

## Research Article

# Saliency Detection via the Improved Hierarchical Principal Component Analysis Method

Yuantao Chen <sup>1</sup>, Jiajun Tao,<sup>1</sup> Qian Zhang,<sup>2</sup> Kai Yang,<sup>2</sup> Xi Chen,<sup>1</sup> Jie Xiong,<sup>3</sup> Runlong Xia,<sup>4</sup> and Jingbo Xie<sup>4</sup>

<sup>1</sup>School of Computer and Communication Engineering and Hunan Provincial Key Laboratory of Intelligent Processing of Big Data on Transportation, Changsha University of Science and Technology, Changsha 410114, China

<sup>2</sup>Department of Electronic Products, Hunan ZOOMLION Intelligent Technology Corporation Limited, Changsha 410005, China

<sup>3</sup>Electronics and Information School, Yangtze University, Jingzhou 434023, China

<sup>4</sup>Hunan Institute of Scientific and Technical Information, Changsha 410001, China

Correspondence should be addressed to Yuantao Chen; [chenyt@csust.edu.cn](mailto:chenyt@csust.edu.cn)

Received 6 April 2020; Revised 14 April 2020; Accepted 17 April 2020; Published 5 May 2020

Academic Editor: Huimin Lu

Copyright © 2020 Yuantao Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aiming at the problems of intensive background noise, low accuracy, and high computational complexity of the current significant object detection methods, the visual saliency detection algorithm based on Hierarchical Principal Component Analysis (HPCA) has been proposed in the paper. Firstly, the original RGB image has been converted to a grayscale image, and the original grayscale image has been divided into eight layers by the bit surface stratification technique. Each image layer contains significant object information matching the layer image features. Secondly, taking the color structure of the original image as the reference image, the grayscale image is reassigned by the grayscale color conversion method, so that the layered image not only reflects the original structural features but also effectively preserves the color feature of the original image. Thirdly, the Principal Component Analysis (PCA) has been performed on the layered image to obtain the structural difference characteristics and color difference characteristics of each layer of the image in the principal component direction. Fourthly, two features are integrated to get the saliency map with high robustness and to further refine our results; the known priors have been incorporated on image organization, which can place the subject of the photograph near the center of the image. Finally, the entropy calculation has been used to determine the optimal image from the layered saliency map; the optimal map has the least background information and most prominently saliency objects than others. The object detection results of the proposed model are closer to the ground truth and take advantages of performance parameters including precision rate (PRE), recall rate (REC), and *F*-measure (FME). The HPCA model's conclusion can obviously reduce the interference of redundant information and effectively separate the saliency object from the background. At the same time, it had more improved detection accuracy than others.

## 1. Introduction

The human's visual attention mechanism had enabled humans to do real-time positioning in complex scene images corresponding to the position of important information, in order to determine the priority sequence of different objectives, which can effectively reduce the range of visual processing, thus greatly saving computing resources. Therefore, to study the human's visual attention mechanism and apply it to the research field of computer vision and image processing

has great significance. Today, researchers at home and abroad have been widely concerned with the saliency areas of the human's visual attention mechanism based on the detection technology. The method has become an important research topic in the research field of computer vision and has been successfully applied to image cropping, multiple object tracking and recognition, and thumbnail generation.

The researchers in computer vision had often used the bottom-up process to simulate the mechanism of visual attention, which is called the bottom-up saliency model.

For example, Itti et al. [1] had simulated the fusion mechanism of human brain visual cortex neurons to color, brightness, and orientation features, built the visual saliency model based on the principle of center periphery, and effectively detected the saliency area. The calculation process of the model is simple, but the detection of the target area is not accurate. Yang et al. [2] had improved Itti's model [1] based on the graph theory model and proposed the GBVS model. The calculation method is similar to Itti's model [1], and the image's color, brightness, and direction are same. The GBVS model [2] can compute and calculate the saliency map by the Markov random field, and it can detect image saliency from a global perspective. But its drawback is its inefficiency and inability to identify the target contour. Hou and Zhang [3] had put forward to the Spectral Residual (SR) algorithm. Liao et al. [4] had considered that the amplitude spectrum of prior knowledge is subtracted from the amplitude spectrum of the image. The rest is the saliency part of the amplitude spectrum, and then, the target saliency map can be obtained through the transformation in the frequency domain. The algorithm is fast, but the accuracy is difficult to guarantee. The Low-Rank (LR) algorithm had been proposed by Zhou et al. [5]. It is able to extract more notable features from the high-level Apriori to the low-rank framework, but the computation is large, and the saliency map obtained is poor. Generally speaking, the bottom-up saliency methods are mostly basic, faster, and simpler, but the saliency testing results are often represented by dense highlights, so they cannot show the outline of saliency objects.

The visual saliency detection method is the top-down model. According to a specific task, Hou et al. [6] had realized the adjustment of shape, size, feature number, threshold, and so on from the bottom-up testing results. Achanta et al. [7] had proposed the Frequency-Tuned (FT) algorithm. It is the Euclidean distance between the average pixel Gaussian low-pass filtering of each image pixel value in the image. The image's value has saliency value of the image point and formed a kind of measurement method based on the comparison of global saliency detection. The Region Contrast (RC) algorithm had been proposed by Cheng et al. [8], by calculating the saliency value of each partition area and building the saliency map based on local contrast. Dalal and Triggs [9] had proposed a method of human body detection based on the feature of the gradient histogram. The method had used gradient direction histogram information to express human characteristics and extracted human shape information and motion information. It had formed rich feature sets. Through local contrast image features, these top-down models [10–14] were characteristically analyzed; due to the extraction of various features, these models' operation speed is slow and is easily affected by the illumination environment and other objective factors, which makes the object detection accuracy greatly reduced.

In recent years, many researchers had applied such methods as machine learning to saliency detection and made great progress. For example, Yu et al. [15] and Chen et al. [16] had built the Deep Convolution Neural Network (DCNN) model based on the principle of human vision. It has combined with the superpixel clustering method to get

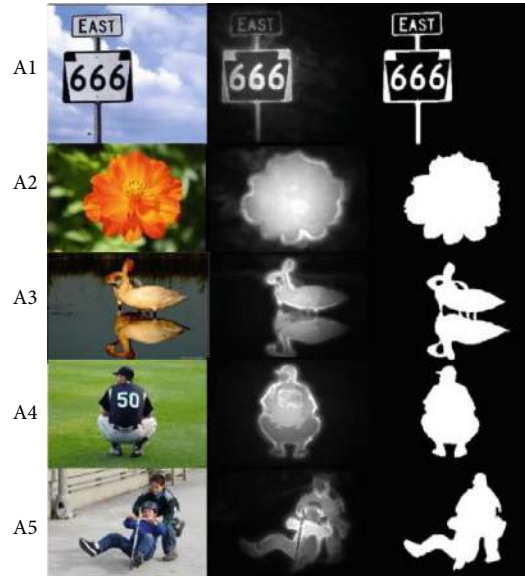


FIGURE 1: The display results of the proposed algorithm.

the image region features. It can achieve effective saliency detection by learning the features. Zhou and Tang [17] had detected the effectiveness and robustness of machine learning and sparse coding. This method has high robustness, but the operation speed is slow. To this end, Principal Component Analysis (PCA) had been applied to the saliency detection; this method had preserved the efficiency of machine learning [18, 19]. However, when the image background information extracted principal components which had represented saliency goals that cannot be effective, it results in greater detection results with background noise [20, 21].

