Reversion Correction and Regularized Random Walk Ranking for Saliency Detection

Yuchen Yuan, *Student Member, IEEE*, Changyang Li, *Member, IEEE*, Jinman Kim, *Member, IEEE*, Weidong Cai, *Member, IEEE*, and David Dagan Feng, *Fellow, IEEE*

***Abstract*— In recent saliency detection research, many graph-based algorithms have applied boundary priors as background queries, which may generate completely "reversed" saliency maps if the salient objects are on the image boundaries. Moreover, these algorithms usually depend heavily on pre-processed superpixel segmentation, which may lead to notable degradation in image detail features. In this paper, a novel saliency detection method is proposed to overcome the above issues. First, we propose a saliency reversion correction (RC) process, which locates and removes the boundary-adjacent foreground superpixels, and thereby increases the accuracy and robustness of the boundary prior based saliency estimations. Second, we propose regularized random walk ranking (RRWR) model, which introduces prior saliency estimation to every pixel in the image by taking both region and pixel image features into account, thus leading to pixel-detailed and superpixel-independent saliency maps. Experiments are conducted on four well-recognized datasets; the results indicate the superiority of our proposed method against 14 state-of-the-art methods, and demonstrate its general extensibility as a saliency optimization algorithm. We further evaluate our method on a new dataset comprised of images that we define as boundary adjacent object saliency (BAOS), on which our method performs better than the comparison methods.**

***Index Terms*— Reversion correction, regularized random walk ranking, saliency optimization, saliency detection.**

I. Introduction

With the rapid increase in smart devices and social networks, we are now immersed in massive amounts of digital media data. Considering the scarcity of attention and time, it is urgent and advantageous to filter out only the most useful messages among all of the available data for further processing. With image data, this equates to the saliency detection process. Saliency detection locates the most conspicuous object (or region) in an image [1], [2], which typically originates from contrasts of objects and their surroundings, such as differences in color, texture, shape, etc. Fig. 1 shows a few examples of natural images, in which the flower, the cookies and the girl attract the most visual attention, and thus are regarded as salient objects. The detections of such salient objects in the images are of vital importance, because they not only improve the subsequent image processing and analyses, but also direct limited computational resources to more efficient solutions [3].

Bottom-up methods [1], [3-8] comprise one of the major branches of saliency detection methods, which focus on low-level image features. Recently, graph-based methods [4], [7-9] have emerged to apply inter-regional relationships in saliency detection, among which the boundary-based feature [6], [10], [11] is one of the highly effective contributors. These methods are generally facilitated by superpixel segmentation and perform well compared with the state-of-the-art. Nevertheless, there still are drawbacks that hinder these methods, with two major issues as described below:

1) It may be implausible to directly apply four image boundaries as the background queries for background saliency detection. More specifically, one or more of the boundaries may contain part of the foreground object, and undesired error may occur if they are still considered background. Examples are shown in Fig. 2, where the salient objects occupy considerable parts of the image boundaries, leading to the failure of the MR [6]. Also note that due to the negative influences of the boundary-adjacent foreground objects, the saliency maps in Fig. 2b look similar to the "reversed" version of the ground truth in Fig. 2d, i.e., most of the background regions are classified as foreground, and most of the foreground regions are classified as background.
2) The superpixel segmentation [12] facilitates the pre-processing of boundary based (and many other graph-based) saliency detection algorithms. However, inaccuracy in the superpixel segmentation itself may directly lead to the failure of the entire algorithm. The operation of assigning the same saliency value to all the pixels within a fix-sized patch unavoidably ignores some detailed information from the original image, making the saliency map appear covered by mosaics and hence decreasing the overall visual quality. It is thus desirable to combine both superpixel and pixel image data in saliency detection where the pixel process can provide better smoothness and hence improve the overall quality and accuracy of the output saliency map.

In this paper, in order to overcome the two issues above, we propose the reversion correction and regularized random walk ranking (RCRR) for saliency detection, a novel graph-based bottom-up saliency detection method. Our key contributions are summarized below:

1) We present the reversion correction (RC) process, which locates and eliminates the boundary-adjacent foreground superpixels, preventing saliency reversions (e.g., Fig. 2b)

Y. Yuan, C. Li, J. Kim, and W. Cai are with the Biomedical & Multimedia Information Technology (BMIT) Research Group, School of Information Technologies, The University of Sydney, Darlington, NSW 2008, Australia. E-mail: {yuchen.yuan, changyang.li, jinman.kim, tom.cai}@sydney.edu.au.

D. D. Feng is with the BMIT Research Group, and also with the Med-X Research Institute, Shanghai Jiaotong University, Shanghai 200030, China. E-mail: dagan.feng@sydney.edu.au.

Fig. 1. Examples of salient objects in natural images. Top row: original images. Bottom row: ground truth.

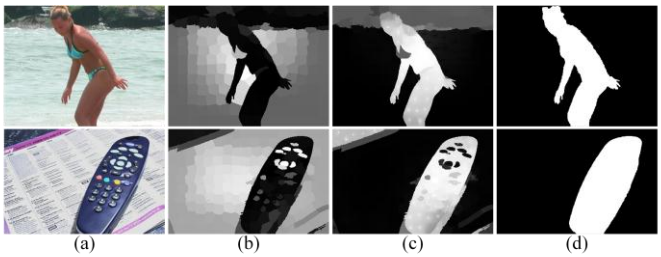


Fig. 2. Examples showing the problem of using boundaries as background queries when the salient objects are boundary-adjacent. (a) input images; (b) results of a boundary-based method MR; (c) results of our proposed method; (d) ground truth. Our method can effectively prevent the saliency reversion

from emerging. This mechanism provides increased robustness, as shown in Fig. 2c.

2) We build the regularized random walk ranking (RRWR) model, which takes both prior saliency estimations and pixel image data into account. RRWR is independent of superpixel segmentation and is able to generate pixel saliency maps that reflect full details of the input images.
3) We explore the extensibility of RC as an optimization algorithm on existing boundary based saliency detection methods, which has the potential of significant performance boosting.
4) We also propose the boundary adjacent object saliency (BAOS) dataset comprising 200 images that have large proportions of salient objects on the image boundaries. This dataset provides an objective evaluation of saliency detection methods' performance on boundary-adjacent salient objects.

This work is an extension to our previous study [11] with marked improvements, specifically the technical contributions above. In addition, we have conducted a more detailed and comprehensive evaluation with 14 state-of-the-art methods, including our previous work [23], on five datasets. The results show the superiority of our proposed RCRR method in terms of both accuracy and robustness.

II. Related Works

In this section, we discuss related saliency detection methods and put our work in proper context. In computer vision, based on data processing mechanisms, saliency detection methods are generally categorized into two groups: bottom-up methods [1], [3-6], [13-15], and top-down methods [7], [8], [16], [17]. The bottom-up methods are often involved in feature properties directly presented in the environment [18]. They simulate instinctive human visual mechanisms, and utilize self-contained features such as color, texture, location, shape, etc. Compared to top-down methods, the bottom-up methods are more attentive and data driven [6], being easier to adapt to various cases, and therefore have been widely applied.

Early bottom-up saliency detection methods are generally feature based, in which only low-level features in the images are used. This includes, but is not limited to, color intensity, color contrast, color distribution, and Euclidean distances. As the pioneers in saliency detection, Itti et al. [1] introduce a conceptually computational model for visual attention detection based on multiple biological feature maps generated by mimicking human visual cortex neurons. Achanta et al. [13] propose a frequency-tuned method based on the low-level features of luminance and color, which is fast to implement and results in pixel-wise saliency maps. In [3], Cheng et al. present a histogram-based contrast method, which exploits pixel-wise color separation to produce saliency maps, which is then improved with a region-based contrast method that takes spatial distances into account at the cost of computational efficiency. To overcome the limitations of color contrast, Fu et al. [14] illustrate the workflow of a combined color contrast and color distribution saliency detection algorithm, together with a refinement process to suppress noise and artifacts. In the work of Hou and Zhang [19], the log-spectrum of the input image is analyzed to extract its spectral residue in the spectral domain, which is then applied to construct the saliency map in the spatial domain.

