Research Article

Saliency Detection in Weak Light Images via Optimal Feature Selection-Guided Seed Propagation

Nan Mu ⬤, Hongyu Wang, Yu Zhang, Hongyu Han, and Jun Yang

School of Computer Science, Sichuan Normal University, Chengdu 610101, China

Correspondence should be addressed to Nan Mu; nanmu@sicnu.edu.cn

Salient object detection has a wide range of applications in computer vision tasks. Although tremendous progress has been made in recent decades, the weak light image still poses formidable challenges to current saliency models due to its low illumination and low signal-to-noise ratio properties. Traditional hand-crafted features inevitably encounter great difficulties in handling images with weak light backgrounds, while most of the high-level features are unfavorable to highlight visually salient objects in weak light images. In allusion to these problems, an optimal feature selection-guided saliency seed propagation model is proposed for salient object detection in weak light images. The main idea of this paper is to hierarchically refine the saliency map by learning the optimal saliency seeds in weak light images recursively. Particularly, multiscale superpixel segmentation and entropy-based optimal feature selection are first introduced to suppress the background interference. The initial saliency map is then obtained by the calculation of global contrast and spatial relationship. Moreover, local fitness and global fitness are used to optimize the prediction saliency map. Extensive experiments on six datasets show that our saliency model outperforms 20 state-of-the-art models in terms of popular evaluation criteria.

1. Introduction

Aiming to mimic *human visual system* (HVS), which has the ability to effortlessly sort out the most attractive things from the scene in front of eyes, the goal of salient object detection is to calculate the most important objects in an image. For the moment, salient object detection can substantially facilitate a series of applications, such as image segmentation [1, 2], object recognition [3], image retrieval [4], image compression [5], and photo cropping [6].

By computing pixel or region uniqueness in either low-level cue or high-level cue, existing salient object detection models can be broadly divided into two types. (1) Bottom-up models are usually unsupervised and based on local contrast or global contrast. These methods tend to suffer from false detections in the context of cluttered background and less effective visual features. (2) Top-down models mainly leverage supervised learning to guide object detection. However, the complexity of the algorithm and the diversity of objectives limit the generality of these methods.

Although a large number of bottom-up and top-down salient object detection models have been proposed, most of them are only designed for normal light scenes. These saliency models are confronted with significant challenges in weak light images due to low signal-to-noise ratio and lack of well-defined features to capture saliency information in low lighting scenarios. The most likely reasons may attribute to two aspects: (1) current hand-crafted visual features can hardly evaluate the objectness in weak light images; (2) most of the high-level features normally present enormous challenges in detecting accurate object boundary information, which can be easily blurred due to multiple levels of convolution layers and pooling layers in common convolutional neural network models.

To address these challenges, this paper proposes an optimal feature selection-based saliency seed propagation model for salient object detection in weak light images (the code of this paper can be downloaded from https://drive. google.com/open?id=1w0qBapNVygh8TOOp7AijFWOxsY xdRcWa). Several hand-crafted visual features are selected to hierarchically refine the saliency map obtained from the high-level cues recursively. The flowchart of our model is presented in Figure 1. The optimal low-level features are first selected to give a robust expression for weak light images,
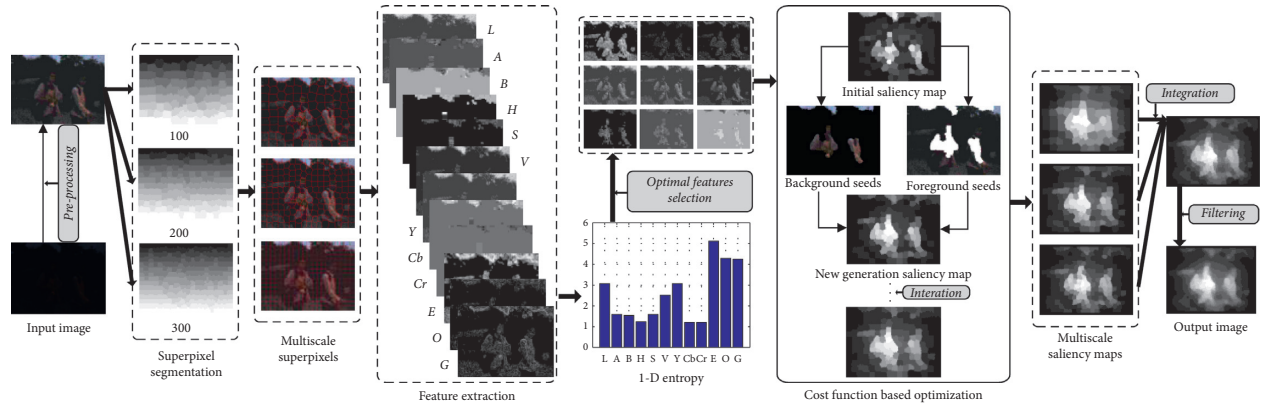
Figure 1: Overview of the proposed salient object detection model in the weak light image.

which aims to capture more objectness information and contributes to the prediction of salient objects under weak light conditions. Next, two cost functions are introduced to iteratively optimize the foreground seeds and background seeds of the initial saliency map, which can continuously compensate for salient information and remove nonsalient information to generate more precise object details. To estimate the overall performance, the proposed model is compared with 20 state-of-the-art salient object detection models on six datasets.

The paper is an extended version of our previously accepted conference paper [7], which provides a more detailed explanation and richer experimental demonstration. To sum up, this research has four main contributions: (1) a bottom-up visual saliency model, which requires no training, is explored toward weak light images, (2) an effective feature selection strategy is put forward to provide a robust representation of saliency information, (3) two cost functions are built to refine the initial saliency seeds recursively, and (4) a *nighttime image* (NI) dataset (the nighttime image (NI) dataset can be downloaded from https://drive.google.com/open?id=0BwVQK2zsuAQwX2hXbnc3ZVMzejQ) is constructed to verify the performance of our model.

The rest of this paper is organized as follows. Section 2 reviews the related works of saliency detection. Section 3 introduces the proposed saliency model. Section 4 presents the experimental results of the state-of-the-art models and the proposed model on six datasets. The conclusion of this paper is given in Section 5.

2. Related Works

Numerous salient object detection models have been proposed recently (see [8] for review); the main task of them is to highlight the most important visual regions for further processing. Depending on whether the task-independent or task-dependent is considered, they can be categorized as the bottom-up models and the top-down models, respectively.

Bottom-up saliency models are stimuli-driven and rely on low-level features. One typical model was presented by Itti et al. [9], which is mainly based on the center-surround difference of multiple features. Following this pioneering work, various bottom-up saliency models were proposed. Goferman et al. [10]

computed the saliency value of image patches by implementing the local and global contrasts. Cheng et al. [11] executed the saliency computation by calculating the histogram and region contrasts. Xu et al. [12] introduced the contrast and spatial distribution strategies to evaluate the image saliency. Kim et al. [13] estimated the local saliency and global saliency based on regression and high-dimensional color transform. Hu et al. [14] performed salient object detection by utilizing the compactness hypothesis of color feature and texture feature. Huang and Zhang [15] presented a minimum directional contrast based salient object detection method. Wang et al. [16] exploited the pyramid attention and salient edges to guide the salient object detection. Sun et al. [17] detected the salient objects by employing a cascaded bottom-up feature aggregation module to capture the detailed information of low-level features. Jiang et al. [18] proposed a task-independent saliency model based on the bidirectional absorbing Markov chains. Molin et al. [19] exploited a neuromorphic dynamic bottom-up saliency detection method, which is feed-forward and requires no training. Typically, these bottom-up saliency models tend to face many difficult problems in handling images of a busy background and struggle to predict the true salient objects, which are in a low-contrast weak light environment.

