

A new oil spill detection algorithm based on Dempster-Shafer evidence theory*

Tianlong ZHANG^{1, 4}, Jie GUO^{1, 2, 3, **}, Chenqi XU^{1, 4}, Xi ZHANG⁵, Chuanyuan WANG^{1, 2, 3}, Baoquan LI^{1, 2, 3}

¹ Yantai Institute of Coastal Zone Research (YIC), Chinese Academy of Sciences (CAS), CAS Key Laboratory of Coastal Environmental Processes and Ecological Remediation, Yantai 264003, China

² Shandong Key Laboratory of Coastal Environmental Processes, Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai 264003, China

³ Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao 266071, China

⁴ University of the Chinese Academy of Sciences, Beijing 100049, China

⁵ First Institute of Oceanography (FIO), Ministry of Natural Resources (MNR), Qingdao 266061, China

Received Jul. 4, 2020; accepted in principle Aug. 31, 2020; accepted for publication Mar. 22, 2021

© Chinese Society for Oceanology and Limnology, Science Press and Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract Features of oil spills and look-alikes in polarimetric synthetic aperture radar (SAR) images always play an important role in oil spill detection. Many oil spill detection algorithms have been implemented based on these features. Although environmental factors such as wind speed are important to distinguish oil spills and look-alikes, some oil spill detection algorithms do not consider the environmental factors. To distinguish oil spills and look-alikes more accurately based on environmental factors and image features, a new oil spill detection algorithm based on Dempster-Shafer evidence theory was proposed. The process of oil spill detection taking account of environmental factors was modeled using the subjective Bayesian model. The Faster-region convolutional neural networks (RCNN) model was used for oil spill detection based on the convolution features. The detection results of the two models were fused at decision level using Dempster-Shafer evidence theory. The establishment and test of the proposed algorithm were completed based on our oil spill and look-alike sample database that contains 1 798 image samples and environmental information records related to the image samples. The analysis and evaluation of the proposed algorithm shows a good ability to detect oil spills at a higher detection rate, with an identification rate greater than 75% and a false alarm rate lower than 19% from experiments. A total of 12 oil spill SAR images were collected for the validation and evaluation of the proposed algorithm. The evaluation result shows that the proposed algorithm has a good performance on detecting oil spills with an overall detection rate greater than 70%.

Keyword: synthetic aperture radar (SAR) data; oil spill detection; subjective Bayesian; Faster-region convolutional neural networks (RCNN); Dempster-Shafer evidence theory

1 INTRODUCTION

Oil pollution is a serious marine problem worldwide (Zhang et al., 2011). Different types of oil and its refined products such as gasoline, kerosene, and diesel oil enter the marine environment during the process of exploitation, refining, storage, transportation, and utilization (Li et al., 2013). The pollution caused by these oil spills greatly threatens the ocean fisheries resource, the ecological environment, and human activities (Leifer et al., 2012). Thus, prompt and efficient oil spill detection is

of great significance to the protection of coastal resources and the marine ecosystem (Guo et al., 2013).

The oil spill detection for polarimetric synthetic aperture radar (SAR) mainly consists of three parts:

* Supported by the National Key R&D Program of China (No. 2017YFC1405600), the National Natural Science Foundation of China (Nos. 42076197, 41576032), and the Major Program for the International Cooperation of the Chinese Academy of Sciences (No. 133337KYSB20160002)

** Corresponding author: jguo@yic.ac.cn

dark spot detection, feature extraction, and discrimination between oil spill and the look-alike (Guo et al., 2018). Many studies have focused on the exploration of new features extracted from the polarimetric SAR data and the innovation of oil-spill detection methods in recent years. The related works indicate that polarimetric features or polarimetric feature sets extracted from different channels can help for oil spill detection (Zhang et al., 2011; Guo et al., 2018; Tong et al., 2019). However, SAR sensors work in different observation models and this has limited the application of these features due to a lack of an appropriate channel. With the development of machine learning and deep learning, many classification algorithms have been introduced to the field of oil spill detection with appropriate adjustments and modifications. The convolutional neural networks (CNN) and the fully convolutional networks (FCN) have also been successfully applied on the discrimination of oil spills and look-alikes in polarimetric SAR images (Huang et al., 2018). Krestenitis et al. (2019) compared oil-spill detection methods based on different CNNs including U-Net, LinkNet, and other CNNs. The authors indicated that the optimal model for classifying oil spills is the deep convolutional neural network (DCNN). It can be found that the CNNs can be used as a model to extract the image features and they have good performance when distinguishing oil spills and look-alikes. Environmental information is also an important source for distinguishing oil spills and look-alikes (Solberg et al., 2003, 2007). However, these studies or methods have not considered environmental information from dark spots, such as wind speed, etc.

To distinguish oil spills and look-alikes more accurately with environmental factors and image features in a more objective way, a new oil spill detection algorithm based on Dempster-Shafer evidence theory (Dempster, 1967; Shafer, 1976) is proposed in this paper. The discrimination of oil spills and look-alikes can also be regarded as a problem of uncertainty in some cases (Solberg et al., 2003, 2007). Many algorithms including Bayes, neural networks and the Dempster-Shafer theory are employed to solve the uncertainty problems. The Bayesian method can make a scientific judgment based on information from multiple sources and it can solve the uncertainty problems successfully (Li et al., 2016). Thus, environmental information including wind speed, the distance from samples to waterways and oil offshore platforms were combined to distinguish oil spills and

look-alikes using the subjective Bayesian method (Duda et al., 1976). Because the Faster-region convolutional neural networks (RCNN) (Ren et al., 2017) has been successfully applied to object detection with convolution features (Manana et al., 2018; Han et al., 2019), Faster-RCNN was used as a model to extract the convolution features and to detect oil spills from the images. Features extracted from the single channel intensity image have been more widely and successfully used in many studies and the use of a single channel intensity image can reduce the channel restrictions (Solberg et al., 2003, 2007). Oil spill detection with features extracted from a single channel intensity image will be more conducive to real-time business application (Solberg et al., 2003, 2007). The Dempster-Shafer theory also shows a great capability to fuse the multiple observational evidence to solve the uncertainty problems (Zeng et al., 2019; Yang et al., 2020). Furthermore, the Dempster-Shafer theory has the characteristics of strong anti-interference, low sensor dependence, and relatively strong fault tolerance. A more robust result with less uncertainty was provided using the feature level and decision-making level Dempster-Shafer theory. Thus, the Dempster-Shafer evidence theory was selected for the fusion of the output results from the subjective Bayesian method and Faster-RCNN model to give the final detection result of oil spills in this paper.