In visual saliency detection tasks, due to the complexity of the image, the saliency graph of the single level detection method is not clear [22]. In order to reduce the impact of image complexity, Wang et al. [23] had proposed the Hierarchical Saliency (HS) algorithm. Chen et al. [24] can effectively suppress the interference of background noise to object detection by stratifying the image and calculating the stratified graph.

Based on the above analysis, in order to weaken the impact efficiency of redundant information on the detection results and retention of machine learning, the paper has proposed the saliency object detection algorithm based on the Hierarchical PCA model, using the layered PCA method which divides the image into multilayer images of a lack of background information in different degrees, so that in the process of extracting principal component information in reducing the amount of calculation and weakening the background information of interference to the detection process, it retains the efficiency of machine learning, to increase the robustness of the algorithm. Figure 1 shows the detection results of the proposed algorithm in the paper.

For this paper, the main contributions are as follows: (1) to attenuate the impact of redundant information on testing results and preserve the efficiency of machine learning; (2) to divide the image into multilayer images with different background information which reduces the computational

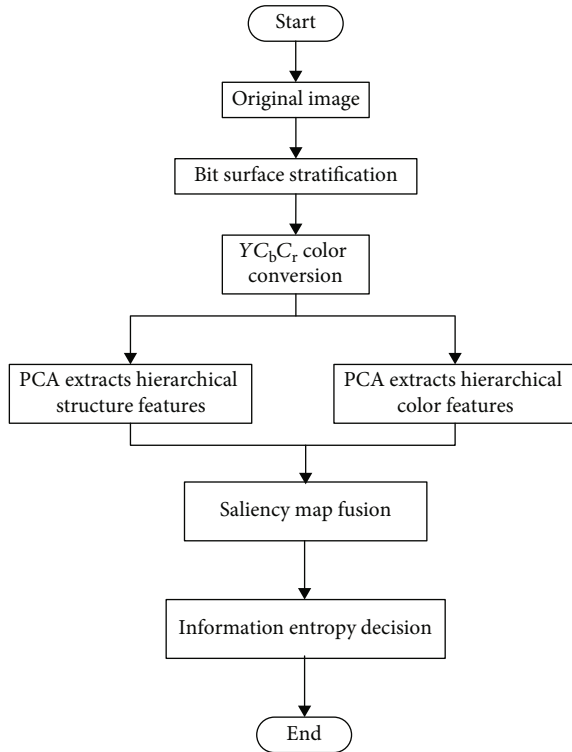


FIGURE 2: The flowchart of the proposed algorithm.

complexity and reduces the interference of background information on the detection process in the process of extracting the principal component information; and (3) to preserve the efficiency of machine learning and increase the robustness of the proposed algorithm.

Section 2 depicts the proposed algorithm details. Section 3 presents the generating saliency graph with Hierarchical PCA. Section 4 describes the experiments on the MRAS-1000, ASD-1000, and ECSSD-1000 datasets; compares; and analyzes with several methods such as IT, GBVS, SR, LC, HS, BSCA, HDCT, and DCRR. Section 5 summarizes the research work and looks forward to the future research works.

## 2. The Proposed Algorithm Details

The Hierarchical PCA visual saliency detection algorithm's flowchart is shown in Figure 2. The image information contained in different bit surface layers is quite different, and the eighth image significantly reduces the information contained in the saliency object, so that the significant object area in the image is missing. Other images, to a certain extent, reduce the background information due to the missing bit layers, highlighting the information contained in the saliency objects.

The basic process procedure is as follows: (1) stratification of the original image, using different bit data reconstruction layers which contain an image thus highlights the saliency object information; (2) in order to integrate multiple features, the original color structure is transferred to the gray-level image after stratification, so that each layer of the image has the corresponding color structure corresponding

to the original image; (3) PCA is used to extract the structure features and color features; (4) the two distinct features are fused to obtain multiple saliency graphs; and (5) the optimal results are selected through the information entropy. Figure 3 shows an example of the proposed algorithm.

*2.1. The Principle of Bit Surface Image Stratification.* The eight-bit gray-level image is considered to be composed of eight planes of one bit, each of which contains saliency information that matches it. Four of the high-order bit planes, especially the last two bit planes, contain most of the information of the saliency object. The low-order bit plane contributes to more detailed gray-level details on the image, which means that we can use the saliency information and more bit levels to build the original image, highlighting the proportion of the saliency target in the whole image. Therefore, different bits of information can be used to represent the layered images. The algorithm steps are as follows:

- (1) The original image has been converted to a gray image, and it is used as the first layer of the image
- (2) The lowest effective bit layer of the first-layer image to zero gets a picture of the image which includes a seven-bit layer as the second-layer image output. The lowest effective bit layer to the second-layer image to zero has the third layer of the image output
- (3) The binary data of different bits are converted into decimal pixel values to obtain the multilayer image matching the number of bits

The way of removing binary data and the bit level has been chosen to achieve image stratification. The purpose is to produce images with multiple objects with the dominant target as the main information and to reduce the interference of background information. The operating results are shown in Figure 4.

In Figure 4, they can be seen that different bit planes contain different image information. The eight images obviously reduce the information contained in the saliency objects, so that the saliency object areas in the images are missing. Other images, due to the missing bit layer, also reduce the background information to a certain extent and highlight the information contained in the visual saliency objects.

*2.2. The Color Conversion.* The image's hierarchy based on the bit surface has been carried out on the basis of the gray-level image. In order to maintain the original color features from layered images, the original image's color structure has been used as the mold to transfer color to the gray-level image after image stratification.

