Pattern Mining Saliency

Yuqiu Kong[1], Lijun Wang[2], Xiuping Liu[1], Huchuan Lu[2(✉)], and Xiang Ruan[3]

[1] Department of Mathematical Sciences,
Dalian University of Technology, Dalian, China
`jinghongkyq@mail.dlut.edu.cn, xpliu@dlut.edu.cn`
[2] Department of Electrical Engineering,
Dalian University of Technology, Dalian, China
`wlj@mail.dlut.edu.cn, lhchuan@dlut.edu.cn`
[3] Tiwaki Corporation, Tiwaki, Japan
`ruanxiang@gmail.com`

Abstract. This paper presents a new method to promote the performance of existing saliency detection algorithms. Prior bottom-up methods predict saliency maps by combining heuristic saliency cues, which may be unreliable. To remove error outputs and preserve accurate predictions, we develop a pattern mining based saliency seeds selection method. Given initial saliency maps, our method can effectively recognize discriminative and representative saliency patterns (features), which are robust to the noise in initial maps and can more accurately distinguish foreground from background. According to the mined saliency patterns, more reliable saliency seeds can be acquired. To further propagate the saliency labels of saliency seeds to other image regions, an Extended Random Walk (ERW) algorithm is proposed. Compared with prior methods, the proposed ERW regularized by a quadratic Laplacian term ensures the diffusion of seeds information to more distant areas and allows the incorporation of external classifiers. The contributions of our method are complementary to existing methods. Extensive evaluations on four data sets show that our method can significantly improve accuracy of existing methods and achieves more superior performance than state-of-the-arts.

Keywords: Saliency detection · Pattern mining · Random walk

1 Introduction

Saliency detection is a fundamental problem that has found wide applications in various computer vision tasks, such as object recognition [1], image segmentation [2] and visual tracking [3]. As a pre-processing step, saliency detection facilitates more sensible assignments of limited processing resource to prominent regions, thus allows more sophisticated subsequent processing stages. Though

much research effort has been made [4–9], it is still a very challenging task to design a saliency model with high performance in real complex scenes.

In their seminal work [10], Itti *et al.* point out that human visual system is sensitive to high-contrast regions and propose to detect saliency by measuring local contrast in multi-scales across different feature channels, including intensity, color, orientation, *etc.* Since then, contrast prior is widely studied and adopted by a variety of saliency models [11–14] from either local or global view. For local methods [10–12,15], saliency is characterized by regional center-surround contrast. Although these methods can highlight salient pixels along object boundaries, they often fail in discovering inner regions of salient objects. Due to the lack of global information, unsatisfactory results are achieved under cluttered scenes. In contrast, global methods [4,13,14] estimate saliency by considering feature contrast over the entire image and thus are more capable of locating salient objects precisely. Since detailed local information is ignored, they have very limited discriminative power to uniformly capture salient objects from background with similar appearance.

Instead of computing local contrast or blindly comparing similarity over the entire image, some saliency methods [6,16,17] propose to explore the boundary prior, *i.e.*, by regarding image boundaries as background and propagating their labels to detect salient foreground. These methods are effective in certain scenarios. Though unlike contrast based models, which suffer from failures in detecting object inner regions or separating foreground from background distractors, these methods also have their own drawbacks. Firstly, boundary prior is mainly utilized in a trivial and heuristic manner, such that salient objects appearing at image boundary will be incorrectly labeled as background. Secondly, most of these methods mainly rely on low-level handcrafted features, which are incapable for high-level cognition and understanding, thus insufficient to highlight semantic objects from complex scenarios. To incorporate high-level concepts, other methods [11,18–20] explore task-driven strategies, which involve supervised learning on image data with pixel-wise annotations. However, obtaining massive amount of manually-labeled data is very expensive and time consuming.

In order to address the above issues of existing methods, we seek an alternative approach for saliency detection. Our first contribution is a novel salient seeds selection method. Saliency maps predicted by combining heuristic saliency cues can sometimes be very noisy, *i.e.,* the saliency maps shown in Fig. 1(b). To improve accuracy of these initial saliency maps, we apply a pattern mining algorithm to recognize saliency rules (feature patterns) which are frequently depicted by foreground regions in the initial saliency map and rarely carried by background regions. Based on these saliency rules, a sufficient number of reliable saliency seeds can be effectively detected (See Fig. 1(c)), which can significantly remove the inaccurate prediction of the initial saliency maps (See Fig. 1(d)). Our second contribution is an Extended Random Walk (ERW) algorithm which incorporates quadratic Laplacian term and an external classifier into traditional approach and achieves significant performance improvement in terms of propagation ability. By exploiting the proposed ERW algorithm, the label information

| (a) | (b) | (c) | (d) | (e) |
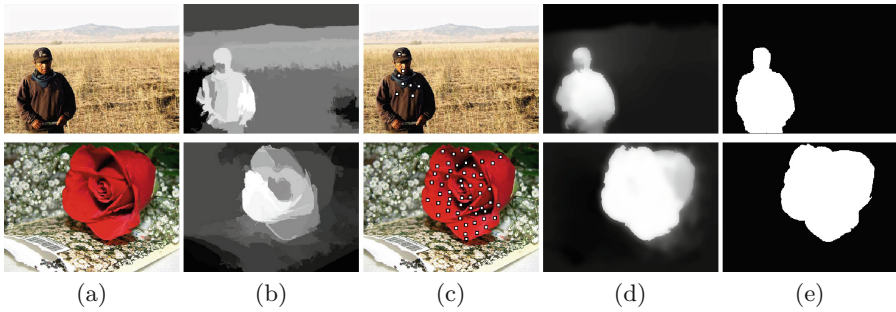
Fig. 1. Intermediate results of the proposed saliency detection algorithm. Brighter pixels indicate higher saliency values. (a) Original images. (b) Saliency maps generated by CB [5] which is used as initial maps. (c) Saliency seeds detected by pattern mining algorithm. (d) The final saliency map via propagating saliency seeds. (e) Ground truth.

of saliency seeds is diffused to more distant areas, which ensures the final saliency map of our model to be more accurate. Taking saliency maps generated by existing methods as initial maps, our algorithm is able to promote the precision of these maps with a considerable margin. Extensive evaluations on four benchmark data sets demonstrate that the promoted results achieve favorable performance against state-of-the-art methods.

