A Remote Sensing Ship Recognition Using Random Forest

Sizhe Huang [1]

College of Information Technology, Harbin Engineering University ,Harbin, 150000, China
Wuhan Digital Engineering Research Institute,Wuhan, 430000, China
E-mail: `nara113@163.com`

Huosheng Xu

Wuhan Digital Engineering Research Institute,Wuhan, 430000, China

Xuezhi Xia

Wuhan Digital Engineering Research Institute,Wuhan, 430000, China

In order to detect the marine targets reliably and timely, a novel ship recognition method by using optical remote sensing data based on random forest is presented. First, in the feature extraction part, in addition to the common features, we introduce the visual saliency features of the target.; second, an improved random forest based on mutual information (MIRF) is utilized to recognize ships in data from the optical remote sensing system; finally, we compare MIRF to classical algorithms. The MIRF has accelerated the operation speed of the algorithm and the classification accuracy remains robust. Theoretical analysis and experiment results show that the proposed method can achieve high recognition rate; therefore, this approach is feasible and efficient in the marine target recognition.

---

[1]Speaker

http://pos.sissa.it/

1. Introduction

The target detection and recognition from remote sensing image play a critical role in various applications of pattern recognition, such as precise guidance, vessel traffic services, search and rescue, anti-terrorism, especial military and civilian fields .

Previous researches on ship target recognition were mainly based on synthetic aperture radar and FLIR's close side view image data, which used the difference of ship superstructure for identification. In recent years, with the development of satellite remote sensing technology, optical remote sensing of ship target recognition has gradually become possible. Zhang Zhen used the support vector machine (SVM) to recognize the Quick Bird satellite data of ship target[1]. Whereas SVM method performs well in dual-class problems and poorly in the multi-class classification tasks. The improved Zernike moment was employed to identify large military vessels in remote sensing image by Lan et al. [2]. Chun used the idea of attributes reduction to compute the significance weight of each candidate feature and select valid recognition features[3]. Hierarchical discriminant regression (HDR) technology is used in target recognition [3-4]. It has the advantage in the process of incremental learning, but needs more training samples to achieve higher accuracy. Antelo et al. designed the Bayesian likelihood function to identify the Quick Bird satellite ship image[5]. Although these methods have improved the recognition accuracy, most of them have the lower time complexity. In view of the above-mentioned facts, these methods mainly have two shortcomings: 1) the selection of feature was seldom considered, which was generally based on the experience to determine the feature vector. The selected vectors are of high dimensional and redundancy; 2) the feature matching is needed to perform the global search, so the algorithm complexity is high.

In order to improve the effectiveness and practicability of optical remote sensing ship target recognition method, this paper takes more than 4 m resolution of optical remote sensing image data as the research object and proposes an universal ship target recognition method. The key point of the method such as feature extraction, feature selection and classifier design are improved. In addition to the commonly used feature extraction methods, the paper adds the visual saliency features and constructs the training set. The training set is classified by using the integrated algorithm. Compared to the traditional algorithm, the MIRF algorithm is fast and has high classification accuracy, which is practical in the ship target recognition.

2. Ship Target Feature Extraction

The distinguishing feature is important for the recognition of ship target. In the optical remote sensing image, there are differences in the size, shape, texture and superstructure of different kinds of ships. In this paper, the ship target feature is extracted from the following six aspects.

First, the rotating target slices to the horizontal direction along the major axis direction. Then, we get the target minimum bounding rectangle slices and divide it into equal K parts. Taking into account the bow and the stern of the ship accounts for the majority $\frac{1}{8} - \frac{1}{3}$ of the whole ship, thus we set the target slice block number range is $3 \le K \le 8$.

In summary, while ensuring the efficiency of the algorithm, we cut the target into five equal parts, .

A$_1$	A$_2$	A$_3$	A$_4$	A$_5$

Figure 1 :Tree Generated by Training

The ship target bow and stern are represented by the features images of A1 and A5. We use the connection point number in A1 and A5 block image as a ship target recognition criteria.