In the color conversion process, the conversional technology has been often used in gray colorful transformation, meaning each image pixel of the black area and white area. They are made of the point of the gray value and sent to the three passages through the implementation of different brightness transformations. It generates the corresponding red value, green value, and blue value, namely, the color image and the pixel corresponding to the color's value, which

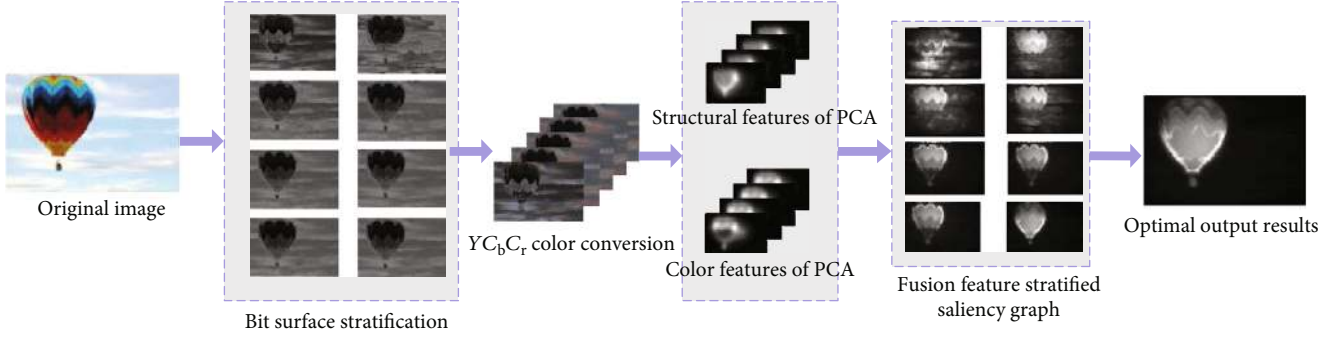


FIGURE 3: The example of the proposed algorithm.

Input: original image

Output: saliency map

Initialization: adjust the size of the input image

Step 1: bit surface stratification

Step 1.1: the original image has been converted to a gray image, and it is used as the first layer of the image

Step 1.2: the lowest effective bit layer of the first-layer image to zero gets a picture of the image which includes a seven-bit layer as the second-layer image output. The lowest effective bit layer to the second-layer image to zero has the third layer of the image output

Step 1.3: the binary data of different bits are converted into decimal pixel values to obtain the multilayer image matching the number of bits

Step 2:  $YC_bC_r$  color conversion

Step 2.1: calculate Mark1 according to formula (3)

Step 2.2: calculate Mark2 according to formula (4)

Step 3: feature extraction using PCA

Step 3.1: calculate  $P(p_i)$  according to formula (7)

Step 3.2: calculate the distance  $d(p_i, p_a)$  between each image block  $p_i$  and average image block  $p_a$  according to formula (6)

Step 3.3: calculate the color feature  $C(r_i)$  of  $r_i$  according to formula (8)

Step 4: saliency map fusion

Step 4.1: calculate the fusion of feature mapping  $D(p_i)$  according to formula (9), and limit the range of fusion features to  $[0, 1]$  by the normalization method

Step 4.2: combine the fusion feature mapping and Gaussian weight mapping to get the prominent visual saliency map  $S(p_i)$  according to formula (10)

Step 5: calculate information entropy  $Ens(x)$  of  $x$  according to formula (11), and calculate the saliency graph  $k_{opt}$  of the best scale according to formula (12)

ALGORITHM 1. Proposed algorithm.

can not only retain the mode difference of the object and background of the original image but also enhance the two-color coded target contrast significantly, making the detection more convenient. The implementation step details of color conversion in the paper are as follows.

Firstly, with the original image as the reference image of  $I_o$  (original image) and segmented images as the image to be processed of  $I_g$  (gray image),  $I_g$  will be extended to three channels, and the expansion of the image and  $I_o$  was transformed into the  $YC_bC_r$  space (where  $Y$  is the luminance component,  $C_b$  is the blue color component, and  $C_r$  is the red color component).

Secondly, the maximum and minimum values of every column from the image matrix constituted by  $I_o$  are assumed. Assuming that the resolution of the image is  $n \times m$ ,  $P(i, j)^n$  ( $1 < i < n$ ,  $1 < j < m$ ) is used to represent the image pixels of the  $m$  column; then, the maximum and minimum values of every column in the matrix can be expressed as follows:

$$P \max^m = \max (P_{1,1}^m, P_{1,2}^m, \dots, P_{1,n}^m), \quad (1)$$

$$P \min^m = \min (P_{1,1}^m, P_{1,2}^m, \dots, P_{1,n}^m). \quad (2)$$

Then, the maximum value of two images,  $I_o^{\max}$  and  $I_g^{\max}$ , and the minimum value of two images,  $I_o^{\min}$  and  $I_g^{\min}$ , are calculated. The two images are normalized to get the reconstructed color image model.

$$\text{Mark1} = \frac{I_g - I_g^{\min}}{I_g^{\max} - I_g^{\min}}, \quad (3)$$

$$\text{Mark2} = \frac{I_o - I_o^{\min}}{I_o^{\max} - I_o^{\min}}. \quad (4)$$

Transfer the colorful map, transfer the image pixel value in Mark2 to the pixel points in the corresponding Mark1, and



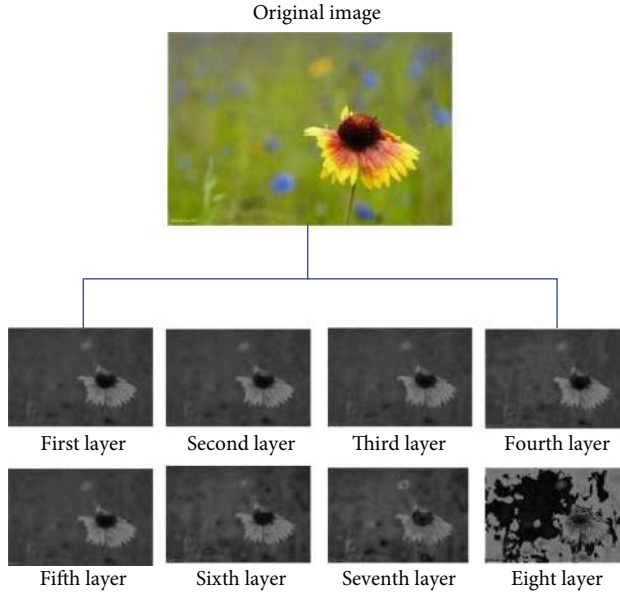


FIGURE 4: The image stratification results of different layers.

make the hierarchical gray image have the same color structure as the original image. The result of the transformation is shown in Figure 5.

### 3. The Generating Saliency Graph with Hierarchical PCA

The PCA is the model used to analyze data in the multivariate statistical analysis procedure. It is a way to describe the sample with a small number of features to reduce the dimension of the feature space. The algorithm proposed in this paper is the reconstruction of the saliency map with two features based on the unique structure and color characteristics of pixels near the hierarchical saliency object.

Due to the hierarchical results, the integration model and color pattern of each layer of the image are different from the other images; by section, that image reduces the outstanding target; a layer of background information always exists in the hierarchical image so by calculating each layer image significantly, the results of the output are then found to be most close to the true value. The experimental results are shown in Figure 6, in which Figures 6(a)–6(h) represent the saliency graph of the corresponding stratified images above them, respectively. The specific calculation process of the algorithm is described as follows in Figure 6.