2   Related Work

Saliency detection can be conducted by either bottom-up computational models or top-down data driven methods. Most bottom-up methods detect salient regions by combining heuristic saliency cues, such as contrast prior [11–14] and boundary prior [6,16,17]. Recently, [8] proposes boundary connectivity to measure the background probability of regions. Although the heuristic saliency cues based methods perform well for images with simple scene, they may fail to capture the true salient regions when the image background is complex or the appearance between objects and background is similar.

Different from bottom-up methods, top-down approaches [11,19–21] are able to automatically learn saliency models in a supervised manner from large number of training samples. While these methods are shown to be more robust in handling complex scenarios, their generalization abilities heavily rely on training data. Moreover, the training process is very computational expensive. In contrast, [9] learns an unique multi-kernel boosting classifier for each input image supervised by an initial saliency map. However, the inaccuracy of initial map will contaminate the saliency labels of training samples and inevitably degrade the performance of the classifier. Different from the above methods, we employ a pattern mining algorithm to detect the common feature patterns of salient regions for each image based on its initial map. The pattern mining algorithm is more robust to noisy initial maps. As a result, the mined patterns can more reliably characterize salient regions and facilitate reliable saliency seeds selection.
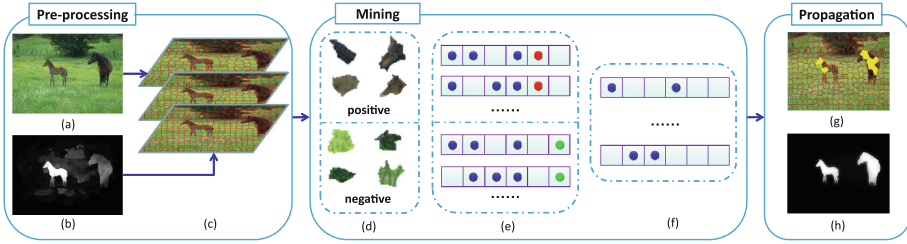
Fig. 2. Framework of the proposed algorithm. (a) Input image. (b) Initial maps generated by other methods. (c) SLIC segmentation results. (d) Sample pool. (e) Transaction database. (f) Saliency rules. (g) Selected saliency seeds. (h) Final saliency map.

Recently, label propagation based saliency detection methods have attracted growing interest from the community. The performance of these methods strongly rely on the quality of the saliency seeds as well as the propagation ability. Some existing methods [6,16,17,22] heuristically treat image boundary as background seeds and use different propagation methods to determine the saliency degree of other image regions. For instance, [17] constructs a graphical model with superpixels as nodes and predicts their saliency according to the hitting time at the equilibrium state. A geodesic distance is defined in [6] to measure the similarity of an image region to the image boundary The proposed method also explores the label propagation scheme for saliency detection. Instead of simply using image boundaries as background seeds, we adopt a pattern mining algorithm to detect more reliable saliency seeds. Compared with random walk based saliency methods [17,23–25], the ERW algorithm incorporates a quadratic Laplacian energy term to explicitly enforce both extensiveness and smoothness of label propagation. In addition, external classifier integrated in ERW algorithm can enable more accurate label assignment.

3   Pattern Mining Algorithm

Pattern mining algorithm is firstly studied for market basket analysis and recently applied in computer vision tasks [26,27]. Given the massive customer transaction database, the aim is to learn the association rules, which indicate the probability of customers buying certain items based on the items they have already bought. In this section, we introduce the terminology and basic concept of pattern mining algorithm.

Frequent itemset. Denote a set of M items as $I = \{i_1, i_2, ..., i_M\}$. A transaction T is a subset of I, namely $T \subseteq I$. A transaction database $D = \{T_1, T_2, ..., T_N\}$ consists of N different transactions. $A \subseteq I$ is called a frequent itemset if A is frequently occurred as a fraction of transactions $T \in D$. The frequency can be described by the support value of A:

$$supp(A) = \frac{|\{T|T \in D, A \subseteq T\}|}{N} \in [0, 1]. \tag{1}$$

If $supp(A) > t_{min}$, A is a frequent itemset, where t_{min} is a pre-defined threshold.

Association rule. An association rule $A \rightarrow p$ describes the situation where item p presents in transactions which contain itemset A. The support value of a rule is defined as:

$$supp(A \rightarrow p) = supp(A \cup \{p\}) = \frac{|\{T|T \in D, A \cup \{p\} \subseteq T\}|}{|D|} \qquad (2)$$

The quality of an association rule $A \rightarrow p$ can be evaluated by a confidence value:

$$conf(A \rightarrow p) = \frac{supp(A \rightarrow p)}{supp(A)} = \frac{|\{T|T \in D, A \cup \{p\} \subseteq T\}|}{|\{T|T \in D, A \subseteq T\}|}. \qquad (3)$$

The association rules with high confidence are regarded as representative rules.

4   Seeds Detection Based on Pattern Mining

In this section, we give details of how to detect sufficient and reliable saliency seeds using pattern mining algorithm. The outline is illustrated in Fig. 2. Given an initial map generated by existing method, we first construct a sample pool (Fig. 2(d)) consisting of both foreground and background regions. A transaction database (Fig. 2(e)) is then created by collecting feature patterns of all the samples in the sample pool. To obtain saliency rules (Fig. 2(f)) that can accurately discriminate foreground from background, we apply an efficient pattern mining algorithm to the transaction database. Finally, the saliency seeds (Fig. 2(g)) can be selected according to the acquired saliency rules.

