*Acta Oceanol. Sin.*, 2022, Vol. 41, No. 7, P. 166–179
https://doi.org/10.1007/s13131-021-1977-x
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

# Detection of oil spill based on CBF-CNN using HY-1C CZI multispectral images

Kai Du[1, 2], Yi Ma[2, 6, 7*], Zongchen Jiang[3], Xiaoqing Lu[4], Junfang Yang[5]

[1] College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China

[2] First Institute of Oceanology, Ministry of Natural Resources, Qingdao 266061, China

[3] School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China

[4] National Satellite Ocean Application Service, Beijing 100081, China

[5] College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao 266580, China

[6] Technology Innovation Center for Ocean Telemetry, Ministry of Natural Resources, Qingdao 266061, China

[7] National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, Xi'an 710072, China

Received 4 March 2021; accepted 12 November 2021



**Abstract**

Accurate detection of an oil spill is of great significance for rapid response to oil spill accidents. Multispectral images have the advantages of high spatial resolution, short revisit period, and wide imaging width, which is suitable for large-scale oil spill monitoring. However, in wide remote sensing images, the number of oil spill samples is generally far less than that of seawater samples. Moreover, the sea surface state tends to be heterogeneous over a large area, which makes the identification of oil spills more difficult because of various sea conditions and sunglint. To address this problem, we used the F-Score as a measure of the distance between forecast value and true value, proposed the Class-Balanced F loss function (CBF loss function) that comprehensively considers the precision and recall, and rebalances the loss according to the actual sample numbers of various classes. Using the CBF loss function, we constructed convolution neural networks (CBF-CNN) for oil spill detection. Based on the image acquired by the Coastal Zone Imager (CZI) of the Haiyang-1C (HY-1C) satellite in the Andaman Sea (study area 1), we carried out parameter adjustment experiments. In contrast to experiments of different loss functions, the F1-Score of the detection result of oil emulsions is 0.87, which is 0.03–0.07 higher than cross-entropy, hinge, and focal loss functions, and the F1-Score of the detection result of oil slicks is 0.94, which is 0.01–0.09 higher than those three loss functions. In comparison with the experiment of different methods, the F1-Score of CBF-CNN for the detection result of oil emulsions is 0.05–0.12 higher than that of the deep neural networks, supports vector machine and random forests models, and the F1-Score of the detection result of oil slicks is 0.15–0.22 higher than that of the three methods. To verify the applicability of the CBF-CNN model in different observation scenes, we used the image obtained by HY-1C CZI in the Karimata Strait to carry out experiments, which include two studies areas (study area 2 and study area 3). The experimental results show that the F1-Score of CBF-CNN for the detection result of oil emulsions is 0.88, which is 0.16–0.24 higher than that of other methods, and the F1-Score of the detection result of oil slicks is 0.96–0.97, which is 0.06–0.23 higher than that of other methods. Based on all the above experiments, we come to the conclusions that the CBF loss function can restrain the influence of oil spill and seawater sample imbalance on oil spill detection of CNN model thus improving the detection accuracy of oil spills, and our CBF-CNN model is suitable for the detection of oil spills in an area with weak sunglint and can be applied to different scenarios of CZI images.

**Key words:** oil spills, CNN, classification, loss function, sunglint, detection

**Citation:** Du Kai, Ma Yi, Jiang Zongchen, Lu Xiaoqing, Yang Junfang. 2022. Detection of oil spill based on CBF-CNN using HY-1C CZI multispectral images. Acta Oceanologica Sinica, 41(7): 166–179, doi: 10.1007/s13131-021-1977-x

## 1 Introduction

Oil spills pollute the ocean and seriously damage marine ecology (Serra-Sogas et al., 2008; Lu et al., 2015; Esbaugh et al., 2016). On April 20, 2010, the Deepwater Horizon oil drilling rig exploded into flames, spilling nearly 400 000 t oil, contaminating approximately 1 773 km coastline, and causing the death of thousands of species (Abbriano et al., 2011; Michel et al., 2013). The Panamanian tanker *Sanchi* collided with the Chinese bulk carrier *CF Crystal* in the East China Sea on January 6, 2018, which caused the release of approximately 113 000 t condensate oil (Sun et al., 2018; Yin et al., 2018). On July 25, 2020, the *Wakashio*, which was sailing from China to Brazil, ran aground off Mauritius, approximately 3.2 km from the coastline. Ten days later, one of the ship's oil tanks burst, spilling thousands of tons of fuel into the sea. Although the oil spill was small, it is the worst ecological disaster in the history of Mauritius because of its fragile ecologic-

Foundation item: The National Natural Science Foundation of China under contract No. 61890964; the Joint Funds of the National Natural Science Foundation of China under contract No. U1906217.
*Corresponding author, E-mail: mayimail@fio.org.cn.

al environment. Therefore, the classification of an oil spill is very important.

Optical sensors can cover a wide area, provide rich spectral information, and have a high temporal resolution. They can distinguish oil spills from suspected targets using optical information characteristics. Because of this, optical sensors have been extensively studied in oil spill classification. Adamo et al. (2009) theoretically described the physical mechanism of remote sensing oil spill monitoring in the visible/near-infrared band. Lu et al. (2013, 2020) pointed out that such a false-color image, where short-wave infrared, near-infrared, and red bands are used as the red, green, and blue channels can qualitatively distinguish oil-in-water and water-in-oil. Hu et al. (2009) used Moderate Resolution Imaging Spectroradiometer data to detect a natural oil spill in the Gulf of Mexico for the first time and pointed out the critical reflectivity and critical angle of the oil spill from dark to bright. Kolokoussis and Karathanassi (2018) developed two object-oriented image analysis methods and applied them to Sentinel 2 multispectral image data to detect oil spills successfully. Corucci et al. (2010) developed several different machine learning techniques, such as simple statistical classifiers, radial basis function and multilayer perceptron neural networks, adaptive neuro-fuzzy reasoning system, and a support vector machine (SVM) to investigate the potential of oil spill classification from optical satellite images. Shen et al. (2020) conducted oil spill classification and comparative analysis using oil spill data from Haiyang-1C (HY-1C) Chinese Ocean Color and Temperature Scanner and Coastal Zone Imager (CZI) sensors, and the results showed that the CZI could not only identify oil spills but also distinguish oil emulsions and non-emulsified oil slicks. The above researches show the potential of optical sensors in oil spill classification; however, the existing methods are mostly traditional and the classification accuracies are still low.

Deep learning methods can extract the inherent regularity and characteristics of an image. In recent years, with the increasing abundance of remote sensing data, the application of deep learning methods in the research of oil spills has improved. Yang et al. (2019) developed a deep convolutional neural network model, which improved the oil spill classification accuracy by approximately 5% when compared with the SVM, Random Forests (RF), and deep belief networks (DBN) methods. Yekeen et al. (2020) developed a deep learning oil spill classification model using the Mask R-CNN. Zhu et al. (2019) implemented an oil spill information extraction method combining spatial information with the stacked autoencoder network model (SAE), thus improving the classification accuracy of traditional SAE neural networks. Moreover, Jiang et al. (2020) conducted a quantitative inversion of the absolute thickness of oil spills using deep learning methods.

Multispectral images have the advantages of high spatial resolution, short revisit period, and wide imaging width, which is suitable for large-scale oil spill monitoring (Niclòs et al., 2013; Chen et al., 2018). However, in wide remote sensing images, the number of oil spill samples and seawater samples are usually imbalanced, that is, the number of oil spill samples is far less than the number of seawater samples. Previous studies have shown that an imbalanced dataset is one of the main reasons for a reduction in the generalization ability of machine learning algorithms (Lin et al., 2017). Moreover, the sea surface state tends to be heterogeneous over a large area, which makes the identification of oil spills more difficult because of various sea conditions and sunglint. To address this problem, we use F-Score as a distance measure between forecast value and true value and propose the Class-Balanced F loss function (CBF loss function), which not only comprehensively considers the precision and recall but also rebalances the loss according to the actual sample numbers of various classes. Using the CBF loss function, we construct a convolution neural network oil spill detection model (CBF-CNN). The experiments were carried out on two images obtained by the HY-1C satellite CZI sensor in the Andaman Sea and the Karimata Strait. It is proved by experiments that the CBF loss function can effectively restrain the influence of oil spill and seawater sample quantity imbalance thus improving the detection accuracy of oil spills, and our CBF-CNN model is suitable for the detection of oil slicks and oil emulsions in the negative contrast area with weak sunglint and can be applied to different observation scenarios of CZI images.