Feature-based methods mostly focus only on the low-level features and therefore ignore the inter-regional relationships, which are composed of rich information stored in the connections between the different pixels and regions in an image. Such problems are resolved in the more recently proposed graph-based methods. Harel et al. [15] offer graph based visual saliency (GBVS), a graph-based bottom-up saliency model with dissimilarity measurements to predict human fixations; it first forms an activation map of certain features and then conducts normalization to highlight the region of interest. In the work of Gopalakrishnan et al. [4], the Markov random walk model is performed on both the complete graph and the sparse *k*-regular graph to detect seed nodes, which are then used in the pop-out graph model to effectively detect the approximate location of the most salient object in an image. The work of Mai et al. [5] explores the application of data aggregation in saliency detection; more specifically, it introduces saliency map aggregation by pixel-wise aggregation, conditional random field (CRF) aggregation and image-dependent aggregation. Jiang et al. [20] introduce discriminative regional feature integration (DRFI), which integrates regional contrasts, regional properties, and background descriptors together to formulate the master saliency map. With graph based manifold ranking (MR), the work of Yang et al. [6] first utilizes the four boundaries of the input image as background queries to obtain background saliency and then applies the background saliency estimation to extract foreground queries for the final saliency map

formulation.

Among the graph-based methods, the boundary-based feature [6], [10], [11] is one of the most effective features for saliency extraction. In this paper, we propose a novel saliency detection method based on refined boundary-based features, which incorporates both prior saliency knowledge and correction of the saliency reversion problem found in [6]. Extensive experiments have shown the high performance of our proposed method in both accuracy and robustness.

III. Preliminaries

In this section, we provide a brief review of *k*-means clustering, manifold ranking, and random walks, which are related to our proposed method. Unless otherwise specified, the given dataset for processing is noted as $\chi = \{x_1, ..., x_n\}$, where n is the element number.

A. k-Means Clustering

The *k*-means clustering partitions the elements in χ into K clusters $S = \{S_1, S_2, ..., S_K\}$ on the condition that the within-cluster sum of squared error is minimized:

$$S = \arg\min_{S} \left(\sum_{k=1}^{K} \sum_{x \in S_k} \| x - m_k \|^2 \right), \tag{1}$$

where m_k is the mean of observations in S_k.

In the proposed algorithm, given its efficiency, robustness and accuracy, the *k*-means clustering is used to group the initial saliency estimation result into foreground / background clusters. The boundary-adjacent foreground superpixels are then recognized and removed. Detailed steps are presented in Section IV.A.

B. Manifold Ranking

Manifold ranking assigns ranks to elements in χ with respect to their intrinsic manifold structure. The first s elements of χ are labeled as queries, while the rest are unknowns. Such identification is recorded in an indication vector $y = [y_1, ..., y_n]^T$, where $y_i = 1$ if x_i belongs to the queries, and $y_i = 0$ otherwise. A graph $G = (V, E)$ with nodes V and edges E is subsequently established, where V corresponds to the dataset χ and E collects the connections between any two nodes in G quantified by a weight matrix:

$$w_{ij} = \exp\left(-\frac{\| c_i - c_j \|^2}{\sigma^2} \right), \tag{2}$$

where c_i and c_j are the mean CIELab colors of the two nodes i and j and σ is a controlling constant. The degree matrix $D = diag(d_1, ..., d_n)$ is generated afterwards, where

$$d_i = \sum\nolimits_j w_{ij}. \tag{3}$$

Let $f : \chi \to \mathbb{R}^n$ be the ranking function assigning ranks $f = [f_1, ..., f_n]^T$ to χ, which is obtained by solving the following minimization problem defined in [21]:

$$f^* = \arg\min_{f} \frac{1}{2} \left(\sum_{i,j=1}^{n} w_{ij} \left\| \frac{f_i}{\sqrt{d_i}} - \frac{f_j}{\sqrt{d_j}} \right\|^2 + \mu \sum_{i=1}^{n} \| f_i - y_i \|^2 \right), \tag{4}$$

where μ is a balance parameter. The optimized solution is given in [6], [22], [23] as:

$$f^* = (D - \alpha W)^{-1} y, \tag{5}$$

where $\alpha = 1/(1+\mu)$.

The manifold ranking model is applied to estimate the approximate saliency in Section IV.A.

C. Random Walks

The purpose of random walks is to group the elements of χ into different classes. We first label m elements from χ as the seed nodes with at least one element of each class. Without loss of generality, we assume that the first m elements of χ are the seeds, so that $\chi = [x_M^T, x_U^T]$, in which x_M are the marked seed nodes and x_U are the unmarked nodes. The graph $G = (V, E)$, weight matrix W, and degree matrix D are constructed similarly to those in Section III.B. We further define the $n \times n$ Laplacian matrix L as:

$$L_{uv} = \begin{cases} d_u & \text{if } u = v, \\ -w_{uv} & \text{if } x_u \text{ and } x_v \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Then, let $p^l = \left[p_1^l, ..., p_n^l \right]^T$ denote the probability vector of χ for class l, which is similarly decomposed as $p^l = \left[\left(p_M^l \right)^T, \left(p_U^l \right)^T \right]$. Here, p_M^l are the seed nodes that have a fixed value 0 or 1. The optimal p^l is achieved by minimizing the Dirichlet integral [24]:

$$\begin{aligned} Dir\left[p^l \right] &= \frac{1}{2} \left(p^l \right)^T L \left(p^l \right) \\ &= \frac{1}{2} \left[\left(p_M^l \right)^T \left(p_U^l \right)^T \right] \begin{bmatrix} L_M & B \\ B^T & L_U \end{bmatrix} \begin{bmatrix} p_M^l \\ p_U^l \end{bmatrix}. \end{aligned} \tag{7}$$

We differentiate $Dir[p^l]$ with respect to p_U^l, and the critical point is found as:

$$p_U^l = -L_U^{-1} B^T p_M^l. \tag{8}$$

In Section IV.B, the random walk model is reformulated as the proposed RRWR for better computation of the final saliency map.

IV. Proposed Algorithm

Our saliency detection algorithm (RCRR) consists of two major steps. The first step comprises the saliency reversion correction (RC) on an initial saliency estimation, which eliminates boundary-adjacent foreground regions from image boundaries; the second step extracts seed references from step one and calculates the pixel-wise saliency map with the proposed regularized random walk ranking (RRWR).

A. Saliency Reversion Correction

As stated in the introduction, it is possible that the

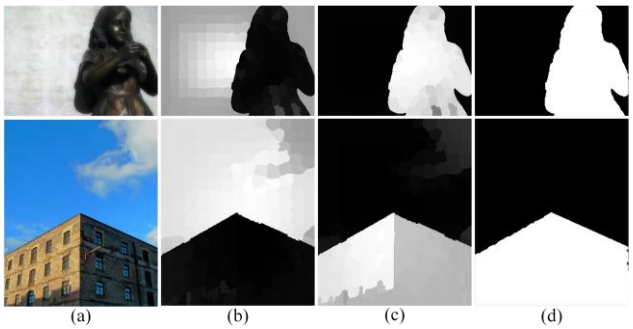


Fig. 3. Examples of RC. (a) input images; (b) saliency estimations without RC; (c) saliency estimations with RC; (d) ground truth. The RC can effectively deal with the saliency reversion problem due to the boundary adjacent objects.