2 MATERIAL AND METHOD

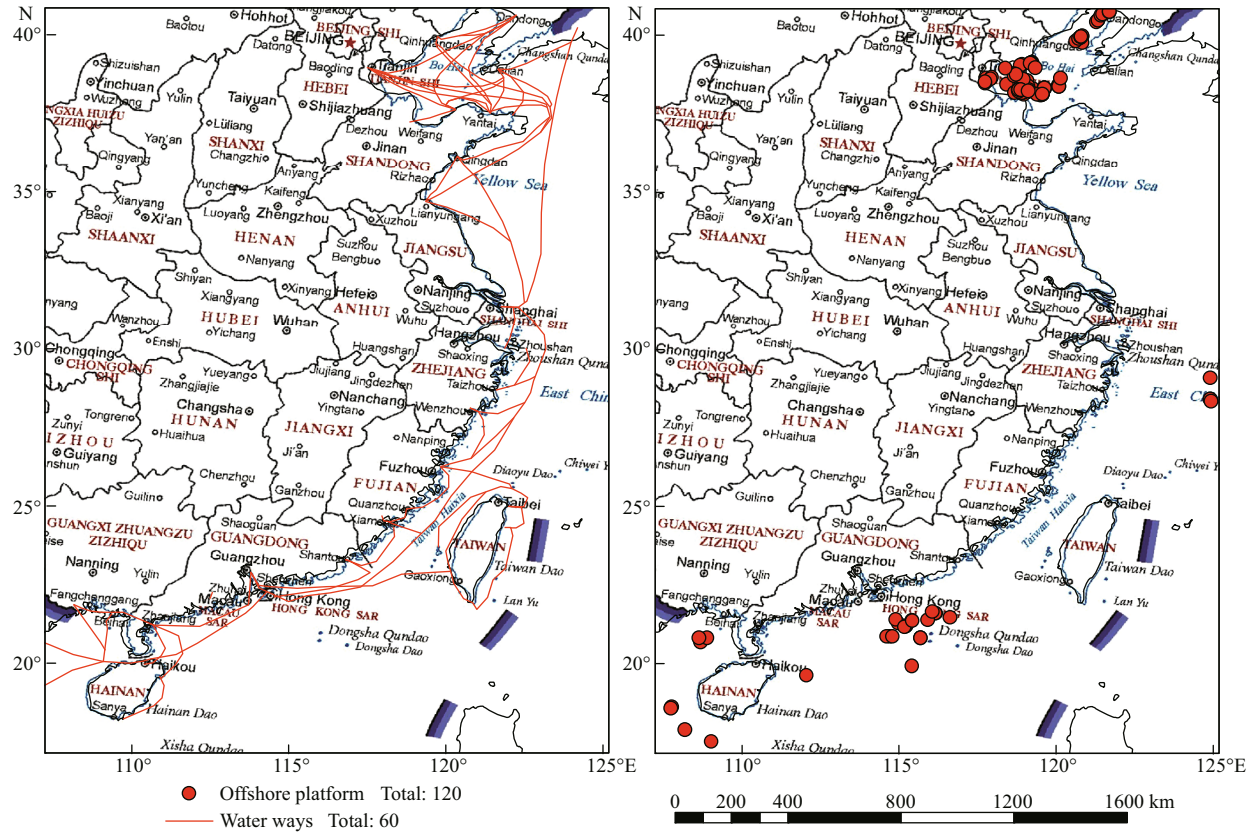
2.1 Data and the preprocessing

2.1.1 SAR data and preprocessing

A total of 190 SAR images from 2008–2018 acquired from ENVISAT-ASAR, Sentinel-1A/B, and RADARSAT-1/2 were collected for the proposed algorithm experiment. These SAR data mainly recorded oil spills and look-alikes in the Bohai Sea, the South China Sea, and the coastal waters of China. The number of SAR images acquired from ENVISAT-ASAR, Sentinel-1A/B, and RADARSAT-2 is 170. The information of the three main SAR systems is listed in Table 1. The C-band VV and HH polarization data from these SAR images were used to provide oil spill and look-alike image samples for our algorithm experiment due to the limitation of SAR images (Krestenitis et al., 2019). The selected data was processed with Sentinel Application Platform (SNAP) software provided by ESA. The processed data were

Table 1 The information of SAR systems

SAR system	Revisit period (d)	Incidence angle (°)	Polarization mode	Resolution (m)	Band/observation mode
ENVISAT-ASAR	35	26°–39°	VV+HH/VV+VH	30 m	C/alternating polarization model, wide swath model
Sentinel-1A/B	12	29°–46°	VV+VH/HH+HV	20 m×2 m	C/interferometric wide swath
RADARSAT-2	24	20°–41°	Quad	25 m×8 m	C/standard strip map

**Fig.1** The distribution of waterways (left) and offshore oil platforms (right) along coast of China

Map review No. (2019)1671.

resampled at a spatial resolution of 30 m. To reduce the interference of speckle noise, all the data was processed by Refined-Lee filter (7×7 window) (Lee et al., 1999).

2.1.2 Supplementary data and the preprocessing

The TRMM Multi-satellite Precipitation Analysis (TMPA) 3B42 data, the Global Precipitation Measurement (GPM) data and European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-interim re-analysis data for the same period as the SAR data were downloaded to provide rainfall and wind speed data for the interpretation of oil spills and look-alikes. The time step and spatial resolution of the TMPA 3B42 data and GPM data are 3 h, 0.5 h and 0.25°, 0.1° respectively. The time interval of the ECMWF ERA-interim data is 6 h and the data are

stored in NetCDF format with a resolution of 0.125°. The TMPA 3B42 data and ECMWF ERA-interim re-analysis data were resampled to 0.125° using bilinear interpolation. In addition, offshore oil platform locations data (Commander Department of the Navy, 2005a, b) and waterway data (China Cartographic Publishing House, 2015) over the coastal waters of China were collected to service the interpretation. The two data were vectorized using ESRI ArcGIS vectorization tools and the distribution of waterways and offshore platforms are shown in Fig.1.

2.2 Establishment of oil spill and look-alike sample database

There are few open-source oil spill and look-alike sample databases and they are difficult to obtain. Therefore, an oil spill and look-alike sample database

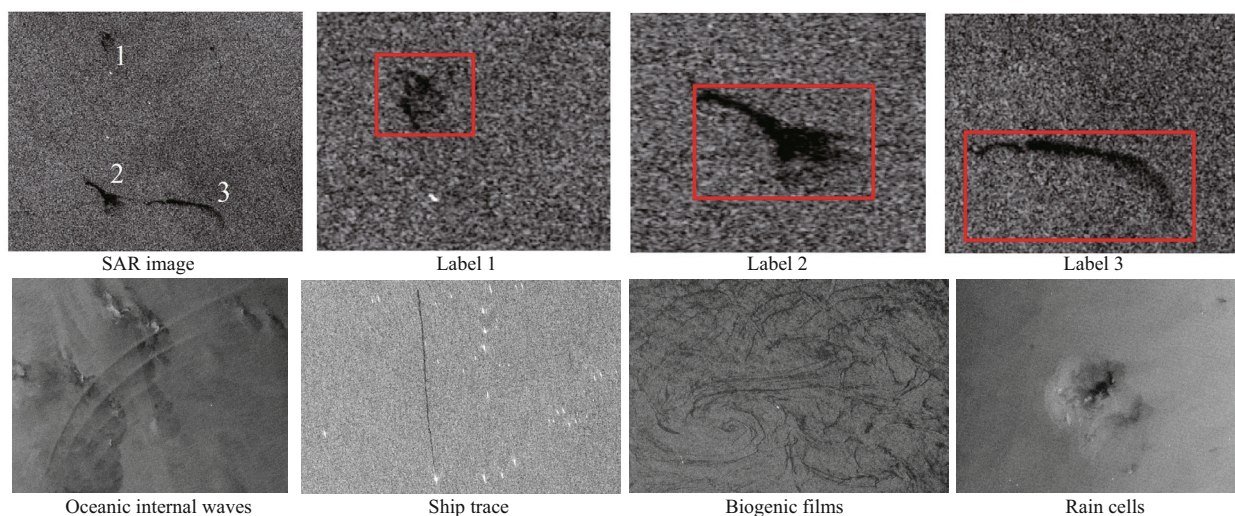


Fig.2 Illustrations of labeled simple images (first row) and look-alikes (second row)

The number represents the label of the oil film, the red frame is the enlarged view of the oil film 1, 2, 3.

(OSLSD) was established with oil spill and look-alike samples that have been interpreted from SAR images. The actual oil spill and look-alike samples were selected using a visual interpretation method based on the prior knowledge of experts (Solberg et al., 2003, 2007; Karathanassi et al., 2006). The visual interpretation followed the rules and descriptions of oil spills with reference to Solberg et al. (2003, 2007) and Karathanassi et al. (2006). The OSLSD is an important basic database for the construction and accurate evaluation of the new proposed algorithm in this paper. It consists of two parts: the first part is the Image Samples Database (ISD) and the second part is the Environmental Information Database (EID). The ISD consists of 1 798 image samples including 828 oil spill samples and 970 look-alike samples in GeoTiff format and the corresponding Extensible Markup Language (XML) files created by the LabelImg (2018). The XML files record the boundary and classification of the corresponding image samples (Fig.2). Seven types of look-alikes including oceanic internal waves, upwelling, ship trace, rain cells, wind sheltered by land, biogenic films, and low wind velocity were interpreted based on the above visually interpreted method.

The EID consists of environmental information such as the wind speed related to the image samples. The data in the two databases were matched by the filename of the image samples. The EID included the wind speed data and the distance between the position of all samples and offshore oil platforms and waterways. The geometric center of image samples was regarded as the optimal location for collecting

wind speed data and distance data. The location data were acquired using a vectorization tool from ESRI ArcGIS software. Using the location data, wind speed data were extracted from the ECMWF ERA-interim data. The distance between the optimal location and the offshore oil platforms or waterways was measured using an ESRI ArcGIS measurement tool. The distance data of oil platforms and waterways are represented by sample-platform data and sample-waterway data for the convenience of writing and expression.