## 2 Materials and methods

### 2.1 *HY-1C CZI remote sensing data*

HY-1C is a sun-synchronous satellite orbiting at an altitude of 782 km. It carries a four-band wide-band satellite CZI, which has a 50 m spatial resolution and a 950 km swath width. The CZI is mainly used for ocean water temperature and coastal zone monitoring. The revisit period is three days and the overpass time is approximately 10:30 local time (Tong et al., 2019). Additional technical specifications of the CZI are shown in Table 1. The CZI has the imaging characteristics of large width, medium resolution, rapid coverage, good spectral response to ocean targets, and high signal-to-noise ratio, and it can meet remote sensing monitoring requirements of marine disasters. It provides services and meets the requirements for marine resource management, marine ecological restoration, land and sea disaster prevention.

HY-1C CZI detected oil spills in two images obtained on April 20 and 21, 2020, in the Karimata Strait and the Andaman Sea. Based on these two CZI images of Rayleigh-corrected reflectance ($R_{rc}$, dimensionless) (Hu, 2009), we carried out oil spill detection experiments.

Atmospheric correction methods that are designed for either ocean color or land applications often result in low-quality or even no surface reflectance data for coastal and inland waters. In contrast, $R_{rc}$ has been used in water applications, although without removing aerosol scattering (Feng et al., 2018). This partial atmospheric correction can be expressed as

$$R_{rc} = \pi L_t^* / (F_0 \times \cos\theta_0) - R_r, \tag{1}$$

where $F_0$ is the extraterrestrial solar irradiance; $\theta_0$ is the solar zenith angle; $L_t^*$ is the radiometrically calibrated at-sensor radiance after removing the gaseous absorption; and $R_r$ is the reflectance due to Rayleigh scattering.

#### 2.1.1 *Oil spill in the Andaman Sea*

As shown in Fig. 1a, the HY-1C CZI detected an oil spill incident in the Andaman Sea on April 21, 2020. The study area (study area 1) is shown in Fig. 1b. The size of the study area is 27.5 km×

**Table 1.** Main technical specifications of the Haiyang-1C Coastal Zone Imager

| Band/nm | Central wavelength/nm | Spatial resolution/m | Signal-noise radio/dB |
|---|---|---|---|
| 420–500 | 460 | 50 | 410 |
| 520–600 | 560 | 50 | 300 |
| 610–690 | 650 | 50 | 248 |
| 760–890 | 825 | 50 | 240 |

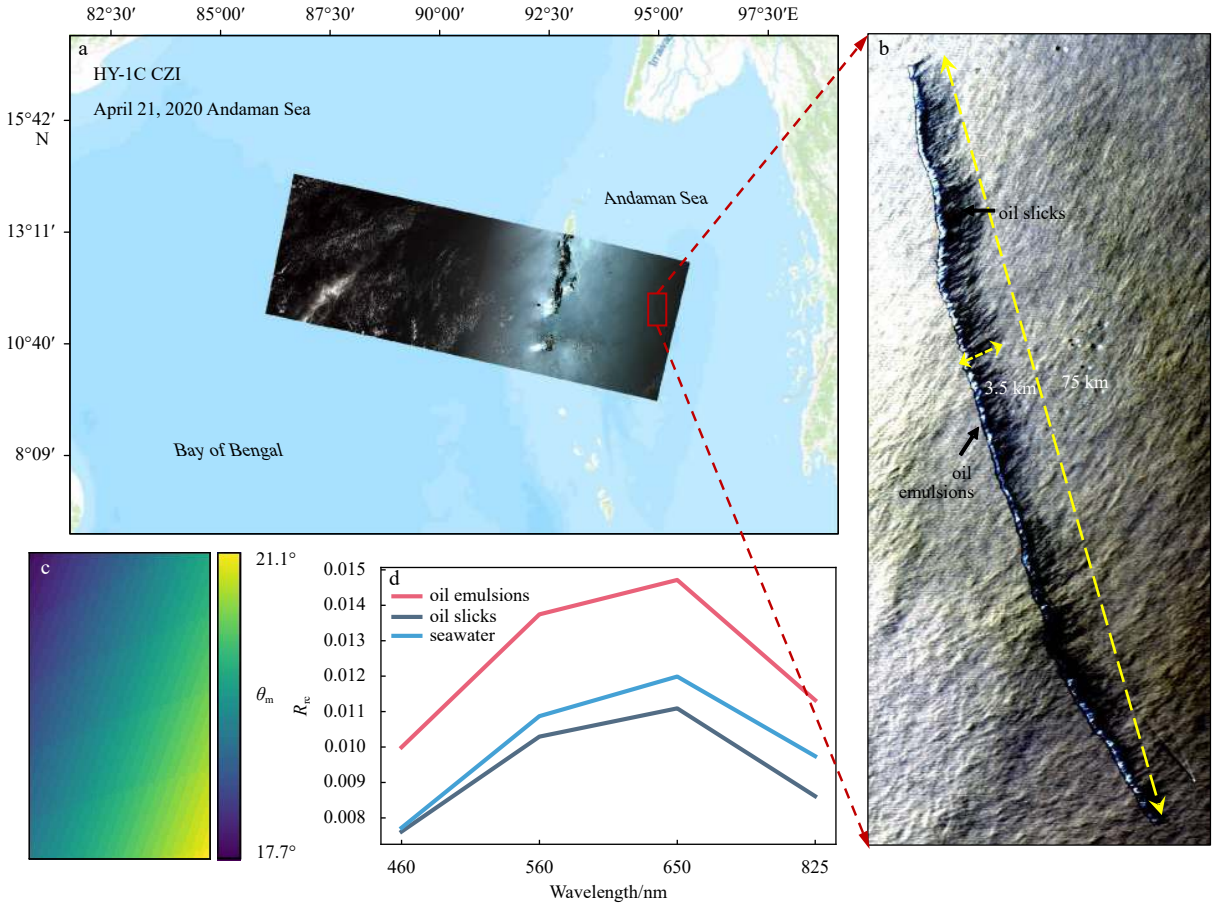


**Fig. 1.** Study area 1. a. Geographical location of study area 1; b. the true color RGB image of study area 1; c. the angle between the viewing direction and the direction of mirror reflection ($\theta_m$) of study area 1; d. the spectra of oil emulsions, oil slicks, and seawater.

70 km. The difference in the surface sunglint reflectance is an important factor to change the reflectance contrast between oil spills and seawater. For the identification of marine oil spill pollution, the first thing is to evaluate the difference between the sunglint reflectance on the surface of an oil spill and oil-free seawater (Wen et al., 2018). We calculated the angle between the viewing direction and the direction of mirror reflection ($\theta_m$) in the three study areas, which directly indicates the intensity of the surface sunglint reflectance in satellite optical images, which can be expressed as Eq. (1) (Hu et al., 2009; Lu et al., 2013; Wen et al., 2018).

$$\cos\theta_m = \cos\theta_0 \cdot \cos\theta - \sin\theta_0 \cdot \sin\theta \cdot \cos\phi, \quad (2)$$

where $\theta_0$ is the solar zenith angle; $\theta$ is the sensor zenith angle; and $\phi$ is the relative azimuth angle.

