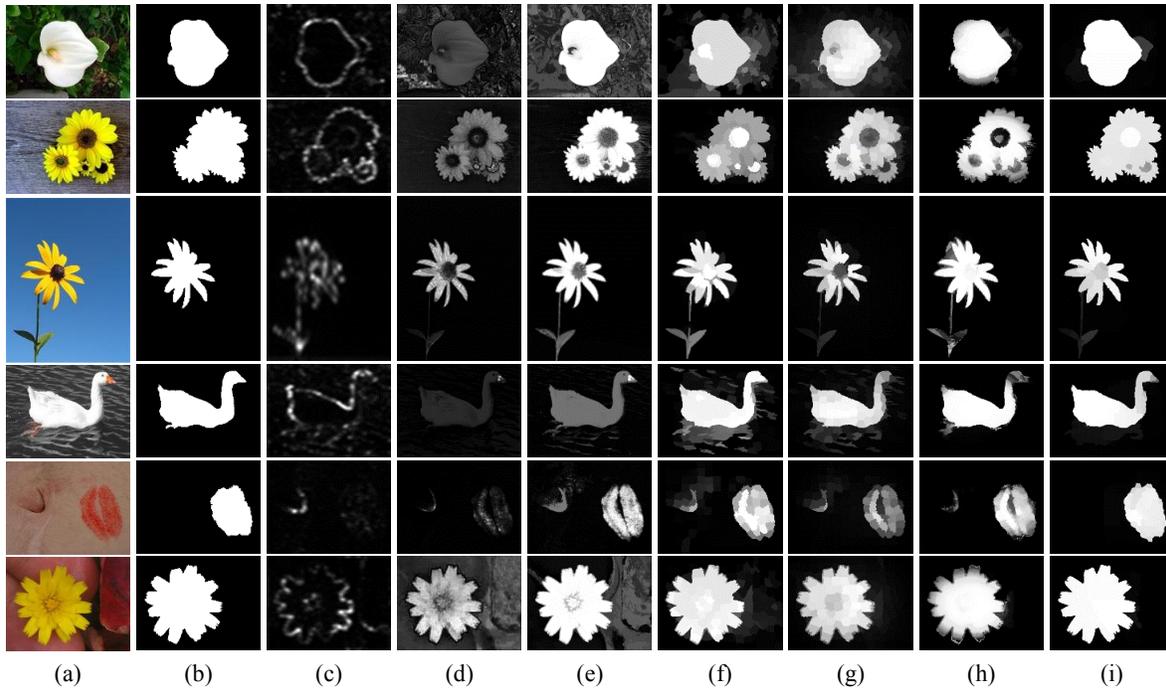


Fig. 2 Visual comparison of saliency maps generated by different methods: (a) Input image, (b) ground truth, (c) SR, (d) FT, (e) HC, (f) GS, (g) BFS, (h) DSR, and (i) Our method.

III. EXPERIMENTS

In the experiments, an image is automatically separated into 200, 300 and 400 superpixels, using the SLIC method [18]. Then, the 50% top-scoring proposals are selected to form the target set. During the saliency information propagation stage with ERW, we set $\sigma^2 = 0.1$ for w_{ij} , and the parameters of the energy function are empirically set to $\alpha = 0.1$ and $\beta = 0.01$. The weighting parameters a and b are set at 0.5. The control parameter for the texture feature σ_t^2 is set at the mean score of the normalized texture.

For quantitative comparison, precision, recall and F -measure are adopted to perform objective evaluation. Similar to the previous works [8, 30], we calculate the precision-recall (PR) curves by comparing the binarized saliency maps with the ground truth. The overall F -measure is a weighted measurement by balancing precision and recall, which is defined as follows:

$$F_{\beta} = \frac{(1+\beta^2)Precision \times Recall}{\beta^2 Precision + Recall}, \quad (9)$$

where β^2 is set to 0.3 to grant more emphasis to precision. In order to verify the effectiveness of the proposed algorithm, extensive experiments for saliency detection are tested to evaluate its performances. The evaluation is carried out on the MSRA-5000 database [36].

We also compare the performance of our proposed algorithm with some representative state-of-the-art algorithms, including FT [30], SR [31], HC [8], GS [31], BFS [33], and DSR [34]. To generate results based on these methods, the

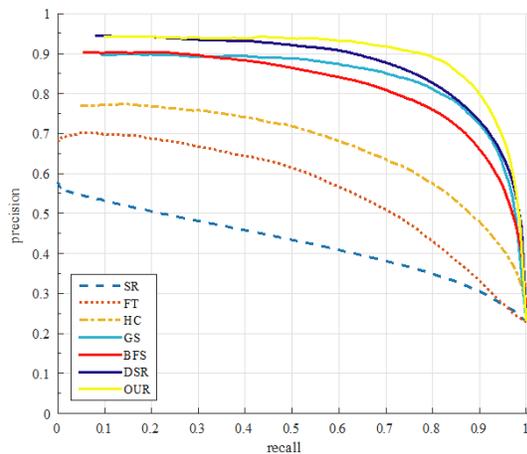
available software or Matlab codes, publicly provided by the authors, are employed. Fig. 2 shows the quantitative comparison results on the MSRA database, in terms of the precision-recall (PR) curves and F -measure. From the PR curves in Fig. 2(a) and the F -measure scores in Fig. 2(b), it can be observed that our proposed method achieves a superior performance to the others. In terms of the overall F -measure, the proposed model gives the most accurate saliency-detection results, compared with other existing algorithms.

