Hot Region Selection Based on Selective Search and Modified Fuzzy C-Means in Remote Sensing Images

Shoulin Yin and Hang Li

***Abstract*—Hot region selection (HRS) is usually performed for larger scenes, especially in military target detection and battlefield situational awareness applications. Many state-of-the-art methods pay more attention on object detection. Although they can detect the region of interest in one image, these images are highly affected by illumination, rotation, and scale changes, which further increases the complexity of analysis compared to those obtained using standard remote sensing platforms. The study of HRS is of great importance in the analysis of remote sensing images. The current research focuses on the specific type of object area detection. A well-developed HRS needs to have three properties: uniform highlighting of the entire HRS, well-defined boundaries, and good robustness. Motivated by these requirements, a HRS method based on the selective search method and modified fuzzy c-means (FCM) in remote sensing images is proposed to address the detection of potential hot regions in large-scale remote sensing images. First, we create a Gaussian curvature filter to preprocess large scale remote sensing images. Second, a modified FCM segmentation method is utilized to segment the image. Third, an enhanced selective search method is adopted to establish well-defined boundaries for the HRS and to improve the immunity to noise. The geographic information is presented in this phase, which is conducive to improve the detection accuracy. In the experimental section, we compare our new method with four other extraction models on three data sets. The experimental results show that compared to the other competing models, the new model better defines the hot regions and obtains more entire boundaries in terms of *Overlap* and *mAP*.**

***Index Terms*—Big scene, fuzzy c-means (FCM), geographic information, hot region selection (HRS), remote sensing, selective search method.**

I. Introduction

THE selection of hot regions, using remote sensing images acquired by various satellites and sensors, has become an attractive issue for research in recent years and is also a significant application of remote sensing [1]–[4]. When faced with complex scenarios or large-scale scenes, humans can rapidly shift their attention to the hot regions in the image in order to prioritize them, which is a visual attention mechanism. In other words, the hot region refers to the area where humans first focus and notice.

Manuscript received July 2, 2020; revised August 6, 2020 and September 10, 2020; accepted September 13, 2020. Date of publication September 22, 2020; date of current version October 8, 2020. *(Corresponding author: Hang Li.)*

Shoulin Yin is with the Software College, Shenyang Normal University, Shenyang 110034, China, and also with the Institute of Image and Information Technology and School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150000, China (e-mail: yslinhit@163.com).

Hang Li is with the Software College, Shenyang Normal University, Shenyang 110034, China (e-mail: lihangsoft@163.com).

Digital Object Identifier 10.1109/JSTARS.2020.3025582

Hot region selection (HRS) technology can distinguish the importance degree of different regions and can further enable to highlight the important content of the images, and remove redundant information. The analysis research of HRS plays an important role in image compression and coding, image retrieval, target detection and recognition, scene analysis [5]–[7], active vision, and other fields. For example, the amount of data has a high requirement for the key diagnosis area in medical image processing [8], and the extraction of an area of interest can more easily locate the critical areas to reduce redundancy. In the scene analysis [9], it can quickly screen out saliency object through the analysis of the HRS in order to facilitate further understanding of the scene, and avoid meaningless image calculations. HRS is often a necessary process to cover a large and full region of interest (ROI) for many remote sensing applications including geographical mapping, resource and environmental monitoring, and disaster monitoring [10], [11].

The traditional HRS technologies are divided into two categories: the artificially labeled HRS and the extracting features to detect the HRS. Nevertheless, the detection accuracy and efficiency of these two types are not high.

Few attempts have been made to detect the HRS in remote sensing images. However, many studies have focused on ROI detection. Wang *et al.* [12] proposed a novel ROI detection model that was based on visually salient regions by utilizing the frequency and space domain features in very-high-resolution remote sensing images. First, the frequency domain features that were based on a multiscale spectrum residual algorithm were extracted to yield the intensity features. Next, the color and orientation features were extracted by generating space dynamic pyramids. Then, the spectral features were obtained by analyzing the spectral information content. In addition, a multiscale feature fusion method was proposed to generate a saliency map. However, it costs long time to process this method. Rojas *et al.* [13] described a computational method based on an optical flow algorithm to detect ROI in near infrared spectroscopy. But this method is not suit for large-scale remote sensing images. Chen and Zhang [14] presented an accurate and efficient approach via superpixel-to-pixel saliency analysis for ROI detection. Initially, the image was downsampled and segmented into superpixels using simple linear iterative clustering (SLIC). Next, the structure tensor and background contrast were used to yield the superpixel feature maps for texture and color. After fusing the feature maps, the overall superpixel saliency map was obtained and then used to achieve the final pixel-level saliency map using superpixel-to-pixel mapping. Due to downsampling and segmentation using

SLIC, it will lose some texture information. Li *et al.* [15] showed a hierarchical task-driven ROI detection method based on the saliency and density to address the detection of the potential object areas in large-scale remote sensing images. The proposed saliency and density-based detection method integrated both bottom-up and top-down strategies. The saliency-based multilevel histogram contrast was presented in the bottom-up phase to obtain the preliminary regions, while the centroid density distribution index was defined in the top-down scheme to refine the previous results. Specifically, the superpixel segmentation was introduced to narrow down the ROI candidates. However, these models are not automatic, and they may not be studied for HRS detection.

Currently, the visual attention model [16], [17] for the calculation model of the region detection can in turn easily obtain the saliency map in remote sensing images to draw viewer's attention. Different from the feature extraction serial processor system of the spatial domain image, the visual features can be processed by super plural or quaternion way, which better satisfies the human visual characteristics. Feng *et al.* [18] proposed the hyper complex Fourier transform model. Through the scale space analysis of the amplitude spectrum of natural images, the model confirmed that the amplitude spectrum contained important information. Combined with the original phase and amplitude spectrum, it reconstructed the 2-D signals and acquired the sequence of the saliency maps with the minimum entropy to choose the optimal saliency map. Experiments confirmed that the model could detect the different sizes of the saliency area, which showed a good performance. Moreover, to the best of our knowledge, the raw data acquired from high-resolution satellites are very large; the processing power of current computers is very difficult to handle these large remote sensing images. Thus, to locate the potential object areas in large-scale remote sensing images, a HRS method based on the selective search method and modified fuzzy c-means (FCM) in remote sensing images is proposed to address the detection of the potential hot regions in large-scale remote sensing images. This is because the selective search method can capture all the scales and can perform rapid calculations. First, the image is segmented using the fuzzy c-means. Then, it is processed with the selective search method, which is conducive to obtain more accurate selection result. In the future, the proposed method can be used in road tracking, unmanned aerial vehicle (UAV) automation, road-following, and traffic analysis using high-resolution remote sensing imagery, and, incident detection algorithm based on radon transform using high-resolution remote sensing imagery [19], [20]. It can further narrow the scope and improve the accuracy of object localization.