2.3 Oil spill detection algorithm based on Dempster-Shafer evidence theory

2.3.1 The subjective Bayesian oil spill detection model

The Support Vector Machine (SVM) and Artificial Neural Networks (ANN) are easy to over-fit if the number of samples is relatively limited when training the classification model. Besides, the uncertainty of samples should be considered when using the Dempster-Shafer model to express the uncertainty of the fusion results (Li et al., 2016; Yang et al., 2020). Thus, the subjective Bayesian method is selected to solve the problem that whether a sample is classified as an oil spill under the influence of single or multiple environmental factors. The environmental information can be regarded as evidence E and the occurrence of an oil spill can be regarded as event H (Duda et al., 1976). Because the probability $P(H|E)$ and the probability $P(E)$ are difficult to determine when applying the subjective Bayesian probabilistic reasoning model, the parameters Likelihood of

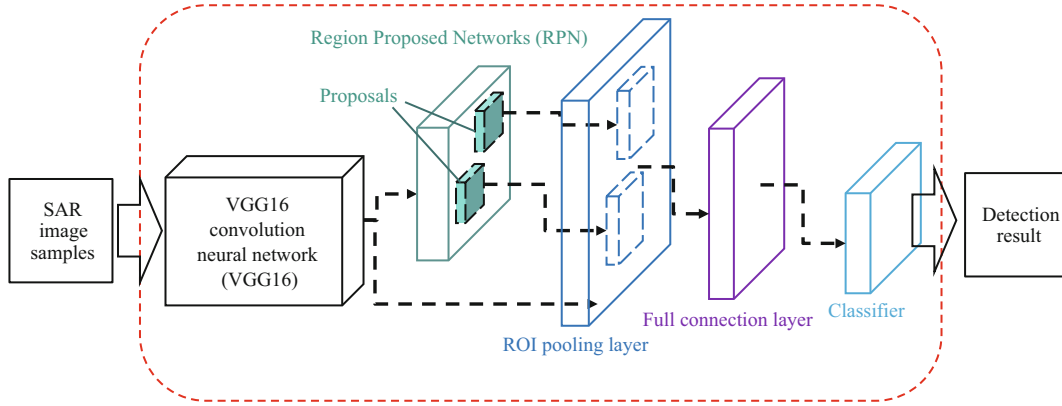


Fig.3 The structure diagram of Faster-RCNN

Sufficiency (LS) and Likelihood of Necessity (LN) were introduced into the subjective Bayesian probabilistic reasoning model to determine the two probabilities (Duda et al., 1976). The LS and LN parameters embody the sufficiency and the necessity of the reasoning rules respectively (Duda et al., 1976). The values of LS and LN can be acquired using the statistical analysis results of the samples, the two parameters can be described as follows:

$$O(H) = \frac{P(H)}{1 - P(H)}; LS = \frac{P(E|H)}{P(E|\sim H)}. \quad (1)$$

If the probability $P(E)$ equals 1, the probability $P(H|E)$ and $P(H|\sim E)$ are defined as follows:

$$P(H|E) = \frac{LS \times P(H)}{(LS - 1) \times P(H) + 1}, P(H) = 1, \quad (2)$$

where H and E are the event and the evidence, respectively. If E' represents the observations affecting evidence E with uncertainty, then $P(H|E')$ is considered in terms of the rules (Eq.3) given by Duda et al. (1976).

$$P(H|E') = P(H|E) \times P(E|E') + P(H|\sim E) \times P(\sim E|E'). \quad (3)$$

It is necessary to combine this evidence E_s to obtain the final posterior probability of event H . The posterior odds of event H can be defined as follows:

$$\begin{aligned} O(H|E_1 \cap E_2 \cdots \cap E_n) \\ &= \frac{O(H|E_1)}{O(H)} \times \frac{O(H|E_2)}{O(H)} \cdots \frac{O(H|E_n)}{O(H)} \times O(H) \\ &= LS_1 \times LS_2 \cdots \times LS_n \times O(H). \end{aligned} \quad (4)$$

The probability $P(H)$, $P(E|H)$, and $P(E|\sim H)$ can be obtained using the statistical analysis result. $P(H|E)$ and $P(H|E_1 \cdots E_n)$ represent the probabilities of a sample if it is an oil spill sample under the influence of a single or multiple environmental factors, respectively. In this study, three environmental factors including wind speed and the distances between

samples and oil platforms or waterways were extracted from the supplementary data from different sensors, respectively. It is not appropriate to detect whether a dark spot is an oil spill using a single environmental factor due to the complexity of the marine environment. Thus, the subjective Bayesian oil spill detection model was employed to product a detection result with the influence of multiple environmental factors.

2.3.2 The Faster-RCNN oil spill detection model

The Faster-RCNN oil spill detection model was established using the sample image features from a different perspective. The Faster-RCNN is a well-known object detection framework that was proposed by Ren et al. (2017) after the introduction of RCNN and Fast-RCNN in 2016. The Faster-RCNN model combines the feature extraction layer, the proposal layer, the bounding box regression layer, and the classification layer into one object detection framework and it is very convenient to complete the process of feature extraction, classification, and position correction (Ren et al., 2017). The structure of the Faster-RCNN model is shown in Fig.3. In this study, the detection result can be regarded as the supportive degree for each detected dark spot in SAR images (Yang et al., 2020).

The VGG16 (Simonyan and Zisserman, 2015) pre-trained convolution neural network was selected as the convolution feature extraction network in this paper. The convolution kernel of window size 7×7 is replaced by three convolution kernels in window size 3×3 . The non-maximum suppression (NMS) algorithm is an important part of the Faster-RCNN and is used to remove the duplicate detection box when the intersection over union (IOU) is greater than a certain threshold (Bodla et al., 2017). The Soft-Non-Maximum Suppression (Soft-NMS) algorithm can solve this

problem effectively by reducing the confidence of those boxes with IOU greater than the threshold rather than deleting them (Bodla et al., 2017).

2.3.3 The Dempster-Shafer evidence theory

To add the environmental information from dark spots to distinguish oil spills and look-alikes, the information for the above two models should be fused using the information fusion method. The method can be divided into three levels: the data level, the feature level, and the decision-making level. The information fusion on the data level and the feature level requires a high precision matching relationship in time and space of the original information. That means the fault tolerance and anti-interference of the above two types of information fusion are not strong. Because of the limitation mentioned above, the data level and the feature-level information fusion are not suitable for oil spill detection in a complex marine environment. Therefore, the decision-making level-information fusion method named Dempster-Shafer theory was employed in this paper due to the characteristics of strong anti-interference, low sensor dependence, and relatively strong fault tolerance. The detection process and the detection result of the Bayesian model and the Faster-RCNN model are independent, and the detection results of the two models can be regarded as the evidence for the occurrence of an oil spill (Dempster, 1967; Shafer, 1976). In Dempster-Shafer theory, the Basic Probability Assignment (BPA), Brief function (Bel) and Plausibility function (Pls) are three important parts of evidence theory (Dempster, 1967; Shafer, 1976). The Pls establishes a bridge between an abstract mathematical model and the actual proposition and it can transform the original logical reasoning problem into a mathematical aggregation problem. BPA represents the initial allocation of believability. Bel and Pls represent the support and suspicion of a proposition and they can be described and calculated with the mass function. The definition of Bel and Pls are as follows:

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B), (\forall A \subseteq \Theta), \quad (5)$$

$$\text{Pls}(A) = \sum_{B \cap A \neq \Phi} m(B) = 1 - \text{Bel}(\bar{A}), (\forall A \subseteq \Theta), \quad (6)$$

where Θ is the discernment frame of proposition A and A is the arbitrary proposition. B and \bar{A} are the subset and supplementary set of A . Φ represents the empty set. $m(B)$, named mass function, represents the BPA and it meets the following conditions (Dempster,

1967; Shafer, 1976):

$$\sum_{B \subseteq \Theta} m(B) = 1; m(\Phi) = 0. \quad (7)$$

It is necessary to fuse the evidence sources to make more accurate decisions on the proposition supported by multiple sets of evidence sources simultaneously. The fusion must meet the Dempster Shafer rules in reference (Dempster, 1967; Shafer, 1976). Because the result of the orthogonal process for the BPA function is still a BPA function (Dempster, 1967; Shafer, 1976), the new Bel function and Pls function can be updated with the updated BPA. The expression is shown as Eq.8:

$$\begin{aligned} \delta_{i=1}^n m_i &= \frac{1}{1-K} \sum_{\substack{C_{i=1}^n \\ A \cap B = A}} \prod_{i=1}^n m_i(A^i), \\ K &= \sum_{\substack{C_{i=1}^n \\ A \cap B = \Phi}} \prod_{i=1}^n m_i(A^i), \end{aligned} \quad (8)$$

where δ represents the fusion result, A^i represents the i^{th} arbitrary proposition for the discernment frame, m_i represents the i^{th} mass function for the corresponding A^i , and K is the conflict level of evidence. The detection result obtained from the subjective Bayesian oil spill detection model and Faster-RCNN oil spill detection model can be regarded as the supportive degree or evidence for a dark spot detected as oil spill. The supportive degrees can be used to construct the BPA function. The Dempster-Shafer theory can also make the final decision with the multiple supportive degrees (Yang et al., 2020) and the fusion result can be considered as the final result for oil spill detection. The implementation chart is shown in Fig.4.