From $\theta_m$, one determines the relative intensity of the sunglint reflectance. The smaller the value of $\theta_m$, the stronger the sunglint. As shown in Fig. 1c, the $\theta_m$ of study area 1 is greater than 17°. Some researchers have pointed out that when $\theta_m \geqslant 13°$, the reflectance of sunglint of an oil spill and oil-free seawater is negative contrast, which means that the sunglint reflectance of oil slicks is less than that of oil-free seawater (Lu et al., 2016; Wen et al., 2018; Hu et al., 2021). Under these circumstances, the light absorption of oil slicks is stronger, so the reflectance is lower than that of seawater; the light scattering characteristic of oil emulsions is stronger, so the reflectance is higher than that of seawater.

We drew the $R_{rc}$ spectra by sampling the oil emulsions, oil slicks, and seawater in study area 1. Through the spectra analysis, we found that the reflectance of oil emulsions is the highest, followed by seawater, and the reflectance of oil slicks is the lowest. At the same time, the spectral similarity between oil slicks and seawater is strong, which makes it difficult to distinguish them effectively.

Study area 1 is where we conducted model parameter adjustment, training data quantity comparison, CBF loss function parameter adjustment, and effect evaluation. To better adjust and optimize our model, we carried out a visual interpretation of study area 1 and used the method of randomly selecting training data for model training. There are a small number of interference factors for oil spill detection such as clouds and their shadows, ships, and their wakes in the study area. Because we randomly selected the training data for model training, we cannot guarantee that the pixels of these interference factors can be selected for training data to make the model learn the difference between them and oil spills. To avoid the influence of these interference factors, we excluded these pixels according to the visual interpretation results before model training. This enabled us (the authors) to focus on solving the problems of strong spectral similarity between oil slicks and seawater and the imbalance in the number of oil spills and seawater samples.

### 2.1.2 *Oil spill in the Karimata Strait*

HY-1C CZI detected an oil spill incident in the Karimata Strait on April 20, 2020. As shown in Fig. 2a, the CZI image contains two

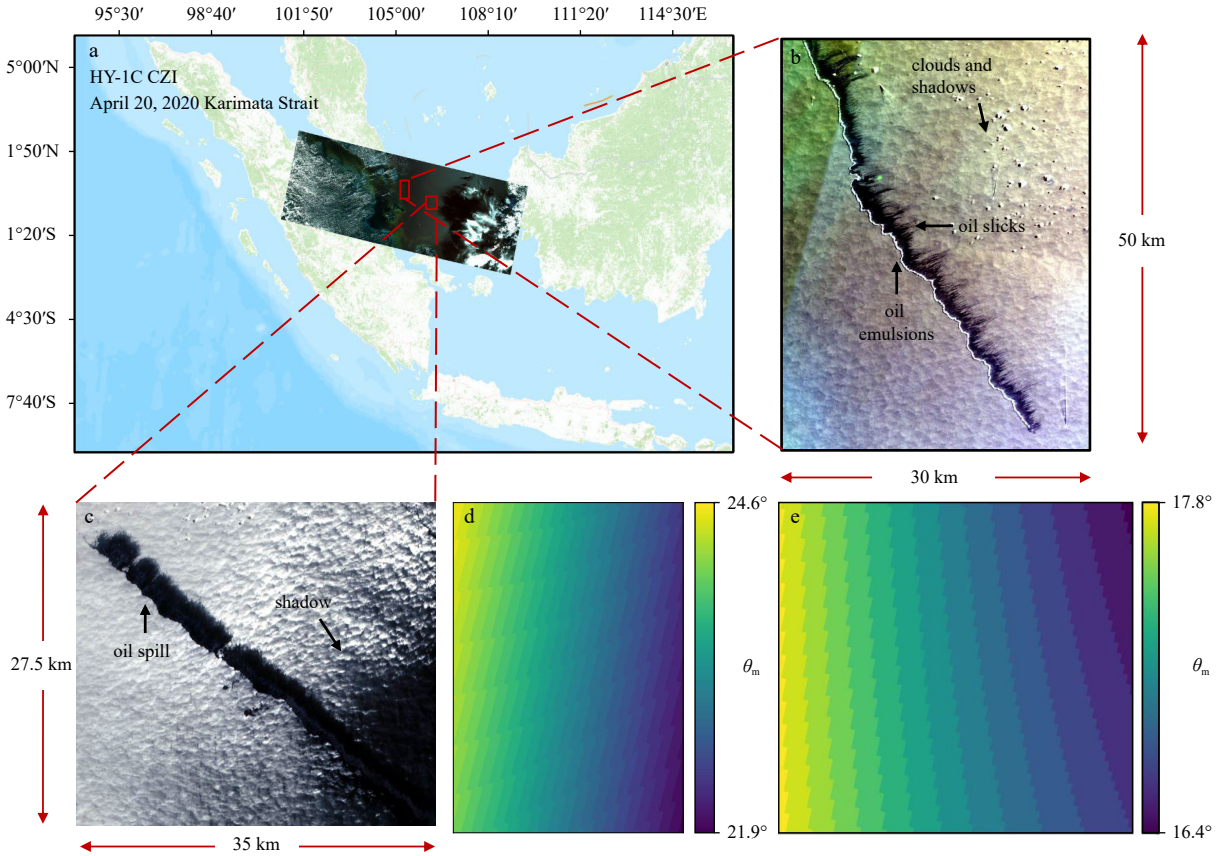


**Fig. 2.** Study area 2 and study area 3. a. Geographical location of study area 1 and study area 2; b. the true color RGB image of area 2; c. the true color RGB image of area 3; d. the $\theta_m$ of study area 2; e. the $\theta_m$ of study area 3.

oil spill areas (study area 2 and study area 3). We tested the applicability of our CBF-CNN oil spill detection model in these two areas. As shown in Figs 2d and e, the $\theta_m$ of the two areas are all greater than 13°, which means that the two study areas all belong to the negative contrast area of weak sunglint reflectance, so the reflectance of oil slicks is lower than that of seawater, and the reflectance of oil emulsions is higher than that of seawater, in the two areas.

There are a large number of clouds and their shadows in study area 2. As shown in Fig. 3a, through the analysis of their spectra, we find that the spectral similarity between oil emulsions and clouds is strong. Clouds and their shadows will seriously interfere with the accuracy of oil spill detection.

In study area 3, the oil spill does not exist in isolation but is connected to a large black background. As shown in Fig. 3b, the spectra of oil spills and black backgrounds are very similar. It is a great challenge to detect oil spills under the interference of this black background.

### 2.2 *Data augmentation*

A CNN model usually needs a large amount of data for training, but in the actual oil spill detection task, little training data are available. In the field of deep learning, there are a large number of data augmentation methods, but because our CBF-CNN mod-

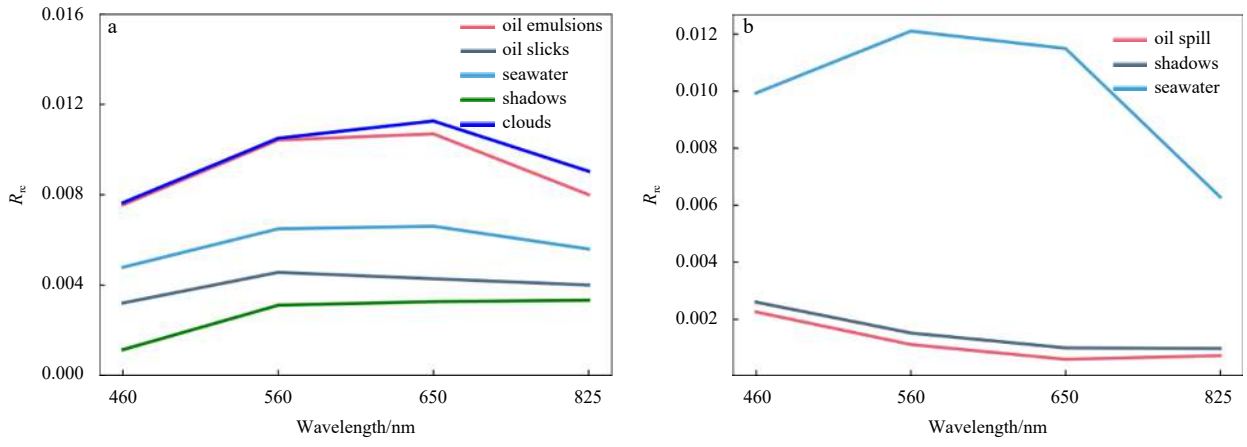


**Fig. 3.** Rayleigh-corrected reflectance ($R_{rc}$) of study area 2 (a) and study area 3 (b).

el is one of patch-based CNN models, so not all methods are suitable for our CBF-CNN model. To address this problem, we adopted a simple sample augmentation method to increase the diversity of the training data, as shown in Fig. 4. For each training sample, three data augmentation operations were performed. The first operation is flipping, which is achieved by flipping the original sample horizontally or vertically; the second operation is to add Gaussian noise to the original sample, and the third operation is to add Gaussian noise to the flipped sample. In this way, the number of training samples can be tripled so that oil spills can be classified by the CBF-CNN model even if only a small number of training samples are available.