The main contributions of this article are summarized as follows.

1) A Gaussian curvature (GC) filter is adopted to preprocess the original large-scale remote sensing image.
2) A modified FCM segmentation method is utilized to segment the image and obtain the separate background region and hot region.
3) An enhanced selective search method is adopted to establish well-defined boundaries for the HRS.

The rest of this article is organized as follows. In Section II, the modules of the proposed methods are introduced. Then, the experiments are explained in Section III. Finally, the concluding remarks are provided in Section IV.

II. HRS Based on Selective Search Method and Fuzzy c-Means

A new method is introduced aiming to select hot regions in large-scale remote sensing images. The proposed approach includes three main modules, as shown in Fig. 1: first, image preprocessing including image input and GC filter; second, image segmentation; and third, region postpressing with the selective search. Each module is described in the following sections.

A. GC Filter

The GC filter is better at preserving the details of an image and efficiently reduces certain variational energies proposed by Gong and Sbalzarini [21] compared with that of other filter methods [22]. Therefore, we adopt the GC filter approach in this article.

Let $\bar{x} = (x, y) \in \Re$ denote the spatial coordinate with $\Re$ as the 2-D image domain. Let $I(i, j) : Z_0^+ \times Z_0^+ \mapsto [0, 1]$ be the given discrete digital image with integer pixel indices *i* and *j* with the continuous intensity $I \in [0, 1]$. Let $U(\bar{x}) \in [0, 1]$ denote the current reconstruction, i.e., the image obtained by reducing the total energy starting from $U_0 = 1$.

We interpret U and I as the geometric surfaces over $\Re$, i.e., $\Phi(\bar{x}) = (\bar{x}, U(\bar{x}))$. From this, GC can be computed by taking partial derivatives. For the GC, we have the following:

$$K(U(\bar{x})) = \frac{U_{xx}U_{yy} - U_{xy}^2}{(1 + U_x^2 + U_y^2)^2} \tag{1}$$

where the subscripts denote the partial derivatives. Recall that the total GC, (i.e., the integral of K) of any surface is related to the surface topology through the Gauss–Bonnet theorem. So the total GC is therefore a topological invariant, one can only minimize the total absolute GC.

Therefore, the total absolute GC regularizer can be displayed as follows:

$$\varepsilon_{\varphi_1}^{\mathrm{GC}}(U) = \int_{\Re} |\kappa_1 \kappa_2| d\bar{x} = \int_{\Re} |K(U)| d\bar{x} \tag{2}$$

where κ_1 and κ_2 are the two principal curvatures of Φ.

For the GC, all minimizers $\hat{S}_i(\bar{x}), i = 1, \ldots, N_s$ in a local 3 $\times$ 3 pixel neighborhood around pixel $\bar{x}$ can be enumerated, as shown below. From this finite set, we choose the element that leads to the smallest change in the intensity of pixel $\bar{x}$

$$S_m(\bar{x}) = \arg\min_{\hat{S}_i(\bar{x})} |\hat{S}_i(\bar{x}) - U(\bar{x})|. \tag{3}$$

And update $\hat{U}(\bar{x}) = S_m(\bar{x})$. When used together with a data-fitting term, an additional condition is added to ensure the total energy is nonincreasing. The reason we select GC is that it only requires 0.0029605 s for one iteration.

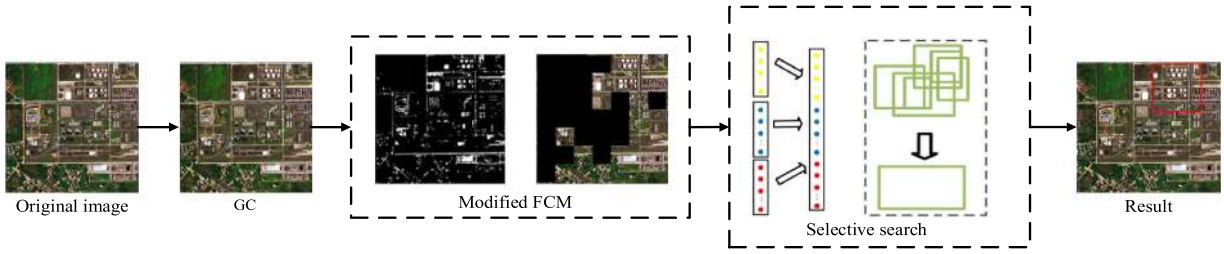


Fig. 1. Framework of our proposed method, which can be divided into three parts (the red region is the detected hot region containing sensitive oil tank region).

B. Modified FCM Segmentation

Clustering can be classified as either hard or fuzzy depending on whether a pattern data set belongs exclusively to a single cluster or to several clusters with different degrees. In the hard clustering, a membership value of zero or one is assigned to each pattern data, whereas in the fuzzy clustering, a value between zero and one is assigned to each pattern with a membership function.

The FCM cluster [23] is used for image segmentation with by iteratively searching the cluster center and the membership relationship between the image pixel and cluster center to minimize the objective function, which realizes the image segmentation.

The FCM needs to initialize the cluster number and member relationship and to calculate the cluster center. The initial cluster center plays a significant role in the convergence speed and precision. Therefore, we utilize the minimum–maximum distance method based on the histogram [24] to acquire the initial cluster center.

If the sample set is the image histogram $p_i, i = 1, 2, \ldots, 255$, the distance expression is shown as follows:

$$d_{i'} = |i - i'| \quad \forall\, p_i, p_{i'} \neq 0 \tag{4}$$

where i' is the known cluster center. The following provides the processes of obtaining the initial cluster center.

1) Step 1. Selecting the biggest pixel gray value as the first cluster center i'_1.
2) Step 2. Based on (4), selecting the sample far from i'_1 as the second cluster center i'_2.
3) Step 3. Gradually calculating the distance $(d_1, \ldots, d_j)$ between each sample i'_1 and i'_2 to determine $d = \min(d_1, \ldots, d_j)$.
4) Step 4. If the results satisfy the required cluster center number, then stop.

The selection of initial clustering center in the abovementioned process uses the image grayscale as the data results. It not only speeds up the search process, but the clustering center search is only related to the number of gray levels without affecting by image size, which can greatly reduce the computational complexity.