IV. CONCLUSIONS

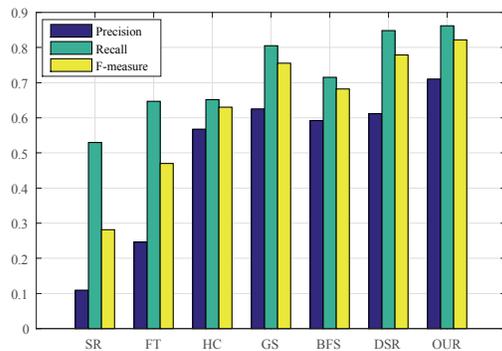
In this paper, we propose a novel saliency-detection model based on Extended Random Walk Ranking and background seeds via selected object proposals. In the proposed model, the texture and global structure in the input image are effectively incorporated simultaneously to refine the saliency maps, which can reliably separate salient regions from image background. Experimental results show that the proposed method outperforms other state-of-the-art saliency-detection algorithms.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (NSFC) (61601427, 61602229, 61771230); Natural Science Foundation of Shandong Province (ZR2016FM40); Shandong Provincial Key Research and Development Program of China (NO. 2017CXGC0701); Fostering Project of Dominant Discipline and Talent Team of Shandong Province Higher Education Institutions.



(a) PR curves



(b) F-measure

Fig. 3 Quantitative comparison results on the MSRA database, in terms of PR curves and F-measure.

REFERENCES

[1] Zhang K., Zhang L., Lam K. M., et al. A level set approach to image segmentation with intensity inhomogeneity, *IEEE transactions on cybernetics*, 2016, 46(2): 546-557.

[2] Rutishauser U., Walther D., Koch C., et al. Is bottom-up attention useful for object recognition? *IEEE CVPR*, Vol. 2, pp. 37-44, 2004.

[3] Itti L., Automatic foveation for video compression using a neurobiological model of visual attention, *IEEE Transactions on Image Processing*, 13(10): 1304-1318, 2004.

[4] M. Jian, Y. Yin, J. Dong, K. M. Lam, Content-based image retrieval via a hierarchical-local- feature extraction scheme, *Multimedia Tools and Applications*, 2018. DOI : 10.1007/s11042-018-6122-2.

[5] Lu S., Mahadevan V., Vasconcelos N., Learning optimal seeds for diffusion-based salient object detection, *IEEE CVPR*, 2014: 2790-2797.

[6] Yang J., Yang M. H., Top-down visual saliency via joint CRF and dictionary learning, *IEEE CVPR*, 2012: 2296-2303.

[7] Itti L., Koch C., Niebur E., A model of saliency-based visual attention for rapid scene analysis, *IEEE Transactions on pattern analysis and machine intelligence*, 1998, 20(11): 1254-1259.

[8] Harel J., Koch C., Perona P., Graph-based visual saliency, *Advances in neural information processing systems*. 2007: 545-552.

[9] Murray N., Vanrell M., Otazu X., et al., Saliency estimation using a non-parametric low-level vision model, *IEEE CVPR*, 2011: 433-440.

[10] Erdem E., Erdem A., Visual saliency estimation by nonlinearly integrating features using region covariances, *Journal of vision*, 2013, 13 (4) :11

[11] Shi J., Yan Q., Xu L., et al. Hierarchical image saliency detection on extended CSSD, *IEEE transactions on pattern analysis and machine intelligence*, 2016, 38(4): 717-729.

[12] Chen S., Zheng L., Hu X., et al. Discriminative saliency propagation with sink points, *Pattern recognition*, 2016, 60: 2-12.

[13] Qin Y., Feng M., Lu H., et al., Saliency detection via cellular automata, *IEEE CVPR*, pp.110-119, 2015.

[14] Huang F., Qi J., Lu H., et al. Salient object detection via multiple instance learning, *IEEE Transactions on Image Processing*, 2017, 26(4): 1911-1922.

[15] Cheng M. M., Mitra N. J., Huang X., et al., Global contrast based salient region detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(3): 569-582.

[16] Yang C., Zhang L., Lu H., et al., Saliency detection via graph-based manifold ranking, *IEEE CVPR*, 2013: 3166-3173.