According to the above definitions, we set the image resolution to R m, length (M pixels) and width (N pixels) of the minimum bounding rectangle of the target slice.

a)Length, Width and Length-Width Ratio:

$$f_1 = M \times R \qquad R \in (0,4) \tag{2.1}$$

$$f_2 = N \times R \qquad R \in (0,4) \tag{2.2}$$

$$f_3 = f_1 : f_2 = (M \times R):(M \times R) = M \times N \qquad R \in (0,4) \tag{2.3}$$

b)Target Area:

$$f_4 = (\sum_{j=1}^{N} \sum_{i=1}^{M} p(i,j)) * R^2 \tag{2.4}$$

In the target slice, $p(i,j)$ is the value "1" point (In the binary image of the target slice, the value of the target area is 1 and the background area is 0).

c)Contour Center Moment

The moment features are invariant to the change of image geometric position and geometric dimension. As to an image with a closed boundary, it can be described by the distance vector sequence of N boundary pixels to the inner feature centroid. $\overline{m_r}$ and a $\overline{\mu_r}$ re moment of one dimensional normalized contour sequence, then the moment m_r and the center moment m_r of the r_{th} contour sequence moment can be expressed as

$$m_r = \frac{1}{N} \sum_{i=1}^{N} [d(i)]^r \tag{2.5}$$

$$\mu_r = \sum_{i=1}^{N} [d(i) - m_l]^r \tag{2.6}$$

The r_{th} normalized sequence moment and the r_{th} normalized sequences central moment are defined as:

$$\overline{m_r} = \frac{m_r}{(\mu_2)^{r/2}} = \frac{\frac{1}{N} \sum_{i=1}^{N} [d(i)]^r}{[\frac{1}{N} \sum_{i=1}^{N} [d(i) - m_1]^2]^{r/2}} \tag{2.7}$$

$$\overline{\mu_r} = \frac{\mu_r}{(\mu_2)^{r/2}} = \frac{\sum_{i=1}^{N} [d(i) - m_l]^r}{[\frac{1}{N} \sum_{i=1}^{N} [d(i) - m_l]]^{r/2}} \tag{2.8}$$

d)The Bow Shape and The Stern Shape:

The features of ship bow shape and ship stern shape mainly describe the shape similarity between the ship bow shape and the ship stern of samples. In this paper, we use the contour discretization parameters for calculation. The contour discretization parameter is a description of the shape of the object contour in the image, which contains the following four factors:

$$r_k = \frac{min|z_k(i)|}{max|z_k(i)|} \tag{2.9}$$

$$s_k = \frac{\sum_{i=1}^{n} [z_k(i) - \overline{z}_k]^3}{\sum_{i=1}^{n} [z_k(i)]^3} \tag{2.10}$$

$$v_k = \frac{\sum_{i=1}^{n} [z_k(i) - \overline{z}_k]^2}{\sum_{i=1}^{n} [z_k(i)]^2} \tag{2.11}$$

$$f_k = \frac{max[z_k(i) - min z_k(i)]}{\overline{z}_k} \tag{2.12}$$

Where $z_k(i)$ represents the distance from the contour discrete points to the centroid of the target.

e)The Connection Points of The Bow and Stern

Bow and stern connection points feature is the connection number of the A1 sub block and A5 sub block. In order to obtain the target connection points feature, we need A1 sub block and A5 sub block for detecting the connection point. As the slice of the bow and stern is the binary image, we use Harris detector operator to complete the connection point detection, and the calculation steps are as follows:

Step one: calculate the correlation matrix M of each pixel;

$$A = w(x, y) \otimes I_x^2 \tag{2.13}$$

$$B = w(x, y) \otimes I_y^2 \tag{2.14}$$

$$C = D = w(x, y) \otimes I_x I_y \tag{2.15}$$

$$M = \begin{bmatrix} A & D \\ C & B \end{bmatrix} \tag{2.16}$$

Step two: calculate the Harris response of each pixel ;

$$R = (AB - CD)^2 - k(A + B)^2 \tag{2.17}$$

Step three: if Harris response value is greater than the threshold, it is considered as the connection point.

f)Spatial Correlation Characteristic Elements

In a given radius of 8 × 8 image window, we get the spatial correlation of the adjacent window and the given window by calculation [6]. The spatial correlation characteristic is the similarity between a given window and its adjacent window. In the spatial correlation significant figure, the lower the spatial correlation is, the easier it is to distinguish the two different goals and describe the significant degree. The correlation between X and Y (two random variables) is:

$$\rho_{X,Y} = cov(X, Y) / \sigma_X \sigma_Y = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X) - E^2(X)}\sqrt{E(Y) - E^2(Y)}} \tag{2.18}$$

g)Image Entropy

In 8 x 8 image window, we use the image entropy to measure the image information content, which is the contribution of multiple pixels in the window and not sensitive to a single point. The image entropy reflects the discrete degree of the image gray level. Privitera and Stark put forward the measuring method of image entropy, indicating the correlation between the image entropy and positioning to the target by human eye . For image entropy $Z_k(i)$, there is:

$$M_E(i,j) = -\sum_{i=1}^{8}\sum_{j=1}^{8} p_{i,j} \log p_{i,j} \tag{2.19}$$

$$p_{(i,j)} = f(i,j)/\sum_{i=1}^{8}\sum_{j=1}^{8} f(i,j) \tag{2.20}$$

Where $f(i,j)$ is the gray of point (i,j) in the image window, $f(i,j) > 0$; $p_{i,j}$ is the probability distribution of the point (i,j) gray at the image window.