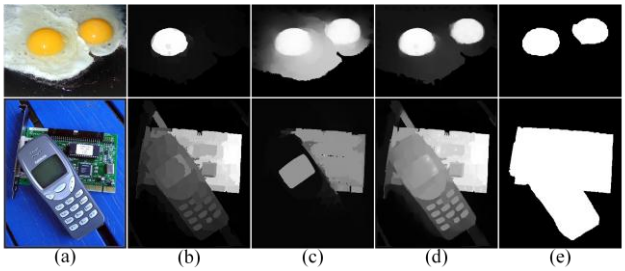


Fig. 4. Examples of RRWR. (a) input images; (b) initial saliency estimations; (c) saliency outputs with classical random walks; (d) saliency outputs with RRWR; (e) ground truth. RRWR is able to further refine the initial saliency estimations; its outputs remarkably outperform the classical random walks.

foreground object is on one or more boundaries of the input image. Using such problematic boundaries as queries in the saliency estimation may lead to undesirable results. Typical examples are illustrated in Fig. 3b, where due to the negative influences of the boundary-adjacent foreground superpixels, the corresponding saliency maps are nearly "reversed" in comparison with the ground truth. To address this issue, it is tempting to directly conduct classification among all the boundary superpixels; however, such classification with the boundary information alone may be too subjective without the global context. We thus propose reversion correction (RC), which functions as a posterior classification based on the initial saliency estimation. The boundary-adjacent foreground regions will then be detected and removed, improving the overall robustness of our algorithm.

For an input image, we first obtain an initial saliency estimation, which can be generated by any boundary-based saliency detection method (e.g., [6], [10], [25]). The graph-based manifold ranking [6] is used in our method due to its relatively high performance and efficiency. From (5):

$$S_{init}(i) = f^*(i), i = 1,...,n, \tag{9}$$

where n is the number of superpixels in the image. S_{init} is then partitioned into background/foreground superpixels by k-means clustering with Lloyd's algorithm [26]:

1) Two uniformly distributed means of S_{init} are generated as $m_k = \frac{k}{3}\max(S_{init}), k = 1,2$;
2) Associate each element of S_{init} with one of the two clusters with the closest mean m_k;
3) Each m_k is then replaced by the mean saliency value of all the elements just assigned to the corresponding cluster;
4) Repeat steps 2) and 3) until a convergence of the two clusters or a desired number of iteration is reached. The labeling map L_{kmeans} is obtained afterwards.

In L_{kmeans}, background superpixels are labeled with 1, and foreground superpixels are labeled with 2. The next step is to recognize if S_{init} is "reversed". Empirically there are less (or no) foreground superpixels on the boundaries of most normal saliency maps; if the majority (or all) of the boundaries of a saliency map are covered with foreground superpixels, we may confidently determine it as "reversed". Therefore, we calculate the average label L_b of all the boundary-adjacent superpixels in L_{kmeans}; if L_b is greater than a pre-defined threshold $t_{reverse}$, we will treat S_{init} as reversed.

Algorithm 1 Saliency Reversion Correction

Input: Initial saliency estimation S_{init}, threshold $t_{reverse}$.

1: Calculate L_{kmeans}. The background and foreground superpixels are labeled with 1 and 2, respectively.

2: Calculate the average boundary label L_b.

3: If $L_b \geq t_{reverse}$, locate and remove the boundary superpixels of S_{init} with label 2 on L_{kmeans}; Repeat the initial saliency estimation with refined boundary to obtain the updated result S_{RC}.

4: If $L_b < t_{reverse}$, directly output $S_{RC} = S_{init}$.

Output: The saliency estimation after RC S_{RC}.

If S_{init} is determined to be reversed, we will find and remove all of the boundary-adjacent superpixels under the guidance of L_{kmeans}, and the initial saliency estimation step is re-performed with the newly formed boundary queries; on the other hand, if S_{init} is not reversed, nothing will be done. The workflow of RC is summarized in Algorithm 1.

The major advantage of RC is that it directly counters the source of the saliency reversion, i.e., the boundary-adjacent foreground superpixels. By locating and eliminating the boundary adjacent foreground superpixels, we can neutralize their negative influences, reversing the "reversed" saliency map back to normal (as shown in Fig. 3c). Nothing will be done if the initial saliency estimation is detected as normal, thus ensuring that RC will not introduce error to good results.

B. Regularized Random Walk Ranking

As introduced in Section II, many state-of-the-art saliency detection algorithms (e.g., [6], [10], [27]) completely depend on the pre-processed superpixel segmentation, which may generate undesirable results if the superpixel segmentation itself is imprecise. Besides that, assigning the same saliency value to all pixels within a superpixel sacrifices the detailed information from the original image. To overcome these drawbacks, we develop the regularized random walk ranking (RRWR) model, which is independent of the superpixel

Algorithm 2 Saliency Detection with Reversion Correction and Regularized Random Walk Ranking (RCRR)

Input: An image and related parameters.

1: Establish superpixel graph; calculate W and D.

2: Conduct initial saliency estimation and obtain S_{init}.

3: Conduct RC in Algorithm 1 and obtain S_{RC}.

4: Compute the pixel-wise saliency p^l with (13).

5: Set $l=2$ and reshape p^2 into S_{final}.

Output: A saliency map with the same size of the input image.

segmentation and can reveal an accurate pixel-wise saliency map of the input image.

RRWR is initially proposed in our previous study [11], which is based on (7), but we suggest a new fitting constraint, which restricts the Dirichlet integral to be as close to the prior saliency distribution as possible:

$$Dir\left[p^l\right]=\frac{1}{2}\left(p^l\right)^T L\left(p^l\right)+\frac{\eta}{2}(p^l-Y)^T(p^l-Y), \tag{10}$$

where the second term is the newly added fitting constraint, η is a controlling parameter, similar to the μ used in (4), and Y is a pixel-wise indication vector inheriting the values of S_{RC} from Section IV.A. Note that RRWR is computed pixel-wisely; hence, both p^l and Y are $N\times 1$ vectors, and L is $N\times N$ matrix, where N is the total pixel number in the image. We define two thresholds, t_{high} and t_{low}:

$$\begin{aligned} t_{high} &= \frac{\text{mean}(S_{RC})+\max(S_{RC})}{2}, \\ t_{low} &= \text{mean}(S_{RC}), \end{aligned} \tag{11}$$

which are used to select pixels with $Y_u > t_{high}$ as foreground seeds, and $Y_u < t_{low}$ as background seeds. The seeds are then combined into p_M^l, $l=1,2$ in Section III.C, where $l=1$ corresponds to the background label, and $l=2$ corresponds to the foreground label. The matrix decomposition of (10) is as follows:

$$\begin{aligned} Dir\left[p^l\right] &= \frac{1}{2}\left[\left(p_M^l\right)^T \left(p_U^l\right)^T\right]\begin{bmatrix} L_M & B \\ B^T & L_U \end{bmatrix}\begin{bmatrix} p_M^l \\ p_U^l \end{bmatrix} \\ &+\frac{\eta}{2}\left(\begin{bmatrix} p_M^l \\ p_U^l \end{bmatrix}-\begin{bmatrix} Y_M^l \\ Y_U^l \end{bmatrix}\right)^T\left(\begin{bmatrix} p_M^l \\ p_U^l \end{bmatrix}-\begin{bmatrix} Y_M^l \\ Y_U^l \end{bmatrix}\right). \end{aligned} \tag{12}$$

Similar to (8), after setting the differentiation of (12) with respect to p_U^l as zero, the optimal solution is obtained as:

$$p_U^l=\left(L_U+\eta I\right)^{-1}\left(-B^T p_M^l+\eta Y_U^l\right). \tag{13}$$

Then, p_U^l and p_M^l are united as p^l. We set $l=2$ to select the foreground possibility p^2, and reshape it to a matrix S_{final} with same size of input image as the final foreground saliency output.

Since the seeds are automatically generated from the result of the RC, unlike classical random walks [24], no user interaction is required in RRWR. The fitting constraint in (10) provides a prior saliency estimation of all pixels instead of the seed pixels alone, which offers better guidance in calculating the final saliency map. The effect of the fitting constraint is shown in Fig. 4, where RRWR (Fig. 4d) improves the saliency map from the initial saliency estimation (Fig. 4b) and remarkably outperforms classical random walk (Fig. 4c), which uses the first term of (10) alone.

The complete workflow of our proposed method is listed in Algorithm 2.