Feature extraction. Given the input image and the corresponding initial saliency map, we first oversegment the image from three different scales and
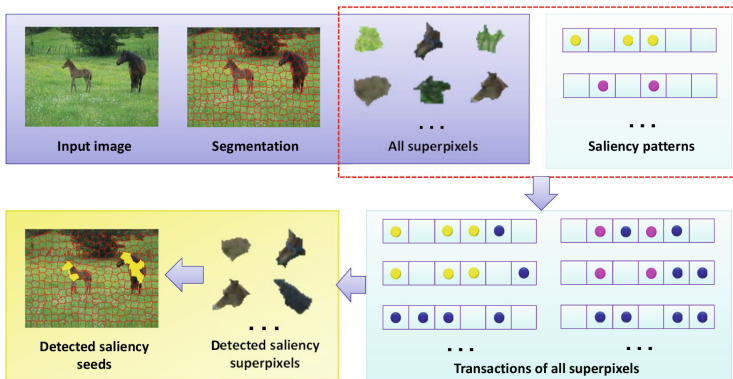

Fig. 3. Detailed process of seeds detection.

obtain a set of superpixels $S = \{s_1, s_2, ..., s_N\}$, serving as the sample pool. By thresholding the saliency maps with a threshold t_0, the image is segmented into foreground and background regions. Superpixels within foreground regions are labeled as positive samples, whereas those within background are labeled as negative samples. Our method can also take multiple initial saliency maps as input by labeling sample superpixels according to each initial map, respectively.

We exploit the bag-of-words representation to character each superpixel considering both global context and local appearance information. Specifically, we apply K-means algorithm to cluster all the superpixels in the RGB color space and obtain a set of cluster centriods W which serves as the visual vocabulary, with each centriod as a visual word. Each superpixel is assigned with the visual word (*i.e.* its cluster centroid) indexed by $w_i \in \{1, 2, ..., |W|\}$, where $|W|$ denotes the total number of visual words. Note that the visual vocabulary contains the global context information of the input image. We then represent each superpixel sample s_i from a local view using the bag-of-words feature which contains three components: its visual word index w_i, the visual word indexes of its K nearest neighbors, and its class label (either *pos* or *neg*).
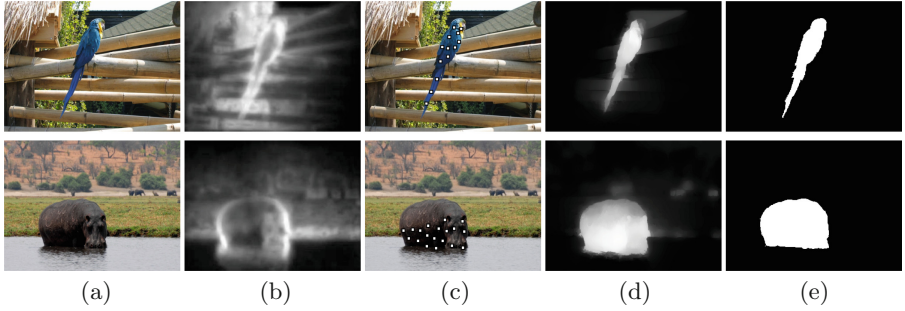
Mining saliency patterns. Pattern mining theories are explored to identify discriminative patterns of bag-of-words features that can accurately distinguish foreground from background. To this end, we regard the set of visual words W as the overall item set with each visual word as an item. Furthermore, the bag-of-words feature of each superpixel can be treated as a transaction with $K + 2$ items, where the first $K + 1$ items are visual words and the last is the label (*pos* or *neg*). The bag-of-words features of all the superpixel samples in the sample pool then creat a transaction database. As a result, to identify discriminative patterns of visual words is then equivalent to find a collection of item sets $\{A\}$ that satisfy the following two conditions:

$$supp(A) > t_1, \tag{4}$$
$$conf(A \rightarrow pos) > t_2, \tag{5}$$

where t_1 and t_2 denote two threshold parameters. Equation 4 indicates that the item set A is a subset of a certain number of transactions, while Eq. 5 enforces that most of these transactions also contain the item *pos*, *i.e.*, most of them are labeled as positive. The item set satisfying the above two conditions represents saliency patterns of bag-of-words features that separate salient foreground regions from background. Considering both efficiency and effectiveness, we exploit the Apriori algorithm [28] for saliency pattern mining.

Detecting saliency seeds. Saliency seeds detection is conducted using the oversegmentation in the first scale (with approximately 300 superpixels). Given a set of saliency patterns $\{A\}$ (*i.e.*, the item sets) accquired by the Apriori algorithm, we select saliency seeds following a straightforward rule. As demonstrated in Fig. 3, a superpixel is selected as a saliency seed only if a subset of its bag-of-words feature belongs to the saliency pattern set $\{A\}$. As illustrated

Fig. 4. Examples of seeds detection results. (a) Original images. (b) Initial maps generated by method CA [4]. (c) Saliency seeds detected by pattern mining algorithm. (d) The final saliency map. (e) Ground truth.

in Fig. 4, with accurate saliency patterns, we can select a sufficient number of reliable saliency seeds with the mined saliency rules. The prediction error of the initial maps can then be significantly removed.

5   Saliency Propagation

We propose an Extended Random Walk (ERW) algorithm on graphic model with superpixels as nodes. Given the selected saliency seeds, the final saliency map of the input image is achieved by propagating the seeds information to other image regions. Both the reliable saliency seeds and the proposed ERW algorithm ensure to more accurately render the final saliency map.