Top-down saliency models are task-driven and rely on high-level perceptual learning. Xu et al. [20] used the *support vector machine* (SVM) model to produce the superpixel-level saliency map. Qu et al. [21] proposed a deep learning-based salient object detection model by combining the superpixel-based Laplacian propagation and the trained *convolutional neural network* (CNN) model. Mu et al. [22] designed a region covariance-based CNN method to learn the saliency value of image patches. Wang et al. [23] employed the top-down process for coarse-to-fine saliency estimation. Mu et al. [24] explored global convolutional and boundary refinement in a top-down manner to guide the learning of salient objects. Qiu et al. [25] introduced an *automatic top-down fusion* (ATDF) saliency model, which utilizes the global information to guide the learning of underlying knowledge. Zhang et al. [26] developed a top-down multilevel fusion method for RGB-D salient object detection. Wang et al. [27] progressively optimized the salient objects by exploiting the fixation map in a top-down mode. Xu et al. [28] utilized a *progressive architecture with a knowledge*

review network (PA-KRN) for salient object detection, which compensates for the important information in a top-down way. Dong et al. [29] presented a *bidirectional collaboration network* (BCNet) for salient object detection, which integrates feature fusion and feature aggregation in an edge-guided top-down progressive pathway. These top-down saliency models generally have high computational complexity and are relatively ineffective in determining accurate boundary and localization of salient objects under weak light conditions.

Since saliency detection in a weak light environment is a challenging problem, there were few studies on the salient object detection of weak light images [30, 31]. Mu et al. [30] proposed an *ant colony optimization* (ACO) based saliency model for predicting the salient objects on weak light images. Xu et al. [31] explored an image enhancement method for salient object detection in weak light images. These saliency models, however, are not robust enough to capture the salient objects in real-time. Different from these previous methods, the proposed model creates a totally unsupervised algorithm by integrating the bottom-up measures and the single-objective optimization cues. Specifically, (1) the proposed saliency model explores low-level features to represent the object properties and selects the most effective ones based on the entropy information; (2) the superpixel-level saliency is directly estimated by the feature dissimilarity and spatial similarity; (3) the prior saliency map, which contains the foreground seeds and the background seeds, is generated by implementing the bottom-up measures; (4) the single-objective optimization cues are formulated by designing the fitness-based cost functions to iteratively optimize the salient and nonsalient seeds; and (5) experimental results indicate that the proposed model can generate a high-performance saliency map in real time.

3. The Proposed Saliency Model

The proposed optimal feature selection-guided saliency seed propagation model is presented in detail in this section. The input image is first segmented into superpixels at three scales. Then, 12 features are extracted from the pre-processing image, and only nine optimal ones are chosen for the next calculation. Next, the initial saliency map is computed by combining the global contrast and center prior. And then, the new saliency map can be obtained by the foreground and background seeds from the previous one. Two cost functions which are based on global fitness and local fitness are defined to control the end of the iteration. At last, the optimal saliency map is obtained, and the results of three scales are integrated to get the final saliency map.

3.1. Multiscale Superpixel Segmentation. To make full use of the midlevel information and preserve the object structure context of the input image, the *simple linear iterative clustering* (SLIC) algorithm [32] is used to divide the input image into N superpixels (denoted as $\{s_i\}$, $i = 1, \ldots, N$). This operation can boost the efficiency of the method by regarding the superpixel as a processing unit. For saliency

detection, the background region is more likely to have semblable superpixels at different scales, while the salient regions may have similar superpixels at some scales. That is, the fusion of the acquired salient superpixels at different scales can more accurately represent the real salient regions. However, as the number of superpixels increases, the time required for superpixel segmentation also increases. For accuracy and efficiency, our model generates the superpixels at three different scales, where the superpixel number N is set to 100, 200, and 300, respectively. The final saliency map is the integration of the obtained multiscale saliency maps.

3.2. Effective Feature Extraction. Given an input image, 12 low-level visual features are extracted, containing nine color features in three color spaces, the texture feature based on local entropy information, the orientation feature fused by the information in four directions, the gradient feature obtained from the horizontal and vertical vectors. Since the effectiveness of these various features varies according to the contrasts of different input images, nine optimal ones are selected from the 12 features, and the adaptive selection strategy is mainly based on the global information entropy of these features. The feature extraction process is introduced in detail as follows.

3.2.1. Color Features. The input image is first normalized to eliminate the interference of shadow and light (see pre-processing in Figure 1). This preprocessing is a general procedure in our model, including processing both normal light images and weak light images. Then, the input image is transformed from RGB color space to LAB, HSV, and YCbCr color spaces to capture nine color features. The L, A, and B components of LAB color space can describe all colors visible to the human eye, which are closer to human visual perception in weak light images. The H, S, and V components of HSV color space can be very intuitive to represent the hue, depth, and bright degree, which have good robustness in low lightness and weak light images. The Y, Cb, and Cr components of YCbCr color space can better perceive the intensity changes and the chromatic differences, which are more conducive to highlight the salient object information in weak light images.

3.2.2. Texture Feature. The 2-dimensional entropy of the original image is mainly used to represent the texture feature. Let I, $(0 \leq I \leq 255)$ denote the gray value of an image pixel, and let J, $(0 \leq J \leq 255)$ denote the average gray value of its neighborhood pixels; the spatial synthesis characteristic of gray distribution can be expressed as follows:

$$p_{IJ} = \frac{f(I, J)}{R^2}, \tag{1}$$

where $f(I, J)$ is the frequency of the characteristic tuple (I, J) and R^2 is the size of the neighborhood region. The discrete 2-dimensional entropy of the input image is defined as follows:

$$E = \sum_{I=0}^{255} p_{IJ} \log p_{IJ}. \tag{2}$$

Since the entropy information has strong resistance against noise interference and geometric deformation, the texture feature changes of salient objects in the weak light image can be well estimated by the variations in entropy.

3.2.3. Orientation Feature.

The orientation feature is computed by executing the Gabor filter of different directions (denoted as $g_\theta(x, y)$) on the grayscale image (denoted as $gray(x, y)$) via

$$O = \sum_{\theta \in \{0°, 45°, 90°, 135°\}} gray(x, y) * g_\theta(x, y). \tag{3}$$

The rotational invariance and the global property of the orientation feature make it have less impact from weak light scenes.

3.2.4. Gradient Feature.

The gradient feature is calculated by averaging the vertical gradient and horizontal gradient via

$$G = |gray(x + 1, y) - gray(x, y)| \\ + |gray(x, y + 1) - gray(x, y)|. \tag{4}$$

Thus, the magnitude information of local grayscale changes can be represented by the gradient feature, which can overcome the interference of a low signal-to-noise ratio in the weak light image.

3.2.5. Optimal Feature Selection.

Feature selection plays an important role in predicting the real salient objects in weak light images. Gopalakrishnan et al. [33] proposed an unsupervised feature selection method, which removes the irrelevant features by maximizing the mixing rate of Markov processes of different features. However, naive inclusion of irrelevant features for a particular image can easily lead to performance degradation. Liang et al. [34] explored feature selection methods in supervised saliency learning, the features utilized in the model are highly redundant. Naqvi et al. [35] selected useful features by measuring the feature quality. However, they use a large number of features trying to explain all possible saliency-related factors, which increases the time cost and ignores some truly effective features. Since the goal of our model is to identify a small set of optimal features, with which the salient object detection in the weak light image can be both efficient and effective, traditional adaptive feature selection techniques are not suitable for us. The proposed model mainly extracts 12 features to participate in the salient object calculation. Due to the fact that the effectiveness of each feature is different when the image contrast changes, which can be seen in Figure 2, nine optimal features (denoted as $\{F_k\}$, $k = 1, \ldots, 9$) that can better describe the attributes of the corresponding weak light image are then selected from the extracted 12 different visual features $\{L, A, B, H, S, V, Y, Cb, Cr, E, O, G\}$ by calculating the 1-dimensional entropy information of these feature maps as follows:

$$entropy = \sum_{I=0}^{255} p_I \log p_I, \tag{5}$$

where p_I denotes the proportion of image pixels and I denotes the grayscale values of these pixels.