After the GC filter, there are still some isolated noise points. Using the standard FCM algorithm in the image segmentation, only considers the current pixel value to calculate the membership degree by ignoring neighborhood pixels. Therefore the algorithm is sensitive to noise. If it directly uses FCM, the results are not satisfactory.

An important feature of image is that the adjacent pixel has a high correlation. These neighborhood pixels have similar grayscale values. In general, if the current grayscale is not significantly different from the grayscale of the adjacent pixel points, so they belong to the same type. In this article, we consider the difference of the grey degree between the current point and its neighborhood point as the weight of the sample point to represent the influence degree of the clustering. For large differences between the gray value of each point and the neighborhood, it is expected to obtain smaller weights to reduce their impact on the clustering results. Instead, it needs to obtain a larger value. The hash function exactly meets these requirements, and can effectively reduce the classification error.

The expression based on the hash function to calculate the pixel weight is as follows:

$$w'(x, y) = \frac{1}{l^2 - 1} \sum_{i=-l}^{l} \sum_{i=-l}^{j} e^{-(h(i,j) - h(x,y))^2} \tag{5}$$

where $h(x, y)$ is the grayscale value of the pixel points in the image, l is the neighborhood side length of (x, y). $w'(x, y)$ is the weighting coefficient of the cluster. This normalizes for w', so the clustering point weight coefficient in this article is as follows:

$$w(x, y) = \frac{1}{\sum_{i=1}^{m} \sum_{j=1}^{n} w'(i, j)} w'(x, y) \tag{6}$$

where *m* and *n* are the length and width of the image, respectively.

Formula (6) is introduced into the object function of the FCM to obtain

$$J = \sum_{i=1}^{c} \sum_{j=1}^{n} w_j u_{ij}^m (x_j - x'_i)^2 \tag{7}$$

where c is the cluster number, n is pixel number of the image, w_j represents the cluster weight of the pixel j, u_{ij} is the membership degree of cluster i for pixel j, m is the fuzzy weight index number, and x'_i is the calculated cluster center. The gray value of the pixels is x_j. And the constraint condition of the membership u_{ij} is

$$\sum_{k=1}^{c} u_{kj} = 1, \ \forall\, j = 1, 2, \ldots, n. \tag{8}$$

Within (8), the FCM can be explained that determining the membership and cluster center provides the minimum object

function. Considering the cluster weight and Lagrangian multiplier method, let

$$F = \sum_{i=1}^{c}\sum_{j=1}^{n} w_j u_{ij}^m (x_j - x'_i)^2 + \sum_{j=1}^{n} \lambda_j \left(1 - \sum_{k=1}^{c} u_{kj}\right). \quad (9)$$

The partial derivative for the input variable is assumed to be zero

$$\frac{\partial F}{\partial u_{ij}} = 0 \Rightarrow m u_{ij}^{m-1} w_j (x_j - x'_i)^2 - \lambda_j = 0. \quad (10)$$

Solving the equation, provides the following:

$$u_{ij} = \left(\frac{\lambda_j}{m w_j (x_j - x'_i)^2}\right)^{1/(m-1)}. \quad (11)$$

Let

$$1 - \sum_{i=1}^{c} u_{ij} = 0. \quad (12)$$

Through (11) and (12), we can obtain

$$\left(\frac{\lambda_j}{m w_j}\right)^{\frac{1}{m-1}} = \frac{1}{\sum_{k=1}^{c}\left(\frac{1}{(x_j - x'_k)^2}\right)^{\frac{1}{m-1}}}. \quad (13)$$

Therefore

$$u_{ij} = \frac{(x_j - x'_i)^{\frac{2}{1-m}}}{\sum_{k=1}^{c}(x_j - x'_k)^{\frac{2}{1-m}}}. \quad (14)$$

Let

$$2\sum_{j=1}^{n} w_j u_{ij}^m (x_j - x'_i) = 0. \quad (15)$$

Finally, we get the x'_i

$$x'_i = \frac{\sum_{j=1}^{n} w_j u_{ij}^m x_j}{\sum_{j=1}^{n} w_j u_{ij}^m}. \quad (16)$$

In the membership space, we further consider the role of the neighborhood. If the membership degree of pixel *j* for class *i* and the surrounding neighborhood pixel of *j* has a larger difference, the membership degree is revised to be consistent with the neighborhood pixel. The formula is described below as follows:

$$u'_{ij} = \text{median}(u_{ij}, j \in \Re). \quad (17)$$

After normalization

$$u'_{ij} = \frac{u'_{ij}}{\sum_{k=1}^{c} u'_{kj}}. \quad (18)$$

Meanwhile, the modified FCM segmentation is proceed as shown in the following.

1) Step 1. Input the image.
2) Step 2. Utilizing the minimum–maximum distance method based on the histogram to determine the initial cluster center.
3) Step 3. Using the pixel space neighborhood information to calculate the cluster weight of the pixel.
4) Step 4. According to (14), computing the membership and using (17) and (18) to modify it.
5) Step 5. Adopting the modified membership to calculate the cluster center.
6) Step 6. With the new cluster center and (14), (17), and (18), calculating the membership matrix.
7) Step 7. Comparing the distance between the two adjacent membership matrices, if it is less than a given threshold, the iteration process will stop, and the outputted image segmentation is obtained; Otherwise, go back to step 5 for the next iteration.

C. Selective Search

The selective search was proposed in 2013 [25] by JR for the object recognition. Considerably, the images contain rich information, such as size, color, scale, and texture. In this article, we use the selective search method due to some advantages.

1) Capture all scales. Exhaustive selective adjustments to the window size to different sizes of objects. In the selective search all object scales must be considered, which is most naturally achieved by using a hierarchical algorithm.
2) Diversification. A single optimal strategy is difficult to group regions together. Furthermore, lighting conditions such as shading and the color of the light may influence how regions form an object. The selective search method can avoid this demerit.
3) Rapid Computation. The goal of the selective search is to yield a set of possible object locations for use in a practical object recognition framework. The creation of this set should not become a computational bottleneck. Hence, our algorithm should be reasonably fast.
4) In the selective search method, four similarities are measured and described as follows.
   a) Color similarity $s_{\text{col}}(r_i, r_j)$. L_1 is adopted to normalize the color histograms. Each region r_i with dimensionality $n = 75$, can have a color histogram $C_i = \{c_i^1, \ldots, c_i^n\}$. The color similarity is measured using the histogram intersection

$$s_{\text{col}}(r_i, r_j) = \sum_{k=1}^{n} \min(c_i^k, c_j^k). \quad (19)$$

   b) Texture similarity $s_{\text{tex}}(r_i, r_j)$. The texture histograms $T_i = \{t_i^1, \ldots, t_i^n\}$ are normalized using the L_1 norm. The texture similarity is measured using the following equation:

$$s_{\text{tex}}(r_i, r_j) = \sum_{k=1}^{n} \min(t_i^k, t_j^k). \quad (20)$$

   c) Size similarity $s_{\text{size}}(r_i, r_j)$. This forces small regions to merge together. The size similarity (here, $\text{size}(\text{im})$ denotes the size of an image in pixels) is defined as follows:

$$s_{\text{size}}(r_i, r_j) = 1 - \frac{\text{size}(r_i) + \text{size}(r_j)}{\text{size}(\text{im})}. \quad (21)$$

