

# Remote Sensing Image Segmentation Based on a Novel Gaussian Mixture Model and SURF Algorithm

Shoulin Yin, College of Information and Communication Engineering, Harbin Engineering University, Harbin, China\*

 <https://orcid.org/0000-0002-5367-1372>

Liguo Wang, College of Information and Communications Engineering, Dalian Minzu University, Dalian, China

Qunming Wang, College of Surveying and Geo-Informatics, Tongji University, Shanghai, China

Jinghui Yang, School of Information Engineering, China University of Geosciences, Beijing, China

Man Jiang, Liaoning Vocational Technical College of Modern Service, China

## ABSTRACT

This paper proposes a novel remote sensing image segmentation method based on Gaussian mixture model and SURF algorithm. Firstly, Gaussian mixture model is used for remote sensing image segmentation. Then the surf matching algorithm is adopted for eliminating misidentified areas. The determinant of Hessian matrix (DoH) is used to detect key points in the image. The non-maximum suppression method and interpolation operation are utilized to search and locate the extreme points. The maximum likelihood method is used to estimate model parameters. Some remote sensing images in THE DOTA data set are selected for experimental verification, and the results show that the new algorithm has obvious improvement in segmentation effect and efficiency. In the background complex image segmentation, the improved algorithm has more obvious advantages compared than state-of-the-art segmentation methods.

## KEYWORDS

Gaussian Mixture Model, Maximum Likelihood, Remote Sensing Image Segmentation, SURF Algorithm

## 1. INTRODUCTION

Image segmentation is an important stage in image processing, which has been widely used in the fields of image recognition, 3D reconstruction and target tracking (Hong et al. 2021; Teng et al. 2022). The areas of interest in the image are called the target or foreground, and the rest are called the background. The essence of image segmentation is to separate and extract the target area in the image (Hatamizadeh et al. 2022).

With the rapid development of remote sensing sensor technology, the spatial resolution of remote sensing image is constantly improved and the information contained is more abundant, so it is widely used in many fields (Tesfamikael et al. 2021). However, due to the improvement of resolution, the

DOI: 10.4018/IJSIR.322301

\*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

spectral heterogeneity of similar features is enhanced, which makes the processing of remote sensing images more difficult. Image segmentation is a key link in the process of image processing, and the results have an important impact on the subsequent work of image processing (Cheng et al. 2021; Minaee et al. 2021). Remote sensing image segmentation covers two basic tasks: 1) determine the number of ground objects (the number of pixels) contained in the region to be segmented in the remote sensing image, and 2) accurately segment the object regions.

In modern remote sensing image segmentation, due to the improvement of spatial resolution, various types of ground objects, complex background environment and difficulties in field investigation, it is difficult to manually interpret the number of land objects covered in remote sensing image (Xu et al. 2021; Liu et al. 2021). If the number of terrestrial species cannot be correctly pointed out, it will lead to wrong segmentation result. In addition, high resolution enhances the ability of remote sensing image to express the details of the covered surface. However, on the other hand, excessive detail expression will lead to the difference increase of the spectral measure of pixels in the similar ground objects (manifested as the increase of noise), which causes great difficulties for the segmentation (Lu 2021; Wang et al. 2022). Therefore, it has always been an important task in the research of high-resolution remote sensing image segmentation to automatically determine the class number and improve the anti-noise performance of the algorithm (Li et al. 2021).

In the many remote sensing image segmentation methods, the statistical segmentation framework is the most effective method. The segmentation method based on hybrid model fits the image to be segmented with the weighted sum of multiple probability distributions. The probability distribution of each mixed component describes the probability statistical distribution law of pixel spectral measure in a specific target class (Yuan et al. 2021). Among them, GMM (Gaussian mixture model) is widely used in remote sensing image segmentation because of its simple principle, intuitive structure and easy realization (Zhu et al. 2022; Chen et al. 2021). However, in the traditional GMM, only the gray information of the image is used, and the spatial relationship between the neighborhood pixels is not introduced, which leads to phenomenon that the segmentation result is very sensitive to the image noise.

At present, image segmentation algorithms can be divided into supervised image segmentation method, unsupervised image segmentation method (Mirsadeghi et al. 2021) and interactive image segmentation method (Chen et al. 2021). The interactive image segmentation is to obtain the region containing the target with the help of artificial interactive operation, and the segmentation results are more accurate. Here, GrabCut algorithm has convenient interactive mode (Salau et al. 2021) and good segmentation effect, it has become a research hotspot in the field of image segmentation. Gupta et al. (2021) introduced the concept of graph theory into the field of image segmentation. The image was mapped to a weighted graph with pixels as nodes. Edges between nodes represented the relationship between pixels, and weights represented the degree of correlation between pixels. Finally, the maximum flow/minimum cut was used to segment the graph. The image segmentation problem was transformed into the graph vertex labeling problem. The graph cuts algorithm based on Graph theory was proposed by Bao et al in (2001), and the interactive GrabCut color image segmentation algorithm (Koresh et al. 2021), both of which belonged to the interactive image segmentation algorithm based on graph theory. In references (Ma et al. 2021; Di et al. 2021), the superpixel method was used to pre-process images to improve the segmentation accuracy and time efficiency. Zhang et al (2021) used significant preprocessing to improve the segmentation performance. In reference (Jin et al., 2021), Bayesian classification and simple linear generation selection clustering were used to improve the algorithm. In reference (Yong et al. 2017), GrabCut texture of wavelet transform was used for image segmentation. However, the segmentation effect of the above algorithms at the foreground background junction is not ideal. The segmentation results have the phenomenon of under-segmentation to some extent. Moreover, the algorithm optimizes the Gaussian mixture model through iterative learning, which will take a certain amount of time, resulting in low segmentation efficiency.