As a statistical feature form, the mean information content contained in the aggregation properties of image grayscale distribution can be well represented by image entropy information. The greater the entropy of the feature map $\{F_k\}$ is, the more efficient this feature will be. Thus, the selected nine optimal features could better account for the visual saliency of the corresponding weak light image.

3.3. Initial Saliency Map Generation.

The global contrast measure and the spatial relationship strategy of the feature map are calculated to estimate the saliency value of each superpixel as follows:

$$Sal(s_i) = \left(\sum_{j=1, j \neq i}^{N} \frac{\sqrt{(F_k(s_i) - F_k(s_i))^2}}{1 + pos(s_i, s_j)} \right) \times c(s_i), \\ c(s_i) = \exp\left(-\frac{(x_i - x')^2}{2v_x^2} - \frac{(y_i - y')^2}{2v_y^2} \right), \tag{6}$$

where $pos(s_i, s_j)$ is the Euclidean distance between superpixels s_i and s_j. $c(s_i)$ denotes the spatial distance between the coordinate (x_i, y_i) and image center (x', y'). v_x and v_y are variables, which are decided by the vertical and horizontal information of the input image.

3.4. Saliency Map Optimization.

To achieve clean and uniform salient objects, optimization strategies are considered to improve detection accuracy. Zhu et al. [36] presented a principled optimization structure to fuse multiple low-level saliency cues, the whole framework mainly relies on the background cues, and it does not work well in weak light images, of which the background information is cluttered. Lu et al. [37] devoted to learning optimal saliency seeds set by utilizing a large margin formulation of discriminant saliency criterion. However, the gradient descent they used is not robust in weak light images and is not efficient for high accuracy salient object detection. In the proposed model, we built two cost functions to refine the generated saliency seeds recursively, which is an effective and straightforward manner to obtain more accurate salient objects in weak light images. The initial saliency map (denoted as $Smap_k$, $k = 0$) is first segmented into the salient region and nonsalient region by utilizing Otsu's thresholding [38]. The salient region and nonsalient region can be seen as the foreground seeds (denoted as FS) and the background seeds (denoted as BS) of the input image, respectively. The larger the difference between the superpixel and the foreground region is, the lower the saliency value of this superpixel is. Conversely, the greater the difference between the superpixel and the background region is, the higher the saliency value of this superpixel will be. Thus, the saliency value of s_i can be updated based on foreground seeds FS and background seeds BS as follows:
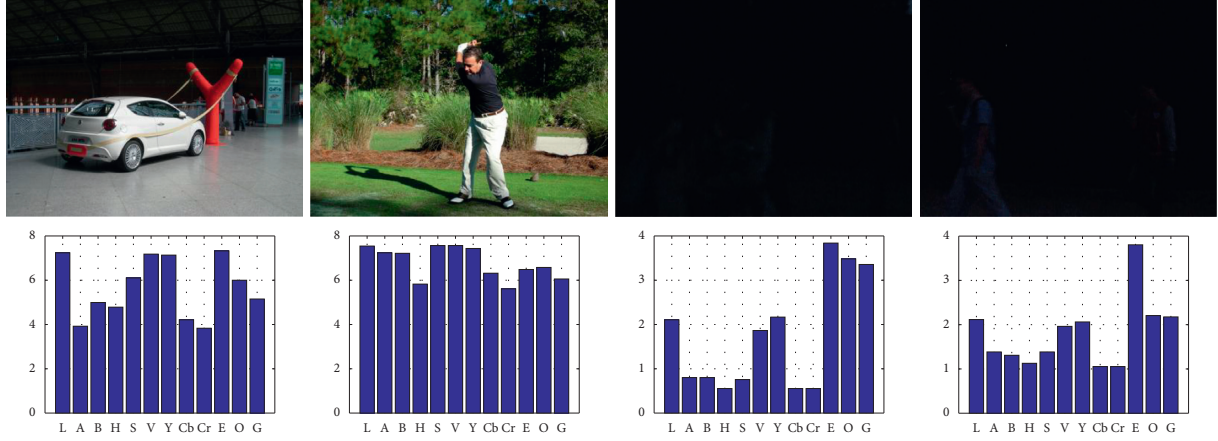
FIGURE 2: The top row contains the test images with different contrasts. The bottom row is the corresponding entropy bar graphs of 12 different features.

$$\text{Sal}_{\text{FS}}(s_i) = \sum_{s_j \in \text{FS}, j \neq i} \frac{1}{\left(\sqrt{\left(F_k(s_i) - F_k(s_j)\right)^2} + \text{pos}(s_i, s_j)\right)}, \tag{7}$$

$$\text{Sal}_{\text{BS}}(s_i) = \sum_{s_j \in \text{BS}, j \neq i} \frac{\sqrt{\left(F_k(s_i) - F_k(s_j)\right)^2}}{\left(1 + \text{pos}(s_i, s_j)\right)}, \tag{8}$$

$$\text{Sal}(s_i) = \left(1 - \exp\left(-\frac{\text{Sal}_{\text{FS}}(s_i) + \text{Sal}_{\text{BS}}(s_i)}{2}\right)\right) \times c(s_i). \tag{9}$$

Then, a new saliency map (denoted as Smap_k, $k = 1$) of the first iteration optimization is obtained. The Otsu's method is reused to generate new FS and BS; the saliency map of the next generation (denoted as Smap_{k+1}) can be computed according to (7-9). Finally, two cost functions are implemented to decide whether the iteration procedures meet the end condition or not:

$$\text{minimize} \begin{cases} f_1(k) = \left(\text{Smap}_k - \text{Smap}_{k-1}\right)^2 \\ f_2(k) = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\left(\text{Sal}(s_i) - \text{Sal}(s_j)\right)^2}{1 + \text{pos}(s_i, s_j)} \end{cases}, \quad \text{where } k \geq 1, \ s_i, s_j \in \text{Smap}_k, \ 1 \leq i, \ j \leq N. \tag{10}$$

The function $f_1(k)$ mainly represents the global fitness, which denotes that the smaller the change between the saliency map of the new generation Smap_k and the previous generation Smap_{k-1} is, the more optimization of the objective can be. The function $f_2(k)$ mainly represents the local fitness, which denotes that the smaller the difference between the superpixel $\text{Sal}(s_i)$ and its neighboring superpixels $\text{Sal}(s_j)$ is, the better the saliency information of each decision variable can be. By minimizing the two functions $f_1(k)$ and $f_2(k)$, the optimal superpixel-level saliency map can be obtained.

4. Experiment Results

Comprehensive experiments are carried out on six datasets to estimate the performance of our model against 20 state-of-the-art salient object detection models.