Graph construction. The saliency propagation procedure is also conducted using the oversegmentation of the input image in the first scale. Given the set of superpixels $S = \{s_1, s_2, ..., s_{N_1}\}$, we construct an undirected graph $G = (V, E)$ with node set V and edge set E, where each node represents a superpixel and is connected to its 2-ring neighbors [16] with undirected edges. The weight matrix $W \in R^{N_1 \times N_1}$ measures the similarity and adjacency relationship between each pair of nodes, with each element $w_{ij} = \exp(-\|g(s_i - g(s_j)\|/2\sigma^2)$, if $j \in \mathcal{N}(i)$, and other positions are 0, where $g(s_i)$ denotes the feature of node s_i and $\mathcal{N}(i)$ indicates the nodes connected to s_i. The Laplacian matrix can be computed by $L = D - W$, where $D = diag(d_1, d_2, ..., d_{N_1})$ is degree matrix with $d_i = \sum_j w_{ij}$.

Extended Random Walk. Let $\mathcal{L}$ denote a labeled node set consisting of all the mined saliency seeds, and $\mathbf{f} = [f_1, f_2, ..., f_{N_1}]^T$ denote the label vector of all the nodes, where f_i is fixed to 1 if $s_i \in \mathcal{L}$, and f_i is initialized to 0 otherwise. Label propagation aims to infer the labels of all the nodes based on the saliency seeds. In this work, we propose an Extended Random Walk algorithm for label propagation by minimizing the following energy function

$$\arg\min_{\mathbf{f}} \frac{1}{2}\sum_{i,j} w_{ij}(f_i - f_j)^2 + \frac{\alpha}{2}\sum_{i=1}^{N_1}(d_i f_i - \sum_{j \in \mathcal{N}(i)} w_{ij} f_j)^2 + \frac{\beta}{2}\sum_{i=1}^{N_1}(f_i - y_i)^2, \quad (6)$$

$$\text{s.t. } f_i = 1, \ \forall s_i \in \mathcal{L},$$

where weight w_{ij} measures the similarity of s_i and s_j; $d_i = \sum_j w_{ij}$ is the degree of node s_i; y_i denotes the output of an external classifier and adopts the mean saliency value of node s_i in initial saliency map; α, β are trade-off parameters.

The first term of Eq. 6 is the traditional random walk formulation which enforces label consistency of nodes with strong affinity. The second term is the quadratic Laplacian. To gain more comprehensive interpretation, we minimize the Laplacian term with respect to $\mathbf{f}$ by setting its derivative to zero and obtain the following solution:
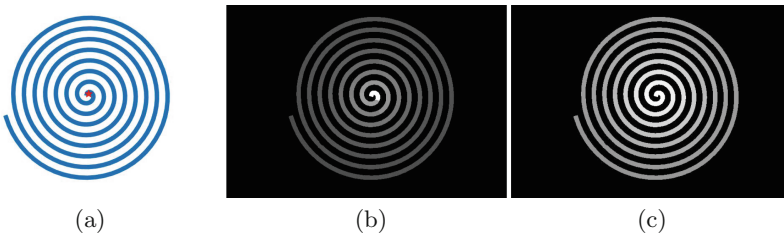
$$f_i = \frac{1}{d_i}\sum_{j \in \mathcal{N}(i)} w_{ij} f_j + \frac{1}{d_i^2}\sum_{j \in \mathcal{N}(i)} w_{ij}\left(\sum_{h \in \mathcal{N}(j)} w_{jh}(f_j - f_h)\right). \quad (7)$$

Apparently, the value of f_i is influenced not only by its direct neighbors $j \in \mathcal{N}(i)$, but also by its neighbors' context $h \in \mathcal{N}(j)$. As a consequence, the seeds information can be more extensively propagated to distant nodes than traditional first-order laplacian diffusion (i.e., the first term of Eq. 6). The third term incorporates the prior knowledge provided by the initial saliency maps into the random walk algorithm, and penalizes saliency predictions that significantly differ from saliency priors. As illustrated in Fig. 5, initialized by the same saliency seed, the proposed ERW algorithm with strong propagation ability achieves more accurate predictions than the traditional method in a challenging setting.

To solve the energy function in Eq. 6, we first re-order the label vector as $\mathbf{f} = [\mathbf{f}_l^T\ \mathbf{f}_u^T]^T$ and the external classifier as $\mathbf{y} = [\mathbf{y}_l^T\ \mathbf{y}_u^T]^T$, where l indicates the labeled nodes set and u corresponds to unlabeled nodes set. The energy minimization problem can then be re-written in the following matrix form

$$\mathbf{f}^* = \arg\min_{\mathbf{f}} \frac{1}{2}\mathbf{f}^T L \mathbf{f} + \frac{\alpha}{2}\mathbf{f}^T L^2 \mathbf{f} + \frac{\beta}{2}(\mathbf{f} - \mathbf{y})^T(\mathbf{f} - \mathbf{y})$$

$$\text{s.t. } \mathbf{f}_l = 1, \quad (8)$$

(a)	(b)	(c)

Fig. 5. (a) Input image and the red pentagram is saliency seeds. (b) Saliency map detected by random walk algorithm. (c) Saliency map detected by random walk with quadratic Laplacian term.

where $L = \begin{bmatrix} L_{ll} & L_{lu} \\ L_{ul} & L_{uu} \end{bmatrix}$ is Laplacian matrix. By setting the derivative of Eq. 8 to zero, the final saliency values of unlabeled nodes are computed as

$$\mathbf{f}_u = M_{uu}^{-1}(-M_{ul}\mathbf{f}_l + \beta\mathbf{y}_u), \tag{9}$$

where $M = L + \alpha L^2 + \beta I$, and I is identity matrix.