**3.1. Extraction of Structural Features.** In order to improve the efficiency of structural feature computation, the PCA model based on Wang et al. [23] has been represented in the paper.

Firstly, the layered color image is analyzed by the PCA model, and each layer is divided into  $9 \times 9$  blocks, and  $N$  is the total number of blocks. For a single-layer image, each image block centered on the pixel point  $i$  is expressed in  $p_i$ , and the average image block  $p_a$  can be defined as

$$p_a = \frac{1}{N} \sum_{i=1}^N p_i. \quad (5)$$

They can calculate the distance between each image block  $p_i$  and average image block  $p_a$ ,  $d(p_i, p_a)$ , along principal component direction. Whether an image block has significant structural characteristics is determined based on the distance. Here, the position coordinates of each image block are represented by its central pixel  $p_i(i_x, i_y)$ , and the position of the average image block is represented by  $p_a(a_x, a_y)$ . The definition of  $d(p_i, p_a)$  is shown as

$$d(p_i, p_a) = \frac{1}{N} \times \sum_{i=1}^N p_i \times \sqrt{\frac{(i_x - a_x)^2 + (i_y - a_y)^2}{S_{(i,a)}^2}}. \quad (6)$$

where  $S_{(i,a)}^2$  is the variance between the two image blocks.

The rule of judgments is as follows: when  $d(p_i, p_a)$  is larger than the threshold of the dataset, the image block is considered to be the saliency area, and the other is a common image block. From the mathematical meaning, the extraction of structural features is attributed to the  $L_1$  norm of  $p_i$  in the PCA coordinate system. Therefore, the structural feature  $P(p_i)$  is further defined as

$$P(p_i) = \|p_i\|_1. \quad (7)$$

In formula (7),  $p_i'$  is the coordinate of  $p_i$  in the PCA coordinate system.  $\|\cdot\|_1$  is the operation symbol of the  $L_1$  norm.

**3.2. The Extraction of Color Features.** Although the extraction of structural features can find the most unique block in the image, it is not suitable for all images. As shown in Figure 7, the structure characteristics of each sphere are the same, but the colors are different. In this case, they are thinking that the color features are more distinctive. So, the extraction of color features is essential.

Here, two steps are used to detect the color difference of the image block. The first step is to divide each layer of the image into several blocks by using the simple linear iterative clustering superpixel segmentation method and then determine which block has unique color characteristics. In the second step, the sum of distance between the image block and the other image blocks in the  $YCbCr$  color space is defined as the color difference of the image block. Here,  $r_i$  is used to represent the position of  $i$  block in the color space, and  $r_j$  is used to represent the location of  $j$  image block. From the mathematical meaning, the color feature extracted for image block  $r_i$  is to calculate its  $L_2$  norm in the PCA coordinate system. Therefore, the color feature  $C(r_i)$  of  $r_i$  is defined as

$$P(p_i) = \|p_i\|_1. \quad (8)$$

In the upper form,  $r_i - r_j$  is represented by the distance between two blocks.  $\|\cdot\|_2$  is the operational symbol of  $L_2$

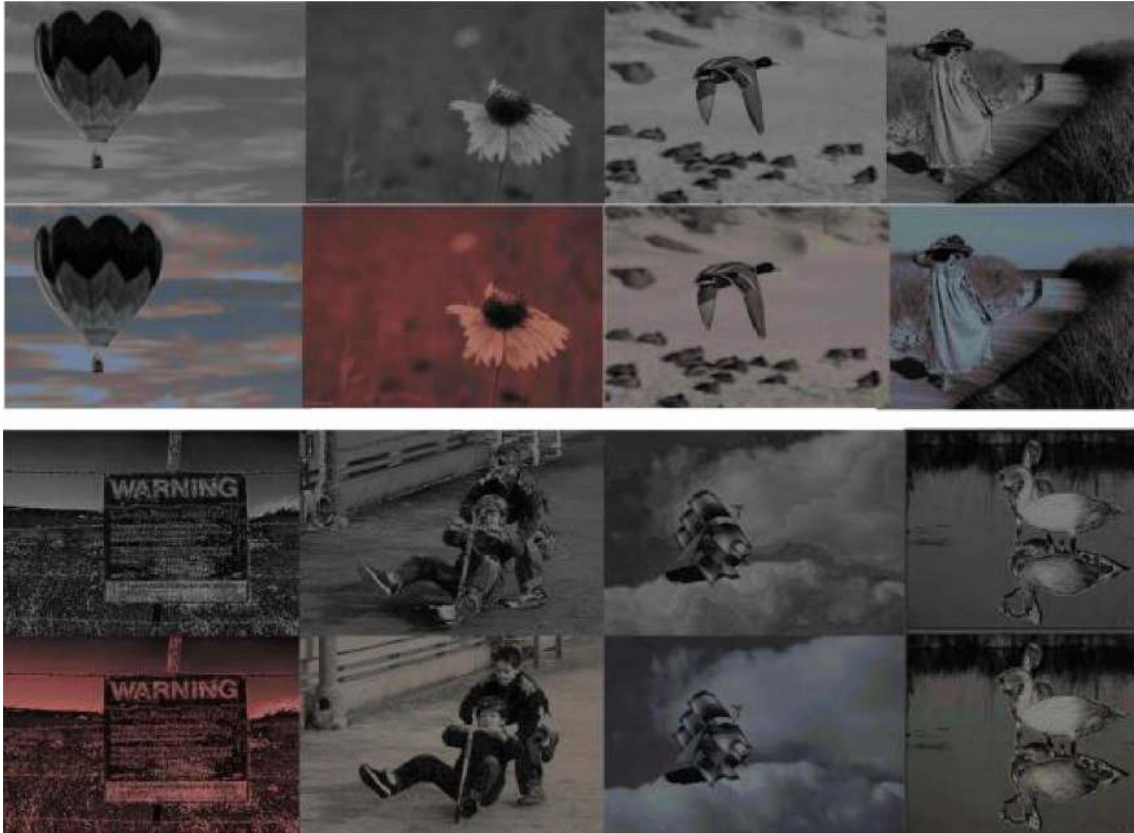


FIGURE 5: The color conversion effect diagram.

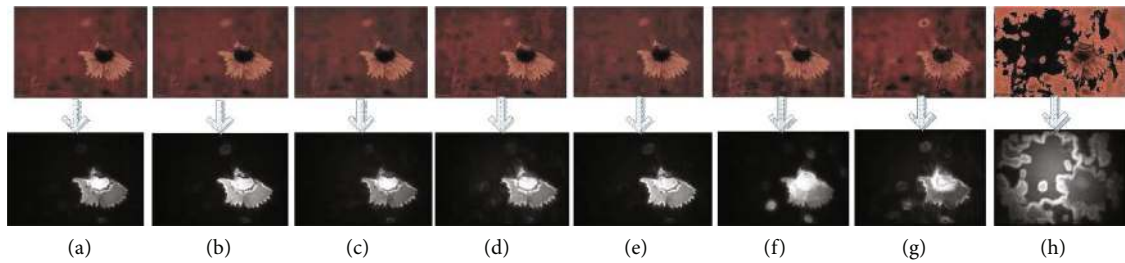


FIGURE 6: The experimental result diagram.