3. Random forest

A. Principle

Random forest is an integrated classification algorithm proposed by Breiman in 2001. This method, as a kind of statistical learning theory, uses bootstrap to extract multiple samples from original sample. For each bootstrap samples, it models the decision tree, and then combines the prediction of the decision trees by voting to get the final prediction results . Plenty of theoretical and empirical studies have proved that the RF has high prediction accuracy and good tolerance to outliers and noise while being not prone to over-fitting.

By constructing different training sets to increase the difference between the classification models, the RF can improve the prediction ability by combining classification model. $\{h_1(x), h_2(x),,, h_k(x)\}$ is a classification sequence model, which was obtained by K times training. Then it constitutes a multi-classification model system, which uses a simple majority voting method to get the final classification results of the system. The final classification decision is:

In the formula, $H(x)$ indicates a combination of classification model, h_i represents a single decision tree classification model, Y represents the output variable (or target variable), and $I(.)$ is the indicator function. The formula illustrates the use of majority voting to determine the final classification.

B. Convergence

Given a set of classification models $\{h_1(x), h_2(x),,, h_k(x)\}$, each training set of every classification model is obtained from a random sample of the original data set (X,Y), then to obtain the margin function:

$$mg(X,Y) = av_k I(h_k(X)=Y) - \max_{i \neq k} av_k I(h_k(X)=j) \tag{3.1}$$

The margin function is used to measure the average value of correct classification over the average value of errors. The greater the margin value is, the more reliable the forecast classification will be. The generalization error can be written as:

$$PE^* = P_{X,Y}(mg(X,Y)<0) \tag{3.2}$$

When the decision tree classification model is enough, $h_k(X) = h(X, \Theta_k)$ obey the law of strong numbers. With increase of the decision tree classification model, it can be proved that all sequences $\Theta_1 ... PE^*$ converge in

$$P_{X,Y}(P_{\Theta}(h(X,\Theta)=Y) - \max_{i \neq Y} P_{\Theta}(h(X,\Theta)=j) < 0) \qquad (3.3)$$

This shows why the RF method does not have the problem of over fitting with the increase of decision tree, but it is possible to produce the generalization error.

C.OOB Generalization Error, The strength and The Correlation Coefficient Estimates

Each training set of random forests is randomly drawn from the original training data set, which is also called Bootstrap sample copy. According to the random sampling, in construction of each tree structure, there is about 36.8% of the training sample which has not been extracted. Some samples of the original training set may appear for several times in the new training set while others sample may not appear for once. Those samples that have not been extracted can be used as an internal estimate calculated OOB (Out-of-Bag) value and also can be used to estimate the strength and relevance of the forest.

4. Mutual Information Random Forest Algorithm

MIRF is an improved algorithm of RF. In MIRF algorithm, each tree classifiers have to vote for the input feature vector. When the input feature variables are classified, the forest contains a number of DT, the feature variables are input to each DT and the results of each DT are classified by Voting. MIRF algorithm selects the most votes as the classification result. MIRF algorithm consists of three steps: 1) the growth of tree; 2) the voting procedure; 3) the estimated error of classification.
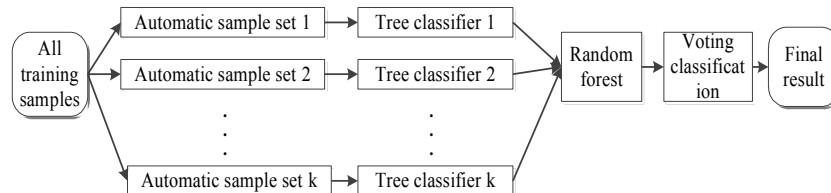

Figure 2: MIRF Algorithm Diagram

A. The Growth of The Tree

Random forests are composed of a number of tree classifiers, which are generated by Bagging sampling technique. Assuming the original training set as N, the new independent sample set can be randomly selected by Bagging method, and it can constitute K trees. The sample can be composed of K Out-of-bag data (OOB). The main operation steps are as follows:

1) From the original training sample set with N capacity, we get the Bagging sample set by adopting the back method, repeat K times.