### 2.3 *CNN model*

Research on CNNs started in the 1980s and 1990s (LeCun and Bengio, 1995). A CNN usually consists of the following five layers: an input layer, a convolutional layer, a pooling layer, a fully connected layer, and a classification layer. The input layer is used to input data into the CNN model and makes its distribution more uniform. The convolutional layer is used to extract different features from the input data. The pooling layer is used to reduce the amount of data. The fully connected layer is the same as that of the normal neural network, and the results are used as the input of the classification layer. The classification layer mainly calculates the probability that the sample belongs to a certain class.

The CNN architecture used in this study is shown in Fig. 5. In the input layer, for each pixel, we extracted the surrounding 11×11 pixels, then performed standardized preprocessing, and finally, sent the data to the other layers. After the input layer, we adopted the batch normalization (BN) layer proposed by Google in 2015 (Ioffe and Szegedy, 2015). It is a technique for deep neural network training that does not only speed up the convergence of models but also alleviates the "disappearing gradient" problem in deep networks to some extent, thus making the training of deep network models easier and more stable. The formula is as follows:

$$\widehat{Z} = \gamma \frac{Z - \mathrm{E}(Z)}{\sqrt{\mathrm{Var}(Z) + \xi}}, \tag{3}$$

where $\widehat{Z}$ is the output of the BN layer; $Z$ is the input of the BN layer; $\mathrm{E}(Z)$and $\mathrm{Var}(Z)$ respectively represent the expectation and variance of $Z$; and $\gamma$ and $\xi$ denote learning parameters.

The output of the BN layer is fed to a series of convolutional layers. Convolution is an effective method for image feature extraction. In general, the square convolution kernel is used to traverse each pixel in the image. The weight of the corresponding point in the convolution kernel is multiplied by the sum of the value of each pixel in the overlapping area of the image and convolution kernel, and the offset is then added to obtain the pixel value in the output image. The formula is as follows:

$$Z = \sum_{i}^{p \times q} w_i \times v_i + b, \tag{4}$$

where $Z$ represents the result of convolution; $p$ and $q$ represent the number of rows and columns of the convolution kernel, respectively; $w_i$ and $v_i$ are the weight and the pixel value of point $i$; and $b$ represents the bias value.

In our CNN model, the output of the BN layer is processed by two convolution kernels of the size of 3×3 pixel and 5×5 pixel, and the obtained features are combined to obtain more features. After the layer with a convolutional kernel of size 5×5 pixel, we connected a layer with a convolutional kernel of size 1×1. In this way, the number of parameters can be reduced and the efficiency of the CNN model can be improved.

After convolution, nonlinear activation functions are usually introduced for activation. In the model proposed in this paper,

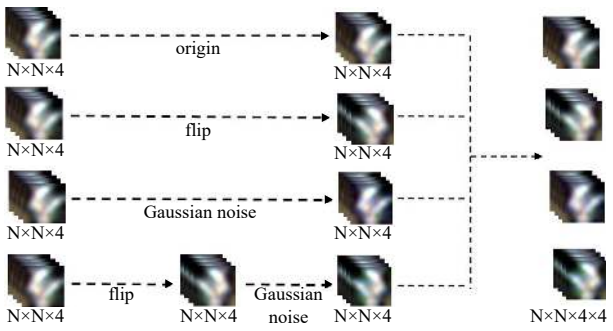


**Fig. 4.** Process of data augmentation.

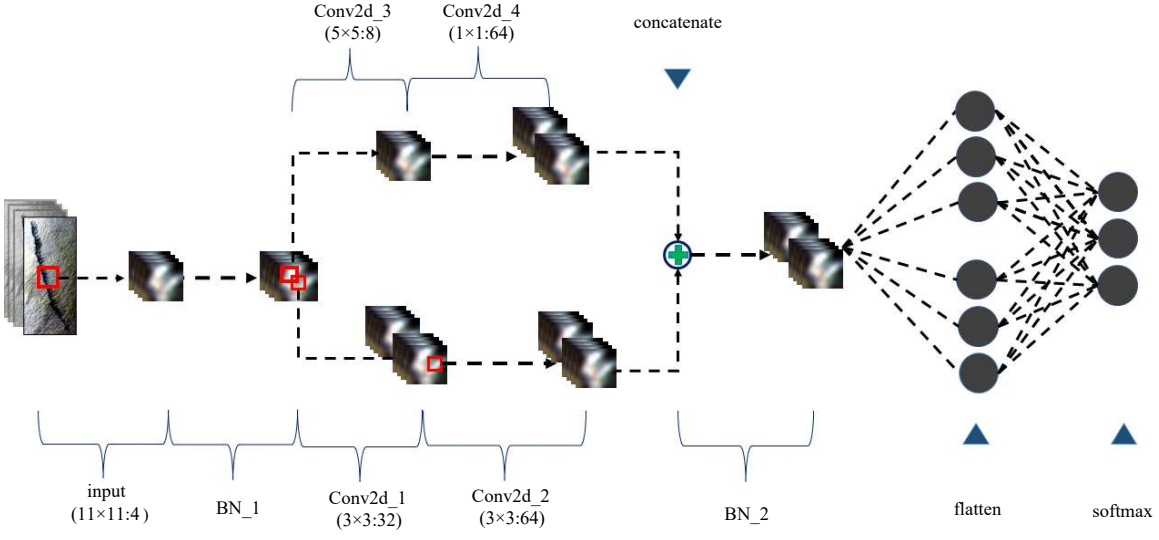


**Fig. 5.** Convolution neural networks structure.

the following Rectified linear unit (ReLU) activation function is used:

$$\tilde{Z} = \max\{0, Z\}, \tag{5}$$

where $\tilde{Z}$ represents the value after activation; and $Z$ is the value before activation.

To avoid the spatial dimension reduction of the feature map, the use of a pooling layer is not considered in this paper. Finally, the output of the convolutional layer is flattened and output through the softmax function. The formula is as follows:

$$\sigma\left(\tilde{Z}\right)_i = \frac{e^{Z_i}}{\sum_{j=1}^{n} e^{Z_j}}, \tag{6}$$

where $\tilde{Z}$ indicates the input vector to the softmax function; $Z_i$ indicates the elements of the input vector to the softmax function; $n$ indicates the number of classes in the multi-class classifier; and $\sum_{j=1}^{n} e^{Z_j}$ is the normalization term. It ensures that all the output values of the function will sum to 1 and each be in the range (0, 1), thus constituting a valid probability distribution.

The training of data was performed in 100 epochs, with a batch size of 450. Adaptive moment estimation (Adam) (Kingma and Ba, 2015) was used as the optimization algorithm with a learning rate of 0.001.

### 2.4 *The Class-Balanced F loss function*

The loss function was used to optimize the parameters of the neural network. To do this, we calculated the loss by matching the output of the neural network with the actual output and then used the gradient descent method to optimize the weight of the network to minimize the loss.