FIGURE 7: The extraction by the color feature.

norm.  $M$  represents the total number of blocks after super-pixel segmentation.

3.3. *The Saliency Fusion by Structural Features and Color Features.* The single image structure feature or color feature cannot effectively characterize all information of the saliency

object. In order to obtain accurate and saliency objects, they can combine the structure and color features of each layer of images to detect the saliency regions of different layers of images. Here, they are using the fusion feature to get

$$D(p_i) = P(p_i) \cdot C(p_i). \tag{9}$$

TABLE 1: Entropy of stratified image information.

	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8
B1	6.34	6.39	6.33	6.35	6.36	5.78	6.05	6.10
B2	7.21	6.73	7.03	7.02	7.01	7.03	7.10	7.13
B3	5.78	5.82	5.63	5.56	5.69	5.78	5.79	5.82



FIGURE 8: The saliency graph using the Hierarchical PCA model.

After that, the fusion features are limited to the  $[0, 1]$  range by normalization. Because visual pixels are usually clustered, they usually correspond to objects in real scenes. In order to further modify the saliency models, people usually use the center prior method to put the target area near the center. The center prior based on the Apriori algorithm usually assumes the target located at the center of the image as a hypothetical condition. By defining the center's prior weight with a peak value-centered Gaussian function, the object saliency in the center of the image is prominently highlighted according to the weight distribution. Here, different target regions are represented by a set of pixels under different thresholds, and the threshold is uniformly distributed in the  $[0, 1]$  interval. Therefore, the process of the center prior calculation is as follows:

- (1) The image pixel sets of different layers are detected by the fusion of feature mapping  $D(p_i)$ , and the center of gravity of each threshold result is calculated
- (2) The center of gravity places a Gaussian distribution with  $\delta$  of 10,000, and the corresponding Gaussian weights are calculated for each threshold
- (3) The Gaussian distribution with weight of five is added to the image center of each layer to improve the weight of the center position

The Gaussian weight mapping  $G(p_i)$  is used to represent the weighted sum of all Gaussian distributions, and different saliency priorities are given according to the difference of weight distribution. Therefore, they can further define the saliency mapping and combine the fusion feature mapping and Gaussian weight mapping to get the prominent visual saliency map  $S(p_i)$ .

$$S(p_i) = D(p_i) \cdot G(p_i). \quad (10)$$

*3.4. The Decision of the Optimal Results.* After the above steps, the saliency image corresponding to each layer can be obtained, and the best detection result diagram will be the final output image.

In the information theory, entropy is a relatively basic concept, which is represented by the average amount of information in random events. The information entropy often implies the distribution of the foreground and background noise in the image signal. Generally speaking, if the saliency area of the image is more obvious, it will be more prominent in the whole image performance, and the repeated background area will also inhibit more. Therefore, the saliency region is also gathered in the value of a particular region on the histogram. It provided the small information entropy. The general rule is that the minimum information entropy corresponds to the best saliency graph. For an image signal  $X$ , its information entropy is defined by

$$\text{Ens}(x) = - \sum_{i=1}^m \sum_{j=1}^n p(X_{i,j}) \log p(X_{i,j}). \quad (11)$$

In formula (11),  $X_{i,j}$  ( $i = 1, 2, \dots, m, j = 1, 2, \dots, n$ ) represents the gray value of image  $X$  in line  $j$  and line  $i$ , and  $p(X_{i,j})$  means the probability of occurrence  $X_{i,j}$  in image  $X$ , and  $\text{Ens}$  represents the entropy of the image. Then, the saliency graph  $k_{\text{opt}}$  of the best scale can be expressed as

$$k_{\text{opt}} = \arg \min_k (\text{Ens}(S_k)), \quad k = 1, \dots, K. \quad (12)$$