Integration. In this paper, we employ the mean CIELab color feature and the Local Binary Pattern (LBP) feature to characterize each superpixel. The above label propagation is independently conducted in the two feature spaces. Color feature is effective when the salient object depicts a distinct color appearance against background. In contrast, the texture feature will be more discriminative when the target object have similar color but different texture compared with background (See the second example of Fig. 8). Based on these observations, we integrate these two feature representations by linearly combining two prediction results to generate the final saliency map,

$$S_f = \lambda S_1 + (1 - \lambda)S_2. \tag{10}$$

where S_1 and S_2 are saliency maps computed in two feature spaces and λ is a weight parameter to balance these two maps. In our experiments, we empirically set $\lambda = 0.5$ to weight these two features.

In this section, we conduct experimental evaluations of the proposed pattern mining based saliency detection method (named as PM) against state-of-the-art methods on benchmark data sets. The contributions of different components of the proposed methods (*i.e.*, seeds selection method and the ERW algorithm) are also analyzed. More results can be found in supplementary material[1].

6 Experiments

6.1 Parameter Setting

In our experiments, we find that the proposed method is insensitive to most of the parameters. Therefore, all the parameters are empirically set through cross-validation and fixed through all the data sets. The threshold t_0 for constructing sample pool is set to 0.5. The size of visual vocabulary $|W|$ is set to 300. The bag-of-words feature for each sample is computed using $K = 20$ nearest neighbors. The thresholds t_1 and t_2 for pattern mining are set to 90 % and 20 %, respectively. The parameter σ to compute weight matrix is set to 10. The trade-off parameters of the ERW algorithm are set as $\alpha = 0.5$ and $\beta = 0.01$, respectively. The proposed method is implemented in MATLAB, and runs at 4 seconds per image on a PC with a 3.4 GHz CPU. The source code will be made publicly available (see Footnote 1).

---

[1] http://ice.dlut.edu.cn/lu/index.html.

6.2   Data Sets and Evaluation Metrics

We evaluate the proposed algorithm on four benchmark data sets. The MSRA-5000 dataset [11] contains 5000 images with complex scenes; The SOD dataset [29] consists of 300 images; The ECSSD dataset [30] incorporates 1000 images and the Pascal-S dataset [31] is composed of 850 images. The later three data sets are very challenging, since most images have cluttered background or more than one salient object.

Precision-Recall curves, F-measure and the mean absolute error (MAE) are employed to evaluate the performance of each detection model, where F-measure is the weighted harmonic mean of precision and recall value, and MAE is the average pixel-wise difference between saliency map and its ground truth.

6.3   Quantitative Analysis

Performance of the proposed framework. We choose 12 existing saliency detection algorithms as baseline methods, including ITTI [10], GBVS [32], CA [4], CB [5], LR [33], DSR [7], UFO [34], HS [30], wCO [8], HDCT [19], BL [9] and RR [25]. Two evaluations are conducted: single model promotion and joint promotion. For single model promotion, we apply the proposed algorithm to promote the performance of each baseline method by taking its predicted saliency map as the initial saliency map. The promoted method is denoted by -PM (*e.g.*, CA-PM denotes the promoted model of baseline CA). For joint promotion, we apply our method to jointly promote a set of baselines (*i.e.*, SET1={CA,CB,LR} and SET2={DSR,UFO,wCO}), by taking their predicted saliency maps as initial maps (See Sect. 4).

Figure 6 compares the P-R curves of the baseline models and their promoted methods on four data sets. Table 1 shows the F-measures and MAE scores, where the baseline results of different methods are shown in the columns of "BS", and the corresponding promoted results are displayed in the columns of "PM". As shown in Fig. 6, our method can effectively promote all the baseline results and achieve state-of-the-art performance regardless of the accuracy of initial maps. The results further verifies that the proposed method has a strong generalization ability across a wide range of baseline methods and is very robust to the noisy prediction of initial maps Especially, initialized by eye fixation results (*e.g.*, ITTI/GBVS), the proposed method is capable to promote their performance with a considerable margin.

In addition, the joint promotion on a set of baselines achieves consistently higher performances than the corresponding single model promotion in most data sets. This may be attributed to the fact that more samples can be acquired for pattern mining in the joint promotion case. However, sometimes labels of samples from different methods may be inconsistent which causes confusion for the pattern mining procedure. Thus the final detection results may be affected by this circumstance.
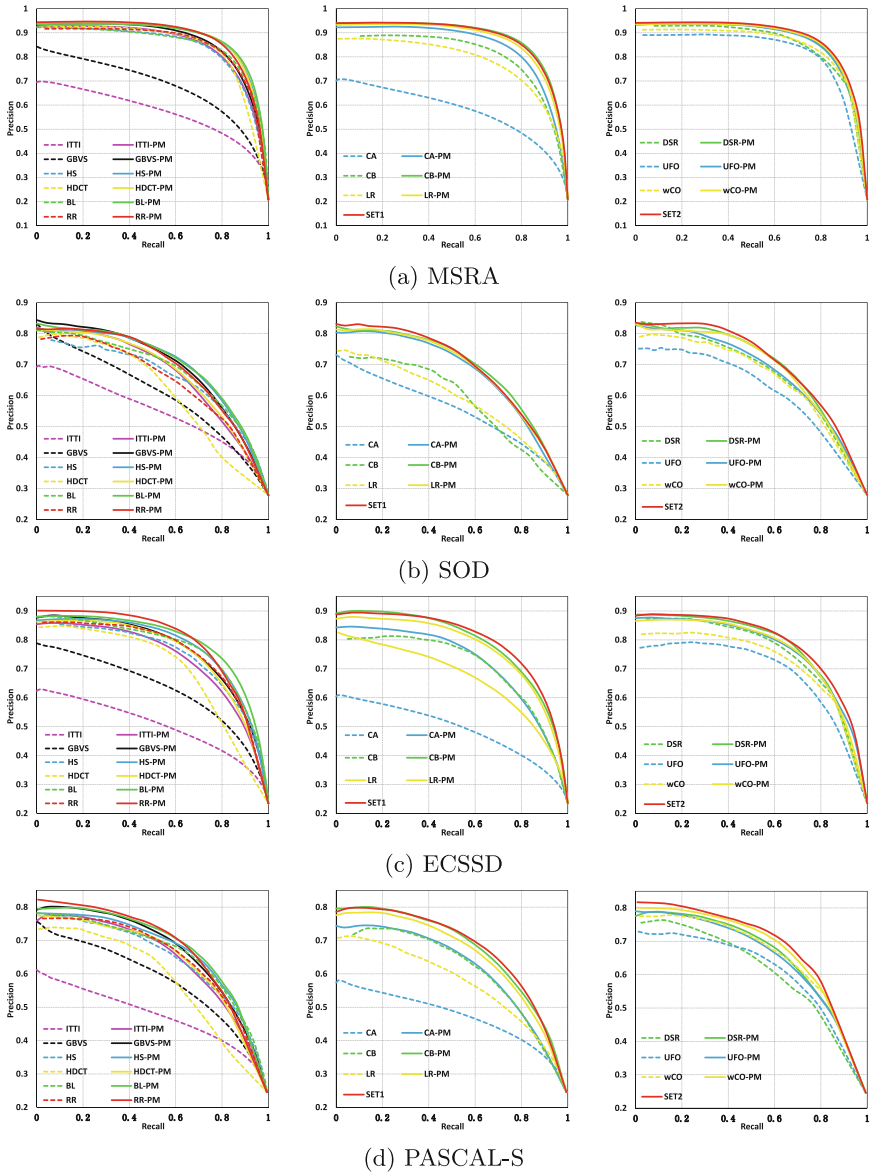
Fig. 6. P-R curve of the state-of-the-art algorithms and their promoted results by our proposed algorithm(PM) on four datasets.