V. Experimental Results

A. Experiment Setup

Datasets: Our method is evaluated on five datasets, including four publicly available datasets and one newly designed dataset. The four public datasets (based on a recent saliency detection benchmark [28]) are MSRA10K [29], which contains 10,000 randomly chosen images from the MSRA dataset [30]; ECSSD [31], which contains 1,000 complex natural images with diversified patterns; SED [32], which contains 100 images with one salient object and 100 images with two salient objects (200 images in total); and PASCAL-S [33], which is derived from the PASCAL VOC [34] segmentation challenge and contains 850 images with complex backgrounds. We also use the new boundary adjacent object saliency (BAOS) dataset, which is specifically designed to evaluate images where large portions (at least 30%) of their boundaries are covered by foreground object(s). It contains 200 images (selected from MSRA10K, ECSSD, and Microsoft Grabcut [35]). All of the datasets come with human-labeled pixel-wise ground truth.

Evaluation Metrics:

We follow the existing metrics in [28] and use precision-recall curve, F-measure, and mean absolute error (MAE) scores as our evaluation metrics. The terms of precision, recall and F-measure are defined in [36] as:

$$precision=\frac{\sum_{i=1}^{N} G(i)\cdot I\left(S_{final}(i)\ge \text{th}\right)}{\sum_{i=1}^{N} I\left(S_{final}(i)\ge \text{th}\right)}, \tag{14}$$

$$recall=\frac{\sum_{i=1}^{N} G(i)\cdot I\left(S_{final}(i)\ge \text{th}\right)}{\sum_{i=1}^{N} G(i)}, \tag{15}$$

$$F_\beta=\frac{(1+\beta^2)\,precision\cdot recall}{\beta^2\,precision+recall}, \tag{16}$$

where G is the ground truth; $I(\cdot)$ is the indicator function that equals 1 if the condition inside is satisfied, and 0 otherwise; S_{final} is the output saliency map corresponding to Algorithm 2; th is the threshold used to binarize S_{final}; and N is the number of pixels in the image. Precision and recall are usually displayed together as precision-recall curves, which are constructed by binarizing the saliency map with thresholds changing from 0 to 255. The F-measure is adopted as a weighted average between precision and recall. As suggested in [37], the average F-measure of a precision-recall curve is computed as its maximal single-point F-measure. We set

$\beta^2 = 0.3$ to grant more importance to the precision, which is consistent with [28]. When used to evaluate a saliency map, the higher the evaluation metric (precision, recall or f-measure) is, the better is the estimation.

MAE is defined as the mean of the difference between the saliency map and the ground truth:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|S_{final}(i) - G(i)\right|. \quad (17)$$

Different from the previous evaluation metrics, smaller MAE means better estimation.

In addition, to evaluate the statistical significance level of RCRR against a comparison method A, we conduct Student's *t*-test between the two methods. We equally divide the images of a particular dataset into 10 subgroups and compute the evaluation metric (F-measure or MAE) in each group. This enables us to obtain the sample mean and sample standard deviation of RCRR and A, namely, $\overline{X}_{RCRR}$, $\overline{X}_{A}$, $s_{X_{RCRR}}$ and s_{X_A}. The *t*-statistic is then computed as:

$$t = \frac{\overline{X}_{RCRR} - \overline{X}_{A}}{s_{X_{RCRR}X_A} \cdot \sqrt{\frac{2}{10}}}, \quad (18)$$

where

$$s_{X_{RCRR}X_A} = \sqrt{\frac{s_{X_{RCRR}}^2 + s_{X_A}^2}{2}}. \quad (19)$$

We then find the one-sided *p*-value corresponding to *t* with 10-1=9 as the degrees of freedom, since our alternative hypothesis is that the metric from RCRR is significantly larger (F-measure) or lower (MAE) than that of A, but not both. The *p*-value is given together with its corresponding evaluation metric in our experiments.

Experimental Parameters:

To objectively compare our algorithm with other algorithms, we use the same parameter settings as in [6], where the superpixel number is set to $n = 200$ and the two controlling parameters in (2) and (5) are set to $\sigma^2 = 0.1$ and $\mu = 0.01$, respectively. The only new parameter in RC is the average boundary label threshold $t_{reverse}$, which is one of the inputs of Algorithm 1. We empirically set $t_{reverse} = 1.5$, which results in the peak performance in Fig. 5. The only new parameter in RRWR is the controlling parameter η. We empirically set $\eta = 0.01$, which results in the peak performance in Fig. 6.

B. Design Option Analyses

We first examine the contributions of our algorithm, RC and RRWR. The red and blue curves in Fig. 7a show the improvements in the precision-recall curves with the use of RC when compared to the saliency output without RC. Similarly, Fig. 7b shows that the F-measure of our method (0.857) is higher than that without using RC (0.850). After that, we generate the saliency maps without the use of RRWR. As shown by the red and brown curves in Fig. 7a, the complete algorithm is superior to the algorithm without RRWR. Fig. 7b shows that our proposed method achieves a higher F-measure than that without RRWR, in which the values are 0.857 in comparison with 0.848.

Based on the analyses above, both RC and RRWR contribute to improving the overall performance.

C. Comparison with State-of-the-art Methods

RCRR is evaluated on the above five datasets in comparison with 14 state-of-the-art saliency detection methods, namely, CA [38], CB [39], DSR [25], FES [40], FT [41], HS [31], IT [1], LR [42], MC [10], MR [6], RR [11], SF [43], SR [19], and wCtr* [27]. All of the algorithms above are evaluated by the corresponding authors' available software codes. The evaluation is conducted both quantitatively and qualitatively. Note that all of the methods above (including our proposed method) are non-training based. Other methods such as DRFI [20] are excluded because they require additional training data, which will significantly affect performance.

Quantitative Evaluation:

A summary of the complete quantitative evaluation results is presented in Table. 1, and detailed analyses of individual datasets are presented in Fig. 8 to Fig. 12. We first conduct our quantitative evaluation on the MSRA10K dataset, which is large enough to cover most types of natural images. The results are shown in Fig. 8. It is clear that the precision-recall curves of our method exceed all other algorithms as in Fig. 8a, where the highest precision value can reach 0.96. Our method also achieves the best F-measure 0.857 in Fig. 8b and the second best MAE score 0.117 in Fig. 8c. In addition, the *p*-values in Fig. 8b and c also indicate that the advantages of RCRR compared to methods in both F-measure and MAE are statistically significant.

Then, we proceed to the ECSSD dataset, which contains 1,000 images with complicated backgrounds. Again, the precision-recall curves of our method outperform all of the other methods, as shown in Fig. 9a. This observation is further validated in Fig. 9b and c, where our method has the highest F-measure 0.711 and the lowest MAE score 0.224, with statistically significant advantages.

Our method behaves similarly on the SED dataset and the PASCAL-S dataset (Fig. 10 and Fig. 11), where it outperforms most of the comparison methods' precision-recall curves and only marginally worse than wCtr* at some points. Our method, DSR, HS and wCtr* have entangled curves in Fig. 10a and Fig. 11a and have close scores in F-measure and MAE. Nevertheless, our method still achieves the best F-measure (0.663) on PASCAL-S, the second best F-measure (0.811) on SED, and the second best MAE scores (0.150 and 0.212) on SED and PASCAL-S, respectively. The statistical *p*-values of our method on SED and PASCAL-S are not as significant as those on MSRA10K and ECSSD, which match the mixed performances we observed above, but our *p*-values are still under 0.1.

Finally, on the BAOS dataset, Fig. 12 displays absolute advantage of our method, which not only has a significantly higher precision-recall curve in Fig. 12a but also obtains the optimal F-measure (0.742) and MAE (0.296) in Fig. 12b and c,

with statistically significant advantages. The dominance of our method on the BAOS dataset demonstrates its elevated robustness to salient objects on the image boundaries.

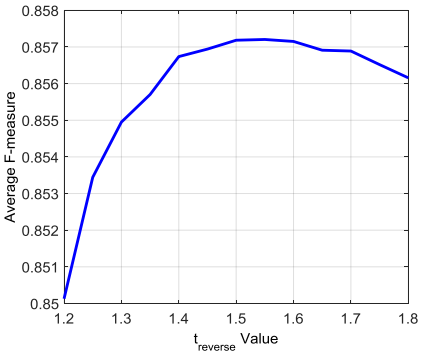


Fig. 5. Average F-measures with different $t_{reverse}$ used in RC on the MSRA10K dataset. The value $t_{reverse}=1.5$, which corresponds to the optimal F-measure, is adopted in our following experiments.