2) Each sample set as the training set to grow into a single tree DT. We randomly select t features from n input features $(t \leq n)$. At each node of the DT, according to the mutual information of n features, we select the best split point for the growth of the tree [7].

a) There is m class of training set S, and the number of per class is $|U_i|(i=1,2,...,m)$ The attribute A is the root of DT, with V value $v_1, v_2, ..., v_v$. Assuming that the probability of

the subset of the U be represented by p_i , $p_i = \dfrac{|U_i|}{|S|}$. U_i represents the number of samples class which belongs to the class $|U_i|$, $|S|$ is the number of sample set S, Then, the information entropy of the subset is defined in the formula:

$$I(e_r) = -\sum_{i=1}^{m} p_i \log_2(p_i), \qquad r = 1, 2, \dots, v \tag{4.1}$$

The conditional entropy of the classification with A as the root. Conditional entropy is calculated as follows:

$$H(A) = \sum_{j=1} P(U_i|V_j) \log_2 \frac{1}{P(U_i|V_j)} \tag{4.2}$$

b) The calculation of mutual information with A as the root [8].

$$I(A) = I(e_r) - H(A) \tag{4.3}$$

By selecting the attribute with the maximum mutual information as the root node and the process iterate, we make every tree DT with sufficient growth and no pruning operation.

B. Voting Process

Since the MIRF algorithm combines a series classification results of DT, which should select the best classification results, we need to vote for predicting the result so that every DT votes on the input feature vector so as to find the largest number of voting category as the category labels of the feature vector.

C. Classification Error Estimates

In the above method, we use the sampling with replacement. When the training set is randomly generated, only about two-thirds of the original sample is drawn as training set of DT and the remaining 1/3 of the data has not been drawn as the data out of the bag (OOB). When a classification tree is generating, we use the tree to classify all the corresponding OOB data. We compare the results of predicting classification with the actual category to get the classification error estimation of MIRF.

By using the mutual information in the feature selection can reduce the redundancy of features, improve the comprehensibility of model，and identify the hidden structures in high-dimensional data．By adding the mutual information, MIRF has many advantages such as not requiring the pretreatment and no over fitting. The sorting method in Bagging can accelerate the speed of operation of the algorithm.

5. Ship Identification based on MIRF Algorithm

Automatic classification based on MIRF has the main two steps: classification model training and testing the model. First, we train the classification model as the left part, which showed in Fig. 2. In the training phase, we segment each image in the training set firstly. In this paper, the image segmentation method uses (Graph Partitioning Active Contours) GPAC image segmentation algorithm [9].
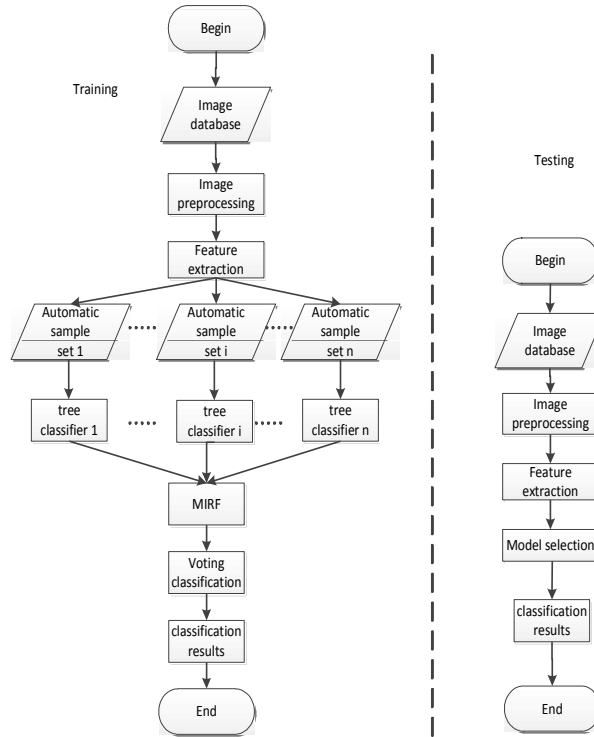
Figure 3: MIRF Algorithm Classification Model

The represents the unlabeled testing images, $Train$ represents the class which has been divided already. We segment the picture and obtain the ship target slices. Each training area $R_i \in Train$ can through a significant feature vector $f_i = x_1, x_2, ..., x_n$ and class label $k_i \in K$ to represent. Each test area $R_i \in Test$ can be represented by $f_i = x_1, x_2, ..., x_n$, but the test image category is unknown.