In order to solve the problem that GMM algorithm is susceptible to image noise, many scholars have introduced spatial constraints into the design of remote sensing image segmentation. GMM combined with MRF (Markov random field) has become a research hotspot in the field of image

segmentation. A spatially variant finitemixture model (SVFMM) image segmentation method was proposed in reference (Juan-Albarracín et al. 2021). Based on GMM, the weight coefficient of Gaussian component was expressed as the prior probability of each pixel. The spatial correlation of the prior probability was modeled by Gibbs distribution to improve the anti-noise performance of the algorithm. Although this method reduced the sensitivity of segmentation results to image noise, the algorithm was not flexible in smoothing noise because the parameter  $\beta$  of smoothing noise control ability in prior distribution was set as a constant. In order to further improve the anti-noise performance of the algorithm and overcome the defects of SVFMM algorithm, A-SVFMM (adaptive spatially variant finite mixture model) was proposed in reference (Ji et al. 2014) based on SVFMM algorithm. The algorithm used MRF to re-model prior distribution. It was assumed that the prior probability obeyed a normal distribution with a mean of 0 and a certain variance. Its variance was set as a random variable related to the class label. The maximum likelihood (ML) function method was used to solve the model parameters. Due to the need to meet the constraints of prior probability, the process of solving the analytic expression of prior probability was complicated. Although this method improved the flexibility of the algorithm against noise, it increased the difficulty of solving the model parameters. At the same time, the accuracy of the segmentation results of high resolution remote sensing images was not high.

Although the above methods improve the anti-noise performance of the algorithms to a certain extent, it does not realize the automatic determination of image class number. To automatically determine the number of class members, many scholars have proposed a variable class image segmentation method based on Gaussian mixture model. Pravitasari et al. (2021) proposed an image segmentation algorithm based on RJMCMC (reversible jump Markov chain Monte Carlo) method with multivariate Gaussian mixture model. The sampling simulation operation of the algorithm included label field sampling, Gaussian distribution parameter sampling, mixed weight coefficient sampling, MRF parameter sampling and class number sampling. The algorithm could determine the most possible class number of the image and estimate the model parameters to achieve the purpose of image segmentation. However, this algorithm was only applied to color image segmentation, not to high resolution remote sensing image. Meanwhile, the RJMCMC method used in this algorithm had many steps and was difficult to implement. Baumgartner et al. (2015) put forward a remote sensing image segmentation algorithm combining GMM and MRF under the framework of Bayesian theory, and took a maximum posterior probability as the parameter to achieve the goal of image segmentation. A new energy function was proposed to describe the difference between pixel spectral measure values and the distance between elements, which improved the anti-noise performance of the algorithm. But the form of energy function was complex and did not have universal applicability. In order to automatically determine the class number, the method specified a number of class numbers in advance within a certain possible range, and realized the corresponding optimal segmentation for each specified class number. The entropy function of optimal segmentation was calculated one by one, and the maximum entropy criterion (Rafique et al. 2022) was adopted to determine the optimal class number, that is, the corresponding entropy value was obtained by calculating the class number one by one, from which the class number corresponding to the convergence point of entropy value was the optimal class number. Although the maximum entropy criterion could determine the class number of images, the process of determining the class number needed to perform fixed class number segmentation for many times, which led to large computation and low efficiency.

Therefore, we propose a novel remote sensing image segmentation method based on Gaussian mixture model and SURF algorithm. The main contributions are as follows. Firstly, Gaussian mixture model is used for remote sensing image segmentation. Then the SURF matching algorithm is adopted for eliminating misidentified areas. The determinant of Hessian matrix (DoH) is used to detect key points in the image. The non-maximum suppression method and interpolation operation are utilized to search and locate the extreme points. The main direction of the key points is determined by gray centroid method. The key points are then described using binary descriptors. The maximum likelihood method is used to estimate model parameters.

## 2. RELATED WORKS

### 2.1 Definition of Image Segmentation

Image segmentation is an important task to understand images and extract effective information from images. It is a hot research problem in image processing and computer vision. Image segmentation technology divides the image into several independent and unique regions, which is the key step of image processing in machine vision. Image segmentation techniques are mainly divided into the following categories: the threshold-based, region-based, edge-based and the segmentation methods based on specific theories (Wang et al. 2021). Image segmentation is defined mathematically as the process of dividing an image into non-intersecting regions. The process of image segmentation is also a marking process, that is, the pixels belonging to the same area are assigned the same number. The mathematical representation is as follows.

Let  $R$  represent the entire space area occupied by an image, image segmentation can be regarded as the process of dividing  $R$  into  $n$  sub-areas  $(R_1, R_2, \dots, R_n)$ . The following conditions should be followed:

1.  $\bigcup_{i=1}^n R_i = R$ ;
2.  $R_i$  is a connected set,  $i = 1, 2, \dots, n$ ;
3.  $R_i \cap R_j = \emptyset$  for all  $i$  and  $j$ ;
4.  $Q(R_i) = true$ ,  $i = 1, 2, \dots, n$ ;
5.  $Q(R_i \cup R_j) = false$  for any  $R_i$  and the adjacent region of  $R_j$ .

where  $Q(R_k)$  is a logical property defined at the point of set  $R_k$ , and  $\emptyset$  represents the empty set. The symbols  $\cup$  and  $\cap$  represent the union and intersection of sets, respectively. If the union of  $R_i$  and  $R_j$  forms a connected set, the two regions are adjacent.

Conditions 1-2 are interpreted as follows: the segmentation must be complete, that is, any pixel belongs to a unique region, and the points in a region are connected in some predefined way (that is, the points must be 4-connected or 8-connected). The regions must be disjoint. The pixels in the segmented area all have the same gray level. The two adjacent regions  $R_i$  and  $R_j$  must be different in the sense of the property  $Q$ .

### 2.2 Image Segmentation Based on Clustering

There is no unified general theory of image segmentation technology. Combining with many new theories and methods, a variety of image segmentation algorithms have been proposed. Clustering is an important method, and the feature space clustering algorithm is used to cluster pixels of similar features in the image. Based on the clustering results of pixel points, the feature space of the cluster is mapped to the original image to get the segmentation results.

GMM is a probabilistic model which is often used in image segmentation. The theoretical background is to aggregate samples into different Gaussian models according to intra-class distance or probabilistic likelihood. In the sample space, if the distance between two samples is closer, the probability of being clustered into the same class is higher.