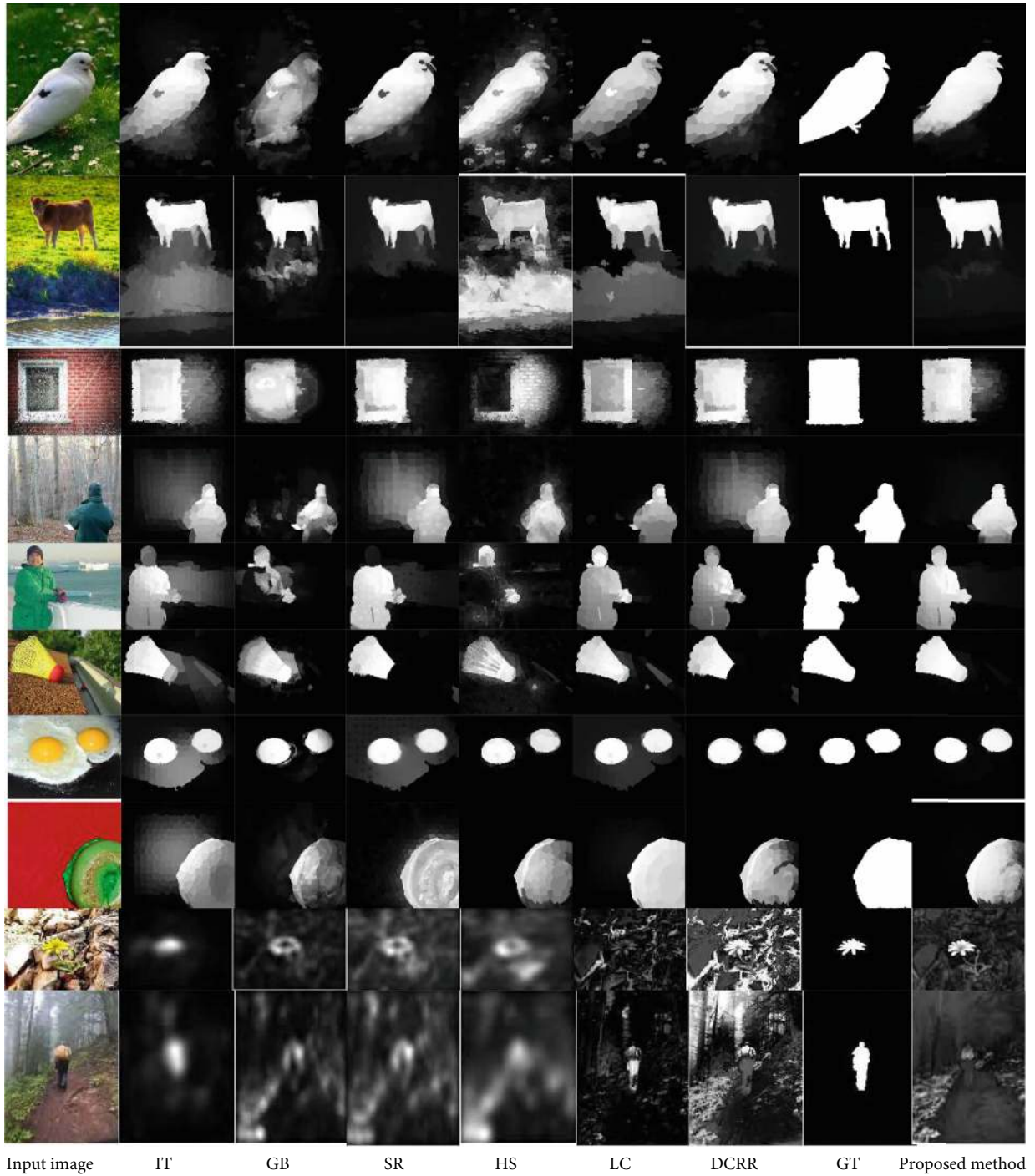


FIGURE 9: The comparison diagram of algorithm results.

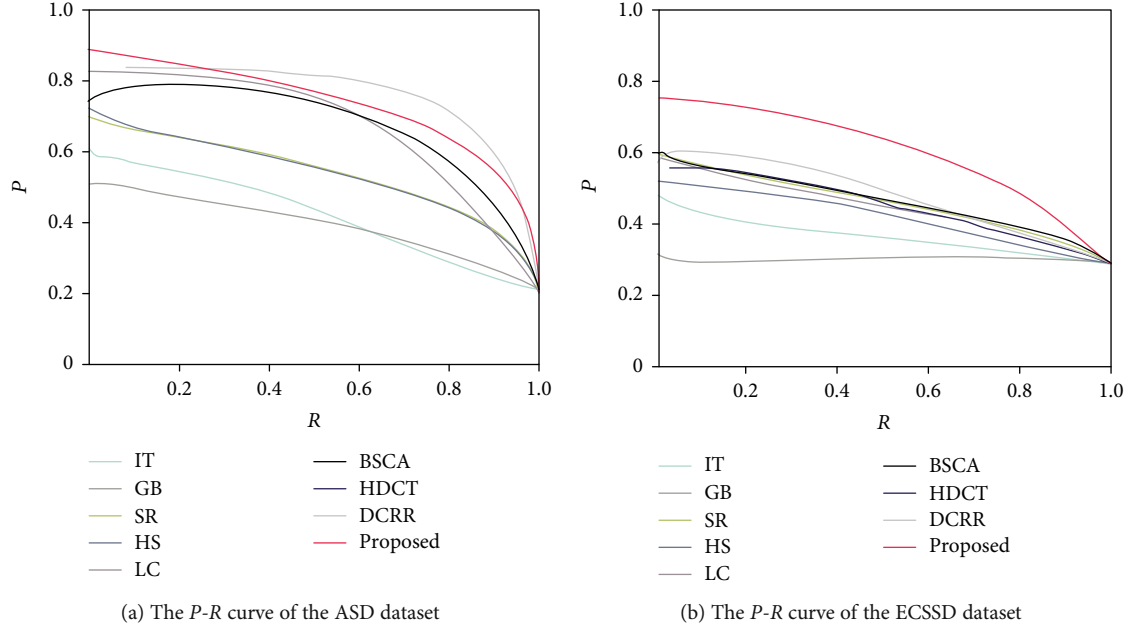
The information entropy is calculated for the multilayer saliency graph after the Hierarchical PCA processing. The information entropy of the stratified image is shown in Table 1.

The data in Table 1 is the information entropy of each image shown in Figure 8. They are using the above information entropy decision rule to decide the eight-level image, select the smallest information entropy image as the output of the optimal result, and get the saliency map with the least background information, which is the final result in Table 1.

#### 4. The Experimental Results and Analysis

The experimental method has used the MATLAB software as the programming platform, and the algorithm is realized on the ThinkPad-E40 laptop. The Hierarchical PCA model in saliency detection is tested on datasets of MRAS-1000, ASD-1000, and ECSSD-1000 and compared with several methods, such as ITTI (IT) [1], GBVS (GB) [2], SR [3], LC [25], HS [23], BSCA [26], HDCT [27], and DCRR [28]. The results of Itti et al. [1] and Yang et al. [2], respectively,



FIGURE 10: The  $P$ - $R$  curve of the ASD and ECSSD datasets.

are provided by Hou and Zhang [3], Fang et al. [25], and Wang et al. [23] in each dataset. The CHS [29] had used the original data that is generated on the ECSSD dataset. The result of the visual contrast is shown in Figure 9. In addition, in order to objectively evaluate the detection results, various algorithms are used such as the precision rate (PRE), recall rate (REC), and  $F$ -measure (FME) to evaluate the performance. The definitions of PRE, REC, and FME are shown in formulas (13)–(15) [30].

$$\text{PRE} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (13)$$

$$\text{REC} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (14)$$

$$\text{FME} = 2 \times \frac{\text{PRE} \times \text{REC}}{\text{PRE} + \text{REC}}. \quad (15)$$

Among them, TP represents the number of image pixels that detect saliency objects. TN means that the background is correctly divided into the number of pixels in the background class. FP indicates the number of pixels that extract the wrong background. FN means that the saliency object error is divided into the number of pixels in the background class. The AUC indicator is defined as the lower area enclosed by the ROC curve and the coordinate axis, and the maximum value is 1. The larger the AUC, the better the prediction performance of the method on the gaze point of the human eye.

Figure 10 shows the  $P$ - $R$  curve [11, 31] of different saliency detection algorithms on three typical common datasets. It can be seen that, because of the high recognition rate of the ECSSD dataset, the accuracy of the HS and LC algorithms is more than 90%, but the precision rate parameter is low. On the ASD dataset, each algorithm reduces the recall

to a certain extent, and the precision rate parameter of the algorithms such as IT, GB, and LC is obviously reduced. In the ASD dataset, the recall rate of the GB algorithm is higher, and the HS algorithm has a certain advantage on the accuracy and the  $F$ -measure value. In general, the accuracy of the algorithm on different datasets is over 90%, and the average  $F$ -measure value is higher than the other algorithms. The detection results are shown as stable and robust.