Then we extract the feature vector of each image slice and then use the Bagging method to generate different training sets. The classification model is obtained by the simple majority voting.

Secondly, test classification model stage. As shown in the right part of the Fig. 2. First, we segment the image to be classified, and then extract feature vector of the area.

According to the above content description and discussion, we obtain the automatic image classification algorithm based on the MIRF method, which are described in the following:

Input: the training image collection with the classification label $Train = \{I_1, I_2, ..., I_T\}$ and the image P which will be classified.

Output: the category of the image P.

Step1: model training. For each training image, the ship target area is obtained by the image segmentation, and the visual features of the region are extracted. The MIRF classification model is used to train the visual feature set. The Bagging method is used to generate the different training subsets.

Step2: the user inputs a new image P which will be classified by using the same method as the training model to get the target area of the ship; the feature is extracted from the region. We input the feature and get the classification result by using decision rules.

8

6. Experimental Results and Analysis

In this paper, the experimental environment is Intel Core 2.0GHz, 2.0GB RAM, Window 7, and Matlab 7. Experimental data are the real remote sensing images. First, segment land, sea and cloud mask pretreatment and then use the connected operator to morphological filtering and then obtain the suspected ship target area, and finally get target slice through GPAC image segmentation algorithm, which is showed in Fig. 4. Experimental data are 120 pieces of image. There are experimental data of six vessels and each vessel has 20 pieces.
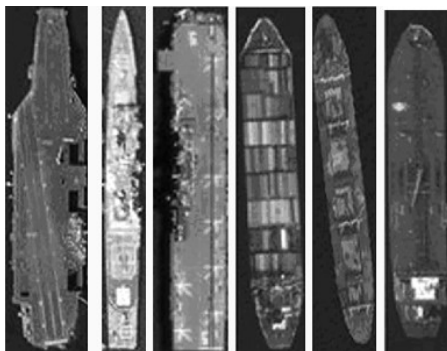

Figure 4: the Image Slice of Ships

In order to verify the effectiveness of the proposed algorithm, the experiment of MIRF and other classic machine learning algorithms, such as SVM,K-nearest neighbor method (KNN) algorithm and HDR and probability generative model(PGM) are compared[10,11]. In this paper, a total of 8 groups of experiments, each group of 6 class ship randomly selected 90, 120, 150, 180, 210, 240, 270 and 300 samples as the training sample, repeat for 10 times. And we use the cross validation (10 - fold cross validation) to evaluate the accuracy of classification algorithm. The test data is as shown in Table 1 and 2:

Method	Performance index	Group 1	Group 2	Group 3	Group 4
KNN	Rate	62.3	56.8	64.2	68.1
	Time/s	0.042	0.031	0.053	0.042
SVM	Rate	62.5	67.7	68.1	71.2
	Time/s	0.284	0.263	0.285	0.292
HDR	Rate	59.2	61.4	65.1	68.9
	Time/s	0.051	0.063	0.064	0.068
PGM	Rate	62.3	62.4	63.6	68. 8
	Time/s	0.054	0.057	0.063	0.065
RF	Rate	62.8	62.1	65.6	69.2
	Time/s	0.082	0.053	0.064	0.055
MIRF	Rate	63.2	64.8	67.2	69.8
	Time/s	0.041	0.044	0.045	0.049

Table 1: The result of the experiment (I)

Method	Performance index	Group 5	Group 6	Group 7	Group 8
KNN	Rate	70.1	71.2	73.4	74.2
	Time/s	0.052	0.063	0.064	0.074
SVM	Rate	71.3	72.1	73.2	75.1
	Time/s	0.314	0.335	0.441	0.464
HDR	Rate	72.8	76.1	80.2	82.2
	Time/s	0.071	0.067	0.076	0.081
PGM	Rate	71.2	75.4	79.5	81.3
	Time/s	0.063	0.057	0.069	0.071
RF	Rate	71.1	72.3	73.9	75.2
	Time/s	0.075	0.089	0.091	0.107
MIRF	Rate	76.2	78.2	83.2	85.3
	Time/s	0.052	0.064	0.065	0.078

Table 2 : The result of the experiment (II)

From the experiment, it can be seen that with the increase of data volume, the average classification accuracy of MIRF algorithm is the highest. On the same sample set by using SVM or the traditional RF algorithm for training, the classification accuracy is quite, but it needs to spend more time. In Tables 1 and 2, under the same experimental conditions, this algorithm compared with KNN, the recognition rate has increased by about 20 percent and the time-consuming is almost same; compared with SVM time-consuming has improved obviously, and recognition rate has slightly improved by 5 percent. The recognition rate of proposed algorithm is slightly higher than HDR and the time-consuming is slightly decreased. Compared with PGM algorithm and RF, its performance is better than the latter no matter in terms of recognition rate or time-consumption.