The advantages of clustering algorithm in image segmentation are simple, fast, high efficiency and strong extensibility for large data sets. The time complexity of this algorithm is linear, which is suitable for the processing of high dimensional images. But an important problem is that the clustering algorithm is sensitive to the initial value and cannot estimate the number of clusters effectively. In addition, clustering algorithms (Yin. et al. 2020) (such as EM) aim at minimizing distance or maximizing likelihood, so their performance is poor in non-convex Gaussian models.

### 3. PROPOSED REMOTE SENSING IMAGE SEGMENTATION METHOD

In this section, we conclude the proposed method as shown in figure 1. It mainly contains three parts: GMM for probabilistic estimation, SURF feature matching and Ratio detection for removing mis-matched points.

#### 3.1 Gaussian Mixture Model

GMM uses Gaussian probability density function to accurately quantify images and decomposes an image into several models based on Gaussian probability density function stroke. The process of GMM image segmentation is as follows. Suppose the image is divided into  $K$  parts, and the pixels in each image region obey the normal distribution with mean  $\mu$  and variance  $\sigma^2$ . Then the feature distribution of the whole image can be described by GMM.

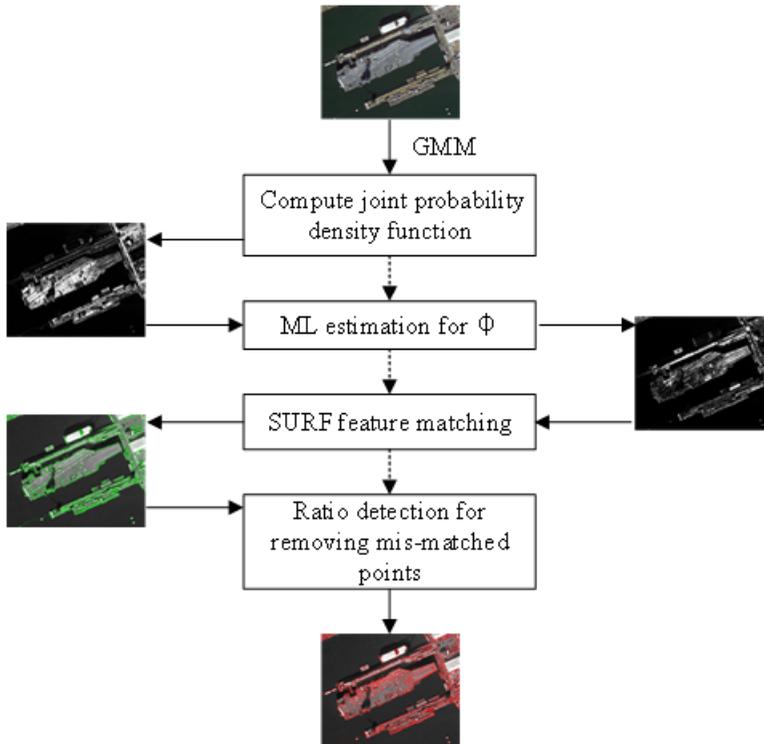
If the Gaussian distribution of the  $k$ -th region is shown in Formula (1), then the Gaussian distribution of the whole image can be expressed by formula (2):

$$f_k(x; \theta_k) = f_k(x; \mu_k; \sigma_k) = (2\pi\sigma_k^2)^{-0.5} \cdot \alpha \left( -0.5(x - \mu_k)^2 / \sigma_k^2 \right) \quad (1)$$

$$f_k(I_{xy}; \Phi) = \sum_{k=1}^K \pi_k f_k(I_{xy}; \theta_k) = \sum_{k=1}^K \pi_k (2\pi\sigma_k^2)^{-0.5} \cdot \alpha \left( -0.5(I_{xy} - \mu_k)^2 / \sigma_k^2 \right) \quad (2)$$

$$\sum_{k=1}^K \pi_k = 1, \text{ and } 0 \leq \pi_k \leq 1, k = 1, 2, \dots, K \quad (3)$$

Figure 1.  
Proposed remote sensing image segmentation structure



$I_{xy}$  is the gray pixel value in row  $x$  and column  $y$ .  $\pi_k$  is the probability weight of  $I_{xy}$  belonging to  $k$ -th class satisfying equation (3).  $f_k(I_{xy}; \theta_k)$  is the probability density function of  $I_{xy}$  belonging to  $k$ -th class.  $\theta = (\mu_1, \mu_2, \dots, \mu_K; \sigma_1^2, \sigma_2^2, \dots, K)$ . The unknown vectors are formed by all parameters  $\Phi = (\pi_1, \pi_2, \dots, \pi_n; \theta_1, \theta_2, \dots, \theta_K)$ .

To achieve image segmentation, it is necessary to obtain the joint probability density function of  $\pi_k$  and  $k$  under the condition of known images  $x$  and  $\theta$ . According to Bayes theorem, the joint probability density function is:

$$p(\pi_k, k | X, \theta) \propto p(X | \pi_k, \theta, k) p(\pi_k | k) p(k) \quad (4)$$

$$= \prod_{i=1}^N \sum_{j=1}^k \pi_{ij} \Phi(x_i | \theta_j) \exp \left[ \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^K G_{ij} \log \pi_{ij} \right] \cdot \frac{\lambda^k}{k!} e^{-\lambda}$$

The parameters in Equation (4) include: class number  $k$ , mean vector  $\mu = \{\mu_j, j = 1, 2, \dots, k\}$ , the standard deviation vector  $\sigma = \{\sigma_j, j = 1, 2, \dots, k\}$ , Gaussian component weight coefficient  $\pi_k = \{\pi_j, j = 1, 2, \dots, k\}$  and  $\pi_j = \{\pi_{ij}, j = 1, 2, \dots, k\}$ . Let the parameter set be  $\varphi = \{\mu, \sigma, \pi_k\}$ .