TABLE I
IMAGE RESOLUTION OF EACH CLASS

Class	Oil tank	Playground	River
Resolution	160~260m	150~250m	100~200m
Location	Canada	Harbin	Puyang

(a) (b) (c)

Fig. 2. Experiment image. (a) Oil tank. (b) Playground. (c) River.

d) Fill similarity $s_{\text{fill}}(r_i, r_j)$. This measures how well the region r_i and r_j fit into each other. The BB_{ij} is defined as the tight bounding box around r_i and r_j

$$s_{\text{fill}}(r_i, r_j) = 1 - \frac{\text{size}(\text{BB}_{ij}) - \text{size}(r_i) - \text{size}(r_j)}{\text{size}(\text{im})}. \tag{22}$$

Finally, the final similarity measure is a combination of the size, texture, fill, and color similarities as follows:

$$s(r_i, r_j) = \varepsilon_1 s_{\text{col}}(r_i, r_j) + \varepsilon_2 s_{\text{tex}}(r_i, r_j) + \varepsilon_3 s_{\text{size}}(r_i, r_j) + \varepsilon_4 s_{\text{fill}}(r_i, r_j) \tag{23}$$

where $\varepsilon \in (0, 1)$ is constant number.

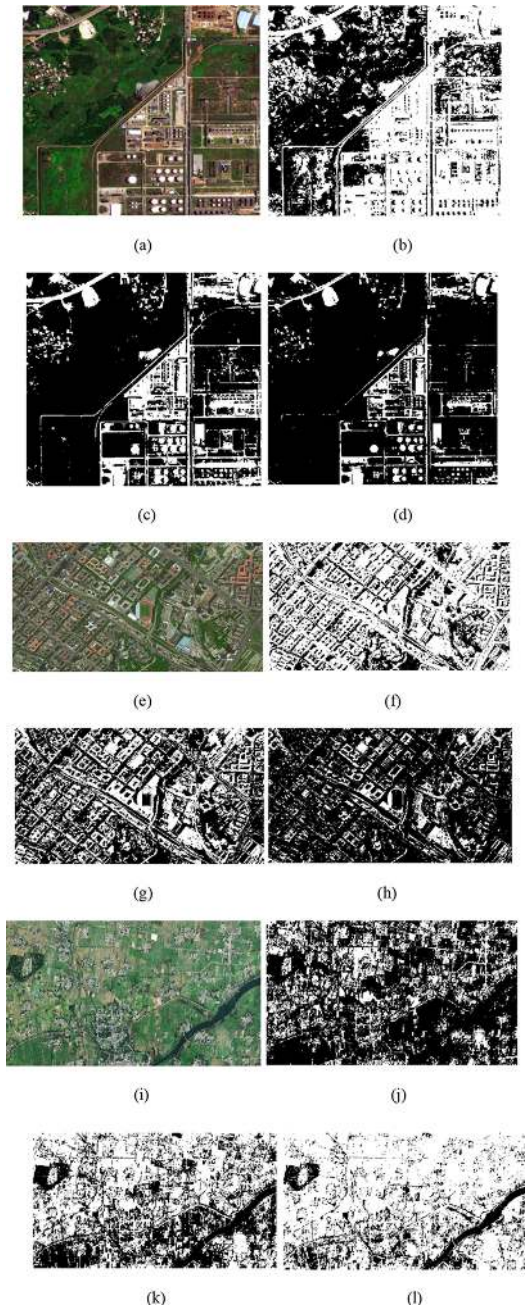
D. Geographic Information

All In this article, we take the geographic information into consideration. For a given image im_κ, the sliding window $W_i = (x_i^p, y_i^p, w_i^p, h_i^p)$ is the pth pixel in the im_κ while the x_i^p and y_i^p are the horizontal and vertical coordinates, respectively. The w_i^p is the width of W_i, and h_i^p is the length of W_i. Finally, the final region is selected to satisfy two criteria: first, x_i^p and y_i^p should be smaller than the upper bound and larger than the lower bound in the ground truth of the image box bound, respectively; second, $w_i^p = \omega \cdot h_i^p$, where $\omega \in [1, 2]$ is a constant number to guarantee that the selected region is regular.

III. EXPERIMENTS AND RESULTS

A. Data Set

Public data sets intended for HRS are very scarce. Therefore, we collect three remote sensing experience data from Google Earth, namely, oil tank, playground, and river. In Table I, we display the image resolution for each class. The size of oil tank, playground, and river are 256 × 256, 700 × 400, 800 × 500 pixels, respectively. Fig. 3 shows the test image sample. Fig. 2(a) describes the oil tank area, Fig. 2(b) depicts the playground area, and Fig. 2(c) exhibits the river area.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k) (l)

Fig. 3. Saliency maps comparison displayed in the order by the (a), (e), and (i) Original images. (b), (f), and (j) Otsu method. (c), (g), and (k) SF method. (d), (h), and (l) MFCM method.

TABLE II
PERFORMANCE COMPARISON OF WITHOUT MFCM AND PROPOSED MODEL

Class	Oil tank	Playground	River	mAP	Time/s
Without MFCM	69.78%	89.65%	70.65%	73.24%	0.157
Proposed	95.34%	93.48%	83.67%	81.22%	0.203