2.4 Experimental setup

The Faster-RCNN for CPU versions based on the Tensorflow framework was used for training the image samples using the Windows 10 operating system in this paper. The CPU was the Intel core i7-6700 and the running memory size was 16 GB. The learning rate and the weight decay were set to 0.005 and 0.0001. The batch size was set to 256 and the maximum iterations were set to 40 000 during training image samples. The step size for reducing the learning rate was set to 5 000. In addition, the number of top proposals to select was set to 2 000 for oil spill detection with the trained network. 70% of the samples (1 259 samples) including oil spill and look-alike samples from the ISD were selected randomly to train the Faster-RCNN oil spill detection model. The remaining 539 samples were selected for the model test.

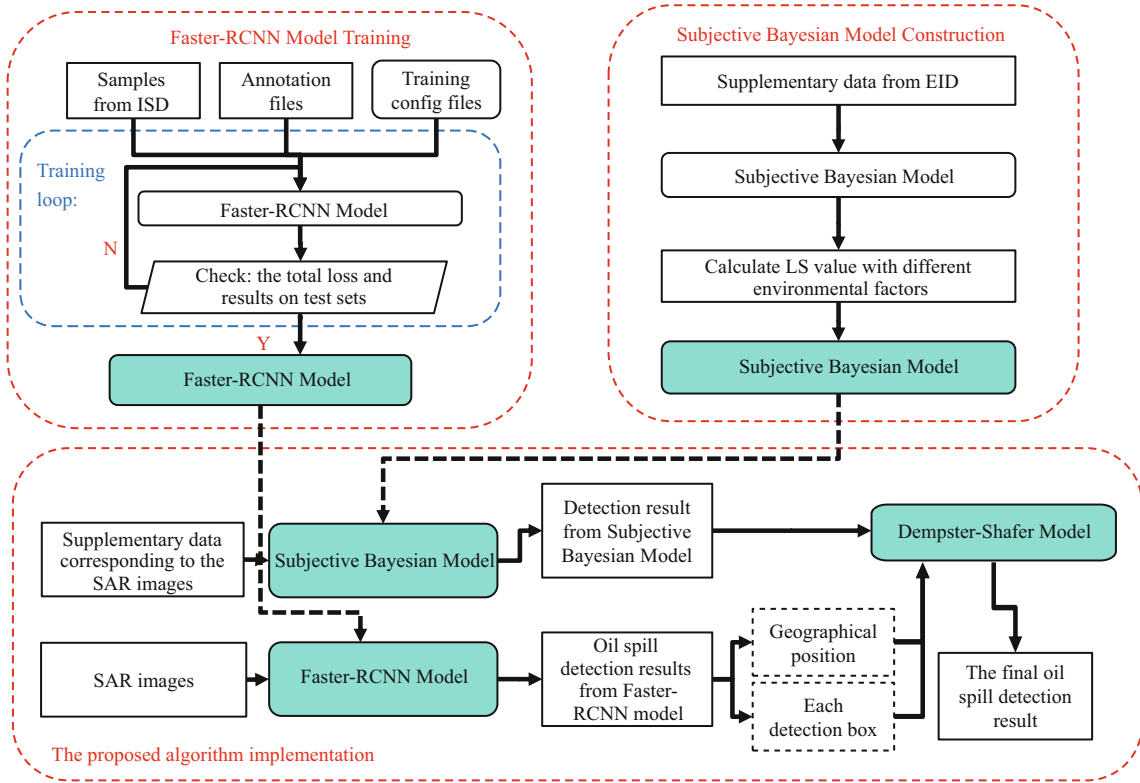


Fig.4 The implementation chart

3 RESULT AND DISCUSSION

3.1 Evaluation method

The detection rate (DR), the false alarm rate (FAR) and the identification rate (IR) were used to obtain objective and accurate evaluation results of the subjective Bayesian detection model, the Faster-RCNN detection model and the proposed algorithm. The expressions for DR, FAR, and IR are as follows:

$$\begin{aligned} DR &= \frac{TT}{TT+TF}, \\ FAR &= \frac{FT}{TT+FT}, \\ IR &= \frac{TT+FF}{TT+FF+TF+FT}, \end{aligned} \quad (9)$$

where TT represents the number of correctly classified oil spill samples and TF represents the number of misclassified oil spill samples. FF represents the number of correctly classified look-alike samples, and FT represents the number of misclassified look-alike samples.

3.2 Analysis and evaluation of subjective Bayesian oil spill detection model

The analysis and evaluation of the subjective Bayesian oil spill detection model was undertaken

with data from the EID. All 1 798 records from the database were used to determine the LS value of the environmental factors. The probability P (oil spill) was 0.461 in terms of the preliminary statistical result of all the records.

To determine the LS value in an objective way, the samples with different wind speed, sample-platform data, and sample-waterway data were counted (Fig.5). The LS value was determined with the certain interval integral method due to the small number of samples. The step interval of the certain interval integral method was set to 1 m/s, 1 km, and 1 km, respectively. The upper limit of the LS value calculation for wind speed data, sample-waterway data, and sample-platform data were set to 5 m/s, 80 km, and 300 km respectively.

The co-ordinates of marked points are 3.27, 0.23 in Fig.6a. When the wind speed is greater than 3.27 m/s, the growth of the LS value is accelerated. The curve in Fig.6a shows that the supportiveness of wind speed on samples belonging to oil spill gradually increases as wind speed increases from 3 m/s. This is consistent with the conclusion of Bern et al. (1993). The increase of supportiveness is mainly due to the decrease in the number of look-alike samples with low wind speeds. Thus, the intervals for the calculation of posteriori probabilities $P(E|H)$ are 0–3 m/s and >3 m/s.