### 3.2 Solving $\varphi$ With the ML Estimation

We construct the  $\pi_k$  logarithmic likelihood estimation function as:

$$L(\pi_k, k | X, \theta) = \log[p(\pi_k, k | X, \theta)] \quad (5)$$

$$= \sum_{i=1}^N \log \left[ \sum_{j=1}^k \Phi(x_i | \theta_j) \right] + \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^k G_{ij} \log \pi_{ij} + \log p(k)$$

Equation (5) is used as the objective function for parameter solving. However, since Equation (5) contains the logarithm of sum, it is difficult to obtain its maximal parametric solution, so the objective function needs to be further simplified.

For a given random variable  $X$ , for any concave random function  $f(X)$ , Jensen inequality can be expressed as  $E[f(Y)] \leq f[E[Y]]$ , where  $E[\cdot]$  is the expected function, and the equation holds if and only if  $Y$  is constant.

For given  $x_i$ , suppose  $Y_{ij} = \frac{\pi_{ij} \Phi(x_i | \theta_j)}{z_{ij}}, j = 1, 2, \dots, k$ .  $\{Y_{i1}, \dots, Y_{ik}\}$  is the  $k$  sampling of random variable  $Y_i$ .  $z_{ij}$  can be regarded as the probability of obtaining  $Y_{ij}$ . Thus, the

$$\log \left[ \sum_{j=1}^k \pi_{ij} \Phi(x_i | \theta_j) \right] \text{ item in (4), can be written as:} \quad (6)$$

$$\log \left[ \sum_{j=1}^k \pi_{ij} \Phi(x_i | \theta_j) \right] = \log \left[ \sum_{j=1}^k z_{ij} \frac{\pi_{ij} \Phi(x_i | \theta_j)}{z_{ij}} \right] = \log \left[ \sum_{j=1}^k z_{ij} Y_j \right] = \log [E[Y_i]]$$

Since  $\log(\cdot)$  is a concave function, according to Jensen inequality rule, we have:

$$\log[E[Y_i]] \geq E[\log[Y_i]] = \sum_{j=1}^k z_{ij} \log\left(\frac{\pi_{ij} \Phi(x_i | \theta_j)}{z_{ij}}\right) \quad (7)$$

To sum up, let  $A = 1$  and  $T = 1$ , and it ignores the terms irrelevant to the parameter  $\varphi$  in Equation (4), then:

$$L(\pi_k, k | X, \theta) \geq \sum_{i=1}^N \sum_{j=1}^k z_{ij} \log\left(\frac{\pi_{ij} \Phi(x_i | \theta_j)}{z_{ij}}\right) + \sum_{i=1}^N \sum_{j=1}^k G_{ij} \log \pi_{ij} \quad (8)$$

If formula (8) is an equality, then  $\frac{\pi_{ij} \Phi(x_i | \theta_j)}{z_{ij}} = c$ , it can derive the  $z_{ij} = \frac{\pi_{ij} \Phi(x_i | \theta_j)}{c}$ . Since the posterior probability meets the condition  $\sum_{j=1}^k z_{ij} = 1$ , then  $c = \sum_{j=1}^k \pi_{ij} \Phi(x_i | \theta_j)$ , and the posterior probability formula is:

$$z_{ij} = \frac{\pi_{ij} \Phi(x_i | \theta_j)}{\sum_{j=1}^k \pi_{ij} \Phi(x_i | \theta_j)} \quad (9)$$

The simplified objective function is:

$$J(\pi_k, k | X, \theta) = \sum_{i=1}^N \sum_{j=1}^k z_{ij}^t \log(\pi_{ij} \Phi(x_i | \theta_j)) + \sum_{i=1}^N \sum_{j=1}^k G_{ij}^t \log \pi_{ij}^{t+1} \quad (10)$$

The partial derivatives of the objective function of Equation (10) with respect to  $\mu_j$  and  $\sigma_j^2$  are calculated respectively. Setting formula (10)=0, it gets:

$$\mu_j^{t+1} = \frac{\sum_{i=1}^N z_{ij}^t x_i}{\sum_{i=1}^N z_{ij}^t} \quad (11)$$

$$\left(\sigma_j^{t+1}\right)^2 = \frac{\sum_{i=1}^N z_{ij}^t (x_i - \mu_j^{t+1})^2}{\sum_{i=1}^N z_{ij}^t} \quad (12)$$

Because the prior distribution needs to satisfy the constraint condition  $\pi_{i1} + \dots + \pi_{ik} = 1$ . Therefore, the Lagrange multiplier method is introduced to find the partial derivative of the prior probability parameters. Set it equal to 0, and it gets:

$$\pi_{ij}^{t+1} = \frac{z_{ij}^t + G_{ij}^t}{\sum_{j=1}^k (z_{ij}^t + G_{ij}^t)} \quad (13)$$

### 3.3 Misidentified Area Removing With SURF

#### 3.3.1 Feature Points Detection

SURF algorithm uses the integral image theory to replace the filtering of images and Gaussian second-order differential templates by simple addition and subtraction of integral images to improve the computing speed. The definition of integral image is described in reference (Wang et al. 2019). Therefore, the gray value integral of S region in Figure 2 can be obtained by  $S = A - B - C + D$ .

The Hessian matrix with scale  $\sigma$  for the point  $(x, y)$  of image  $J$  is:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (14)$$

where  $L_{xx}(x, \sigma)$  is the second partial derivative of the Gaussian function at point  $x$  and the convolution of the image  $J(x)$ .  $L_{xy}(x, \sigma)$  and  $L_{yy}(x, \sigma)$  are similar to it. The convolution operation is transformed into a box filtering operation, which is a simplification of the Gaussian second order differential template. A box filter with  $9 \times 9$  is an approximation of a Gaussian filter with  $\sigma = 1.2$ . It is used as the latest scale space value in image filtering and spot detection. The values after convolution operation are  $D_{xx}$ ,  $D_{xy}$  and  $D_{yy}$ , respectively. They will replace  $L_{xx}$ ,  $L_{xy}$  and  $L_{yy}$ , as shown in Figure 3.

Therefore, the determinant of the Hessian matrix can be abbreviated as:

$$Det(H_{approx}) = D_{xx}D_{yy} - (\omega D_{xy})^2 \quad (15)$$

where  $Det(H_{approx})$  represents the box filtering response value in the area near point  $x$ , which can be used to detect extreme points.  $\omega$  is a coordination parameter whose size is related to  $\sigma$ , and  $\omega$  is usually valued at 0.9.