Figure 11 is a contrast histogram of the  $F$ -measure value results of various algorithms shown in Figure 12. It can be seen that due to the high image recognition rate of the ECSSD dataset, the accuracy of the HS and LC algorithms exceeds 90%, but the  $F$ -measure value is low. On the more complex MRAS dataset, each algorithm reduces the recall rate to a certain extent, and the  $F$ -measure values of IT, GB, LC, and other algorithms are significantly reduced. On the relatively simple ASD dataset, the GB algorithm has a higher recall rate, and the HS algorithm has certain advantages in accuracy and  $F$ -measure values. In general, the accuracy of the algorithm in different datasets exceeds 90%, and the average  $F$ -measure value is higher than the other algorithms. The detection effect is stable, and the robustness is better than the other solutions.

The results in Table 2 can show the AUC score of each method. It can be seen that our method has the highest AUC score, indicating that our method still has a good detection effect in natural images with complex backgrounds, and can effectively label saliency targets. At the same time, it shows that the method in this paper has higher accuracy and the saliency map obtained is closer to the ground truth.

The calculation speed is an important index for evaluating the superiority of the method. The calculation speed of the method determines whether it can be applied to a real-time system. As a preprocessing process in various image

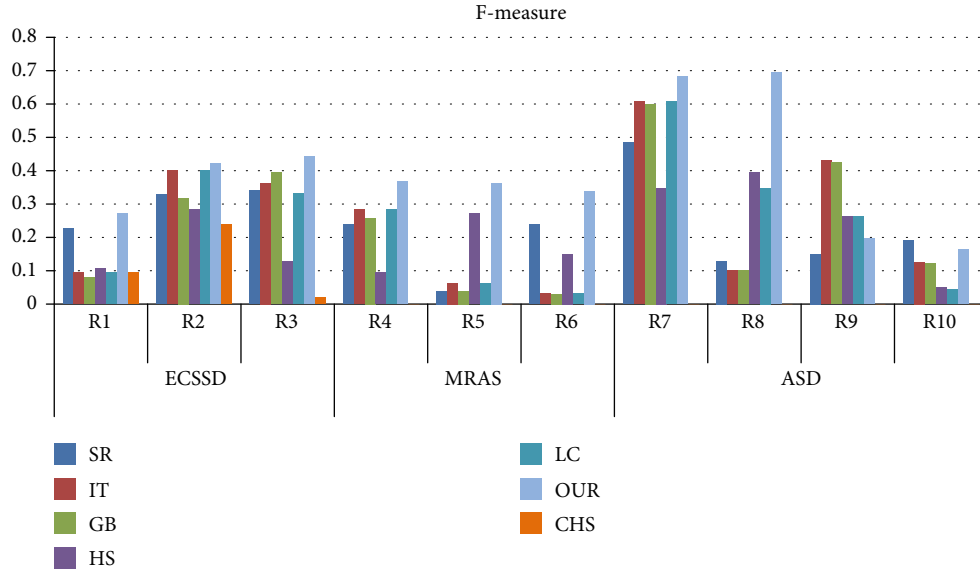


FIGURE 11: The contrast histogram of the  $F$ -measure value.

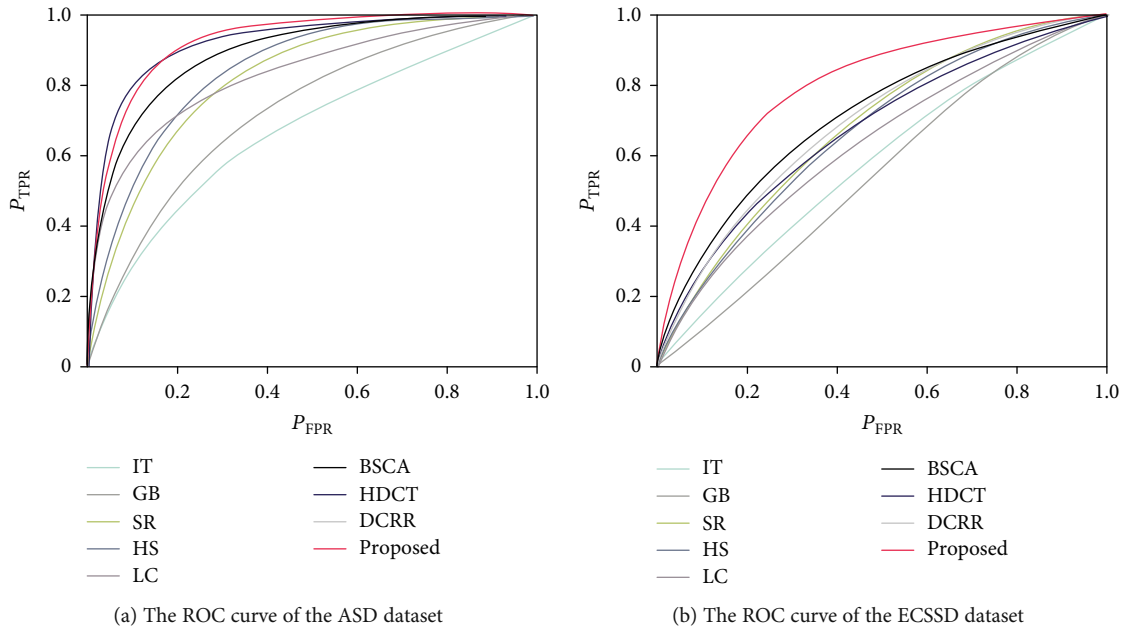


FIGURE 12: The ROC curve of the ASD and ECSSD datasets.

processing fields, the calculation speed is very important. On the premise that the accuracy of the method meets the expected requirements, the faster the calculation speed, the better the overall performance of the method. The average calculation time of each method is shown in Table 3. The method in this paper has fast calculation speed and can meet the basic application requirements.

### 5. Conclusions

In this paper, the saliency object detection algorithm based on the Hierarchical PCA model was proposed. The experi-

mental results had shown that the proposed algorithm can reduce the interference of background noise, and the background and target separation has certain advantages in precision, recall, and  $F$ -measure parameters, while retaining the excellent characteristics of machine learning methods in order to improve the saliency detection effect. Therefore, the Hierarchical PCA saliency detection is an effective method for object detection under complex backgrounds. The Hierarchical PCA algorithm cannot analyze all the information in the image at the same time. When the objects in the background have the same level of brightness and resolution, it is difficult to extract the complete object information.

TABLE 2: The AUC value of the ASD and ECSSD datasets.

Method	ASD dataset	ECSSD dataset
IT	0.7252	0.5493
GB	0.8207	0.6681
SR	0.6736	0.5805
HS	0.8232	0.6813
LC	0.8451	0.6954
BSCA	0.8302	0.6755
HDCT	0.8894	0.6954
DCRR	0.9212	0.7091
Proposed	0.9242	0.7990

TABLE 3: Calculating the time of different methods.