[17] Çağlar Aytekin, et al., "Learning Graph Affinities for Spectral Graph-based Salient Object Detection." *Pattern Recognition*, 64: 159-167, 2016.

[18] Achanta R., Shaji A., Smith K., et al., SLIC superpixels compared to state-of-the-art superpixel methods, *IEEE transactions on pattern analysis and machine intelligence*, 2012, 34(11): 2274-2282.

[19] Zitnick C. L., Dollár P., Edge boxes: Locating object proposals from edges, *European Conference on Computer Vision*. Springer, 2014: 391-405.

[20] Alexe, Bogdan, Thomas Deselaers, and Vittorio Ferrari. "Measuring the objectness of image windows." *IEEE transactions on pattern analysis and machine intelligence*, 34.11 (2012): 2189-2202.

[21] Cheng M. M., et al. "BING: Binarized normed gradients for objectness estimation at 300fps." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014.

[22] Xi T., Zhao W., Wang H., et al. Salient Object Detection with Spatiotemporal Background Priors for Video, *IEEE Transactions on Image Processing*, 2017, 26(7): 3425-3436.

[23] Zhu W., Liang S., Wei Y., et al., Saliency optimization from robust background detection, *IEEE CVPR*, 2014: 2814-2821.

[24] M. Jian, J. Dong, J. Ma. "Image retrieval using wavelet-based salient regions", *Imaging Science Journal*, The. Vol. 59, No. 4, pp. 219-231, 2011.

[25] Kong Y., Wang L., Liu X., et al., Pattern mining saliency, *European Conference on Computer Vision*. Springer, 2016: 583-598.

[26] Gopalakrishnan V., Hu Y., Rajan D., Random walks on graphs for salient object detection in images, *IEEE Transactions on Image Processing*, 2010, 19(12): 3232-3242.

[27] Grady L., Random walks for image segmentation, *IEEE transactions on pattern analysis and machine intelligence*, 2006, 28(11): 1768-1783.

- [28] Li C., Yuan Y., Cai W., et al. Robust saliency detection via regularized random walks ranking, *IEEE CVPR*, 2015: 2710-2717.
- [29] Kim T. H., Lee K. M., Lee S. U., Learning full pairwise affinities for spectral segmentation, *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(7): 1690-1703.
- [30] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection, *IEEE CVPR*, pp. 1597–1604, 2009.
- [31] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. *IEEE CVPR*, pp. 1–8, 2007.
- [32] Wei Y., Wen F., Zhu W., et al. Geodesic saliency using background priors, *Computer Vision–ECCV 2012*, 2012: 29-42.
- [33] Wang J., Lu H., Li X., et al. Saliency detection via background and foreground seed selection, *Neurocomputing*, 2015, 152: 359-368.
- [34] Li X., Lu H., Zhang L., et al., Saliency detection via dense and sparse reconstruction, *IEEE ICCV*, 2013: 2976-2983.
- [35] Q. Wang, J. Wan, Y. Yuan, "Deep metric learning for crowdedness regression", *IEEE Trans. Circuits System and Video Technology*, DOI: 10.1109/TCSVT.2017.2703920.
- [36] Liu T., Yuan Z., Sun J., et al., Learning to detect a salient object, *IEEE Transactions on Pattern analysis and machine intelligence*, 2011, 33(2): 353-367.
- [37] J. Yan, M. Zhu, H. Liu, et al., "Visual saliency detection via sparsity pursuit," *IEEE Signal Processing Letters*, vol. 17, no. 8, pp. 739-742, 2010.
- [38] Y. Li, J. Yan, Y. Zhou, et al., "Optimum subspace learning and error correction for tensors," *ECCV 2010*, pp. 790-803, 2010.
- [39] M. Jian, Q. Qi, J. Dong, Y. Yin, K. M. Lam, Integrating QDWD with pattern distinctness and local contrast for underwater saliency detection, *Journal of Visual Communication and Image Representation*, Vol. 53, 2018, pp. 31-41.
- [40] M. Jian, K. M. Lam, J. Dong, et al., "Visual-patch-attention-aware Saliency Detection," *IEEE Transactions on Cybernetics*, vol. 45, no. 8, pp. 1575-1586, 2015.
- [41] X. Yan, Y. Wang, Q. Song, K. Dai, Salient object detection via boosting object-level distinctiveness and saliency refinement, *Journal of Visual Communication and Image Representation*, Vol. 48, 2017, pp. 224-237.
- [42] Q. Wang, J. Wan, Y. Yuan, Locality constraint distance metric learning for traffic congestion detection, *Pattern Recognition*, vol. 75, pp. 272-281, 2018.
- [43] M. Jian, et al., Assessment of feature fusion strategies in visual attention mechanism for saliency detection, *Pattern Recognition Letters*, <https://doi.org/10.1016/j.patrec.2018.08.022>, 2018.