4.1. Experimental Setup

4.1.1. Testing Datasets.
The six test datasets contain five public datasets and the proposed weak light image dataset as follows: (1) the MSRA dataset [39] includes 10000 images which have relatively high contrast and only simple background; (2) the SOD dataset [40] includes various images of multiple objects and complex background; (3) the CSSD dataset [41] includes complex natural scenes; (4) the DUT-OMRON dataset [42] includes complex and challenging images; (5) the PAS-CAL-S dataset [43] includes images of cluttered background; and (6) our NI dataset includes 200 weak light images, which are captured at night with a stand camera. The resolution of these images is 640×480, and the human-annotated *ground-truths* (GTs) are also given.
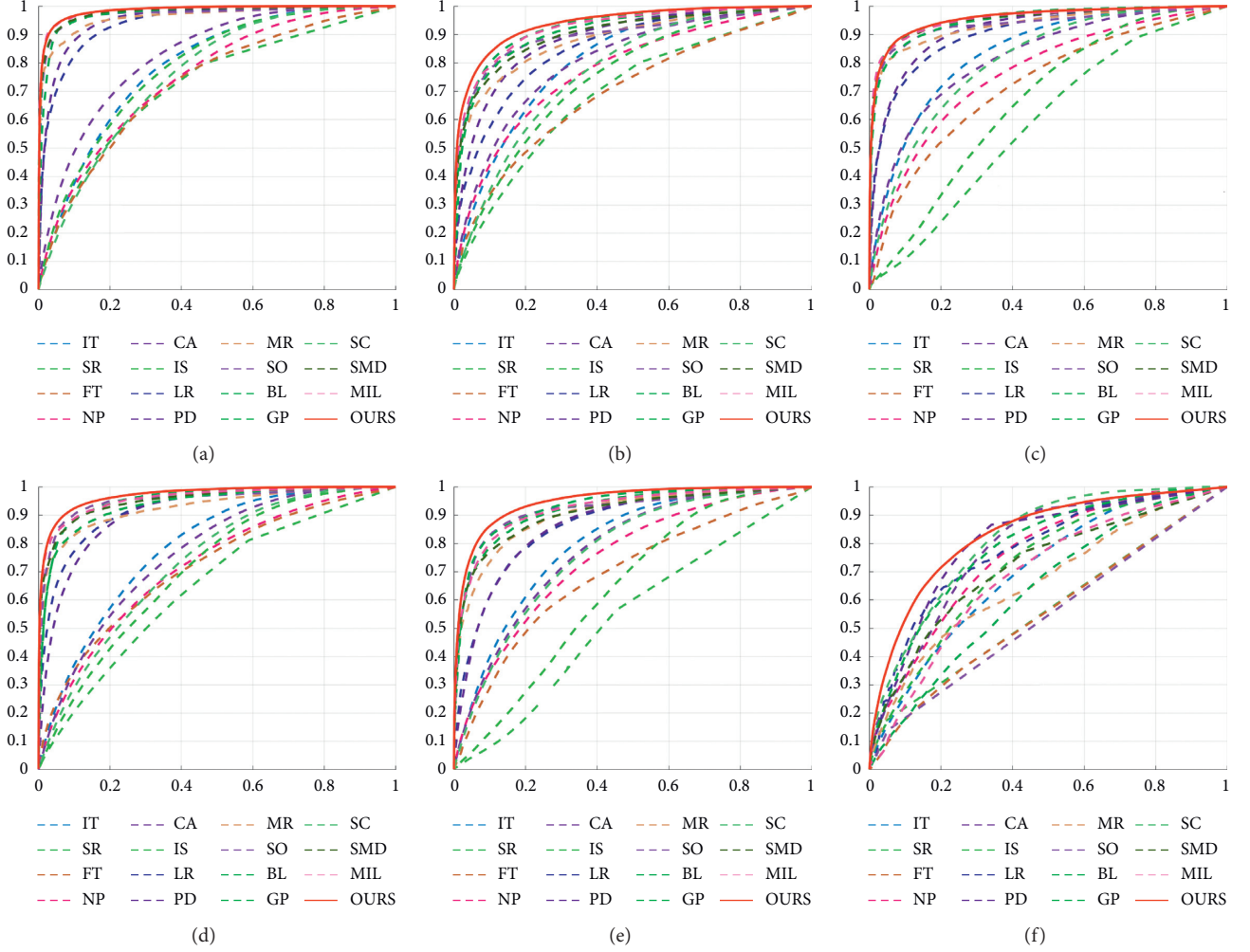
FIGURE 3: The TPRs-FPRs curves performance comparisons of different saliency models on the six datasets. (a) MSRA dataset. (b) SOD dataset. (c) CSSD dataset. (d) DUT-OMRON dataset. (e) PASCAL-S dataset. (f) NI dataset.

4.1.2. Comparison Models. The first 15 state-of-the-art saliency models include: *Itti*'s (IT) model [9], *spectral residual* (SR) model [44], *frequency-tuned* (FT) model [45], *nonparametric* (NP) model [46], *context-aware* (CA) model [10], *image signature* (IS) model [47], *low rank matrix recovery* (LR) model [48], *patch distinct* (PD) model [49], *graph-based manifold ranking* (MR) model [42], *saliency optimization* (SO) model [36], *bootstrap learning* (BL) model [50], *generic promotion* (GP) model [51], *spatiochromatic context* (SC) model [52], *structured matrix decomposition* (SMD) model [53], and *multiple-instance learning* (MIL) model [54]. All these experiments are performed by MATLAB software on an Intel i5-5250 CPU (1.6 GHz) PC with 8 GB RAM.

4.1.3. Evaluation Criteria. To estimate the overall performance of various saliency models, seven criteria are used, including the *true positive rates* and *false positive rates* (TPRs-FPRs) curve, the *precision-recall* (PR) curve, the *area under the curve* (AUC) score, the *mean absolute error* (MAE) score, the *weighted F-measure* (WF) score, the *overlapping ratio* (OR) score, and the average execution time per image (in seconds).

The TPR is defined as the ratio of salient pixels that are correctly detected to all the true salient pixels, and FPR corresponds to the ratio of falsely detected salient pixels to all the true nonsalient pixels. The precision is computed as the ratio of correctly detected salient pixels to all the detected salient pixels, and the recall is the same as TPR, which measures the comprehensiveness of the detected salient pixels. By varying the threshold over the obtained saliency map, different TPRs, FPRs, precisions, and recalls can be calculated by comparing the generated different binary images with GT via

$$
\begin{aligned}
\text{TPR} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\
\text{FPR} &= \frac{\text{FP}}{\text{FP} + \text{TN}}, \\
\text{precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\
\text{recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}},
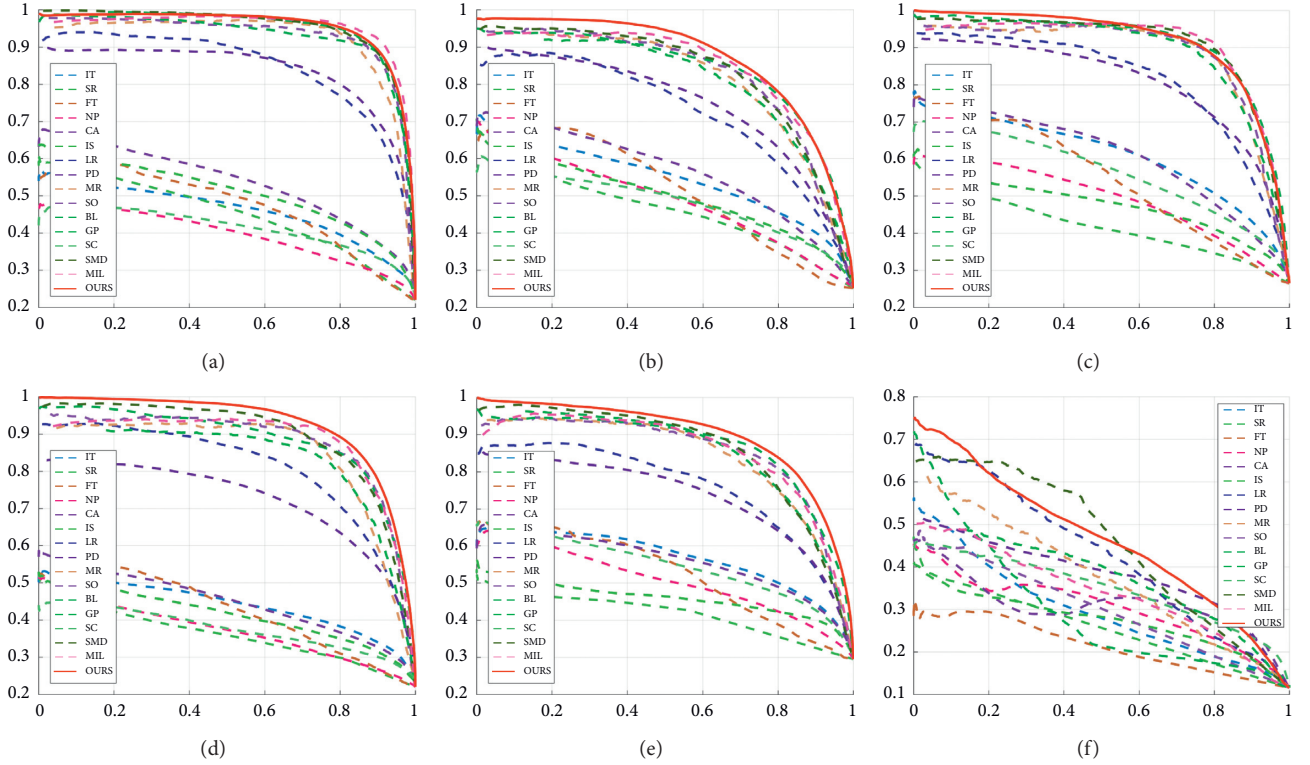\end{aligned}
\tag{11}
$$