Validation of pattern mining based seeds detection. Saliency seeds plays a very critical role in our saliency algorithm. To verify the effectiveness of the proposed seeds selection method, we report the mean precision rate of the saliency seeds selected by our method by taking saliency maps of each baseline method
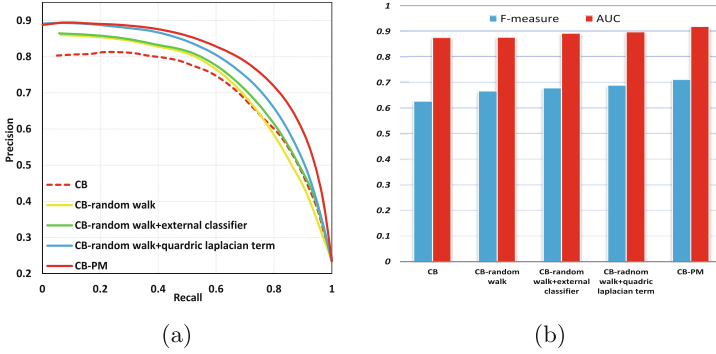
Fig. 7. Quantitative evaluation of four different propagation strategies on the ECSSD data set. (a) P-R curve. (b) AUC and F-measure scores.

as initial maps. The precision rate of the selected saliency seeds is computed as the number of true positive selected seeds over the number of all the saliency seeds (foreground superpixels). For comparison, we further evaluate the precision rate of initial maps of each baseline. Specifically, we firstly compute the average saliency value of each superpixel according to the initial map and select the superpixels whose saliency values are higher than an adaptive threshold [35] as saliency seeds. As demonstrated in Table 2, the proposed seeds selection methods (denoted by PM) can consistently outperform baseline methods (denoted by BS) in terms of mean precision rate on four data sets, which verifies that our seeds selection methods can effectively remove noise and preserve accurate prediction of initial maps. We further demonstrate the effectiveness of our seed selection strategy over two baseline methods in the supplementary material.

Effectiveness of extended random walk. To analyze the effectiveness of the proposed ERW based propagation strategy, we evaluate the performance of each term of the ERW formulation (Eq. 6) on the ECSSD data set by taking saliency maps of CB [5] as initial maps. The P-R curves are shown in Fig. 7(a). The AUC and F-measure scores are illustrated in Fig. 7(b). It can be observed that both the proposed quadratic Laplacian regularization and the incorporated external classifier can improve the propagation ability over the traditional random walk algorithm. By combining these two techniques together, the optimal performance is obtained. More detailed analysis can be found in supplementary material.

6.4   Qualitative Analysis

Figure 8 illustrates some example saliency maps generated by baseline methods and the corresponding (jointly) promoted results. In the first example, the appearance between salient object and image background is unconspicuous in color space. Due to the adopted LBP texture feature, our algorithm can accurately capture the foreground object. The background regions in the second

Table 1. F-measures and MAE scores of baseline methods and their promoted methods on MSRA, SOD, ECSSD and PASCAL-S data sets. The promoted results are marked as blue if they out-perform their baseline methods. If the jointly promoted results (SET1, SET2) are higher than the corresponding single promotion, the scores are marked as red.