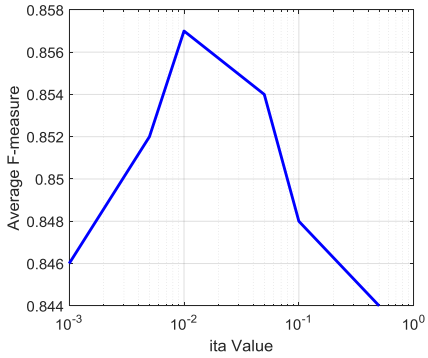
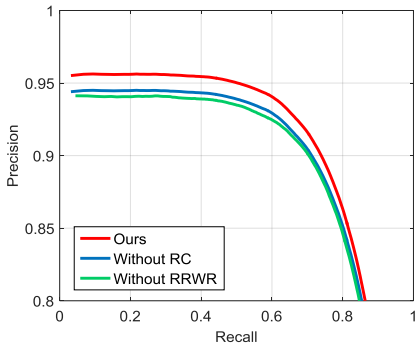
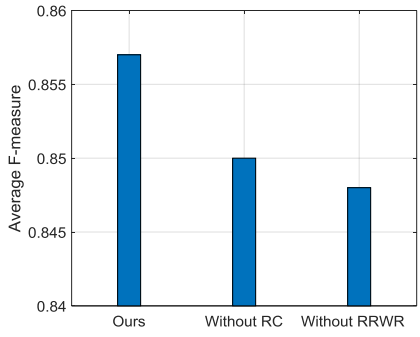


Fig. 6. Average F-measures with different η used in RRWR on the MSRA10K dataset. The value $\eta=0.01$, which corresponds to the optimal F-measure, is adopted in our following experiments.



(a)

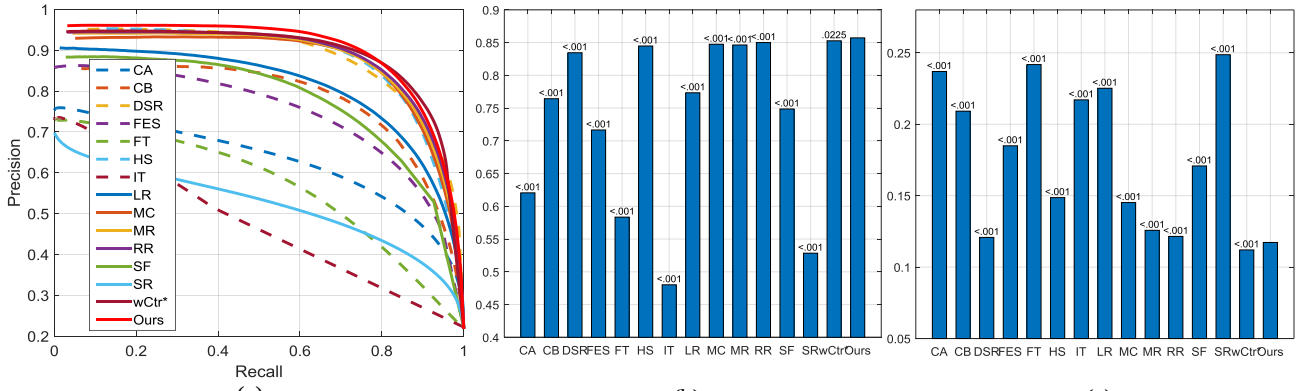


(b)

Fig. 7. Examine results of our method with different design options on the MSRA10K dataset. (a) the precision-recall curves of our method, our method without using RC, and our method without using RRWR; (b) the average F-measures of the three conditions in (a). Using the full configuration leads to the best performance.

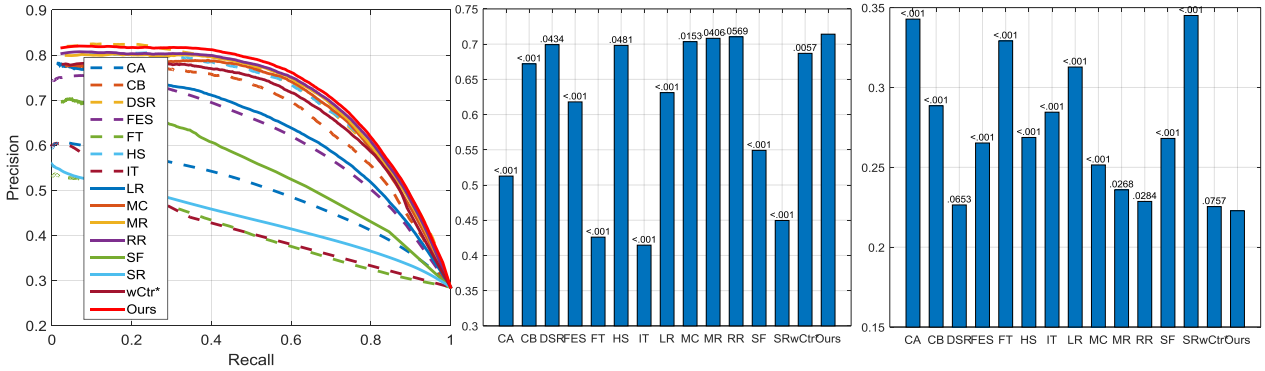
Table. 1. F-measure and MAE evaluation results. The best and second best results are marked in red and blue, respectively.

		CA	CB	DSR	FES	FT	HS	IT	LR	MC	MR	RR	SF	SR	wCtr*	Ours
F-measure	MSRA10K	0.621	0.764	0.834	0.717	0.583	0.845	0.480	0.773	0.847	0.846	0.850	0.749	0.528	**0.853**	**0.857**
	ECSSD	0.513	0.672	0.699	0.618	0.426	0.698	0.415	0.631	0.703	0.708	**0.710**	0.549	0.450	0.687	**0.714**
	SED	0.603	0.693	0.806	0.672	0.605	0.806	0.507	0.720	0.810	0.802	0.806	0.719	0.541	**0.815**	**0.811**
	PASCAL-S	0.496	0.625	0.651	0.624	0.406	0.645	0.421	0.580	0.658	0.612	0.639	0.496	0.454	**0.659**	**0.663**
	BAOS	0.605	0.665	0.693	0.613	0.545	0.729	0.483	0.691	0.684	0.711	**0.737**	0.670	0.570	0.724	**0.742**
MAE	MSRA10K	0.237	0.209	0.121	0.185	0.242	0.149	0.217	0.225	0.145	0.126	0.121	0.171	0.249	**0.112**	**0.117**
	ECSSD	0.343	0.289	0.226	0.265	0.329	0.269	0.285	0.313	0.251	0.236	0.229	0.268	0.345	**0.225**	**0.223**
	SED	0.246	0.254	0.151	0.207	0.247	0.179	0.233	0.247	0.172	0.154	0.151	0.202	0.253	**0.147**	**0.150**
	PASCAL-S	0.301	0.286	0.215	0.223	0.316	0.264	0.246	0.288	0.232	0.259	0.232	0.241	0.294	**0.208**	**0.212**
	BAOS	0.392	0.362	0.335	0.389	0.412	**0.303**	0.413	0.365	0.346	0.330	0.306	0.382	0.407	0.330	**0.296**



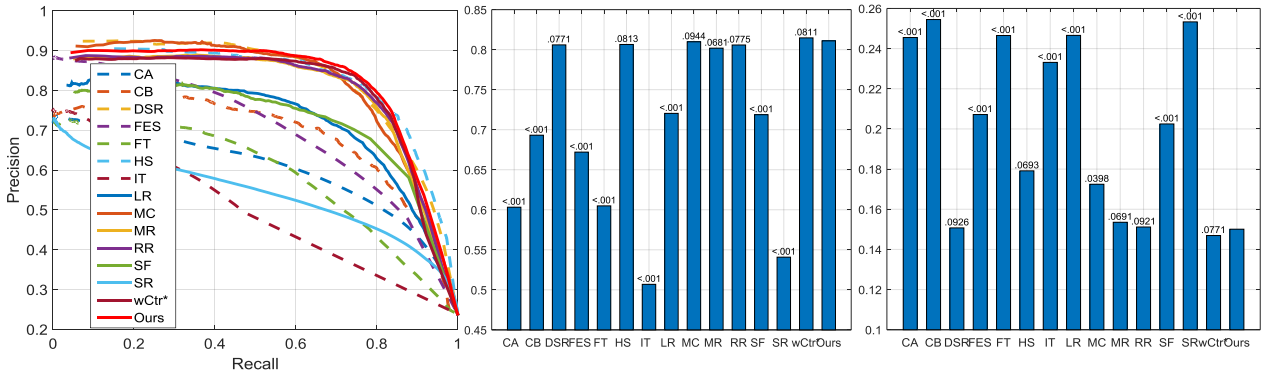
(a) (b) (c)

Fig. 8. Evaluation results on MSRA10K. (a) precision-recall curves; (b) F-measures; (c) MAE scores. Our method performs favorably against comparison methods.