In different practical application scenarios, customized loss functions should be designed according to the scenario's requirements to ensure that the obtained neural network model is suitable for the actual situation (Jiang and Ma, 2020). To solve the problems caused by the imbalanced numbers of an oil spill and seawater pixels in wide remote sensing images and heterogeneity in marine environments, we used the F-Score as a distance measure between the predicted and real values to propose a loss function, the CBF loss function. Our proposed function comprehensively considers the accuracy and recall rates and rebalances the loss according to the actual sample numbers of various classes. The formula of the loss is as follows:

$$\text{CBF loss} = \min\left\{\sum_{i}^{m}\left[\frac{1-\beta}{1-\beta^{n_i}}\left(1-F_{\alpha i}\right)\right]\right\}, \tag{7}$$

where $F_{\alpha i}$ refers to the F-Score of class $i$; and $\frac{1-\beta}{1-\beta^{n_i}}$ is the class-balanced loss term, which is based on the weighted scheme designed by Cui et al. (2019). Parameter $\beta$ is the class equilibrium parameter; $n_i$ is the number of $i$ class; and $m$ is the number of categories.

F-Score is a statistical quantity. F-Score is the weighted and harmonic average of precision and recall. It is a common evaluation standard in the information retrieval field and is often used to evaluate the quality of classification models. Among them, precision refers to the proportion of correctly recognized pixels in the predicted pixels of this category, and recall refers to the proportion of correctly recognized pixels in the real pixels of this category. Accuracy and recall rates are equal to user accuracy and producer accuracy, respectively. The formulas for precision and recall are as follows:

$$P = \frac{\text{TP}}{\text{TP}+\text{FP}}, \tag{8}$$

$$R = \frac{\text{TP}}{\text{TP}+\text{FN}}, \tag{9}$$

here, TP represents the number of samples that have been correctly predicted to be positive; FP represents the number of samples that have been wrongly predicted to be positive; and FN represents the number of samples that have been wrongly predicted to be negative.

Then, the formula of the F-Score is

$$F_\alpha = \frac{\left(\alpha^2+1\right)PR}{\alpha^2 P + R}, \tag{10}$$

where $P$ is the precision rate; $R$ is the recall rate; and $\alpha$ is the harmonic parameter of the precision and recall rates. CBF loss can then be expressed as

$$\text{CBF loss} = \min\left\{\sum_{i}^{m}\left[\frac{1-\beta}{1-\beta^{n_i}}\left(1-\frac{\left(\alpha^2+1\right)P_i R_i}{\alpha^2 P_i + R_i}\right)\right]\right\}. \tag{11}$$

In different cases, the emphasis on precision rate or recall rate should be different. We can adjust the value of the parameter to make the F-Score meet the requirements of the classification problem. When $\alpha$>1, the recall rate has a greater effect, and when it is infinity, only the recall rate is used; when 0<$\alpha$<1, the precision has a greater effect; when $\alpha$ is infinitely close to 0, and only the precision is considered in the loss.

In an oil spill emergency response, our main objective is to accurately identify the oil spill area, so we pay more attention to the loss of the oil spill in the model training, which can be achieved by adjusting the parameter $\beta$. Here, parameter $\beta \in (0, 1)$, and when $\beta$= 0, this is the equivalent of no weight, and when $\beta$ tends to be 1, this corresponds with a valid sample of the inverse ratio of the loss function weighted to balance each contribution of the category to the loss function.

The experiment shows that when $\alpha$=1.1 and $\beta$=0.99, the CBF loss function can effectively suppress the influence of the imbalanced number of seawater and oil spill samples on the oil spill classification. The experimental analysis is detailed in Section 3.3.

### 2.5 *Comparison methods*

The comparison methods selected for the evaluation in this paper are the classical deep neural networks (DNN), SVM, and RF.

The DNN is a set of connected input-output networks in which a weight is associated with each connection. It is generally composed of three parts: the input layer, the hidden layer, and the output layer. Each layer is composed of neurons, in which there is only one input layer and one output layer, and the number of hidden layers can be arbitrary. The DNN improves performance by updating the weights of the connections through forward propagation and backpropagation.

The SVM was proposed in 1964 and has been developed rapidly since the 1990s. It is widely used owing to its effectiveness with small datasets. The fundamental concept of an SVM is to determine a separation plane for splitting a spatial neighborhood into two parts; for an $N$-dimensional spatial neighborhood, we need an ($N$–1)-dimensional separation plane to separate it. By comparing SVM models with different kernel functions, we found that the SVM model based on the radial basis function (RBF) is more suitable for the task of oil spill classification. The parameters that need to be adjusted for SVM with the RBF kernel are penalty coefficient $C$ and kernel function parameter Gamma. Penalty coefficient $C$ represents the tolerance of the model to error: a small $C$ value means the penalty for misclassification is small and the generalization ability of the model is enhanced, but the model may be under fitted. In contrast, with a large $C$ value, the penalty for misclassification is large and the accuracy of model classification is improved, but overfitting is more likely to occur. Kernel function parameter Gamma mainly determines the effect of a single sample on the overall classification hyperplane. When Gamma is small, the influence of a single sample on the entire classification hyperplane is relatively small, and the number of support vectors of the model is also small. In contrast, when Gamma is large, the influence of a single sample on the entire classification hyperplane is relatively large, and it is easier for a vector to be selected as a support vector, so the number of support vectors of the whole model will also be large.

RF classification is one of the most commonly used machine learning algorithms which is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest (Breiman, 2001). RF is an excellent ensemble learning method, because thanks to the law of large numbers, it can avoid overfitting.

## 3 Results and discussion

### 3.1 *Spatial neighborhood-scale adjustment*

The results were generated on a personal computer equipped with an AMD Ryzen-7 4800H at 2.9 GHz and Nvidia GeForce RTX 2060 graphics card. The study image is a 550×1 400 pixels image with an 8-bit depth.

To determine the optimal spatial neighborhood of input data, we followed the principle of control variables. First, we randomly selected 10% of pixels in each class as training data, then the training data were augmented, and the parameters $\alpha$ and $\beta$ of the CBF loss were also set to 1.0 and 0.9, respectively. Under such conditions, an experiment was carried out for each of the following spatial neighborhood sizes: 5×5, 7×7, 9×9, 11×11, 13×13, 15×15, and 17×17.

Figure 6 illustrates the influence of the size of the spatial neighborhood on the classification accuracy of oil emulsions and oil slicks. As the space scale increases, the classification accuracy of the oil spills is improved. For oil emulsions, when the spatial scale is 11×11, the classification result has the highest accuracy, whereas when the spatial scale is greater than 15×15, the accuracy begins to decline. This indicates that the spatial neighborhood of 11×11 can fully provide the spatial context information needed for oil emulsion identification. In contrast, when the spatial neighborhood is too large, the redundant information will interfere with the classification of oil emulsions, resulting in a decrease in the classification accuracy. For oil slicks, when the spatial neighborhood scale is between 11×11 and 15×15, the classification accuracy is high and tends to be stable, indicating that

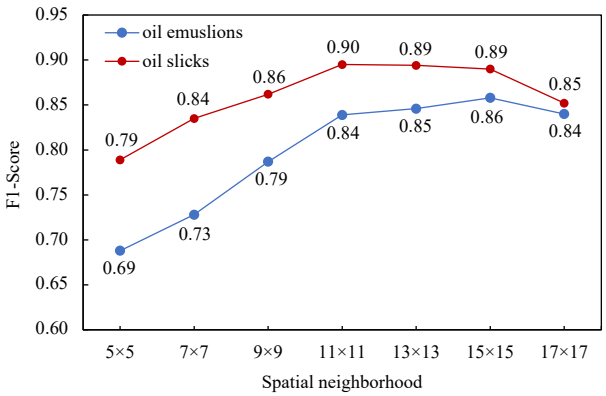


**Fig. 6.** Classification accuracy of oil spill under different sizes of spatial neighborhoods.

the spatial scales of 11×11 to 15×15 can meet the needs of oil slick classification.