Method	Metric	MSRA		SOD		ECSSD		PASCAL-S	
		BS	PM	BS	PM	BS	PM	BS	PM
ITTI	F-measure	0.515	0.779	0.433	0.573	0.428	0.664	0.391	0.568
	MAE	0.249	0.151	0.307	0.259	0.290	0.203	0.296	0.234
GBVS	F-measure	0.608	0.792	0.508	0.593	0.549	0.684	0.496	0.589
	MAE	0.227	0.150	0.292	0.257	0.263	0.203	0.273	0.230
CA	F-measure	0.537	0.779	0.447	0.563	0.429	0.653	0.402	0.552
	MAE	0.250	0.149	0.313	0.259	0.310	0.210	0.301	0.237
CB	F-measure	0.737	0.815	0.490	0.592	0.626	0.711	0.511	0.590
	MAE	0.185	0.128	0.294	0.250	0.240	0.185	0.269	0.220
LR	F-measure	0.694	0.803	0.484	0.584	0.562	0.695	0.476	0.581
	MAE	0.221	0.138	0.308	0.257	0.274	0.195	0.287	0.228
DSR	F-measure	0.784	0.814	0.596	0.615	0.690	0.721	0.554	0.590
	MAE	0.117	0.123	0.234	0.239	0.171	0.161	0.214	0.218
UFO	F-measure	0.775	0.810	0.548	0.594	0.644	0.698	0.549	0.579
	MAE	0.146	0.128	0.257	0.248	0.205	0.187	0.232	0.224
HS	F-measure	0.767	0.814	0.521	0.600	0.635	0.693	0.528	0.594
	MAE	0.162	0.124	0.283	0.247	0.228	0.185	0.264	0.209
wCO	F-measure	0.796	0.816	0.598	0.606	0.676	0.698	0.596	0.612
	MAE	0.111	0.120	0.245	0.225	0.178	0.161	0.201	0.192
HDCT	F-measure	0.773	0.820	0.561	0.594	0.644	0.694	0.532	0.583
	MAE	0.144	0.121	0.244	0.243	0.192	0.183	0.232	0.210
BL	F-measure	0.784	0.815	0.577	0.601	0.683	0.713	0.571	0.609
	MAE	0.169	0.125	0.267	0.246	0.217	0.182	0.249	0.217
RR	F-measure	0.807	0.815	0.567	0.590	0.698	0.716	0.587	0.607
	MAE	0.126	0.124	0.265	0.255	0.185	0.162	0.231	0.204
SET1	F-measure	-	0.811	-	0.598	-	0.708	-	0.597
	MAE	-	0.136	-	0.237	-	0.179	-	0.206
SET2	F-measure	-	0.823	-	0.614	-	0.717	-	0.620
	MAE	-	0.121	-	0.222	-	0.158	-	0.195

example depict different features which causes failure in the most existing methods. Our model succeeds to highlight the entire salient object even with inaccurate initial maps, which attributes to the robustness of our seeds selection methods against noisy initial maps. When there exists small scale noise in the

Table 2. Precision rate of saliency seeds on the MSRA, SOD, ECSSD and PASCAL-S data sets.

	ITTI		CA		CB		LR		DSR		UFO		SET1
	BS	PM	BS	PM	BS	PM	BS	PM	BS	PM	BS	PM	
MSRA	.536	.843	.549	.876	.806	.893	.733	.890	.878	.891	.860	.895	.891
SOD	.519	.708	.515	.751	.637	.776	.569	.752	.731	.772	.686	.764	.780
ECSSD	.488	.749	.469	.774	.714	.828	.610	.812	.810	.811	.746	.829	.835
PASCAL-S	.468	.680	.479	.697	.637	.734	.565	.723	.687	.729	.688	.730	.744
	GBVS		HS		wCO		HDCT		BL		RR		SET2
	BS	PM	BS	PM	BS	PM	BS	PM	BS	PM	BS	PM	
MSRA	.568	.855	.821	.890	.844	.888	.828	.889	.815	.891	.859	.889	.898
SOD	.525	.712	.672	.778	.714	.748	.705	.756	.696	.761	.677	.751	.793
ECSSD	.525	.767	.723	.816	.750	.809	.751	.806	.734	.815	.775	.804	.832
PASCAL-S	.514	.709	.645	.719	.680	.720	.665	.723	0.651	.728	.686	.721	.750

background (such as the third example), most saliency models detect yellow flowers as salient regions, while our algorithm is effective in suppressing the response of noise regions. In the case that saliency object presents various features as is shown in the last example (the cattle with dark brown and light brown hair),

(a)   (b)   (c)   (d)   (e)   (f)   (g)   (h)   (i)   (j)   (k)   (l)
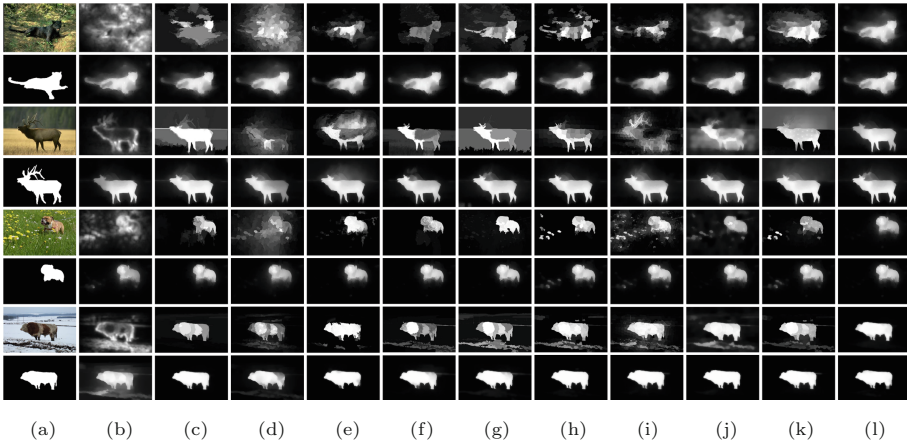
Fig. 8. Saliency maps of four examples. Every two rows correspond to one example. In each example, saliency maps at the first row of (b)-(k) are generated by existed saliency models and the second row present their promoted results. (l) is the jointly boosted results generated by SET1 and SET2, respectively. From left to right: (a) input image and its ground truth (b) CA [4] (c) CB [5] (d) LR [33] (e) DSR [7] (f) UFO [34] (g) HS [30] (h) wCO [8] (i) HDCT [19] (j) BL [9] (k) RR [25] (l) SET1 and SET2. Our model is able to highlight foreground uniformly and suppress the response of cluttered background.

some initial saliency maps fail to highlight the entire object uniformly. Based on pattern mining algorithm, our method is able to detect saliency seeds with a variety of features. Thus all saliency regions can be consistently detected by the proposed method.