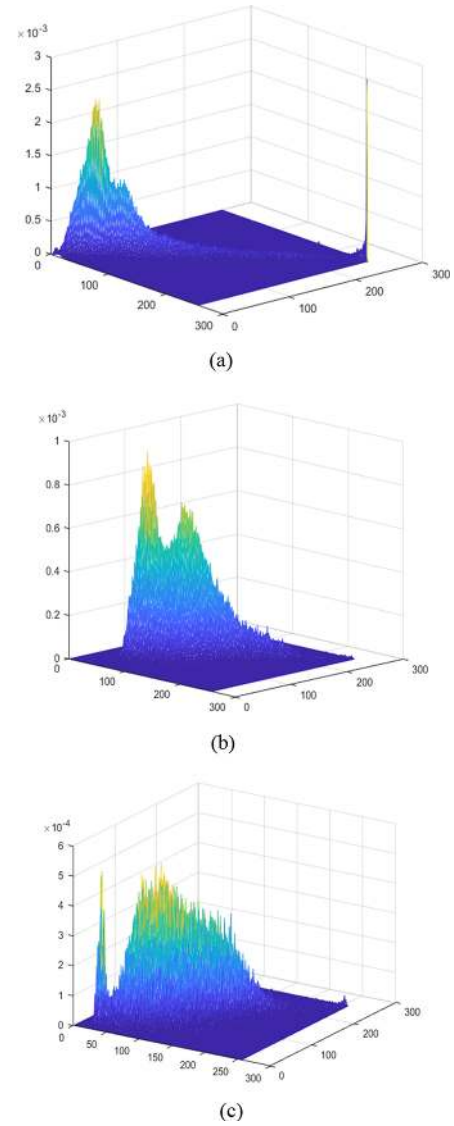


Fig. 4. 2-D histogram. (a) Oil tank. (b) Playground. (c) River.

TABLE III
SELECTION RESULTS FOR OIL TANK, PLAYGROUND, AND RIVER IN TERMS OF OVERLAP

Method	Oil tank	Playground	River
FCMK	75.9%	88.7%	79.7%
STPP	81.6%	89.4%	82.5%
STDI	85.7%	91.1%	90.8%
Proposed	94.2%	93.1%	92.5%

B. Evaluation of HRS

To evaluate the quality of our proposed method, we use the mean average best overlap (MABO) scores to quantitatively showcase the performance. To calculate the MABO for a specific class c, we calculate the best overlap between each ground truth annotation $g_i^c \in G^c$ and the object hypotheses l generated for the corresponding image. The overlap score is taken from and measurements of the intersection area of the two regions divided by their union

$$\text{Overlap}(g_i^c, l_j) = \frac{\text{area}(g_i^c) \cap \text{area}(l_j)}{\text{area}(g_i^c) \cup \text{area}(l_j)}. \tag{24}$$

In this article, mean average precision (mAP) is also used as the criteria for hot region detection. If the mAP value is larger, the hot region detection accuracy of the framework is greater. The mAP is calculated as the following formula.

$$\text{mAP} = \frac{1}{n} \sum_{i}^{n} \text{AP}_i \tag{25}$$

where n is the number of region classes and i is the label of a class. AP_i is the average precision of label i class and is the area under the precision-recall curve (PRC).

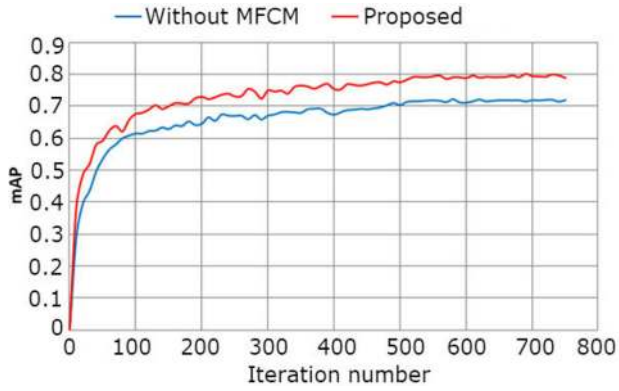


Fig. 5. mAP trend diagram of without MFCM and proposed model.

TABLE IV
SSIM COMPARISONS WITH DIFFERENT METHODS LARGER VALUES INDICATE BETTER PERFORMANCE

Method	Oil tank	Playground	River
FCMK	0.617	0.622	0.638
STPP	0.652	0.685	0.692
STDI	0.693	0.728	0.716
Proposed	0.786	0.792	0.774

TABLE V
PSNR COMPARISONS WITH DIFFERENT METHODS LARGER VALUES INDICATE BETTER PERFORMANCE

Method	Oil tank	Playground	River
FCMK	60.885	57.645	59.834
STPP	62.521	60.582	62.071
STDI	63.413	61.225	62.843
Proposed	66.327	67.541	66.296