Figure 7 shows the time required for model training and global classification at various spatial neighborhood sizes. As the size of the sample spatial neighborhood increases, the time needed for CNN model training and global classification increases exponentially. When the sample spatial neighborhood is 5×5, the training time is only 110 s and the global classification time is 15 s; when the sample spatial neighborhood is 11×11, the training time is 306 s and the global classification time is 26 s; when the sample spatial neighborhood is 17×17, the training time reaches 683 s and the global classification time reaches 45 s. Therefore, a larger sample spatial neighborhood may improve the classification accuracy, but the time cost will also increase exponentially.

When the sample spatial neighborhood is 11×11, the oil slicks classification accuracy is the highest, and the oil emulsions classification accuracy is also relatively good, and the time needed for model training and global classification is also within an acceptable range. Considering both oil spill classification accuracy and time cost, we chose 11×11 as the final spatial neighborhood size for the input data.

### 3.2 *Training sample expansion*

To prove the effectiveness of the training sample augmentation method, we again followed the principle of controlling variables. The spatial neighborhood of the input data was set as 11×11, and the parameters in the CBF loss were set to 1.0 and 0.9. Under these conditions, we first randomly selected 1%, 2%, 3%, 5%, 8%, 10%, 20%, 30%, and 40% of the pixels in each class as training data, and the influence of training data augmentation on

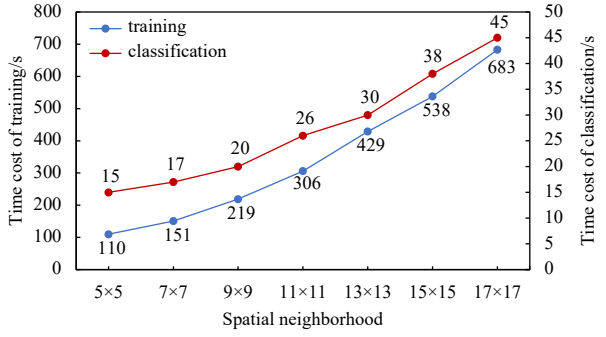


**Fig. 7.** The time cost of different spatial neighborhood sizes.

the classification accuracy of oil emulsions and oil slicks was compared under conditions with different numbers of training samples.

Figure 8 shows both the influence of training sample augmentation on the classification accuracy of oil emulsions and oil slicks under different training sample quantities. For oil emulsions, when the number of training samples is between 1% and 8%, the classification accuracy is substantially higher when sample augmentation is used than when it is not used; when the number of training samples is over 8%, the effect of sample augmentation is weakened. For black oil slicks, when the number of training samples is between 1% and 8%, the classification accuracy is substantially higher when sample augmentation is used than when it is not used; when the number of training samples is more than 8%, the benefit of sample augmentation is reduced. This analysis shows that the training sample augmentation method described can enrich the training samples to a certain extent. When the number of training samples is relatively small, the sample augmentation can effectively improve the accuracy of the model so that the CNN model can achieve better classification results with only a small number of training samples.

### 3.3 *CBF loss parameter adjustments*

We conducted extensive studies on the CBF loss with various parameters. The search space of hyperparameters was $\alpha \in \{0.5, 0.7, 0.9, 1.1, 1.3\}$ and $\beta \in \{0.9, 0.99, 0.999\}$. Figures 9a and b show that when we use different values of $\alpha$ and $\beta$, the classification results for oil emulsions and oil slicks vary.

As Fig. 9a shows, the classification accuracy of oil emulsions varies greatly for different values of $\beta$. When $\beta$=0.99, the classification accuracy of oil emulsions is the highest and most stable. When $\beta$=0.9 or $\beta$=0.999, the classification accuracy of oil emulsions is low and first increases with the increase of $\alpha$ and then decreases.

As shown in Fig. 9b, for oil slicks, the classification accuracy is generally improved as $\alpha$ increases. It can be seen that when $\beta$=0.999, the classification result for oil slicks is the worst, and when $\beta$=0.9 or $\beta$=0.99, the classification accuracy of oil slicks is substantially higher. In general, when $\alpha$=1.1, $\beta$=0.99 or $\alpha$=0.9, $\beta$=0.9, the classification result for oil slicks is the best. On the whole, the highest classification accuracy of oil emulsions and oil slicks is achieved when $\alpha$=1.1 and $\beta$=0.99.

These results demonstrate that the influence of $\alpha$ and $\beta$ on the classification accuracy of oil spills is obvious and predictable. Setting appropriate values of $\alpha$ and $\beta$ helps eliminate the imbalance between the number of oil spill samples and seawater samples in large-scale oil spill images as well as the effects of heterogeneity in a marine environment. Considering the classification accuracy of oil emulsions and oil slicks together, we finally chose $\alpha$=1.1 and $\beta$=0.99 as the parameters of the CBF loss.

Under the same conditions, we compared the CBF loss with cross-entropy loss, hinge loss, and focal loss functions.

The cross-entropy loss function is commonly used in machine learning, which can promote the training of CNN more effectively. The cross-entropy loss function is defined as

$$L_{\mathrm{CE}} = -\sum_{i=1}^{n} t_i \ln(p_i), \tag{12}$$

where $t_i$ is the truth label; $p_i$ is the softmax probability for the $i$ class; and $n$ is the number of classes.

The hinge loss function is generally used in SVM, and its advantage is that it has a good effect with small sample size and

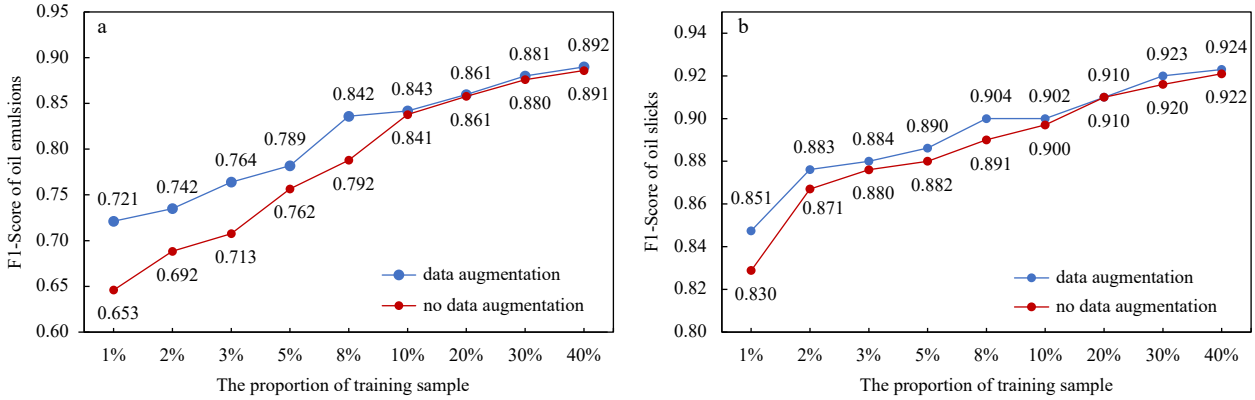


**Fig. 8.** Influence of data augmentation on the classification accuracy of oil emulsions and oil slicks. a. F1-Score of oil emulsions. b. F1-Score of oil slicks.