In order to obtain the spots of different scales, the Gaussian pyramid of the image needs to be constructed first. The basic idea of building the image Gaussian pyramid is to keep the original image size unchanged, and the scale space can be obtained by changing the filter size. The SURF algorithm continuously increases the size of the box filter template, uses the box filter template with different sizes and integral images to obtain the response value of the Hessian matrix, and calculates the feature points of different scales on the response images.

$3 \times 3$  neighborhood non-maximum suppression is carried out on the response images of feature points at different scales, so that the pixels on each layer of the image are compared with 8 pixels on the same scale and 9 pixels on the upper and lower layers of the image adjacent to it, as shown in figure 4.

Figure 2.  
Integral image method

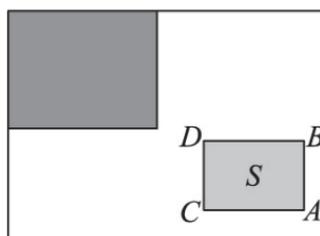


Figure 3.  
 Box filter

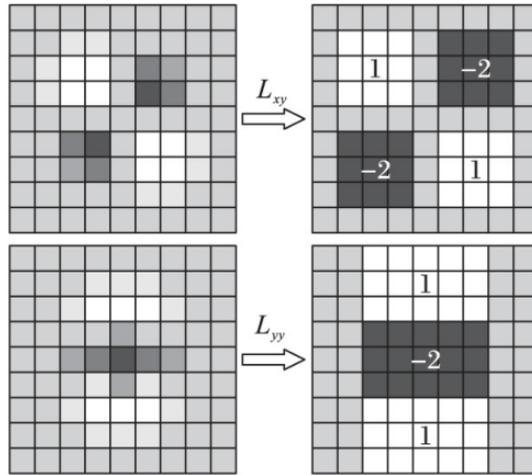
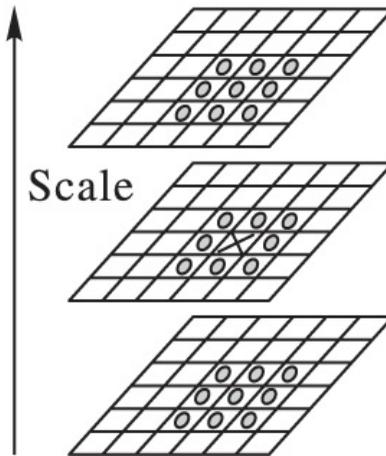


Figure 4.  
 Non-maximum suppression

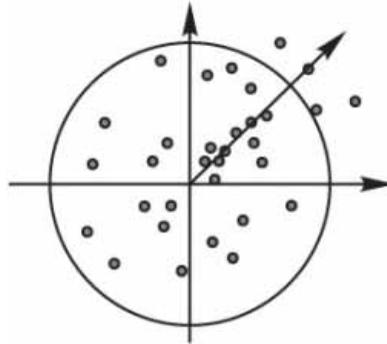


If the extreme value is greater than or less than the extreme value of the 26 points, this point is used as a candidate feature point. Then, the quadratic function fitting method is used to locate the interest points, and the positioning accuracy reaches sub-pixel and sub-scale level to obtain the scale and position information of stable interest points and complete the positioning.

### 3.3.2 Feature Point Descriptor

In order to make the target image rotationally invariant, the main direction of the feature points must be determined. Firstly, taking the above feature points as the center, the Harr wavelet responses of the points in the neighborhood with the radius of  $6s$  ( $s$  is the scale value where the feature points are located) in the  $x$  and  $y$  directions are calculated. The Harr wavelet side length is set to  $4s$ , and the response value is weighted by Gaussian. Then, the response values in the  $60^\circ$  sector area are accumulated to form a new vector, and the whole circular area is traversed. The direction of the vector with the longest length is selected as the main direction of the feature point, as shown in Figure 5.

Figure 5.  
Determination of main direction



The above operations have determined the position of the feature points and the main direction, and then the descriptor is computed for the local region. Firstly, it rotates the coordinate axis to the direction of the feature points to ensure rotation invariance. With feature points as the center, it is carried out in a rectangular region parallel to the main direction. The size of this area is  $20s \times 20s$  and it is divided into  $4 \times 4$  sub-blocks. Then the response value of each sub-block is calculated by using the Harr wavelet template with the size of  $2s$ , and the vector of each sub-block is obtained. Since there are  $4 \times 4$  sub-blocks in this region, the generated feature descriptors are composed of  $4 \times 4 \times 4 = 64$ -dimensional feature vectors. Then, the ratio detection method (Chekanov et al. 2022) is used to remove the false matching points of those multi-point matching.

#### 4. EXPERIMENTS AND ANALYSIS

This section designs a comparison experiment on the DOTA data set (Xia et al. 2018), and selects four remote sensing images to verify the effectiveness of the proposed segmentation algorithm in this paper. We also make comparison with KAZE (Pourfard et al. 2021), CPLF (Su et al. 2022). These experiments are conducted on MATLAB 2017a.

Figure 6 displays the segmentation result with different methods. We can see that the segmentation effect is better with proposed method. KAZE and CPLF have the problem of incoherent segmentation of the object edge. Over-segmentation is also a serious problem. If the background is too complex, segmentation is very inefficient. The method in this paper can effectively improve the above segmentation problems.

Segmentation accuracy and segmentation time are used as evaluation indexes. The detailed results are shown in table 1.

As can be seen from Table 1, for image1, the segmentation accuracy of the proposed method reaches 97.5%, while the accuracy of KAZE and CPLF reaches 89.6% and 92.1% respectively, which are 7.9% and 5.4% lower than that of the proposed method. For image2, the segmentation accuracy of the proposed method reaches 96.4%, which is 6.0% higher than that of AKAZE and 4.7% higher than that of CPLF. Similarly, the results are similar for the other two images. The segmentation time of the proposed method is also the shortest. Therefore, comprehensive analysis shows that the efficiency of the proposed method is the best.