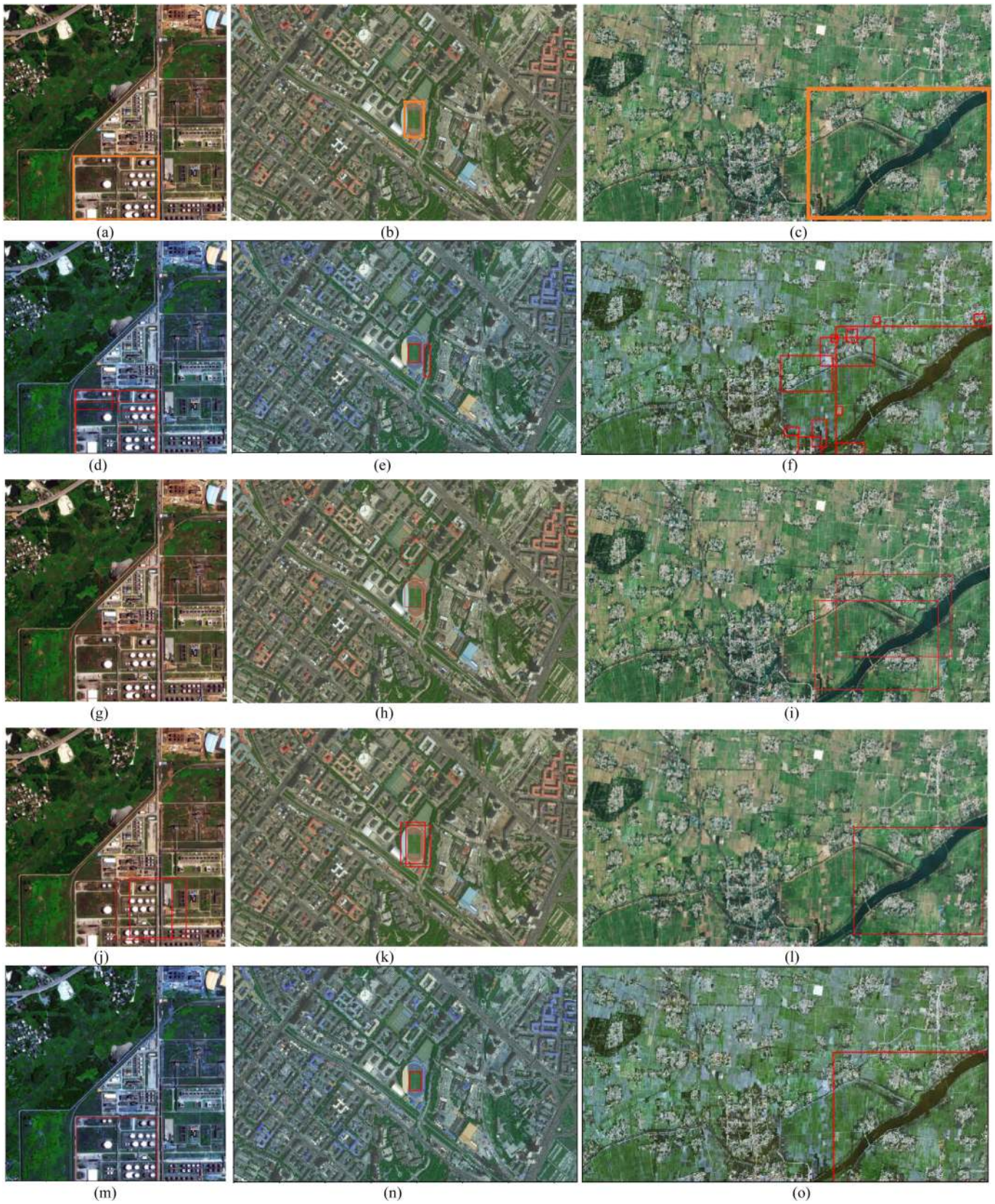


Fig. 6. Selection results for oil tank, playground, and river. From left to right (a), (b), and (c) Ground truth. (d), (e), and (f) FCMK method. (g), (h), and (i) STPP method. (j), (k), and (l) STDI method. (m), (n), and (o) Proposed method.

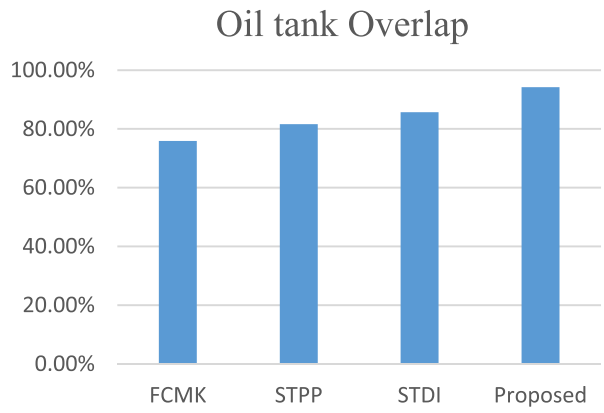


Fig. 7. Overlap comparison of FCMK, STPP, STDI, proposed method for oil tank.

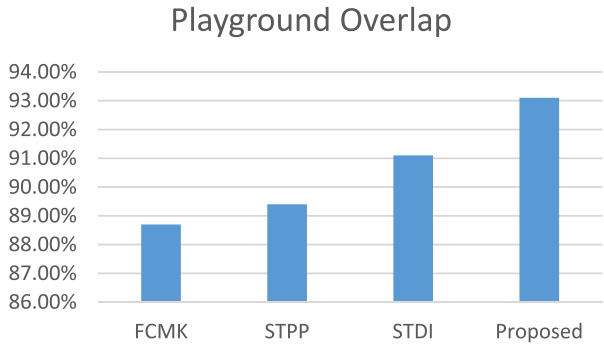


Fig. 8. Overlap comparison of FCMK, STPP, STDI, proposed method for playground.

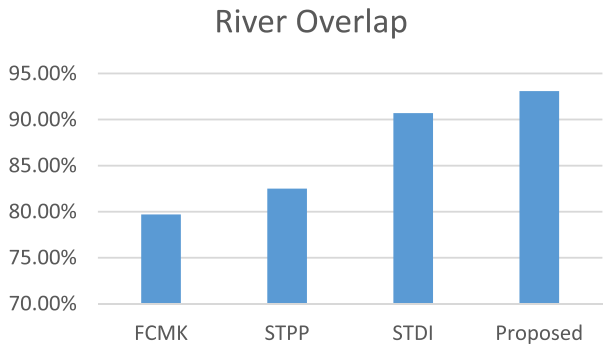


Fig. 9. Overlap comparison of FCMK, STPP, STDI, proposed method for river.

Structural similarity index measure (SSIM): The SSIM is to measure the level of similarity between images. Its value varies between -1 and 1, where one indicates that the compared images are identical. The SSIM is calculated as follows:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_x^2 + c_1)(\sigma_x^2 + \sigma_x^2 + c_2)} \tag{26}$$

where μ_x, μ_y are the mean of x and y, respectively. σ_x^2, σ_y^2 are the variances of x and y, respectively. σ_{xy} is the covariance of both x and y. Two variables c_1 and c_2 are used to stabilize the division with a weak denominator. Having calculated the SSIM between the detected image and the source image, the result is given in the form of a weighted sum. Also, peak to signal ratio (PSNR) is used for evaluating the effect, the detailed information is displayed in reference [26].

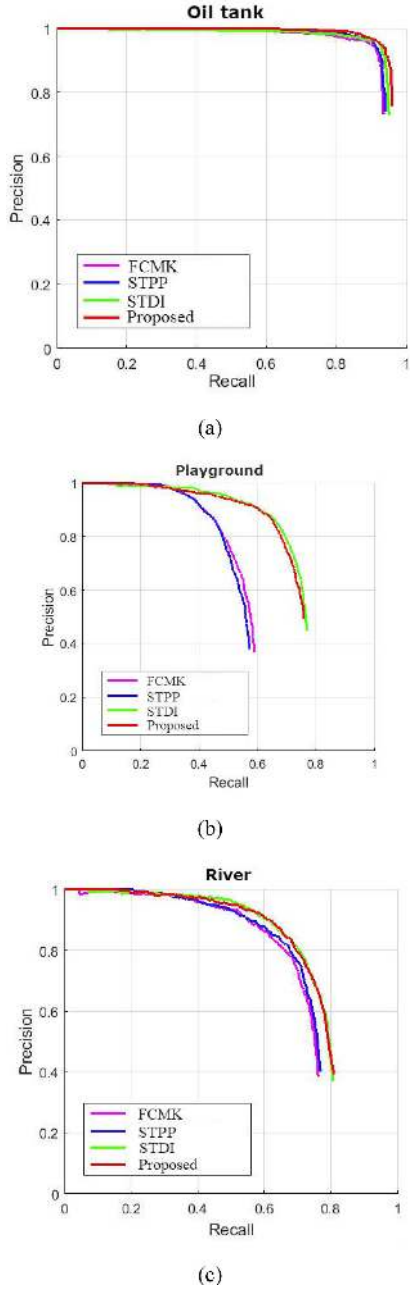


Fig. 10. PRCs of FCMK, STPP, STDI, and proposed models. (a) Oil tanks. (b) Playgrounds. (c) Rivers.