FIGURE 4: The PR curves performance comparisons of different saliency models on the six datasets. (a) MSRA dataset. (b) SOD dataset. (c) CSSD dataset. (d) DUT-OMRON dataset. (e) PASCAL-S dataset. (f) NI dataset.

where the *true positive* (TP) is the collection of pixels that correctly identify the salient object; the *false positive* (FP) is the collection of pixels that falsely identify the salient object; the *true negative* (TN) is the collection of pixels that correctly identify the nonsalient pixels; and the *false negative* (FN) is the collection of pixels which falsely identify the nonsalient pixels.

The TPRs-FPRs curve and the PR curve can be generated by plotting the corresponding ratios. The AUC score is calculated by measuring the proportion of the area under the TPRs-FPRs curve, which can give an intuitive indication of how well the obtained saliency map represents the real salient objects. The MAE score is calculated as the average absolute difference between the generated saliency map (denoted as Salmap) and the ground-truth (denoted as GT) via

$$MAE = mean\left(|Salmap\left(x, y\right) - GT\left(x, y\right)|\right). \quad (12)$$

The smaller the MAE value is, the higher the similarity between Salmap and GTis. The F-measure score is computed as the weighted harmonic mean of the precision and the recall via

$$F_{\beta} = \frac{\left(1 + \beta^2\right)precision \cdot recall}{\beta^2 \cdot precision + recall}, \quad (13)$$

where $\beta^2 = 0.3$ is the parameter to weigh the precision and recall. The WF score is calculated by adding a weighting function to the detection errors [55].

The OR score is measured by computing the overlapping ratio of salient pixels between the binary saliency map (denoted as Bmap) and GT via

$$OR = \frac{|Bmap\left(x, y\right) \cap GT\left(x, y\right)|}{|Salmap\left(x, y\right) \cup GT\left(x, y\right)|}. \quad (14)$$

4.2. Experimental Results. The quantitative performances of our salient object detection method against the other 15 saliency models on the six datasets are presented in Figure 3 and 4 and Tables 1–6. The best three experimental results of Table 1–6 are highlighted in the red, blue, and green fonts, respectively. In particular, the up-arrow ↑ denotes the larger the value is, the better the performance of the saliency model is. At the same time, the down-arrow ↓ indicates the opposite meaning. As shown in the quantitative results, our salient object detection model performs the first or second performance on the five public datasets in most cases and obtains the best performance on the NI dataset in a relatively low time-consuming.

On MSRA, DUT-OMRON, and PASCAL-S datasets ((a), (d) and (e) of Figures 3 and 4 and Tables 1–3), our model achieves the best performance on the TPRs-FPRs curve, PR curve, and AUC score, while the saliency model SO obtains the best MAE score and WF score, and the saliency model MIL obtains the best OR score. The main reason is that the SO model used boundary connectivity and global optimization to increase its robustness, and the MIL model introduced a multiple-instance learning approaches to

TABLE 1: Quantitative performance of various salient object detection models using five criteria on the MSRA dataset.

Criteria	IT	SR	FT	NP	CA	IS	LR	PD	MR	SO	BL	GP	SC	SMD	MIL	OURS
AUC↑	0.7261	0.5556	0.6682	0.6942	0.7580	0.7375	0.8462	0.8557	0.8544	0.8639	**0.8729**	0.8714	0.6922	0.8712	**0.8751**	*0.8793*
MAE↓	0.3475	0.2204	0.2659	0.4211	0.2700	0.3042	0.2240	0.1869	0.1067	*0.0934*	0.1548	0.1099	0.2841	**0.0956**	0.0963	**0.0945**
WF↑	0.2681	0.1032	0.2886	0.2668	0.3316	0.2860	0.4327	0.4772	0.6935	*0.7314*	0.5765	0.6853	0.2451	**0.7250**	0.7125	0.7274
OR↑	0.1551	0.2490	0.3002	0.0703	0.3324	0.2569	0.5515	0.6043	0.7688	**0.8044**	0.7541	0.7861	0.2109	0.8021	*0.8182*	**0.8068**
TIME↓	1.902	1.797	0.172	1.480	78.61	0.231	29.95	4.462	0.684	0.316	12.95	1.428	31.57	2.130	101.5	7.437

The italic values and bold values indicate the best performance.

Table 2: Quantitative performance of various salient object detection models using five criteria on the DUT-OMRON dataset.

Criteria	IT	SR	FT	NP	CA	IS	LR	PD	MR	SO	BL	GP	SC	SMD	MIL	OURS
AUC↑	0.7263	0.5325	0.6611	0.6669	0.7007	0.6605	0.8300	0.8165	0.8203	0.8441	**0.8523**	0.8291	0.6700	0.8441	**0.8489**	*0.8657*
MAE↓	0.3486	0.2301	0.2675	0.4315	0.2898	0.3461	0.2322	0.2149	0.1343	*0.1141*	0.1772	0.1441	0.2882	**0.1209**	0.1255	**0.1194**
WF↑	0.2663	0.0869	0.2758	0.2572	0.2892	0.2546	0.4107	0.4128	0.6085	*0.6645*	0.5297	0.6040	0.2291	**0.6443**	0.6273	**0.6560**
OR↑	0.1434	0.1810	0.2825	0.0543	0.2566	0.1662	0.4953	0.4650	0.6669	**0.7048**	0.6580	0.6401	0.1905	0.6975	0.7085	**0.6992**
TIME↓	1.99	1.962	0.140	1.385	54.16	0.238	30.36	4.305	0.687	0.437	18.53	1.411	35.28	1.861	114.4	8.749

The italic values and bold values indicate the best performance.

TABLE 3: Quantitative performance of various salient object detection models using five criteria on the PASCAL-S dataset.

Criteria	IT	SR	FT	NP	CA	IS	LR	PD	MR	SO	BL	GP	SC	SMD	MIL	OURS
AUC↑	0.7220	0.4942	0.6363	0.6736	0.6985	0.6011	0.7715	0.7525	0.7672	*0.7820*	**0.8078**	**0.7949**	0.6880	0.7817	0.7888	*0.8177*
MAE↓	0.3529	0.3074	0.3096	0.4083	0.3187	0.3925	0.2754	0.2580	0.1963	*0.1787*	0.2096	0.1920	0.3162	**0.1810**	0.1890	**0.1801**
WF↑	0.3289	0.0453	0.2977	0.3358	0.3511	0.2750	0.4071	0.4064	0.5514	*0.6037*	0.5459	0.5827	0.2825	**0.5917**	0.5714	**0.5966**
OR↑	0.1451	0.1358	0.2859	0.0777	0.2672	0.1419	0.3975	0.4330	0.5596	*0.5982*	0.5953	0.5752	0.2423	**0.6009**	0.6058	**0.6021**
TIME↓	2.57	2.564	0.189	1.487	99.38	0.383	60.78	7.810	0.833	0.556	25.40	3.104	27.70	2.232	140.9	12.07

The italic values and bold values indicate the best performance.

TABLE 4: Quantitative performance of various salient object detection models using five criteria on the SOD dataset.