And then we conduct a set of ablation experiments with SURF and without SURF for showing the advantage of removing mismatched points. The purpose is to prove the effectiveness of the matching algorithm for the segmentation algorithm. It is obvious that the segmentation effectiveness is greatly improved after adding SURF. Figure 7 clearly shows that after adding SURF, the segmentation effect is significantly improved. The minimum efficiency is above 96%.

Figure 6.  
 Segmentation with different methods

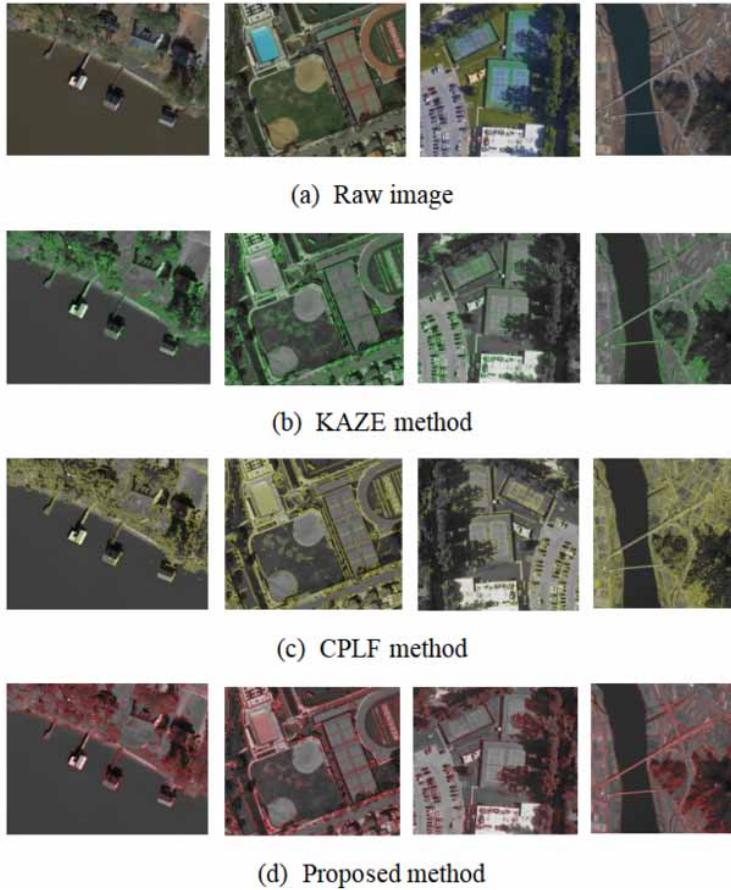


Table 1.  
 Segmentation results with different methods

Method	Accuracy/%				Time/s			
	Image 1	Image 2	Image 3	Image 4	Image 1	Image 2	Image 3	Image 4
KAZE	89.6	90.4	88.1	87.9	6.9	7.6	8.5	6.8
CPLF	92.1	91.7	92.6	92.4	5.7	5.5	4.2	5.7
Proposed	97.5	96.4	98.1	97.2	4.1	5.3	3.8	4.6

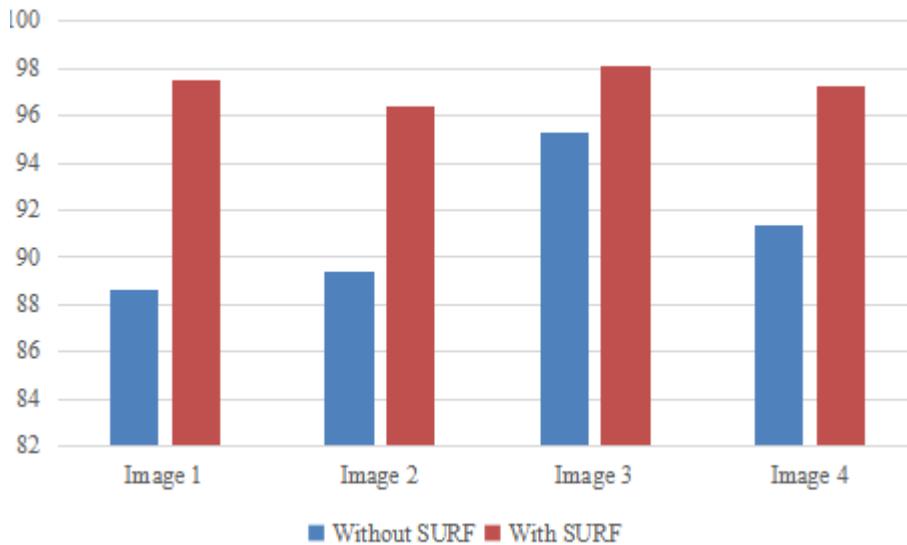
## 5. CONCLUSION

Gaussian mixture model image segmentation is one of the main research tasks in remote sensing image segmentation. Therefore, this paper proposes a novel remote sensing image segmentation method based on Gaussian mixture model and SURF algorithm. By using a new energy function to change the iteration process, the new method reduces the large amount of time spent by the traditional GMM algorithm in the segmentation iteration, and improves the situation that the segmentation results are still

**Table 2.**  
**Ablation experiment %**

Module	Image 1	Image 2	Image 3	Image 4
Without SURF	88.6	89.4	95.3	91.4
With SURF	97.5	96.4	98.1	97.2
Average (with SURF)	97.3			

**Figure 7.**  
**The visualization of table 2**



under segmentation or over segmentation when the iteration converges. In addition, the SURF matching algorithm is used to eliminate mismatched points, reducing the complexity of feature description and improving the algorithm computing speed. Finally, the experiments are carried out on the open remote sensing DOTA data. From the analysis of subjective and objective evaluation indexes, it is concluded that the new method in this paper has significantly improved the segmentation accuracy and segmentation time. In the next step, each object category is modeled as a GMM component to accurately depict the statistical distribution characteristics of pixel spectral measurement.

## CONFLICTS OF INTEREST

I declare that the guest editor of the special issue being an author for this article.