7 Conclusions

In this paper, we propose a novel saliency detection model based on pattern mining algorithm. Given an initial saliency map generated by any existing method, our method can effectively recognize discriminative and representative saliency patterns. According to these saliency patterns, sufficient and reliable saliency seeds are detected. Subsequently, we propose an Extended Random Walk (ERW) algorithm to further propagate saliency labels of saliency seeds to other image regions. Compared with prior methods, ERW constrained by a quadratic Laplacian term allows the propagation of saliency seeds to more distant areas and incorporates external classifiers at the same time. Quantitative and qualitative experiments on four benchmark data sets demonstrate that our method is able to improve the performance of existing algorithms and performs favorably against the state-of-the-arts.

References

1. Yang, L., Zheng, N., Yang, J., Chen, H.: A biased sampling strategy for object categorization. In: ICCV (2009)
2. Rahtu, E., Kannala, J., Salo, M., Heikkilä, J.: Segmenting salient objects from images and videos. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part V. LNCS, vol. 6315, pp. 366–379. Springer, Heidelberg (2010)
3. Mahadevan, V., Vasconcelos, N.: Saliency-based discriminant tracking. In: CVPR (2009)
4. Goferman, S., Zelnik-Manor, L., Tal, A.: Context-aware saliency detection. In: CVPR (2010)
5. Jiang, H., Wang, J., Yuan, Z., Liu, T., Zheng, N., Li, S.: Automatic salient object segmentation based on context and shape prior. In: BMVC (2011)
6. Wei, Y., Wen, F., Zhu, W., Sun, J.: Geodesic saliency using background priors. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part III. LNCS, vol. 7574, pp. 29–42. Springer, Heidelberg (2012)
7. Li, X., Lu, H., Zhang, L., Ruan, X., Yang, M.H.: Saliency detection via dense and sparse reconstruction. In: ICCV (2013)
8. Zhu, W., Liang, S., Wei, Y., Sun, J.: Saliency optimization from robust background detection. In: CVPR (2014)
9. Tong, N., Lu, H., Ruan, X., Yang, M.H.: Salient object detection via bootstrap learning. In: CVPR (2015)
10. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. TPAMI **20**(11), 1254–1259 (1998)

11. Liu, T., Sun, J., Zheng, N.N., Tang, X., Shum, H.Y.: Learning to detect a salient object. In: CVPR (2007)
12. Xie, Y., Lu, H., Yang, M.H.: Bayesian saliency via low and mid level cues. TIP **22**(5), 1689–1698 (2013)
13. Achanta, R., Hemami, S., Estrada, F., Süsstrunk, S.: Frequency-tuned salient region detection. In: CVPR (2012)
14. Cheng, M., Zhang, G., Mitra, N.J., Huang, X.H., S.: Global contrast based salient region detection. In: CVPR (2011)
15. Klein, D.A., Frintrop, S.: Center-surround divergence of feature statistics for salient object detection. In: ICCV (2011)
16. Yang, C., Zhang, L., Lu, H., Yang, M.: Saliency detection via graph-based manifold ranking. In: CVPR (2013)
17. Gopalakrishnan, V., Hu, Y., Rajan, D.: Random walks on graphs for salient object detection in images. TIP **19**(12), 3232–3242 (2010)
18. Jiang, H., Wang, J., Yuan, Z., Wu, Y., Zheng, N., Li, S.: Salient object detection: A discriminative regional feature integration approach. In: CVPR (2013)
19. Kim, J., Han, D., Tai, Y.W., Kim, J.: Salient region detection via high-dimensional color transform. In: CVPR (2014)
20. Lu, S., Mahadevan, V., Vasconcelos, N.: Learning optimal seeds for diffusion-based salient object detection. In: CVPR (2014)
21. Wang, L., Lu, H., Ruan, X., Yang, M.H.: Deep networks for saliency detection via local estimation and global search. In: CVPR (2015)
22. Li, N., Sun, B., Yu, J.: A weighted sparse coding framework for saliency detection. In: CVPR (2015)
23. Chang, K.Y., Liu, T.L., Chen, H.T., Lai, S.H.: Fusing generic objectness and visual saliency for salient object detection. In: ICCV (2011)
24. Jiang, B., Zhang, L., Lu, H., Yang, C., Yang, M.H.: Saliency detection via absorbing markov chain. In: ICCV (2013)
25. Li, C., Yuan, Y., Cai, W., Xia, Y., Feng, D.D.: Robust saliency detection via regularized random walks ranking. In: CVPR (2015)
26. Fernando, B., Fromont, E., Tuytelaars, T.: Effective use of frequent itemset mining for image classification. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part I. LNCS, vol. 7572, pp. 214–227. Springer, Heidelberg (2012)
27. Li, Y., Liu, L., Shen, C., van den Hengel, A.: Mid-level deep pattern mining. In: CVPR (2015)
28. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. In: VLDB (1994)
29. Movahedi, V., Elder, J.H.: Design and perceptual validation of performance measures for salient object segmentation. In: POCV, pp. 49–56 (2010)
30. Yan, Q., Xu, L., Shi, J., Jia, J.: Hierarchical saliency detection. In: CVPR (2013)
31. Li, Y., Hou, X., Koch, C., Rehg, J., Yuille, A.: The secrets of salient object segmentation. In: CVPR (2014)
32. Schölkopf, B., Platt, J., Hofmann, T.: Graph-based visual saliency. In: NIPS (2007)
33. Shen, X.S., Wu, Y.: A unified approach to salient object detection via low rank matrix recovery. In: CVPR (2012)
34. Jiang, P., Ling, H., Yu, J., Peng, J.: Salient region detection by ufo: Uniqueness, focusness and objectness. In: ICCV (2013)
35. Otsu, N.: A threshold selection method from gray-level histograms. SMC **9**(1), 62–66 (1979)