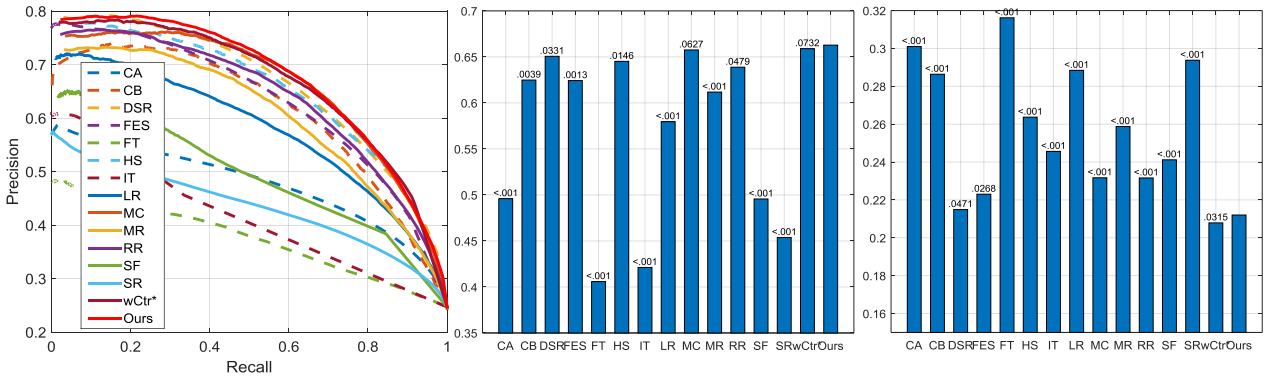
(a) (b) (c)

Fig. 9. Evaluation results on ECSSD. (a) precision-recall curves; (b) F-measures; (c) MAE scores. Our method performs favorably against comparison methods.



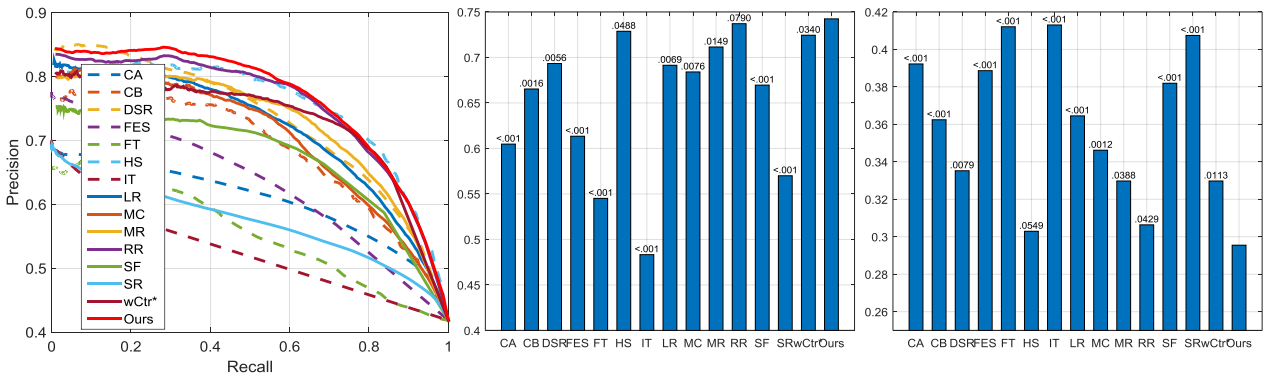
(a) (b) (c)

Fig. 10. Evaluation results on SED. (a) precision-recall curves; (b) F-measures; (c) MAE scores. Our method performs favorably against most comparison methods.



(a) (b) (c)

Fig. 11. Evaluation results on PASCAL-S. (a) precision-recall curves; (b) F-measures; (c) MAE scores. Our method performs favorably against comparison methods.



(a) (b) (c)

Fig. 12. Evaluation results on the newly proposed BAOS dataset. (a) precision-recall curves; (b) F-measures; (c) MAE scores. Our method performs favorably against comparison methods.

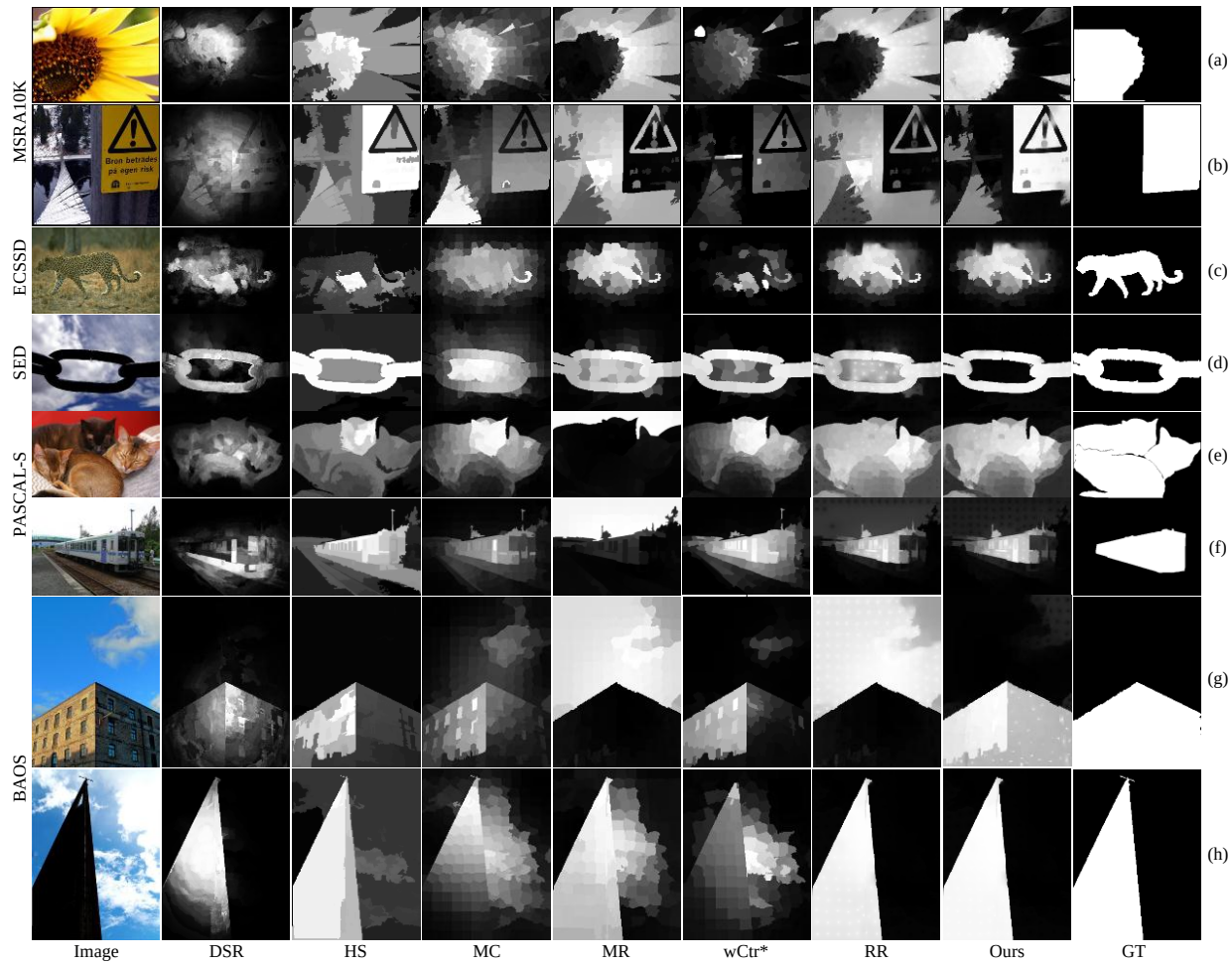


Fig. 13. Visual saliency map samples on the five datasets with different methods. Our method produces more accurate and robust saliency maps.