## ACKNOWLEDGEMENTS

This work was supported in part by the National Natural Science Foundation of China under Grant 62071084 and National Natural Science Foundation of China (62001434). Also supported by Talents Project of the State Ethnic Affairs Commission. Supported by Scientific Research Funds of Education Department of Liaoning Province in 2021 (General Project) (LJKZ1311).

## REFERENCES

- Bao, Y., Song, K., Liu, J., Wang, Y., Yan, Y., Yu, H., & Li, X. (2021). Triplet-graph reasoning network for few-shot metal generic surface defect segmentation [J]. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1–11. doi:10.1109/TIM.2021.3083561
- Baumgartner, J., Gimenez, J., Scavuzzo, M., & Pucheta, J. (2015). A new approach to segmentation of multispectral remote sensing images based on MRF [J]. *IEEE Geoscience and Remote Sensing Letters*, *12*(8), 1720–1724. doi:10.1109/LGRS.2015.2421736
- Chekanov, M., Shipitko, O., & Skoryukina, N. (2022). Study of Keypoints Detectors and Descriptors Performance on X-Ray Images Compared to the Visible Light Spectrum Images [J]. *IEEE Access : Practical Innovations, Open Solutions*, *10*, 38964–38972. doi:10.1109/ACCESS.2022.3159650
- Chen, D., Zhu, J., Zhang, X., Shu, M., & Cohen, L. D. (2021). Geodesic paths for image segmentation with implicit region-based homogeneity enhancement. *IEEE Transactions on Image Processing*, *30*, 5138–5153. doi:10.1109/TIP.2021.3078106 PMID:34014824
- Chen, M., & Ludtke, S. J. (2021). Deep learning-based mixed-dimensional Gaussian mixture model for characterizing variability in cryo-EM. *Nature Methods*, *18*(8), 930–936. doi:10.1038/s41592-021-01220-5 PMID:34326541
- Cheng, B., Girshick, R., & Dollár, P. (2021) Boundary IoU: Improving object-centric image segmentation evaluation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 15334-15342). IEEE.
- Di, S., Liao, M., Zhao, Y., Li, Y., & Zeng, Y. (2021). Image superpixel segmentation based on hierarchical multi-level LI-SLIC *Optics & Laser Technology*, *135*, 106703. doi:10.1016/j.optlastec.2020.106703
- Gupta, A., Watson, S., & Yin, H. (2021). Deep learning-based aerial image segmentation with open data for disaster impact assessment *Neurocomputing*, *439*, 22–33. doi:10.1016/j.neucom.2020.02.139
- Hatamizadeh, A., Tang, Y., & Nath, V. (2022) Unetr: Transformers for 3d medical image segmentation. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, (pp. 574-584). IEEE.
- Hong, D., Han, Z., & Yao, J. (2021). SpectralFormer: Rethinking hyperspectral image classification with transformers. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–15. doi:10.1109/TGRS.2022.3172371
- Ji, Z., Xia, Y., Sun, Q., Chen, Q., & Feng, D. (2014). Adaptive scale fuzzy local Gaussian mixture model for brain MR image segmentation *Neurocomputing*, *134*, 60–69. doi:10.1016/j.neucom.2012.12.067
- Jin, X., Che, J., & Chen, Y. (2021). Weed identification using deep learning and image processing in vegetable plantation. *IEEE Access : Practical Innovations, Open Solutions*, *9*, 10940–10950. doi:10.1109/ACCESS.2021.3050296
- Juan-Albarracín, J., Fuster-García, E., Juan, A., & García-Gómez, J. M. (2021). Non-local spatially varying finite mixture models for image segmentation. *Statistics and Computing*, *31*(1), 1–10. doi:10.1007/s11222-020-09988-w
- Koresh, H. J. D., Chacko, S., & Periyayagi, M. (2021). A modified capsule network algorithm for oct corneal image segmentation. *Pattern Recognition Letters*, *143*, 104–112. doi:10.1016/j.patrec.2021.01.005
- Li, A., Jiao, L., & Zhu, H. (2021). Multitask semantic boundary awareness network for remote sensing image segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–14.
- Liu, Y., Zhu, Q., Cao, F., Chen, J., & Lu, G. (2021). High-Resolution Remote Sensing Image Segmentation Framework Based on Attention Mechanism and Adaptive Weighting. *ISPRS International Journal of Geo-Information*, *10*(4), 241. doi:10.3390/ijgi10040241
- Lu, H. (2021). Graph Theory-based Fast Linear Iterative Clustering Multi-feature Fusion Model for Art Image Segmentation. *Journal of Applied Science and Engineering*, *25*(4), 607–614.
- Ma, F., Zhang, F., & Yin, Q. (2021). Fast SAR image segmentation with deep task-specific superpixel sampling and soft graph convolution. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–16.
- Minaee, S., Boykov, Y. Y., Porikli, F., Plaza, A. J., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(7), 3523–3542. doi:10.1109/TPAMI.2021.3059968 PMID:33596172