Criteria	IT	SR	FT	NP	CA	IS	LR	PD	MR	SO	BL	GP	SC	SMD	MIL	OURS
AUC↑	0.7377	0.5537	0.6436	0.7038	0.7337	0.7044	0.7820	0.7820	0.7648	0.7725	**0.8181**	**0.8063**	0.6969	0.7923	0.8006	0.8292
MAE↓	0.3316	0.2555	0.2674	0.4102	0.2867	0.3490	0.2569	0.2268	0.1732	*0.1630*	0.1981	0.1839	0.2912	**0.1664**	0.1695	**0.1655**
WF↑	0.2976	0.1026	0.2754	0.2939	0.3296	0.2775	0.3745	0.4056	0.5494	**0.5659**	0.5019	0.5426	0.2568	**0.5734**	0.5486	0.5739
OR↑	0.1878	0.2241	0.2958	0.0961	0.3135	0.1801	0.4093	0.4553	0.5438	0.5817	0.5516	0.5366	0.2316	**0.5909**	**0.5934**	0.5938
TIME↓	2.450	2.465	0.162	1.397	81.03	0.533	39.65	5.888	0.768	0.527	20.49	2.023	30.18	2.420	139.2	9.967

The italic values and bold values indicate the best performance.

Table 5: Quantitative performance of various salient object detection models using five criteria on the CSSD dataset.

Criteria	IT	SR	FT	NP	CA	IS	LR	PD	MR	SO	BL	GP	SC	SMD	MIL	OURS
AUC↑	0.7606	0.5087	0.6570	0.6967	0.7466	0.6259	0.8062	0.8088	0.8023	0.8044	**0.8384**	0.8212	0.7187	**0.8256**	0.8215	*0.8393*
MAE↓	0.3336	0.2905	0.2706	0.4183	0.2960	0.4170	0.2449	0.2335	0.1479	**0.1366**	0.1730	0.1519	0.2943	*0.1314*	0.1428	**0.1407**
WF↑	0.3239	0.1059	0.2931	0.3096	0.3459	0.2745	0.4260	0.4380	0.6283	**0.6628**	0.5651	0.6253	0.2821	*0.6749*	0.6366	**0.6568**
OR↑	0.2086	0.1775	0.2978	0.0690	0.3267	0.1138	0.4795	0.4944	0.6425	0.6654	0.6388	0.6327	0.2658	0.6851	**0.6712**	**0.6701**
TIME↓	1.97	1.878	0.159	1.370	65.28	0.363	21.95	4.292	0.668	0.431	21.78	1.277	42.80	2.293	109.4	7.052

The italic values and bold values indicate the best performance.

TABLE 6: Quantitative performance of various salient object detection models using five criteria on the NI dataset.

Criteria	IT	SR	FT	NP	CA	IS	LR	PD	MR	SO	BL	GP	SC	SMD	MIL	OURS
AUC↑	0.6721	0.5117	0.5072	0.7079	0.6143	0.6433	**0.7545**	0.6123	0.6726	0.5379	0.6253	**0.7534**	0.7095	0.6419	0.5840	*0.7931*
MAE↓	0.2407	**0.1162**	0.1252	0.2178	0.1329	0.1864	0.2193	**0.1291**	0.3609	0.1357	0.3993	0.2778	0.1552	0.1580	0.1746	*0.1049*
WF↑	0.1392	0.0269	0.0359	0.1983	0.1432	0.1565	0.2114	0.1722	0.1671	0.0741	0.1398	**0.2226**	0.1647	**0.2177**	0.1551	*0.2298*
OR↑	0.1617	0.0983	0.0968	0.2295	0.2370	0.1931	**0.2636**	0.2187	0.1431	0.0624	0.1037	0.2475	**0.2599**	0.2413	0.1837	*0.2957*
TIME↓	4.279	4.744	0.293	4.941	107.1	0.807	179.6	13.64	1.387	1.274	94.297	6.520	34.16	4.978	191.1	13.20

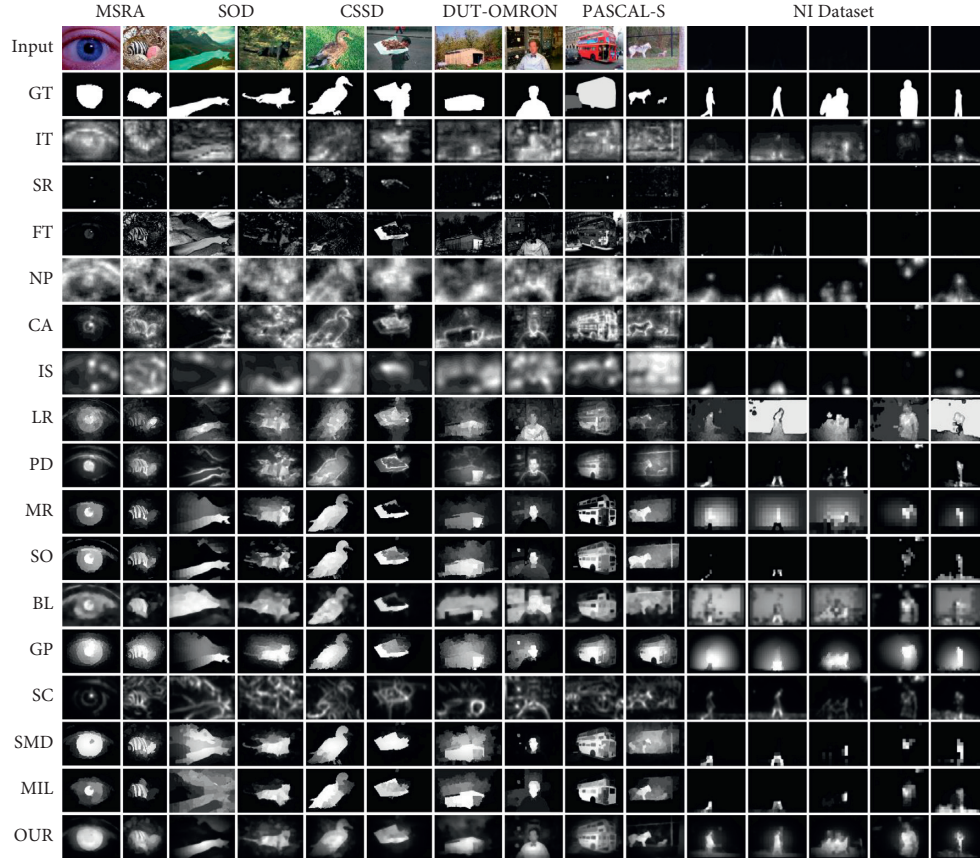The italic values and bold values indicate the best performance.

FIGURE 5: Visual comparisons of saliency maps construction using different approaches on the six datasets.

increase precision. These two saliency models take full advantage of the background measures, which can be of some effect in detecting the salient object under complex background conditions. Although the MAE score, WF score, and OR score of the proposed saliency model are slightly lower than the two models SO and MIL, our detection results are more competitive than the other models. The average time consumption of the MIL model is more than 100 seconds per image, which is not efficient in generating the saliency map.

On the SOD dataset (Figures 3(b) and 4(b), and Table 4), our saliency model has the best performance on the TPRs-FPRs curve, PR curve, AUC score, WF score, and OR score. In terms of the MAE criterion, the proposed model performs the second-best performance, which only has a small gap (0.0025) with the best MAE score of the SO model.

On the CSSD dataset (Figures 3(c) and 4(c), and Table 5), our saliency model performs the best performance on the TPRs-FPRs curve, PR curve, and AUC score. The MAE, WF, and OR scores of the proposed model are slighter than the best results achieved by the SMD model. The SMD model is based on the structured matrix decomposition with two regularizations, which has a strong potential in detecting the image of complex environments. The main reason for the poor performance of the proposed model on these metrics is that the selected optimal features contain less useful information that can effectively distinguish the salient objects.

On the NI dataset (Figures 3(f) and 4(f), and Table 6), our saliency model is superior, as it achieves the best performance on these criteria with a relatively short time-consuming.