KNN method has taken the nearest Euclidean distance samples classes as output, it can't make use of the training sample distribution, which causes the lower recognition performance. SVM method is proposed based on statistical learning theory, which is suitable for two classes' recognition so as to expand multi-class recognition, the performance has declined. HDR has unified the classification and regression problems to return to the issue of regression, which is a tree classifier and has the ability of incremental learning. We can find out in Group 5, 6, 7 and 8. With the samples increasing, it has a higher recognition rate, but it takes long time. PGM method has an efficient recognition performance because of direct iteration method without the training process and not all the neighbor nodes participating in computation; but it is difficult in estimating the parameter of the model.

From the experiment result, it shows MIRF has better performance on both recognition rate and the time-consuming than the other algorithms. To sum up, we can conclude that MIRF algorithm can effectively improve the training speed and maintain the stability of the classification accuracy. MIRF algorithm also has its disadvantage of large demand of space storage. In the future work, we need to find a way to solve this problem.

7. Conclusion

In this paper, an improved ensemble classification algorithm MIRF is proposed and tested on the real remote sensing images. The experimental results show that the algorithm has robustness. With the increase of the amount of data, the classification accuracy is improved. The MIRF has a wide application in the ship classification. In this paper, we make some improvements in terms of the random forest method and improve the classification accuracy of random forest. What we do in this paper is a preliminary exploration, and these problems in the future need to be studied:

(1)If better methods of statistical analysis and random forest algorithm are combined, which derive a new random forest algorithm will be a good research direction.

(2)With the advent of the era of big data, the traditional random forest algorithm will become obsolete. How to realize the random forest algorithm for distributed and concurrent computing is an important research direction.

References

[1] Z .Zhang, A Study on Harbor Target Recognition in High Resolution Optical Remote Sensing Image[D]. Hefei: University of Science and Technology of China, 2009. (In Chinese)

[2] J. Lan, L.Wan. Automatic ship target classification based on aerial images[J], Proceedings of SPIE. Bellingham Wash:SPIE. 1-10 (2009).

[3] C. Du, J. X. Sun, Z.Y. Li, S. H. Teng, Method for ship recognition using optical remote sensing data [J], Journal of Image and Graphics,17(4): 591-592(2012). (In Chinese)

[4] J. Y. Weng, W. H. wang, Incremental hierarchical discriminant regression[J], IEEE Transactions on Neural Networks, 18(2): 397-415(2007).

[5] J. Antelo，G. Ambrosio, J. Gonzalez, Ship detection and recognition in high resolution satellite images, IEEE International Geoscience and Remote Sensing Symposium．Washington DC，USA: IEEE Computer Society, 514-517 (2009)．

[6] Y. An, X. F. Wang, X. Z Xia, L. Li, Detection and recognition of sea battlefield's targets[J], J.Huazhong Univ.of Sci.&Tech.(Natual Science Edition),40(10):9-12(2012). (In Chinese)

[7] L. Rokach, Ensemble-based classifiers[J], Artificial Intelligence Review , 33(1):1−39(2010).

[8] J.L. Xu, Y.M. Zhou, L .Chen, B.W. Xu, An Unsuper Vised Feature Selection Approach Based on Mutual Information[J], Journal of computer Research and Development,49(2):372-382(2012). (In Chinese)

[9] B. S.Manjunath, B. Sumengen, Graph partitioning active contours (GPAC) for image segmentation [J], IEEE Transactions on Pattern Analysis and Machine Intelligence,28(4):509-521(2006)．

[10] G.C. Anagnostopoulo, SVM-based target recognition from synthetic aperture radar images using target region outline descriptors[J], Nonlinear Analysis,71(12):2934-2939(2009).

[11] W. Y.Guo,X. Z.Xia,X. F.Wang, A remote sensing ship recognition method based on dynamic probability generative model[J], Expert Systems with Applications,41(14):6452-6456(2014).