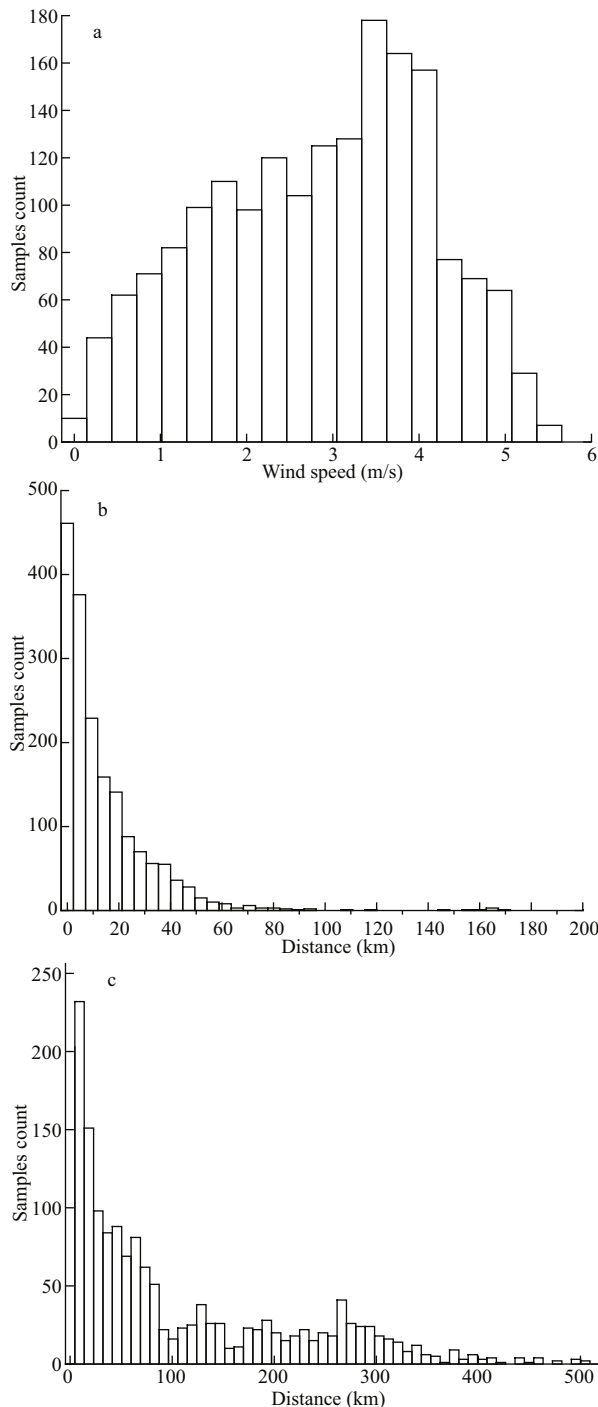


Fig.5 Histograms of wind speed data (a), the distance to waterways (b), and offshore oil platforms (c)

The curve of the LS values and the sample-waterway data is shown in Fig.6b. The fluctuation of the curve with the sample-waterway data ranges from 0 to 10 km and this is due to the small number of samples and the variation of the proportion of oil spill and look-alike samples. The coordinates of the marked turning point are 19.3 and 1.72. When the distance is greater than 19.3 km, the rate of decrease

Table 2 The detailed information and calculation result of posteriori probabilities $P(E|H)$

Environmental factor	$P(E H)$	Qualified samples/ total samples	The value of $P(E H)$
Wind speed	$P(0-3 \text{ m/s} \text{oil spill})$	178/828	0.215
	$P(>3 \text{ m/s} \text{oil spill})$	650/828	0.785
	$P(0-3 \text{ m/s} \text{look-alike})$	589/970	0.607
	$P(>3 \text{ m/s} \text{look-alike})$	391/970	0.393
Offshore oil platform	$P(0-30 \text{ km} \text{oil spill})$	155/828	0.187
	$P(>30 \text{ km} \text{oil spill})$	673/828	0.813
	$P(0-30 \text{ km} \text{look-alike})$	103/970	0.106
	$P(>30 \text{ km} \text{look-alike})$	867/970	0.894
Waterway	$P(0-20 \text{ km} \text{oil spill})$	600/828	0.676
	$P(>20 \text{ km} \text{oil spill})$	228/828	0.324
	$P(0-20 \text{ km} \text{look-alike})$	331/970	0.342
	$P(>20 \text{ km} \text{look-alike})$	639/970	0.658

in the LS value becomes increasingly slow. Although in general the LS decreases in line with the increase of distance, all the LS values are greater than 1. This trend means that the supportiveness of distance of oil spill samples increases with a decreasing distance. This is consistent with the fact that an oil spill from a ship is an important source of marine oil spills. The intervals for the calculation of posteriori probabilities $P(E|H)$ are 0–20 km and >20 km.

The relationship between LS values and sample platform data is shown in Fig.6c. The marked turning point is the peak of the curve, and its coordinates are 32.8, 1.78. In general, the LS value decreases with increasing distance from 32.8 km. The supportiveness of distance on samples belonging to oil spill decreases as distance increases. This trend means that the result of the LS value is consistent with the fact that oil spill from offshore oil platforms is another important source of oil spill. The intervals for the calculation of posteriori probabilities $P(E|H)$ are 0–30 km and >30 km.

For the definition of posteriori probabilities $P(E|H)$, the situation whereby a sample belongs to an oil spill can be regarded as event H . Similarly, the intervals of wind speed data, sample-waterway data and sample-platform data are regarded as evidence E . The result of posteriori probabilities $P(E|H)$ with different intervals under different environmental factors were calculated and the detailed information and calculated results are shown in Table 2.

The posterior probability $P(H|E_1 \cdots E_n)$ can be regarded as the final result with the influence of multiple environmental factors and it can be obtained with the calculation rules in Section 2.3.1 and the

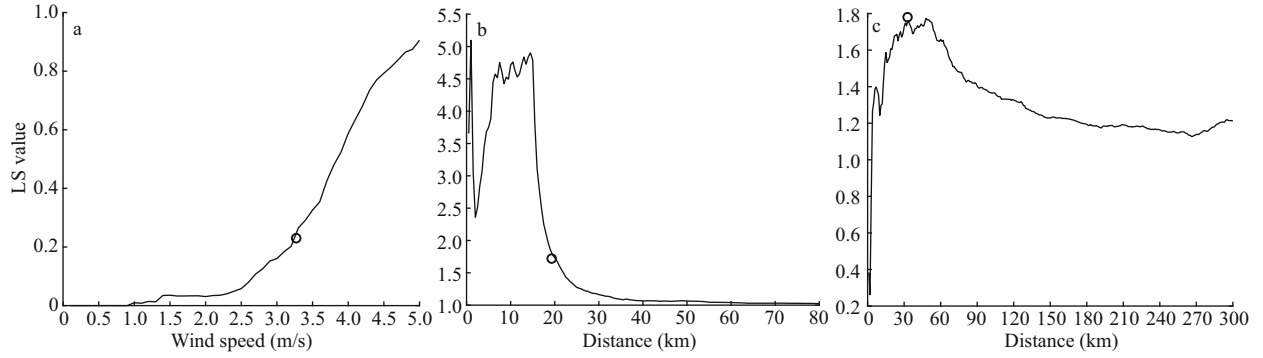


Fig.6 The curve of LS with wind speed (a), sample-waterway data (b), and sample-platforms data (c)

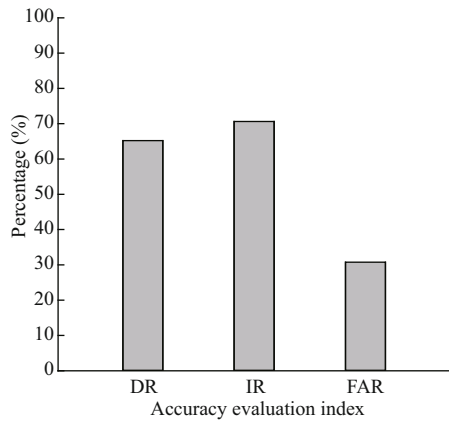


Fig.7 The evaluation result of subjective Bayesian oil spill detection model

above computation result. The method introduced in Section 3.1 was selected to complete the evaluation. The detection rule is that if the process result for a sample is greater than the threshold value, the sample will be classified as oil spill; otherwise, the sample is classified as look-alike. Based on the prior knowledge of experts, the above statistical results in Section 3.2 and the suggestion given by Duda et al. (1976), the threshold value is set to 0.5. The evaluation result with 1 798 samples is shown in Fig.7. The evaluation

result indicates that the subjective Bayesian oil spill detection model effectively distinguished oil spill and look-alike and achieved DR and IR values greater than 60%.