The qualitative comparisons of saliency maps generated by the various salient object detection models on the six datasets are shown in Figure 5, indicating that our saliency model can detect the real salient object accurately in complex and/or weak light images (more detected saliency maps can be downloaded from https://drive.google.com/open?id=0BwVQK2zsuAQwQjZHeUJ1dlBsQms).

Since the standard real-world images and the weak light images have different properties, the proposed framework employs a feature selection strategy over the candidate feature set to pick out the most relevant features that apply to different types of images, which ensures that our model can be adapted to both standard saliency datasets and the weak light image dataset. In addition, we further optimize the saliency results through iteration to ensure robustness.

To further verify the effectiveness of our model, we have added some experiments with other five state-of-the-art deep learning-based saliency models (NLDF [56], LPS [57], BAS [58], F^{3}Net [59], and LDF [60]) to better illustrate the advantages of the proposed flowchart. The subjective performance comparisons of the proposed model with the latest deep saliency models are shown in Figure 6.
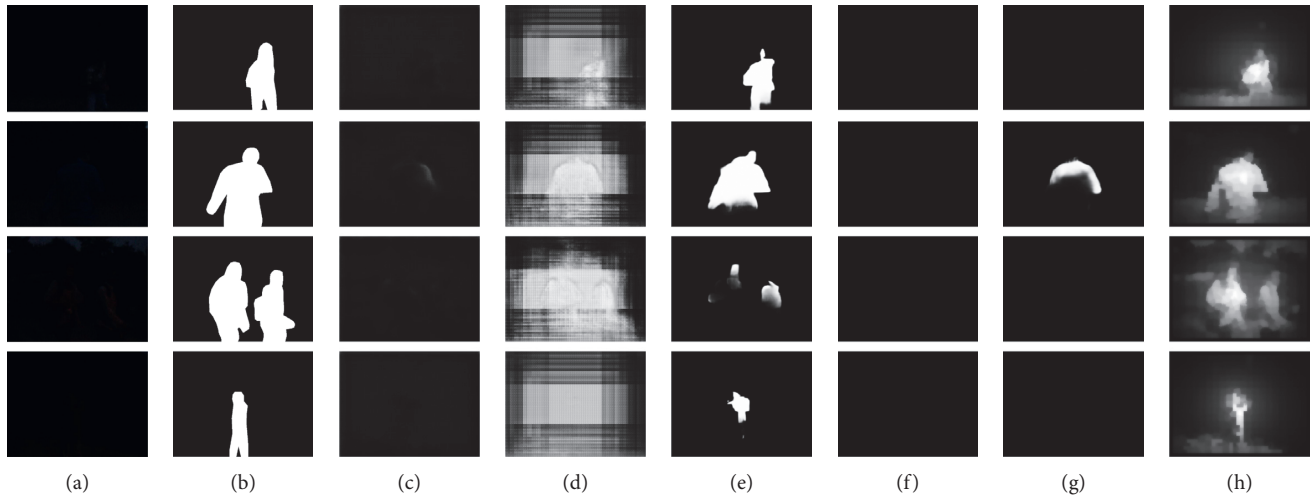
FIGURE 6: Visual comparisons of our saliency map with state-of-the-art deep learning-based saliency models on weak light image dataset. (a) Input. (b) GT. (c) NLDF. (d) LPS. (e) BAS. (f) F^{3}Net. (g) LDF. (h) OURS.

As can be seen in Figure 6, the saliency maps of the NLDF and F^{3}Net models cannot capture the effective salient objects in weak light images. The saliency results of the LPS model are seriously interfered with by the background noise. The saliency maps generated by the BAS model can highlight the salient objects with less noise, but the detected salient objects are incomplete. The LDF model has difficulty in detecting the whole objects and is prone to failure. Relatively speaking, the proposed model can accurately detect the real salient objects from the background on weak light images.

5. Conclusion

In this paper, we propose an optimal feature selection-based saliency seed propagation model to detect the salient object in weak light images. The main idea of the proposed saliency model is to execute saliency calculation by learning the optimal hand-crafted visual features and refining the foreground seeds and background seeds recursively. Guided by the optimized saliency seeds, the final saliency map can be achieved by fusing the multiple superpixel-level saliency maps at three different scales. Comprehensive experiments demonstrate that our saliency model performs satisfactory results against 20 state-of-the-art saliency models on five public datasets and a weak light image dataset.

Serving as a preprocessing step, salient object detection can efficiently focus on the most interesting area associated with the current visual task and it facilitates various computer vision applications such as image classification, object segmentation, visual tracking, etc. The proposed salient object detection model can be used to optimize correlational vision applications under weak light conditions, and it is of great application value to the monitoring system. In the future, we will further improve the time performance of the proposed saliency model and explore more potential applications.

Data Availability

The proposed nighttime image (NI) dataset, the code of the proposed model, and the experimental result data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] L. Wang, G. Hua, R. Sukthankar, J. Xue, Z. Niu, and N. Zheng, "Video object discovery and co-segmentation with extremely weak supervision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 10, pp. 2074–2088, 2017.

[2] W. Wang, J. Shen, and F. Porikli, "Saliency-aware video object segmentation," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 3395–3402, Boston, MA, USA, June 2015.

[3] H. Chenini, "An embedded FPGA architecture for efficient visual saliency based object recognition implementation," in *Proceedings of the International Conference On Systems And Control*, pp. 187–192, Batna, Algeria, May 2017.

[4] X. Xiyu Yang, X. Xueming Qian, and Y. Yao Xue, "Scalable mobile image retrieval by exploring contextual saliency," *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1709–1721, 2015.

[5] S. Li, M. Xu, Y. Ren, and Z. Wang, "Closed-form optimization on saliency-guided image compression for HEVC-MSP," *IEEE Transactions on Multimedia*, vol. 20, no. 1, pp. 155–170, 2018.

[6] W. Wang, J. Shen, and H. Ling, "A deep network solution for attention and aesthetics aware photo cropping," *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 7, pp. 1531–1544, 2019.

[7] N. Mu, X. Xu, and X. Zhang, "Optimal feature selection for saliency seed propagation in low contrast images," in *Proceedings of the Pacific Rim Conference On Multimedia*, pp. 35–45, Hefei, China, September 2018.

[8] W. Wang, Q. Lai, H. Fu, J. Shen, H. Ling, and R. Yang, "Salient object detection in the deep learning era: an in-depth survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1, 2021.

[9] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.

[10] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 10, pp. 1915–1926, 2012.

[11] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S.-M. Hu, "Global contrast based salient region detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 569–582, 2015.

[12] X. Xu, N. Mu, L. Chen, and X. Zhang, "Hierarchical salient object detection model using contrast-based saliency and color spatial distribution," *Multimedia Tools and Applications*, vol. 75, no. 5, pp. 2667–2679, 2016.

[13] J. Kim, D. Han, Y.-W. Tai, and J. Kim, "Salient region detection via high-dimensional color transform and local spatial support," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 9–23, 2016.

[14] P. Hu, W. Wang, C. Zhang, and K. Lu, "Detecting salient objects via color and texture compactness hypotheses," *IEEE Transactions on Image Processing*, vol. 25, no. 10, pp. 4653–4664, 2016.

[15] X. Huang, Y.-J. Zhang, and ", "FPS salient object detection via minimum directional contrast," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4243–4254, 2017.

[16] W. Wang, S. Zhao, J. Shen, S. C. H. Hoi, and A. Borji, "Salient object detection with pyramid attention and salient edges," in *Proceedings of the IEEE International Conference On Computer Vision And Pattern Recognition*, pp. 1448–1457, Long Beach, CA, USA, June 2019.

[17] F. Sun, L. Huang, X. Yuan, and C. Zhao, "Salient object detection via attention-aware cascaded bottom-up feature aggregation," in *Proceedings of the IEEE International Conference On Multimedia And Expo*, Shenzhen, China, July 2021.

[18] F. Jiang, B. Kong, J. Li, K. Dashtipour, and M. Gogate, "Robust visual saliency optimization based on bidirectional Markov chains," *Cognitive Computation*, vol. 13, no. 1, pp. 69–80, 2021.