Experiments are conducted on Python3.1 and Opencv3.0 with Windows8.1, CPU2.20 GHz, and RAM 8.0 G. We define that an overlap score that is larger than 50% is considered to be found.

C. Hot Region Detection

In the preliminary extraction, the modified FCM (abbreviated to MFCM) is adopted to generate a saliency map for the raw

image. We compare the MFCM with the Otsu [27] and SF models [28], and the results are displayed in Fig. 4.

It is noted that Fig. 3 demonstrates that the saliency maps obtained from the Otsu model and the SF model are redundant textures, and background information, which may largely lead to fuzzy edges and increase the computational cost in the subsequent processing. The MFCM method suppresses the background texture effectively.

During the selection process, we utilize the selective search method to extract all the possible regions. Next, the geographic information feature is used to merge the regions to obtain the hot region in the remote sensing image. For a visual representation, we exhibit their 2-D histograms in Fig. 4.

Fig. 5 shows the mAP trend of without MFCM and proposed model (with MFCM) as the number of iterations increasing. In Fig. 5, with an increase in the number of iterations, the mAP values of proposed model are larger than without MFCM. The experimental results show that proposed model can attain better hot region detection accuracy than without MFCM for remote sensing images.

Table II shows the quantitative comparison results of without MFCM and proposed model measured by AP values when the mAP values of without MFCM and proposed are greatest in their respective mAP trend diagrams. In Table II, the AP values of proposed model for oil tanks, playgrounds, and rivers are greater than those of without MFCM. In particular, the AP values of without MFCM for oil tanks, playgrounds, and rivers are 0.6978, 0.8965, and 0.7065, respectively. The AP values of proposed model for oil tanks, playgrounds, and rivers are 0.9534, 0.9348, and 0.8367, respectively. We obtain significant performance improvement for the hot regions. Furthermore, the mAP values of without MFCM and proposed model are 0.7324 and 0.8122, respectively. Those shows that proposed model can obtain better hot region detection results than without MFCM in remote sensing images.

Table II also shows the average consuming time of hot region detection for per image using the two methods in the test set. In Table II, it is efficient for hot region detection using the two methods, and hot region detection using without MFCM takes less time than proposed model due to additional MFCM process.

The region selection results obtained by our proposed model are compared with two competing models, namely, fuzzy c-means and K-means algorithm (FCMK) [29], superpixel-to-pixel process (STPP) [30], and statistical distinctiveness (STDI) [31] qualitatively and quantitatively. Fig. 6 shows the region selection results with the different methods.

Fig. 6(d)–(f) is the selection results of the FCMK method, Fig. 6(g)–(i) shows the selection results of the STPP method, Fig. 6(g)–(i) shows the selection results of the STDI method, and Fig. 6(m)–(o) shows the results of the proposed method. As listed in the results, the FCMK method suffers from many false selections. The STPP and STDI method achieves a better detection performance than the FCMK, but it is worse than the proposed method. The region selection performance of the proposed method is significantly better than the performances of the other two methods. Table III shows the selection results for the oil tank, playground, and river using the overlap radio with the proposed method. Furthermore, Figs. 8–10 are the comparison curves with different methods to determine *overlap*.

Figs. 7–9 (*y*-axis is the overlap percentage) show the overlap curves for the selection methods as applied to the three test sequence. As seen from the figures, the overlap of the proposed method is always in a higher-ranking position than the other tested sequences. Hence, it achieves a higher selection accuracy on the HRS.

Fig. 10 shows the PRCs of FCMK, STPP, STDI, and proposed model. As seen, for oil tanks, playgrounds, and rivers, the AP values of proposed model are greater than those of FCMK, STPP, and STDI. This can be easily explained because the selective search method in proposed model is more suitable for hot regions than those of other methods. Better hot region detection results are obtained using our method.

The SSIM and PSNR results of the quantitative comparison of the abovementioned three images are shown in Tables IV and V. According to the data, no matter the circumstance, our method obtained the best visual results. It yielded the best performance in terms of SSIM and PSNR. This shows that our method is universal and effective, which is superior to other enumeration methods.

IV. Conclusion

In this article, we propose a HRS method based on the modified FCM and selective search approach in remote sensing images. In the first stage, a GC filter is used to preprocess large-scale remote sensing images. In the second stage, a modified fuzzy c-means segmentation method is utilized to segment the image. Then, an enhanced selective search method is adopted to establish well-defined boundaries for the HRS and improve the immunity to noise. Additionally, geographic information is presented in this phase, which is conducive to improve the detection accuracy. In addition, we investigated the influences of different segmentation methods on the detection performance. These results can help guide other researchers to obtain good results. The results of the experiments indicate that the proposed HRS method is both simple and robust. In further work, we will continue to enhance this method and improve its selection performance.

References

[1] L. Onapa and A. Somjit, "Determination of the appropriate parameters for K-means clustering using selection of region clusters based on density DBSCAN (SRCD-DBSCAN)," *Expert Syst.*, vol. 34, no. 3, May 2017, Paper e12204.

[2] J. Lv, L. Zhang, and S. Wang, "Region of interest extraction based on saliency detection and contrast analysis for remote sensing images," *Proc. SPIE*, vol. 10004, Oct. 2016, Art. no. 1000424.

[3] L. Wen, J. Zhu, J. Pei, and L. Wang, "Thick clouds region detection base on double scale background Gauss model in large scale high resolution ocean remote sensing images," in *Proc. Int. Conf. Signal Process.*, 2017, pp. 783–787.

[4] L. Zhang, B. Qiu, X. Yu, and J. Xu, "Multi-scale hybrid saliency analysis for region of interest detection in very high resolution remote sensing images," *Image Vis. Comput.*, vol. 35, pp. 1–13, 2015.

[5] H. Grailu, "Textual image compression for maintaining or improving the recognition performance," *Circuits Syst. Signal Process.*, vol. 36, no. 2, pp. 1–17, 2016.