3.3 Analysis and evaluation of Faster-RCNN oil spill detection model

The Faster-RCNN oil spill detection model was evaluated using the method described in Section 3.1 based on the remaining 539 samples. The total loss curve with iterations is shown in Fig.8a. It can be found that the network converged and the total loss was lower than 0.7 when the iterations reached 40 000. It also can be seen that the total loss curve dropped several times after protrusions in Fig.8a. The reason for this phenomenon is that the larger initial detection box in early iteration gradually returns with the increase of iterations during training. In order to see the sensitivity and the specificity of the oil-spill detection model, a receiver operating characteristic (ROC) curve was plotted with a true positive rate (TPR) and a false positive rate (FPR). TPR indicates the proportion of TT and the sum of TT and FT. FPR indicates the rate of FT and the sum of FT and TF. The area under the curve (AUC) is also calculated with

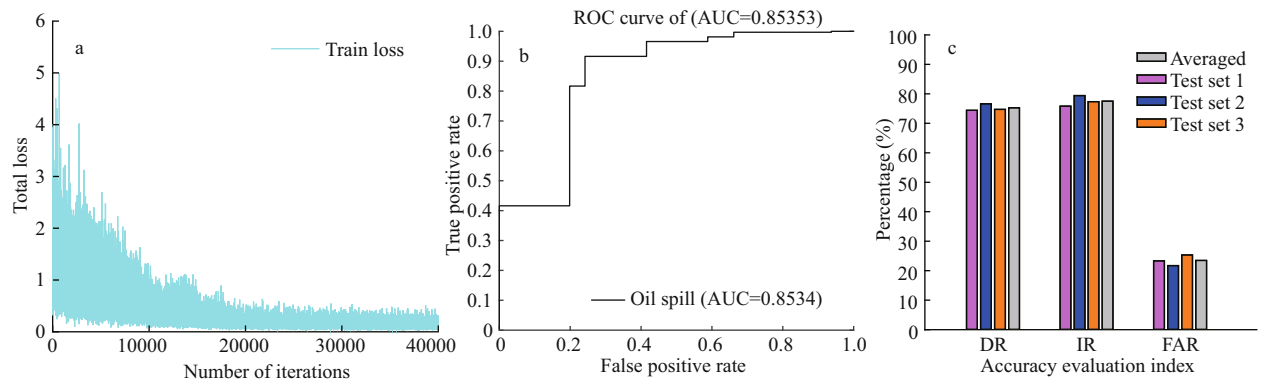


Fig.8 The Faster-RCNN total loss curve during model training (a), the ROC curve of the detection model (b), and the evaluation result of Faster-RCNN oil spill detection model (c)

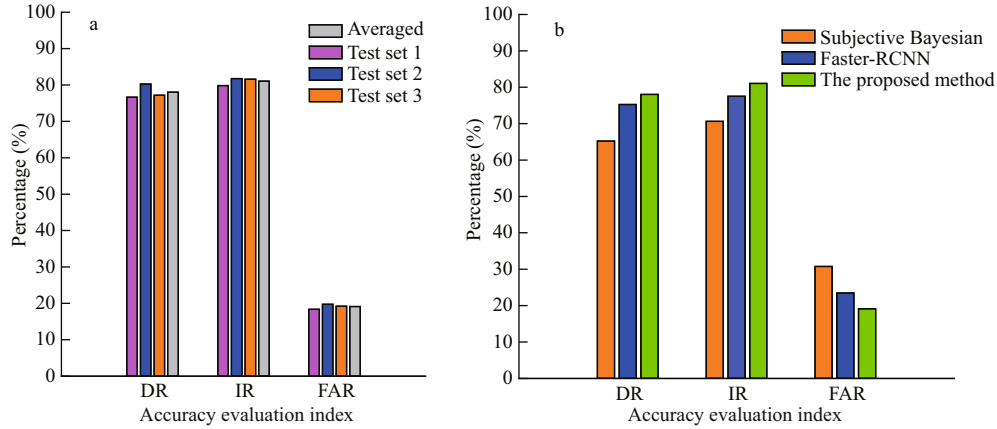


Fig.9 The evaluation result of the proposed algorithm with the test sets (a) and the comparison result of the subjective Bayesian method, Faster-RCNN method, and the proposed method (b)

TPR and FPR using the certain interval integral method. The ROC curve is shown in Fig.8b. The ROC curve and the AUC are helpful to assess the performance of a classifier. In general, a good classifier corresponds to a higher AUC value. The closer the ROC curve is to the upper left corner, the better the classifier is. It can be found that the detection model shows a good performance with AUC rates higher than 0.8. In addition, the ROC curve is relatively close to the upper left corner. The ROC curve and AUC shows that the classifier can distinguish oil spills and look-alikes effectively. According to all of the above, the detection model was used for oil spill detection of the remaining samples. A total of 459 samples randomly selected from the remaining samples were used to build a test set. The above process was repeated three times to build 3 test sets. The evaluation result is shown in Fig.8c.

3.4 Analysis and evaluation of the proposed algorithm

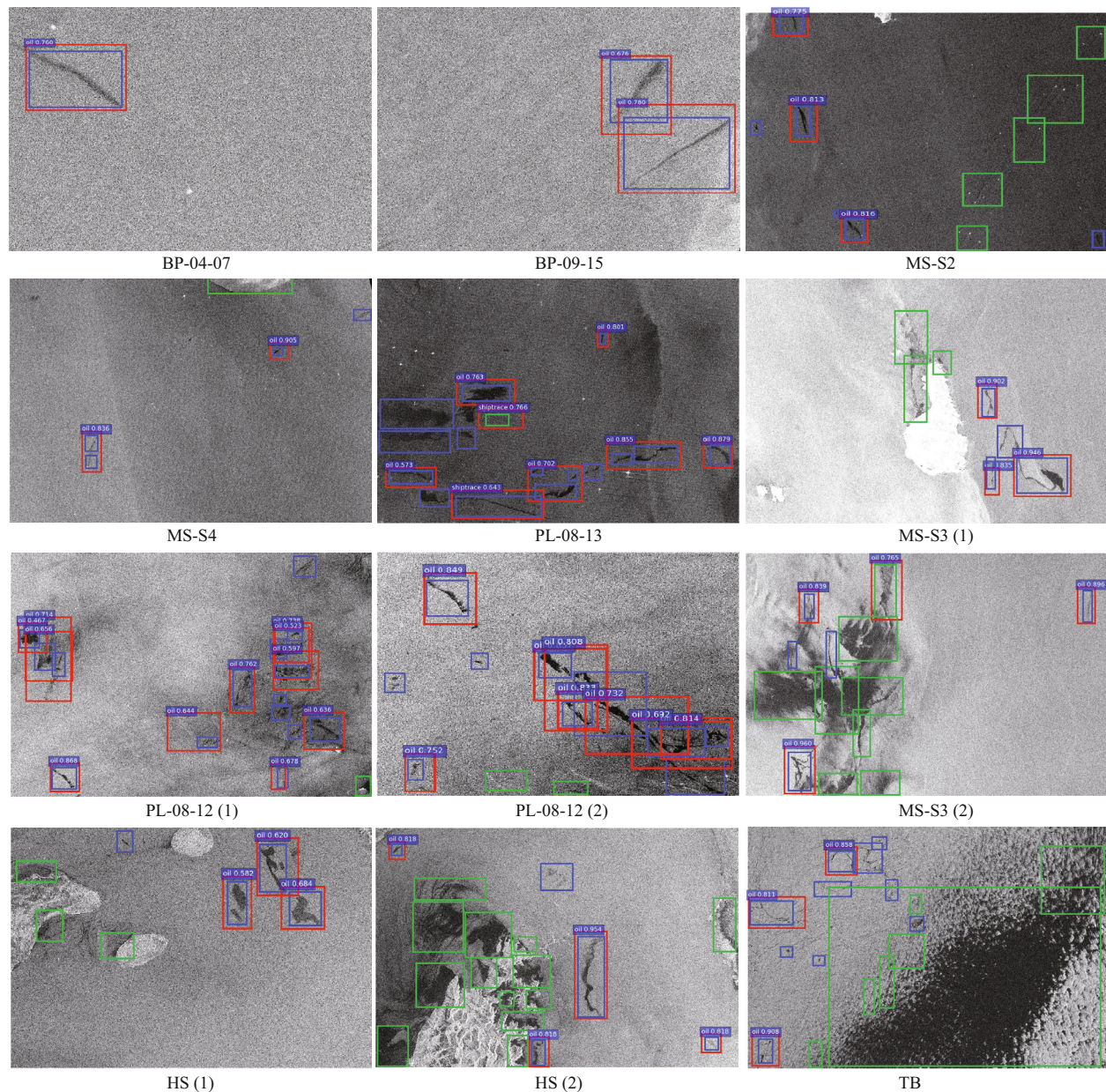
According to the principle introduced in Section 2.3.3, the discernment frame \mathcal{A} consists of three elements: A_1 indicates that the target can be classified as oil spill; A_2 indicates that the target can be classified as look-alike; and A_3 indicates that the target cannot be classified as oil spill or look-alike. The detection results of the above two oil spill detection models can be regarded as two independent pieces of evidence in Dempster-Shafer evidence theory (Dempster, 1967; Shafer, 1976). The IR for the two models were 0.70 and 0.78 (average of three test sets), respectively. Thus, the mass function for the two pieces of evidence can be defined as Eqs.10 & 11 as follows (Dempster, 1967; Shafer, 1976):