Method	Time (s)
IT	0.2224
GB	0.0163
SR	0.0109
HS	0.0147
LC	0.0288
BSCA	0.0956
HDCT	0.1532
DCRR	0.0752
Proposed	0.0282

Therefore, the future work for the proposed technique is to study the problem of incomplete object information and further improve the information utilization of the whole image to get more accurate and saliency object information.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work is supported by grants of the National Natural Science Foundation of China (Nos. 61972056, 61972212, 61402053, and 61981340416), the Open Research Fund of Hunan Provincial Key Laboratory of Intelligent Processing of Big Data on Transportation (No. 2015TP1005), the Changsha Science and Technology Planning (Nos. KQ1703018, KQ1706064, KQ1703018-01, and KQ1703018-04), the Research Foundation of Education Bureau of Hunan Province (Nos. 17A007 and 19B005), the Changsha Industrial Science and Technology Commissioner (No. 2017-7), the Natural Science Foundation of Jiangsu Province (No. BK20190089), and the Junior Faculty Development Program

Project of Changsha University of Science and Technology (No. 2019QJCZ011).

## References

- [1] 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.
- [2] C. Yang, L. Zhang, H. Lu, X. Ruan, and M. Yang, "Saliency detection via graph-based manifold ranking," in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3166–3173, Portland, OR, USA, June 2013.
- [3] X. Hou and L. Zhang, "Saliency detection: a spectral residual approach," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, Minneapolis, Minnesota, USA, June 2007.
- [4] Z. Liao, R. Zhang, S. He, D. Zeng, J. Wang, and H.-J. Kim, "Deep learning-based data storage for low latency in data center networks," *IEEE Access*, vol. 7, pp. 26411–26417, 2019.
- [5] S. Zhou, M. Ke, and P. Luo, "Multi-camera transfer GAN for person re-identification," *Journal of Visual Communication and Image Representation*, vol. 59, no. 2, pp. 393–400, 2019.
- [6] X. 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.
- [7] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1597–1604, Miami, Florida, USA, June 2009.
- [8] 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.
- [9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 886–893, San Diego, CA, USA, June 2005.
- [10] Y. Chen, J. Xiong, W. Xu, and J. Zuo, "A novel online incremental and decremental learning algorithm based on variable support vector machine," *Cluster Computing*, vol. 22, no. S3, pp. 7435–7445, 2019.
- [11] Y. Luo, J. Qin, X. Xiang, Y. Tan, Q. Liu, and L. Xiang, "Coverless real-time image information hiding based on image block matching and dense convolutional network," *Journal of Real-Time Image Processing*, vol. 17, no. 1, pp. 125–135, 2020.
- [12] F. Yu, L. Liu, S. Qian et al., "Chaos-based application of a novel multistable 5D memristive hyperchaotic system with coexisting multiple attractors," *Complexity*, vol. 2020, Article ID 8034196, 19 pages, 2020.
- [13] F. Yu, Z. Zhang, L. Liu et al., "Secure communication scheme based on a new 5D multistable four-wing memristive hyperchaotic system with disturbance inputs," *Complexity*, vol. 2020, Article ID 5859273, 6 pages, 2020.
- [14] G. Sheng, X. Tang, K. Xie, and J. Xiong, "Hydraulic fracturing microseismic first arrival picking method based on non-subsampled shearlet transform and higher-order-statistics," *Journal of Seismic Exploration*, vol. 28, no. 6, pp. 593–618, 2019.
- [15] Y. Yu, S. Tang, K. Aizawa, and A. Aizawa, "Category-based deep CCA for fine-grained venue discovery from multimodal

- data,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 4, pp. 1250–1258, 2019.
- [16] Y. Chen, J. Wang, S. Liu et al., “Multiscale fast correlation filtering tracking algorithm based on a feature fusion model,” *Concurrency and Computation: Practice and Experience*, no. - article e5533, 2019.
- [17] L. Zhou and J. Tang, “Fraction-order total variation blind image restoration based on L1-norm,” *Applied Mathematical Modelling*, vol. 51, pp. 469–476, 2017.
- [18] Y. Yu, S. Tang, F. Raposo, and L. Chen, “Deep cross-modal correlation learning for audio and lyrics in music retrieval,” *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 15, no. 1, pp. 1–16, 2019.
- [19] W. Li, H. Xu, H. Li et al., “Complexity and algorithms for superposed data uploading problem in networks with smart devices,” *IEEE Internet of Things Journal*, p. 1, 2019.
- [20] Z. Liu, Z. Lai, W. Ou, K. Zhang, and R. Zheng, “Structured optimal graph based sparse feature extraction for semi-supervised learning,” *Signal Processing*, vol. 170, article 107456, 2020.
- [21] H. Lu, Y. Li, S. Mu, D. Wang, H. Kim, and S. Serikawa, “Motor anomaly detection for unmanned aerial vehicles using reinforcement learning,” *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2315–2322, 2018.
- [22] K. Gu, N. Wu, B. Yin, and W. Jia, “Secure data query framework for cloud and fog computing,” *IEEE Transactions on Network and Service Management*, vol. 17, no. 1, pp. 332–345, 2020.
- [23] J. Wang, H. Jiang, Z. Yuan, M.-M. Cheng, X. Hu, and N. Zheng, “Salient object detection: a discriminative regional feature integration approach,” *International Journal of Computer Vision*, vol. 123, no. 2, pp. 251–268, 2017.
- [24] Y. Chen, J. Wang, R. Xia, Q. Zhang, Z. Cao, and K. Yang, “The visual object tracking algorithm research based on adaptive combination kernel,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 12, pp. 4855–4867, 2019.
- [25] Y. Fang, C. Zhang, J. Li, J. Lei, M. Perreira da Silva, and P. le Callet, “Visual attention modeling for stereoscopic video: a benchmark and computational model,” *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4684–4696, 2017.
- [26] Y. Qin, H. Lu, Y. Xu, and H. Wang, “Saliency detection via cellular automata,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 110–119, Boston, MA, USA, June 2015.
- [27] J. Kim, D. Han, Y. Tai, and J. Kim, “Salient region detection via high-dimensional color transform,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 883–890, Columbus, OH, USA, June 2014.
- [28] H. Lu, D. Wang, Y. Li et al., “CONet: a cognitive ocean network,” *IEEE Wireless Communications*, vol. 26, no. 3, pp. 90–96, 2019.
- [29] J. Zhang, W. Wang, C. Lu, J. Wang, and A. K. Sangaiah, “Light-weight deep network for traffic sign classification,” *Annals of Telecommunications*, 2019.
- [30] Y. Chen, J. Wang, X. Chen et al., “Single-image super-resolution algorithm based on structural self-similarity and deformation block features,” *IEEE Access*, vol. 7, pp. 58791–58801, 2019.
- [31] F. Yu, L. Liu, L. Xiao, K. Li, and S. Cai, “A robust and fixed-time zeroing neural dynamics for computing time-variant nonlinear equation using a novel nonlinear activation function,” *Neurocomputing*, vol. 350, pp. 108–116, 2019.