[6] M. Abozahhad, R. Gharieb, S. Ahmed, and M. Abd-Ellah, "Huffman image compression incorporating DPCM and DWT," *J. Signal Inf. Process.*, vol. 6, no. 2, pp. 123–135, 2015.
[7] P. Roy, A. Bhunia, and U. Pal, "Date-field retrieval in scene image and video frames using text enhancement and shape coding," *Neurocomputing*, vol. 274, pp. 37–49, 2018.
[8] I. Chaabouni, W. Fourati, and M. Bouhlel, "Using ROI with ISOM compression to medical image," *Int. J. Comput. Vis. Robot.*, vol. 6, no. 1/2, pp. 65–76, 2016.
[9] Q. Wang and X. Hu, "Scene analysis for effective visual search in rough three-dimensional-modeling scenes," *J. Electron. Imag.*, vol. 2, no. 6, 2016, Art. no. 061622.
[10] X. Li, R Feng, X Guan, H. Shen, and L. Zhang, "Remote sensing image mosaicking: Achievements and challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 4, pp. 8–22, Dec. 2019.
[11] H. Ma and Y. Yang, "Two specific multiple-level-set models for high-resolution remote-sensing image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 558–561, Jul. 2009.
[12] L. Wang, S. Wang, and L. Zhang, "An efficient visual saliency analysis model for region-of-interest extraction in high-spatial-resolution remote sensing images," *Proc. SPIE*, vol. 9988, Oct. 2016, Art. no. 99880W.
[13] R. Rojas, X. Huang, and K. Ou, "Region of interest detection and evaluation in functional near infrared spectroscopy," *J. Near Infrared Spectroscopy*, vol. 24, no. 4, pp. 317–326, 2016.
[14] J. Chen and L. Zhang, "Joint multi-image saliency analysis for region of interest detection in optical multispectral remote sensing images," *Remote Sens.*, vol. 8, no. 6, 2016, Art. no. 461.
[15] T. Li, J. Zhang, X. Lu, and Y. Zhang, "SDBD: A hierarchical region-of-interest detection approach in large-scale remote sensing image," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 699–703, May 2017.
[16] A. Ablavatski, S. Lu, and J. Cai, "Enriched deep recurrent visual attention model for multiple object recognition," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2017, vol. 1, pp. 971–978.
[17] X. Wang and H. Duan, "Hierarchical visual attention model for saliency detection inspired by avian visual pathways," *IEEE/CAA J. Autom. Sinic a*, vol. 6, no. 2, pp. 540–552, Mar. 2019.
[18] H. Feng, G. Chang, and X. Guo, "Hash algorithm for color image based on super-complex fourier transform coupled with position permutation," *J. Frontier Comput. Sci. Technol.*, vol. 11, no. 11, pp. 1837–1848, 2017.
[19] S. M. M. Kahaki, M. Fathy, and M. Ganj, "Road-following and traffic analysis using high-resolution remote sensing imagery," in *Proc. Int. Workshop Intell.Vehicle Controls Intell.Transp. Syst.*, 2015, pp. 1–10.
[20] S. M. M. Kahaki, J. Nordin, and A. H. Ashtari, "Incident detection algorithm based on radon transform using high-resolution remote sensing imagery," in *Proc. Int. Conf. Elect. Eng. Informat.*, Jul. 2011, 1–5.
[21] Y. Gong and I. Sbalzarini, "Curvature filters efficiently reduce certain variational energies," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1786–1798, Apr. 2017.
[22] L. Teng, H. Li, and S. Yin, "Modified pyramid dual tree direction filter-based image de-noising via curvature scale and non-local mean multigrade remnant multi-grade remnant filter," *Int. J. Commun. Syst.*, vol. 31, no. 16, Nov. 2018, Art. no. e3486.
[23] Q. Zhang, C. Zhu, T. Yang, Z.i Chen, L. Zhao, and P. Li, "An incremental CFS algorithm for clustering large data in industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1193–1201, Jun. 2017.
[24] G. Liu, X. Liang, and J. Zhang, "Contourlet transform and improved fuzzy c-means clustering based infrared image segmentation," *Syst. Eng. Electron.*, vol. 33, no. 2, pp. 443–448, 2011.
[25] J. Uijlings, K. Sande, T. Gevers, and A. Smeulders, "Selective search for object recognition," *Int. J. Comput. Vis.*, vol. 104, no. 2, pp. 154–171, 2013.
[26] S. M. M. Kahaki, H. Arshad, M. J. Nordin, and W. Ismail. "Geometric feature descriptor and dissimilarity-based registration of remotely sensed imagery," *PLoS One*, vol. 13, no. 7, 2018, Art. no. e0200676.
[27] X. Yuan, L. Wu, and Q. Peng, "An improved Otsu method using the weighted object variance for defect detection," *Appl. Surf. Sci.*, vol. 349, pp. 472–484, 2015.
[28] X. Zhang, X. Sun, C. Xu, and G. Baciu, "Multiple feature distinctions based saliency flow model," *Pattern Recognit.*, vol. 54, pp. 190–205, 2016.
[29] R. Mandal, M. Gupta, and C. Kar, "Automated ROI detection for histological image using fuzzy c-means and K-means algorithm," in *Proc. Int. Conf. Elect., Electron., Optim. Technol.*, Nov. 2016, pp. 1173–1178.
[30] L. Ma, B. Du, H. Chen, and Q. Soomro, "Region-of-interest detection via superpixel-to-pixel saliency analysis for remote sensing image," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1752–1756, Dec. 2016.
[31] G. Liu, L. Qi, Y. Tie, and L. Ma, "Region-of-interest detection based on statistical distinctiveness for panchromatic remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 2, pp. 271–275, Feb. 2019.

Shoulin Yin received the B.S. and M.S. degrees in image proccessing from Software College, Shenyang Normal University, Shenyang, China, in 2013 and 2015, respectively. He is currently working toward the Ph.D. degree with the School of Electronic and Information Engineering, Harbin Institute of Technology, Harbin, China.

He is a Lecturer with Software College, Shenyang Normal University. His research interests include image fusion, target detection, and image recognition.

Hang Li received the B.S., M.S., and Ph.D. degrees in computer applications technology from Dalian Fisheries University, Dalian, China, Shenyang Institute of Technology, Shenyang, China, Northeastern University, Shenyang, China, in 1999, 2002, and 2005, respectively.

In 2002, he joined Software College as a Teacher. He is a full Professor and Master Supervisor. He is an Outstanding Young Backbone Teacher of Liaoning General Institutions of Higher Learning (Liaoning Education Department 200612225). His research interests include image analysis and processing, big data, and cloud computing.

Dr. Li won the Second Prize of National Defense Science and Technology Award of The Commission of Science, Technology and Industry for National Defense (National Defense Science, Technology and Industry Commission 2005GFJ2126-6) and First Prize of Science and Technology Award of China North Industries Group Corporation (2005-BQJ-1-0019-6).