$$m_{\text{subBa}} \{A_1, A_2, A_3\} = (m_{\text{subBa}}^{\text{oil spill}} = P_{\text{subBa}}^{\text{oil spill}} \times 0.70, m_{\text{subBa}}^{\text{look-alike}} = P_{\text{subBa}}^{\text{look-alike}} \times 0.70, m_{\text{subBa}}^{\text{unknown}} = 0.30), \quad (10)$$

$$m_{\text{FstRCN}} \{A_1, A_2, A_3\} = (m_{\text{FstRCN}}^{\text{oil spill}} = P_{\text{FstRCN}}^{\text{oil spill}} \times 0.78, m_{\text{FstRCN}}^{\text{look-alike}} = P_{\text{FstRCN}}^{\text{look-alike}} \times 0.78, m_{\text{FstRCN}}^{\text{unknown}} = 0.22), \quad (11)$$

where $P_{\text{subBa}}^{\text{oil spill}}$ and $P_{\text{subBa}}^{\text{look-alike}}$ are the probability of oil spill and look-alike, respectively, calculated using the subjective Bayesian oil spill detection model. The $P_{\text{FstRCN}}^{\text{oil spill}}$ and $P_{\text{FstRCN}}^{\text{look-alike}}$ are the probability of oil spill and look-alike, respectively, provided by the Faster-RCNN oil spill detection model. The two detection results of both models were fused based on the Dempster fusion rules, with one assumption that the probabilities of oil spill and look-alike would be 1 (Dempster, 1967; Shafer, 1976). The detection rule of the proposed algorithm was based on the comparison of probabilities of oil spill and look-alike and the larger one is regarded as the final result. Three test sets were created in the same way as the test sets in Section 3.3. The environmental information records corresponding to the image samples in the test sets were selected from the EID. An evaluation for the proposed algorithm was completed and the result is shown in Fig.9a.

The evaluation result shows that the proposed algorithm has good ability to classify oil spill and look-alike with DR and IR values greater than 75% and lower FAR values ($\text{FAR} < 19\%$). Compared with the subjective Bayesian oil spill detection model (Fig.9b), the proposed algorithm shows an improved capacity to distinguish between oil spill and look-alike with increases of 13% and 11% in DR and IR,



Detected oil spills are labeled with red rectangle, the interpretation results of oil spills and look-alikes are labeled with blue and green rectangles respectively.

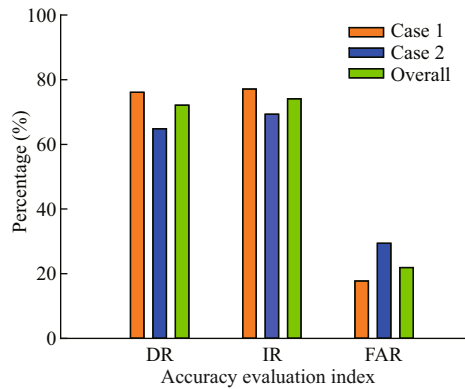
3.5 Application cases validation

There are few open-source SAR images with simultaneous instrumental observations related to oil spill accidents and they are difficult to obtain. To complete the validation, the oil spills and the look-alikes in the selected images were labeled using the visual interpretation method in Section 2.2. In Fig.10, the detection result is marked with red rectangles and the interpreted oil spills and look-alikes are marked

To evaluate the performance of the proposed algorithm, a total of 12 oil spill images were selected for the oil spill detection experiments. SAR images

Table 3 The validation SAR images in this study

Date	Sensor	Polarization mode	Code	Resolution	Incidence angle	Accident-prone area
12-Aug-2011	ENVISAT ASAR	VV	PL-08-12	30 m	26°–39°	Penglai 19-3 oil spill accident (Guo et al., 2013)
13-Aug-2011		VV	PL-08-13			
28-Aug-2011		VV	PL-08-28			
07-Apr-2010	RADARSAT2	HH	BP-04-07	25 m×8 m	20°–41°	BP Deepwater Horizon accident (Leifer et al., 2012)
15-Sep-2010		HH	BP-09-15			
29-Sep-2010		HH	BP-09-29			
11-Apr-2013		VV	HS			
11:16:43-21-Aug-2017	Sentinel-1	VV	MS-S1	20 m×22 m	29°–46°	Malacca Strait (Vaezzadeh et al., 2017)
11:17:08-21-Aug-2017		VV	MS-S2			
11:17:33-21-Aug-2017		VV	MS-S3			
11:17:58-21-Aug-2017		VV	MS-S4			
27-Dec-2002	ERS-2	VV	TB	30 m	23°	Gulf of Thailand (Gullaya, 2012)

**Fig.11 The evaluation result of case 1, case 2, and overall**

with blue and green rectangles, respectively. The complex degree of SAR image background is classified by the texture of SAR images and the number and area of oil spills and look-alikes (Karathanassi et al., 2006). The first two rows show the detection result when the SAR image background is slightly complicated. Twenty-three out of 31 oil spill dark spots were detected and the detection rate for oil spills was 74.19%. The third row and forth row show the detection result when the SAR image background is relatively complicated. The total number of labeled oil spills was 49 and the detected oil spills was 31. The detection rate for oil spill was 63.26%. Most oil spills in the first two rows were detected correctly although a few oil spills with small areas were not detected. The detection result shows that the irregular line-like oil spills PL-08-13 in Fig.10 failed to be detected by the proposed algorithm. In addition, a look-alike dark spot MS-S3(2) in Fig.10 was mistakenly classified as oil spill. The oil spill detection result is also affected by the difference of NRCS between oil and seawater. The difference of

Table 4 The evaluation result of the proposed method in different cases

	Oil spill	Undetected oil spill
Case 1: oil spill detection at relatively slightly complicated sea states		
Labeled oil spill	36	11
Labeled look-alike	8	33
Case 2: oil spill detection at relatively complicated seawater states		
Labeled oil spill	38	19
Labeled look-alike	15	43

NRCS between oil spill dark spots and seawater in the 12 validation SAR images were calculated and the minimum threshold (the difference of NRCS) for the oil spill dark spots detected by the proposed method is 4.9 dB. The minimum threshold appears in the lower right corner red rectangle of HS (2) in Fig.10. If the difference of NRCS between oil spill dark spots and seawater is lower than the minimum threshold value, the oil spill dark spots will not be detected by the proposed algorithm. The detection result indicates that the proposed method shows good ability to detect oil spills when the SAR image background is not complicated. Meanwhile, the ability and accuracy for detecting oil spills with complicated SAR image background needs to be improved.

Two cases were considered in this paper according to reference Karathanassi et al. (2006), Solberg et al. (2007) and the above analysis in this section. The undetected oil spills were regarded as misclassified labeled oil spills. The look-alike unclassified as oil spills were regarded as correctly classified look-alikes. The evaluation result of oil spill detection at different sea states is shown in Table 4. The DR, FAR, and IR for case 1 and case 2 are shown in Fig.11.

Compared with the evaluation results of case 1, the DR and IR of case 2 were reduced by about 10% and 8% and the FAR of case 2 was increased by about 10%. When comprehensively considering case 1 and case 2, the number of labeled oil spills, detected oil spills, undetected oil spills, and look-alikes misclassified as oil spills are 104, 74, 30, and 23 respectively. The overall DR, FAR, and IR are calculated by the formula proposed in section 3.1 and their values are 71.15%, 23.71%, and 73.89% respectively, as shown in Fig.11 in Green.

4 CONCLUSION

To distinguish oil spills and look-alikes more accurately with environmental factors and image features in a more objective way, a new oil spill detection algorithm based on the Dempster-Shafer evidence theory is proposed to give the optimal detection result of oil spills. The proposed algorithm was applied to the test sets to conduct detection experiments. The evaluation results indicate that the proposed algorithm has good ability to detect oil spills in SAR images showing DR and IR values greater than 75% and lower FAR values ($FAR < 19\%$). In addition, a total of 12 SAR oil spill images were selected to validate the effectiveness of the proposed algorithm. The evaluation result shows that the proposed algorithm can detect oil spills effectively with overall DR greater than 70% and FAR lower than 25%. The subjective Bayesian model is built with environmental information of image samples extracted from multi-source remote sensing data along China's coastline. This means that the detection results of the model depend partly on the geographical position of samples. This requires us to expand the sample database, add environmental factors and pixel factors of true or suspicious oil spills in different sea areas of the world, to increase the model adaptability. The Dempster-Shafer method and the Faster-RCNN model in the proposed algorithm can also be improved with appropriate adjustment to achieve a higher detection accuracy.