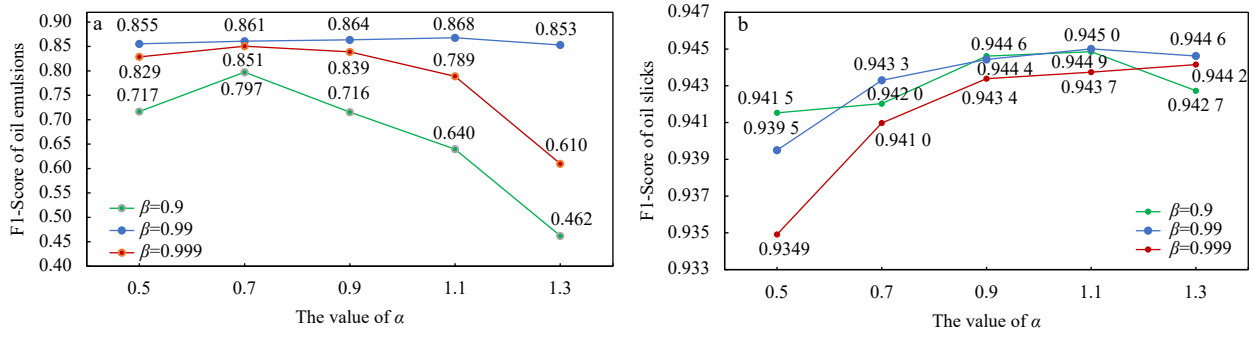


**Fig. 9.** Influence of the values of $\alpha$ and $\beta$ on oil spill classification accuracy. a. The F1-Score of oil emulsions. b. The F1-Score of oil slicks.

strong tolerance to noise. The mathematical formulation of hinge loss is as follows:

$$L_{\text{Hinge}} = \sum_{W_i \neq T_i} \max\left(0, S_{W_i} - S_{T_i} + 1\right), \tag{13}$$

where for the $i$ class, $W_i$ indicates the wrong label; $T_i$ represents the true label; $S_{W_i}$ indicates the output value of the model on the wrong label; and $S_{T_i}$ indicates the output value on the true label.

The focal loss adds a factor $(1-p_i)^{\gamma}$ to the standard cross-entropy criterion. Setting $\gamma > 0$ reduces the relative loss for well-classified examples, putting more focus on hard, misclassified examples (Lin et al., 2017). The focal loss is defined as

$$L_{\text{Focal}} = -\sum_{i=1}^{n} \alpha_i \left(1-p_i\right)^{\gamma} t_i \ln\left(p_i\right), \tag{14}$$

where $t_i$ is the truth label; $p_i$ is the softmax probability for the $i$ class; $n$ is the number of classes; $\gamma$ is the tunable focusing parameter; and $\alpha_i$ is the class balancing factor for the $i$ class. After sufficient experimental adjustment, we found that $\gamma = 1.5$, $\alpha_{\text{seawater}} = 0.1$, $\alpha_{\text{oil emulsions}} = 0.6$, and $\alpha_{\text{oil slicks}} = 0.3$ works best.

Figure 10 shows the oil spill detection results of the CNN models with different loss functions. It can be seen that the F1-Scores of the oil emulsions and oil slicks in the detection result of the CNN model with CBF loss are higher than other loss functions, which indicates that the CBF loss function can effectively avoid the influence of the imbalanced number of sample categories on model training, and thus improve the classification accuracy of the CNN model on the oil spill.

### 3.4 *Comparison of different methods*

To evaluate the effectiveness of the CBF-CNN oil detection model, it was experimentally compared with DNN, SVM, and RF, which have been commonly used in oil spill detection in recent years. The parameters of different models are shown in Table 2. The confusion matrix of the classification results of various methods is shown in Table 3, and the classification accuracy results are shown in Fig. 11.

The histogram of the F1-Score of oil emulsions and oil slicks in comparison with different methods is shown in Fig. 11. For the oil slicks, the F1-Score of the CBF-CNN is 0.94, whereas the F1-Scores of the other methods range from 0.72 to 0.79. It is difficult to distinguish oil slicks from seawater using any method other than the CBF-CNN model, which is mainly caused by the similarity of the spectral information in oil slicks and seawater. The CBF-CNN can automatically mine the spatial context information of oil spills, learn more essential features, and thus obtain a higher classification accuracy. For oil emulsions, the F1-Score of CBF-CNN of 0.87 is higher than those of the DNN, SVM, and RF classification results. Moreover, Table 3 shows that the CBF-CNN model mainly misidentifies a small number of oil emulsions as oil slicks. In oil spill emergency response, the most important thing is to distinguish the clean seawater and oil spill, so in an actual application, this kind of misclassification costs less than misclassification of oil film as seawater.

Figure 12 shows the classification results of different methods. It can be seen that the CBF-CNN model has a high classification accuracy and avoids the misclassification of oil slicks and seawater to a large extent. In contrast, the results of other classification methods are relatively coarse. In particular, in the classi-

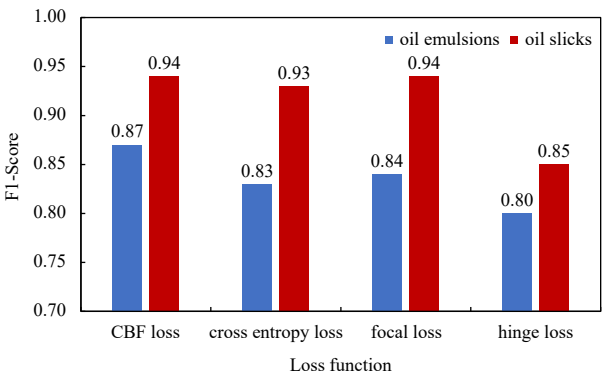


**Fig. 10.** Comparison of the results of the different loss functions. CBF: Class-Balanced F.

**Table 2.** Parameters of different models

| Models | Parameters |
|---|---|
| CBF-CNN | Epoch: 100; Batch size: 450; Optimizer: Adam; Loss function: CBF; Learning rate: 0.001; Spatial neighborhood-scale: 11×11 |
| CNN | Epoch: 100; Batch size: 450; Optimizer: Adam; Loss function: cross entropy; Learning rate: 0.001; Spatial neighborhood-scale: 11×11 |
| DNN | Hidden layer sizes: 50; Activation: ReLU; Optimizer: Adam; Learning rate: 0.2; Epoch: |
| SVM | Kernel: RBF; C: 700; Gamma: 0.25; Degree: 3 |
| RF | The number of the trees: 90; The minimum number of samples on leaf: 50; The maximum number of elements: 5 |

**Table 3.** Confusion matrices of different methods

| Methods | Class | Seawater | Oil emulsions | Oil slicks |
|---|---|---|---|---|
| CBF loss-CNN | seawater | 699 123 | 1 | 1 667 |
| | oil emulsion | 170 | 8 704 | 1 576 |
| | oil slick | 2 857 | 849 | 54 021 |
| DNN | seawater | 700 303 | 25 | 463 |
| | oil emulsion | 527 | 7 943 | 1 980 |
| | oil slick | 23 060 | 869 | 33 798 |
| SVM | seawater | 682 360 | 97 | 18 334 |
| | oil emulsion | 281 | 8 054 | 2 115 |
| | oil slick | 9 689 | 1 124 | 46 914 |
| RF | seawater | 688 620 | 1 898 | 10 273 |
| | oil emulsion | 198 | 8 387 | 1 865 |
| | oil slick | 10 726 | 1 664 | 45 337 |

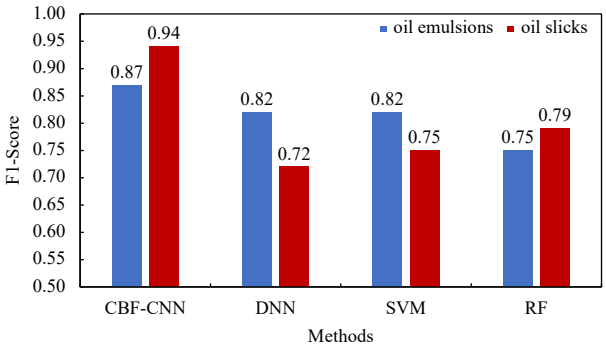


**Fig. 11.** F1-Score of oil emulsions and oil slicks comparisons in different methods. CBF-CNN: Class-Balanced F convolution neural network; DNN: deep neural network; SVM: support vector machine; RF: Random Forests.