- Mirsadeghi, S. E., Royat, A., & Rezatofghi, H. (2021). Unsupervised image segmentation by mutual information maximization and adversarial regularization. *IEEE Robotics and Automation Letters*, 6(4), 6931–6938. doi:10.1109/LRA.2021.3095311
- Pourfard, M., Hosseini, T., Saeidi, R., Motamedi, S. A., Abdollahifard, M. J., Mansoori, R., & Safabakhsh, R. (2021). KAZE-SAR: SAR image registration using KAZE detector and modified SURF descriptor for tackling speckle noise [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–12. doi:10.1109/TGRS.2021.3084411
- Pravitasari, A. A., Iriawan, N., & Fithriasari, K. (2021). Unsupervised Learning for MRI Brain Tumor Segmentation with Spatially Variant Finite Mixture Model in Reversible Jump MCMC Algorithm[C]//Journal of Physics: Conference Series. *IOP Publishing*, 1776(1), 012041.
- Rafique, A. A., Gochoo, M., & Jalal, A. (2022). Maximum entropy scaled super pixels segmentation for multi-object detection and scene recognition via deep belief network. *Multimedia Tools and Applications*, 1–30.
- Salau, A. O., Yesufu, T. K., & Ogundare, B. S. (2021). Vehicle plate number localization using a modified GrabCut algorithm. *Journal of King Saud University-Computer and Information Sciences*, 33(4), 399–407. doi:10.1016/j.jksuci.2019.01.011
- Su, J., Meng, J., Hou, W., Wang, R., & Luo, X. (2022). Multi-Angle Optical Image Automatic Registration by Combining Point and Line Features. *Sensors (Basel)*, 22(3), 739. doi:10.3390/s22030739 PMID:35161485
- Teng, L., & Qiao, Y. (2022). BiSeNet-oriented context attention model for image semantic segmentation. *Computer Science and Information Systems*, 19(3), 1409–1426. doi:10.2298/CSIS220321040T
- Tesfamikael, H. H., Fray, A., Mengsteab, I., Semere, A., & Amanuel, Z. (2021). Simulation of Eye Tracking Control based Electric Wheelchair Construction by Image Segmentation Algorithm. [JIIP]. *Journal of Innovative Image Processing*, 3(01), 21–35. doi:10.36548/jiip.2021.1.003
- Wang, L., & Yin, S. (2022). A novel deep learning-based single shot multibox detector model for object detection in optical remote sensing images. *Geoscience Data Journal*, 2022, gdj3.162. doi:10.1002/gdj3.162
- Wang, R., Shi, Y., & Cao, W. (2019). GA-SURF: A new speeded-up robust feature extraction algorithm for multispectral images based on geometric algebra. *Pattern Recognition Letters*, 127, 11–17. doi:10.1016/j.patrec.2018.11.001
- Wang, W., Zhou, T., & Yu, F. (2021). Exploring cross-image pixel contrast for semantic segmentation. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, (pp. 7303-7313).
- Xia, G. S., Bai, X., & Ding, J. (2018). DOTA: A large-scale dataset for object detection in aerial images. *Proceedings of the IEEE conference on computer vision and pattern recognition*. (pp. 3974-3983). IEEE.
- Xu, Z., Zhang, W., Zhang, T., Yang, Z., & Li, J. (2021). Efficient transformer for remote sensing image segmentation. *Remote Sensing*, 13(18), 3585. doi:10.3390/rs13183585
- Yin, S., & Li, H. (2020). Hot Region Selection Based on Selective Search and Modified Fuzzy C-Means in Remote Sensing Images *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5862–5871. doi:10.1109/JSTARS.2020.3025582
- Yong, Z., Jiazheng, Y., Hongzhe, L., & Qing, L. (2017). GrabCut image segmentation algorithm based on structure tensor. *Journal of China Universities of Posts and Telecommunications*, 24(2), 38–47. doi:10.1016/S1005-8885(17)60197-3
- Yuan, X., Shi, J., & Gu, L. (2021). A review of deep learning methods for semantic segmentation of remote sensing imagery. *Expert Systems with Applications*, 169, 114417. doi:10.1016/j.eswa.2020.114417
- Zhang, H., Liu, J., Yu, Z., & Wang, P. (2021). MASG-GAN: A multi-view attention superpixel-guided generative adversarial network for efficient and simultaneous histopathology image segmentation and classification. *Neurocomputing*, 463, 275–291. doi:10.1016/j.neucom.2021.08.039
- Zhu W, McBrearty I W, Mousavi S M, et al. (2022) Earthquake phase association using a Bayesian Gaussian mixture model. *Journal of Geophysical Research: Solid Earth*, 127(5), e2021JB023249.

*Shoulin Yin received the B.S. and M.S. degrees in image processing from Software College, Shenyang Normal University, Shenyang, China, in 2013 and 2015, respectively. He is currently a PHD with School of Information and Communication Engineering, Harbin Engineering University. He has published more than 10 technical articles in scientific journals and conference proceedings. His research interests include image fusion, object detection, and image recognition. Email: yslin@hit.edu.com.*

*Liguo Wang received his M.A. degree in 2002 and Ph.D. degree in signal and information processing in 2005 from Harbin Institute of Technology, Harbin, China. He held postdoctoral research position from 2006 to 2008 in the College of Information and Communications Engineering, Harbin Engineering University. He is currently a Professor with Information and Communications Engineering, Dalian Minzu University, Dalian, China. His research interests are remote sensing image processing and machine learning. He has published three books, 27 patents, and more than 200 papers in journals and conference proceedings.*

*Qunming Wang biography. Qunming Wang received the Ph.D. degree from The Hong Kong Polytechnic University, Hong Kong, in 2015. He is currently a Professor with the College of Surveying and Geo-Informatics, Tongji University, Shanghai, China. From 2017 to 2018, he was a Lecturer (Assistant Professor) with Lancaster Environment Centre, Lancaster University, Lancaster, U.K., where he is currently a Visiting Professor. His 3-year Ph.D. study was supported by the hypercompetitive Hong Kong Ph.D. Fellowship and his Ph.D. thesis was awarded as the Outstanding Thesis in the Faculty. He has authored or coauthored more than 70 peer-reviewed articles in international journals such as Remote Sensing of Environment, IEEE Transactions on Geoscience and Remote Sensing, and ISPRS Journal of Photogrammetry and Remote Sensing. His research interests include remote sensing, image processing, and geostatistics. Prof. Wang is an Editorial Board Member for Remote Sensing of Environment, and serves as an Associate Editor for Science of Remote Sensing (sister journal of Remote Sensing of Environment) and Photogrammetric Engineering & Remote Sensing. He was an Associate Editor for Computers and Geosciences (2017-2020).*

*Jinghui Yang received the B.S., M.S. and Ph.D. degrees from Harbin Engineering University, Harbin, China, in 2010, 2013 and 2016, respectively. She is currently an Associate Professor with the School of Information Engineering, China University of Geosciences, Beijing, Beijing, China. Her research interests include hyperspectral imagery, remote sensing, pattern recognition and signal processing.*

*Man Jiang is with the Liaoning Vocational Technical College of Modern Service. His major is computer science.*