In future work, the ability of oil spill detection shall be enhanced for more complex environments, and more comprehensive and detailed improvement works including the computational analysis for our algorithm will be put forward based on more samples, more features such as the polarimetric features, and more accurate waterway and oil offshore data. In addition, the proposed method should be compared with the previous oil spill detection methods built

with machine learning methods (such as ANNs and SVMs), deep-learning methods, or information fusion methods, to evaluate the proposed method in a more objective way.

5 DATA AVAILABILITY STATEMENT

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

6 ACKNOWLEDGMENT

We thank the European Space Agency (ESA) for sharing and providing the Sentinel-1 data, the NASA's Global Precipitation Measurement Project (GPM) for providing the TRMM and GPM data, the ECMWF (European Centre for Medium-Range Weather Forecasts) for the ERA-interim re-analysis data, and ENVISAT-ASAR, and RADARSAT-2 for agreements or business data. The interpretation data are from the authors.

References

- Bern T I, Wahl T, Anderssen T, Olsen R. 1993. Oil spill detection using satellite based SAR: experience from a field experiment. *Photogrammetric Engineering and Remote Sensing*, **59**(3): 423-428.
- Bodla N, Singh B, Chellappa R, Davis L S. 2017. Soft-NMS—improving object detection with one line of code. *In: Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy. p.5 562-5 570.
- China Cartographic Publishing House. 2015. The World Port Traffic Atlas (2015 Edition). China Cartographic Publishing House. Beijing, China. (in Chinese)
- Commander Department of the Navy (Navigation Guarantee Department). 2005a. Guide to Chinese Port: Bohai Sea and Yellow Sea. Chinese Navigation Publications Press, Tianjin, China. (in Chinese)
- Commander Department of the Navy (Navigation Guarantee Department). 2005b. Guide to Chinese Ports: South China Sea. Chinese Navigation Publications Press, Tianjin, China. (in Chinese)
- Dempster A P. 1967. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, **38**(2): 325-339, <https://doi.org/10.1214/aoms/1177698950>.
- Duda R O, Hart P E, Nilsson N J. 1976. Subjective Bayesian methods for rule-based inference systems. *In: Proceedings of National Computer Conference and Exposition*. ACM, New York. p.1 075-1 082.
- Gullaya W. 2012. Petroleum pollution in the Gulf of Thailand: a historical review. *Coastal Marine Science*, **35**(1): 234-245.
- Guo H, Wei G, An J B. 2018. Dark spot detection in SAR

- images of oil spill using Segnet. *Applied Sciences*, **8**(12): 2 670, <https://doi.org/10.3390/app8122670>.
- Guo J, Liu X, Xie Q. 2013. Characteristics of the Bohai Sea oil spill and its impact on the Bohai Sea ecosystem. *Chinese Science Bulletin*, **58**(19): 2 276-2 281, <https://doi.org/10.1007/s11434-012-5355-0>.
- Han C, Gao G Y, Zhang Y. 2019. Real-time small traffic sign detection with revised Faster-RCNN. *Multimedia Tools and Applications*, **78**(10): 13 263-13 278, <https://doi.org/10.1007/s11042-018-6428-0>.
- Huang H S, Deng J Z, Lan Y B, Yang A Q, Deng X L, Zhang L. 2018. A fully convolutional network for weed mapping of unmanned aerial vehicle (UAV) imagery. *PLoS One*, **13**(4): e0196302, <https://doi.org/10.1371/journal.pone.0196302>.
- Karathanassi V, Topouzelis K, Pavlakis P, Rokos D. 2006. An object-oriented methodology to detect oil spills. *International Journal of Remote Sensing*, **27**(23), 5 235-5 251, <https://doi.org/10.1080/01431160600693575>.
- Krestenitis M, Orfanidis G, Ioannidis K, Avgerinakis K, Vrochidis S, Kompatsiaris L. 2019. Oil spill identification from satellite images using deep neural networks. *Remote Sensing*, **11**(15): 1 762, <https://doi.org/10.3390/rs11151762>.
- LabelImg. 2018. Available online: <https://github.com/tzutalin/labelImg> (accessed on 18 April 2018).
- Lee J S, Grunes M R, de Grandi G. 1999. Polarimetric SAR speckle filtering and its implication for classification. *IEEE Transactions on Geoscience and Remote Sensing*, **37**(5): 2 363-2 373, <https://doi.org/10.1109/36.789635>.
- Leifer I, Lehr W J, Simecek-Beatty D, Bradley E, Clark R, Dennison P, Hu Y X, Matheson S, Jones C E, Holt B, Reif M, Roberts D A, Svejksky J, Swayze G, Wozencraft J. 2012. State of the art satellite and airborne marine oil spill remote sensing: application to the BP *Deepwater Horizon* oil spill. *Remote Sensing of Environment*, **124**: 185-209, <https://doi.org/10.1016/j.rse.2012.03.024>.
- Li M M, Stein A, Bijker W, Zhan Q M. 2016. Urban land use extraction from Very High Resolution remote sensing imagery using a Bayesian network. *ISPRS Journal of Photogrammetry and Remote Sensing*, **122**: 192-205, <https://doi.org/10.1016/j.isprsjprs.2016.10.007>.
- Li X F, Li C Y, Yang Z Z, Pichel W. 2013. SAR imaging of ocean surface oil seep trajectories induced by near inertial oscillation. *Remote Sensing of Environment*, **130**: 182-187, <https://doi.org/10.1016/j.rse.2012.11.019>.
- Manana M, Tu C L, Owolawi P A. 2018. Preprocessed Faster RCNN for Vehicle Detection. 2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC). p.1-4, <https://doi.org/10.1109/ICONIC.2018.8601243>.
- Ren S Q, He K M, Girshick R, Sun J. 2017. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**(6): 1 137-1 149, <https://doi.org/10.1109/TPAMI.2016.2577031>.
- Shafer G. 1976. A Mathematical Theory of Evidence. Princeton University Press, NJ, USA. 314p.
- Simonyan K, Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition. In: Proceedings of the International Conference on Learning Representations (ICLR). San Diego, CA, USA.
- Solberg A H S, Brekke C, Husoy P O. 2007. Oil spill detection in Radarsat and Envisat SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, **45**(3): 746-755, <https://doi.org/10.1109/TGRS.2006.887019>.
- Solberg A H S, Dokken S T, Solberg R. 2003. Automatic detection of oil spills in ENVISAT, RADARSAT and ERS SAR images. In: Proceedings of the IEEE IGARSS. Toulouse, France. p.2 747-2 749.
- Tong S W, Liu X G, Chen Q H, Zhang Z J, Xie G Q. 2019. Multi-feature based ocean oil spill detection for polarimetric SAR data using random forest and the self-similarity parameter. *Remote Sensing*, **11**(4): 451, <https://doi.org/10.3390/rs11040451>.
- Vaezzadeh V, Zakaria M P, Bong C W. 2017. Aliphatic hydrocarbons and triterpane biomarkers in mangrove oyster (*Crassostrea belcheri*) from the west coast of Peninsular Malaysia. *Marine Pollution Bulletin*, **124**(1): 33-42, <https://doi.org/10.1016/j.marpolbul.2017.07.008>.
- Yang F B, Wei H, Feng P P. 2020. A hierarchical Dempster-Shafer evidence combination framework for urban area land cover classification. *Measurement*, **151**: 105916, <https://doi.org/10.1016/j.measurement.2018.09.058>.
- Zeng H, Yang B, Wang X Q, Liu J W, Fu D M. 2019. RGB-D object recognition using multi-modal deep neural network and DS evidence theory. *Sensors*, **19**(3): 529, <https://doi.org/10.3390/s19030529>.
- Zhang B, Perrie W, Li X, Pichel W G. 2011. Mapping sea surface oil slicks using RADARSAT-2 quad-polarization SAR image. *Geophysical Research Letters*, **38**(10): L10602, <https://doi.org/10.1029/2011GL047013>.
- Zhao J, Temimi M, Al Azhar M, Ghedira H. 2015. Satellite-based tracking of oil pollution in the Arabian Gulf and the Sea of Oman. *Canadian Journal of Remote Sensing*, **41**(2): 113-125, <https://doi.org/10.1080/07038992.2015.1042543>.