[19] J. L. Molin, C. S. Thakur, E. Niebur, and R. Etienne-Cummings, "A neuromorphic proto-object based dynamic visual saliency model with a hybrid FPGA implementation," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 15, 2021.

[20] X. Xu, N. Mu, H. Zhang, and X. Fu, "Salient object detection from distinctive features in low contrast images," in *Proceedings of the IEEE International Conference On Image Processing*, pp. 3126–3130, Quebec City, Canada, September 2015.

[21] L. Qu, S. He, J. Zhang, J. Tian, Y. Tang, and Q. Yang, "RGBD salient object detection via deep fusion," *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2274–2285, 2017.

[22] N. Mu, X. Xu, X. Zhang, and H. Zhang, "Salient object detection using a covariance-based CNN model in low-contrast images," *Neural Computing & Applications*, vol. 29, no. 8, pp. 181–192, 2018.

[23] W. Wang, J. Shen, M.-M. Cheng, and L. Shao, "An iterative and cooperative top-down and bottom-up inference network for salient object detection," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 5968–5977, Long Beach, CA, USA, June 2019.

[24] N. Mu, X. Xu, and X. Zhang, "Salient object detection in low contrast images via global convolution and boundary refinement," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition Workshops*, pp. 1–9, Long Beach, CA, USA, June 2019.

[25] Y. Qiu, Y. Liu, H. Yang, and J. Xu, "A simple saliency detection approach via automatic top-down feature fusion," *Neurocomputing*, vol. 388, pp. 124–134, 2020.

[26] M. Zhang, Y. Zhang, Y. Piao, B. Hu, and H. Lu, "Feature reintegration over differential treatment: a top-down and adaptive fusion network for RGB-D salient object detection," in *Proceedings of the ACM International Conference On Multimedia*, pp. 4107–4115, Seattle, WA USA, October 2020.

[27] W. Wang, J. Shen, X. Dong, A. Borji, and R. Yang, "Inferring salient objects from human fixations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 8, pp. 1913–1927, 2020.

[28] B. Xu, H. Liang, R. Liang, and P. Chen, "Locate globally, segment locally: a progressive architecture with knowledge review network for salient object detection," in *Proceedings.of the AAAI Conference On Artificial Intelligence*, pp. 1–9, Vancouver, Canada, February 2021.

[29] B. Dong, Y. Zhou, C. Hu, K. Fu, and G. Chen, "BCNet: bi-directional collaboration network for edge-guided salient object detection," *Neurocomputing*, vol. 437, pp. 58–71, 2021.

[30] N. Mu, X. Xu, and X. Zhang, "Ant colony optimization based salient object detection for weak light images," in *Proceedings of the IEEE International Conference On Ubiquitous Intelligence And Computing*, pp. 1432–1437, Guangzhou, China, October 2018.

[31] X. Xu, S. Wang, Z. Wang, X. Zhang, and R. Hu, "Exploring image enhancement for salient object detection in low light images," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 17, no. 8, pp. 1–19, 2021.

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

[33] V. Gopalakrishnan, Y. Hu, and D. Rajan, "Unsupervised feature selection for salient object detection," in *Proceedings of the Asian Conference On Computer Vision*, pp. 15–26, Queenstown, New Zealand, November 2010.

[34] M. Ming Liang and X. Xiaolin Hu, "Feature selection in supervised saliency prediction," *IEEE Transactions on Cybernetics*, vol. 45, no. 5, pp. 914–926, 2015.

[35] S. S. Naqvi, W. N. Browne, and C. Hollitt, "Feature quality-based dynamic feature selection for improving salient object detection," *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, vol. 25, no. 9, pp. 4298–4313, 2016.

[36] W. Zhu, S. Liang, Y. Wei, and J. Sun, "Saliency optimization from robust background detection," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 2814–2821, Columbus, OH, USA, June 2014.

[37] S. Lu, V. Mahadevan, and N. Vasconcelos, "Learning optimal seeds for diffusion-based salient object detection," in

Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition, pp. 2790–2797, Columbus, OH, USA, June 2014.

[38] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[39] T. Liu, J. Sun, N.-N. Zheng, X. Tang, and H.-Y. Shum, "Learning to detect a salient object," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 1–8, Minneapolis, MN, USA, June 2007.

[40] V. Movahedi and J. H. Elder, "Design and perceptual validation of performance measures for salient object segmentation," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition Workshops*, pp. 49–56, San Francisco, CA, USA, June 2010.

[41] Q. Yan, L. Xu, J. Shi, and J. Jia, "Hierarchical saliency detection," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 1155–1162, Portland, OR, USA, June 2013.

[42] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 3166–3137, Portland, OR, USA, June 2013.

[43] Y. Li, X. Hou, C. Koch, J. Rehg, and A. Yuille, "The secrets of salient object segmentation," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 280–287, Columbus, OH, USA, June 2014.

[44] X. Hou and L. Zhang, "Saliency detection: a spectral residual approach," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 1–8, Minneapolis, MN, USA, June 2007.

[45] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 1597–1604, Miami, FL, USA, June 2009.

[46] N. Murray, M. Vanrell, X. Otazu, and C. A. Parraga, "Saliency estimation using a non-parametric low-level vision model," in *Proceedings of the IEEE Computer Conference On Computer Vision And Pattern Recognition*, pp. 433–440, Colorado Springs, CO, USA, June 2011.

[47] X. Xiaodi Hou, J. Harel, and C. Koch, "Image signature: highlighting sparse salient regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 194–201, 2012.

[48] X. Shen and Y. Wu, "A unified approach to salient object detection via low rank matrix recovery," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 853–860, Providence, RI, USA, June 2012.

[49] R. Margolin, L. Zelnik-Manor, and A. Tal, "What makes a patch distinct?" in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 1139–1146, Portland, OR, USA, June 2013.

[50] N. Tong, H. Lu, and M. Yang, "Salient object detection via bootstrap learning," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 1884–1892, Boston, MA, USA, June 2015.

[51] P. Jiang, N. Vasconcelos, and J. Peng, "Generic promotion of diffusion-based salient object detection," in *Proceedings of the IEEE International Conference On Computer Vision*, pp. 217–225, Santiago, Chile, December 2015.

[52] J. Zhang, M. Wang, S. Zhang, X. Li, and X. Wu, "Spatiochromatic context modeling for color saliency analysis," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 6, pp. 1177–1189, 2016.

[53] H. Peng, B. Li, H. Ling, W. Hu, W. Xiong, and S. J. Maybank, "Salient object detection via structured matrix decomposition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 818–832, 2017.

[54] F. Huang, J. Qi, H. Lu, L. Zhang, and X. Ruan, "Salient object detection via multiple instance learning," *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1911–1922, 2017.

[55] R. Margolin, L. Zelnik-Manor, and A. Tal, "How to evaluate foreground maps?" in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 248–255, Columbus, OH, USA, June 2014.

[56] Z. Luo, A. Mishra, A. Achkar, J. Eichel, S. Li, and P.-M. Jodoin, "Non-local deep features for salient object detection," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 6609–6617, Honolulu, HI, USA, July 2017.

[57] Y. Zeng, H. Lu, L. Zhang, M. Feng, and A. Borji, "Learning to promote saliency detectors," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 1644–1653, Salt Lake City, UT, USA, June 2018.

[58] X. Qin, Z. Zhang, C. Huang, C. Gao, M. Dehghan, and M. Jagersand, "BASNet: boundary-aware salient object detection," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 7479–7489, Long Beach, CA, USA, June 2019.

[59] J. Wei, S. Wang, and Q. Huang, "F³Net: fusion, feedback and focus for salient object detection," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 7, pp. 12321–12328, 2020.

[60] J. Wei, S. Wang, Z. Wu, C. Su, Q. Huang, and Q. Tian, "Label decoupling framework for salient object detection," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 13025–13034, Seattle, WA, USA, June 2020.