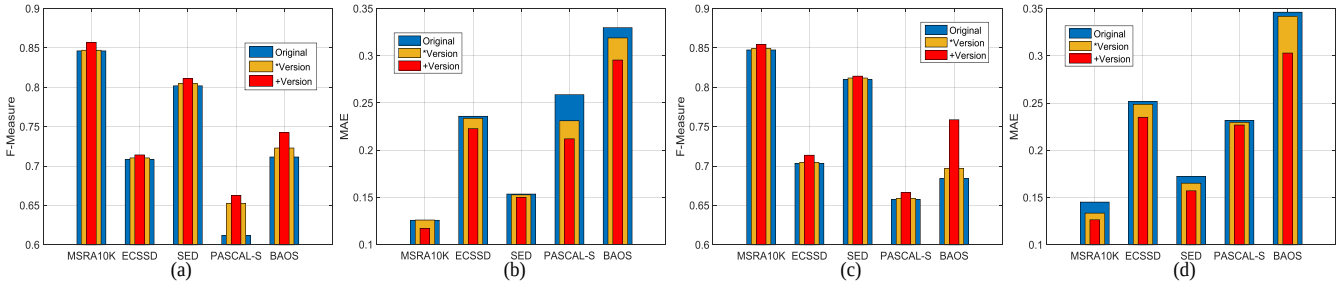


Fig. 14. Evaluation results of MR and MC on different datasets. The (*) and (+) versions stand for optimizations with RBD and our method, respectively. (a) F-measures of MR; (b) MAEs of MR; (c) F-measures of MC; (d) MAEs of MC. Our method is observed to effectively boost the performances of the baseline methods in terms of both F-measure and MAE.

Qualitative Evaluation:

We select the best six methods in Table. 1, namely, DSR, HS, MC, MR, wCtr* and RR, in the qualitative evaluation with our proposed method. Several visual samples on the five datasets are drawn in Fig. 13. We observe that in general, our method achieves the best performance among the chosen images. The DSR [25] method computes saliency via multi-scale reconstruction errors followed by an object-based Gaussian refinement. However, since the saliency map boundaries are frequently suppressed by the Gaussian refinement, DSR will always produce dark boundaries, which is clearly visible on all of the chosen images. The HS [31] method is ideal in dealing with small-scale high-contrast regions by using a tree model. Yet, since it depends on the extraction of cue maps of low-level features such as color and position, it does not work well with images that have low contrast between the foreground and the background, e.g., Fig. 13b and c. The MC [10] model applies absorbed time of the Markov chain in calculating the saliency value and provides fair enough estimations in most cases. Nevertheless, it tends to highlight the center due to its longer distance to the boundaries and will frequently fail in detecting boundary adjacent salient objects, which is seen in Fig. 13a, b and h. The MR [6] method evaluates superpixel saliency via graph-based manifold ranking, which functions well in images with centered salient objects. However, it completely relies on the image boundaries as background queries, which creates saliency reversions when boundary-adjacent salient objects are presented, as shown in Fig. 13a, b, e, f and g. The wCtr* [27]

method optimizes saliency detection by exploiting the proportion of a region connecting to the boundaries, which shows good results on centered salient objects. However, its core idea, the boundary connectivity, is still based on image boundaries as background references, which has similar drawbacks as MR. Finally, the RR [11] method is a former version of our proposed method, but instead of applying RC, it uses 3 of the 4 image boundaries as background queries. Fig. 13a, b, d and g demonstrate that the proposed RC can provide even higher robustness than RR [11].

Our method generates saliency maps that visually correlate better with the ground truth. It exhibits high robustness under various cases, even in cases with complex backgrounds such as in Fig. 13b and f. With the improvement from the proposed RC, it shows marked advantage in handling boundary-adjacent salient object images, thus minimizing the emergence of saliency reversion. The proposed RRWR helps to provide elevated accuracy and smoothness to the output saliency maps, which are seen in Fig. 13c, d and h. We further note that our method is good at suppressing background regions that share similar patterns to the salient object, such as Fig. 13b.

D. Extensibility as a Saliency Optimization Algorithm

As stated in Section IV.A, the initial saliency estimation can also be generated by other boundary-based methods. In this case, we suggest that our method functions as a saliency optimization algorithm. To evaluate its optimization performance, we compare our method with RBD [27], which is a state-of-the-art saliency optimization algorithm that can be widely applied to different saliency detection methods for performance improvements. The evaluation is conducted on two boundary-based methods MR [6] and MC [10] and the same five datasets from the previous section are used. The results are listed in Fig. 14. It is obvious that our method outperforms RBD on each F-measure bars in Fig. 14a and c. Our method also has the lowest MAE scores on all five datasets in Fig. 14b and d, when compared to both the original methods and their RBD-optimized versions. Our method is better than RBD because RBD relies on image boundaries as the background references, which inevitably suffers from saliency reversion cases. Our method shows especially high performance on the BAOS dataset, which further proves its high robustness on boundary-adjacent salient objects.

RC and RRWR can also be independently exploited as two separate saliency optimization algorithms, which provide further flexibility of our algorithm in practical applications.

E. Efficiency

Our method is implemented on MATLAB R2014b. The computational efficiency is tested on a 64-bit PC with Intel 6-Core i7-5820K CPU @ 3.3 GHz and 64GB RAM. The average time per image of our method is 0.408 s (excluding the time for superpixel generation and initial saliency estimation) where the RC step takes less than 0.01s, and the RRWR step takes 0.358 s. The source code of our method, together with the BAOS dataset, is available at: https://github.com/yuanyc06/rcrr/.

F. Limitation

One limitation of our proposed method, as observed in the experiments, is that in images where the salient object occupies more than half (or even all) of the image boundaries, the originally correct initial saliency estimation S_{init} will be mistakenly detected as "reversed" and thus be unnecessarily processed by the RC. Considering that such cases only appear occasionally, our method still prevails over other state-of-the-art methods in overall performance.

VI. Conclusion

In this paper, we propose RCRR, a novel saliency detection method with reversion correction (RC) and regularized random walk ranking (RRWR). The significant contributions of our method have two aspects: first, the RC process can effectively neutralize the negative influences of the boundary adjacent foreground regions and thereby reverse the "reversed" saliency maps back to normal and lead to more accurate and robust saliency estimations; second, the RRWR can provide a prior saliency estimation to all of the pixels in an image, resulting in smoother and more detailed saliency map outputs. We also distribute the BAOS image dataset, which can be used to evaluate the performance on boundary adjacent salient objects. Our method is fully automatic without any user supervision. The results of experiments on five datasets show that our method significantly outperforms 14 state-of-the-art saliency detection methods in both accuracy and robustness. We further demonstrate the extensibility of our method as a saliency optimization algorithm.

References

[1] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254-1259, Nov. 1998.

[2] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2003, pp. II-264-II-271 vol. 2.

[3] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 409-416.

[4] V. Gopalakrishnan, Y. Hu, and D. Rajan, "Random walks on graphs for salient object detection in images," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3232-3242, Dec. 2010.

[5] L. Mai, Y. Niu, and F. Liu, "Saliency aggregation: A data-driven approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1131-1138.

[6] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3166-3173.

[7] D. Gao and N. Vasconcelos, "Discriminant saliency for visual recognition from cluttered scenes," in *Adv. Neural Inf. Process. Syst.*, Dec. 2004, pp. 481-488.

[8] S. Frintrop, G. Backer, and E. Rome, "Goal-directed search with a top-down modulated computational attention system," in *Pattern Recognit.*, ed: Springer, 2005, pp. 117-124.

[9] F. Yang, H. Lu, and M.-H. Yang, "Robust superpixel tracking," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1639-1651, Apr. 2014.

[10] B. Jiang, L. Zhang, H. Lu, C. Yang, and M.-H. Yang, "Saliency detection via absorbing markov chain," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1665-1672.

[11] C. Li, Y. Yuan, W. Cai, Y. Xia, and D. D. Feng, "Robust saliency detection via regularized random walks ranking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 2710-2717.

[12] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274-2282, Nov. 2012.
[13] R. Achanta, F. Estrada, P. Wils, and S. Süsstrunk, "Salient region detection and segmentation," in *Comput. Vis. Syst.*, ed: Springer, 2008, pp. 66-75.
[14] K. Fu, C. Gong, J. Yang, Y. Zhou, and I. Yu-Hua Gu, "Superpixel based color contrast and color distribution driven salient object detection," *Signal Process.: Image Comm.*, vol. 28, no. 10, pp. 1448-1463, Jul. 2013.
[15] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Adv. Neural Inf. Process. Syst.*, Dec. 2006, pp. 545-552.
[16] J. Yang and M.-H. Yang, "Top-down visual saliency via joint crf and dictionary learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2296-2303.
[17] A. Borji, D. N. Sihite, and L. Itti, "Probabilistic learning of task-specific visual attention," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 470-477.
[18] J. Theeuwes, "Top–down and bottom–up control of visual selection," *Acta Psycho.*, vol. 135, no. 2, pp. 77-99, May. 2010.
[19] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1-8.
[20] H. Jiang, J. Wang, Z. Yuan, Y. Wu, N. Zheng, and S. Li, "Salient object detection: A discriminative regional feature integration approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2083-2090.
[21] D. Zhou and B. Schölkopf, "Learning from labeled and unlabeled data using random walks," in *Pattern Recognit.*, ed: Springer, 2004, pp. 237-244.
[22] D. Zhou, J. Weston, A. Gretton, O. Bousquet, and B. Schölkopf, "Ranking on data manifolds," *Adv. Neural Inf. Process. Syst.*, vol. 16, pp. 169-176, Dec. 2004.
[23] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Comput.*, vol. 13, no. 7, pp. 1443-1471, Nov. 2001.
[24] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768-1783, Nov. 2006.
[25] X. Li, H. Lu, L. Zhang, X. Ruan, and M.-H. Yang, "Saliency detection via dense and sparse reconstruction," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 2976-2983.
[26] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129-137, Mar. 1982.
[27] W. Zhu, S. Liang, Y. Wei, and J. Sun, "Saliency Optimization from Robust Background Detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2814-2821.
[28] A. Borji, M.-M. Cheng, H. Jiang, and J. Li, "Salient object detection: A survey," *arXiv preprint arXiv:1411.5878*, 2014.
[29] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. Torr, and S. Hu, "Global contrast based salient region detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 569-582, May. 2015.
[30] A. Borji, D. N. Sihite, and L. Itti, "Salient object detection: A benchmark," in *Comput. Vis.–ECCV 2012*, ed: Springer, 2012, pp. 414-429.
[31] Q. Yan, L. Xu, J. Shi, and J. Jia, "Hierarchical saliency detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1155-1162.
[32] S. Alpert, M. Galun, R. Basri, and A. Brandt, "Image segmentation by probabilistic bottom-up aggregation and cue integration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1-8.
[33] Y. Li, X. Hou, C. Koch, J. Rehg, and A. Yuille, "The secrets of salient object segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 4321-4328.
[34] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303-338, Sep. 2010.
[35] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309-314, Sep. 2004.
[36] G. Hripcsak and A. S. Rothschild, "Agreement, the f-measure, and reliability in information retrieval," *J. Ame. Med. Inf. Asso.*, vol. 12, no. 3, pp. 296-298, May. 2005.
[37] D. R. Martin, C. C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 5, pp. 530-549, May. 2004.
[38] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 1915-1926, Oct. 2012.
[39] H. Jiang, J. Wang, Z. Yuan, T. Liu, N. Zheng, and S. Li, "Automatic salient object segmentation based on context and shape prior," in *Brit. Mach. Vis. Conf.*, Aug. 2011, p. 7.
[40] H. R. Tavakoli, E. Rahtu, and J. Heikkilä, "Fast and efficient saliency detection using sparse sampling and kernel density estimation," in *Image Analysis*, ed: Springer, 2011, pp. 666-675.
[41] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1597-1604.
[42] X. Shen and Y. Wu, "A unified approach to salient object detection via low rank matrix recovery," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 853-860.
[43] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 733-740.

Yuchen Yuan received the B. E. degree in biomedical engineering from Tsinghua University, Beijing, China, in 2010, and the M. S. degree in biomedical engineering from Washington University in St. Louis, St. Louis, MO, USA, in 2012. He then worked as a senior system research engineer at Mindray Co., Ltd, Shenzhen, China, from 2012 to 2014.
He is currently a Ph. D. candidate at the School of Information Technologies, the University of Sydney. His research interests include deep learning, saliency detection, image segmentation and bioinformatics.

Changyang Li received the Ph.D. degree from the University of Sydney, Sydney, Australia, in 2012.
He is currently a Research Associate with the School of Information Technologies, the University of Sydney. His current research interests include image segmentation, pattern recognition, processing for biomedical images, and multimodality data registration.

Jinman Kim received the B. S. (honours) and Ph. D. degrees in computer science from the University of Sydney, Sydney, Australia, in 2001 and 2006, respectively. Since his Ph. D., he has been a research associate at the leading teaching hospital, the Royal Prince Alfred. From 2008 to 2012, he was an ARC postdoc research fellow, one year leave (2009-2010) to join the MIRALab research group, Geneva, Switzerland, as a Marie Curie senior research fellow. Since 2013, he has been with the School of IT, the University of Sydney, where he was a senior lecturer, and became an associate professor in 2016.
His research interests include medical image analysis and visualization, computer aided diagnosis, and Telehealth technologies.

Weidong Cai received the Ph. D. degree in Computer Science from the Basser Department of Computer Science, The University of Sydney, Sydney, Australia, in 2001.
He is currently an Associate Professor in the School of Information Technologies, Director of the Multimedia Laboratory in The University of Sydney. He has been a Lead Investigator / Visiting Professor on medical image analysis and medical computer vision at Surgical Planning Laboratory (SPL), Harvard Medical School during his 2014 SSP. His research interests include multimedia computing, medical image analysis, image / video processing and retrieval, computer vision & pattern recognition, big data mining, bioimaging informatics, and computational neuroscience.

Dagan Feng received the M. E. degree in Electrical Engineering & Computer Science (EECS) from Shanghai Jiao Tong University, Shanghai, China, in 1982, and the M. S. in Biocybernetics and Ph. D. in Computer Science from the University of California, Los Angeles, CA, USA, in 1985 and 1988 respectively, where he received the Crump Prize for Excellence in Medical Engineering.

He is currently a Professor in the School of Information Technologies, Director of Biomedical & Multimedia Information Technology (BMIT) Research Group and Director of the Institute of Biomedical Engineering and Technology at the University of Sydney. He has published over 800 scholarly research papers, pioneered several new research directions, and made a number of landmark contributions in his field. Prof. Feng's research in the areas of biomedical and multimedia information technology seeks to address the major challenges in “big data science” and provide innovative solutions for stochastic data acquisition, compression, storage, management, modeling, fusion, visualization and communication. Prof. Feng is Fellow of ACS, HKIE, IET, IEEE and Australian Academy of Technological Sciences and Engineering.