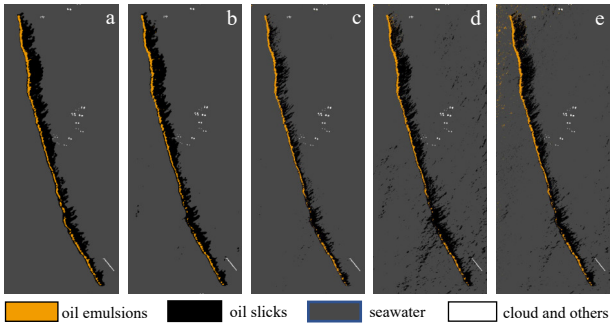


**Fig. 12.** Classification results of different methods. a. Result of visual interpretation; b. result of the CBF-CNN model; c. result of the DNN model; d. result of the SVM; e. result of the RF model.

fication results of SVM and RF, there are a large number of misclassified fragments of oil slicks that are due to the influence of wind and sea waves. These methods identify a part of the seawater samples as oil slicks. The DNN model, in contrast, tends to mistake the oil slicks as seawater, which seriously affects the determination of the amount of oil spill.

### 3.5 *Application of the CBF-CNN model*

In order to prove the applicability of the CBF-CNN model in different observation scenarios, we also carried out oil spill detection experiments in study area 2 and study area 3, which are located in the Karimata Strait.

There are a large number of clouds and their shadows in study area 2, and their spectra are very similar to those of oil spills (the spectra of clouds and oil emulsions are similar, and the spectra of shadows and oil slicks are similar). Clouds and shadows are common interference factors in oil spill detection, as look-alikes, they will seriously interfere with the accurate detection of oil spills.

As shown in Fig. 13, using uniformly distributed selection, 13 142 pixels were selected as training data, and 33 235 pixels were selected as test data in study area 2. These data include three categories: oil emulsions, oil slicks, and background, in which the background includes seawater, clouds, and their shadows. The detailed information of training data and test data are listed in Table 4.

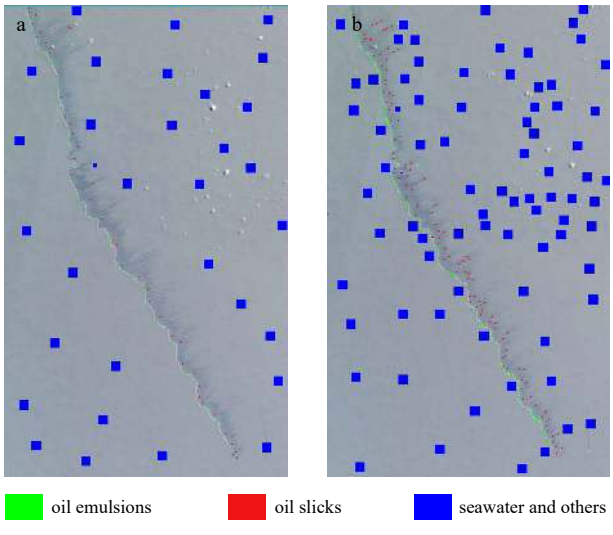


**Fig. 13.** Spatial distribution of training data (a) and test data (b).

**Table 4.** Statistics of training data and test data

| Class | Oil emulsions | Oil slicks | Background | Total |
|---|---|---|---|---|
| Train | 288 | 369 | 12 485 | 13 142 |
| Test | 924 | 1 000 | 31 311 | 33 235 |
| Total | 1 212 | 1 369 | 43 796 | 46 377 |

To evaluate the effectiveness of the CBF-CNN oil detection model, it was experimentally compared with DNN, SVM, and RF. The parameters of different models are shown in Table 5. The histogram of the F1-Score of oil emulsions and oil slicks in comparisons with different methods is shown in Fig. 14. It can be seen that the CBF-CNN model has higher detection accuracy for oil emulsions and oil slick than other models.

The classification results of different methods are shown in Fig. 15. It can be seen that there is a large amount of salt-and-pepper noise in the classification results of DNN, SVM, and RF models. This is because the spectra of clouds and oil emulsions are very similar, so the models cannot effectively distinguish them. However, The CBF-CNN model, with its powerful feature extraction capabilities, can fully explore the spectral differences between oil emulsions and clouds, and extract their texture information, so as to achieve accurate distinctions between oil emulsions and clouds.

In study area 3, the oil spill does not exist in isolation but is connected to a large black background and cannot be accurately distinguished. In order to improve the contrast between oil spills and black background, we used the logarithmic transformation method to enhance the image of study area 3, as shown in Fig. 16. Logarithmic transformation is a commonly used remote sensing image processing method. When we want to expand the low brightness area of the image, we can use this method. The transformation method is given by the following formula:

**Table 5.** Parameters of different models

| Models | Parameters |
|---|---|
| CBF-CNN | Epoch: 100; Batch size: 450; Optimizer: Adam; Loss function: CBF; Learning rate: 0.001; Spatial neighborhood-scale: 11×11 |
| DNN | Hidden layer sizes: (100, 50); Activation: ReLU; Optimizer: Adam; Learning rate: 0.2; Epoch: 200 |
| SVM | Kernel: RBF; C: 500; Gamma: 0.5; Degree: 3 |
| RF | The number of the trees: 100; The minimum number of samples on leaf: 50; The maximum number of elements: 5 |

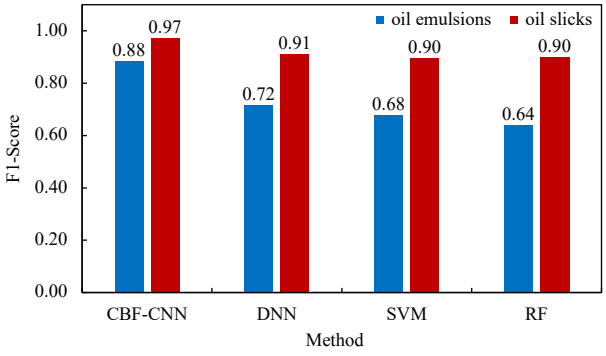


**Fig. 14.** F1-Score of oil emulsions and oil slicks comparisons in different methods. CBF-CNN: Class-Balanced F convolution neural network; DNN: deep neural network; SVM: support vector machine; RF: Random Forests.

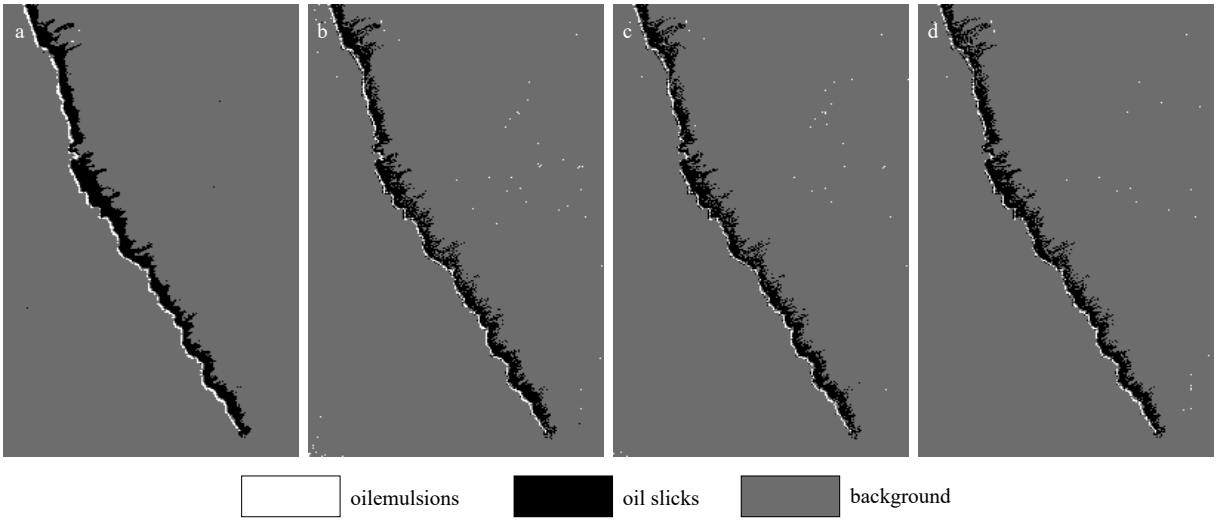


**Fig. 15.** The results of different classification methods. a. Classification result of the CBF-CNN; b. classification result of the DNN; c. classification result of the SVM; d. classification result of the RF.

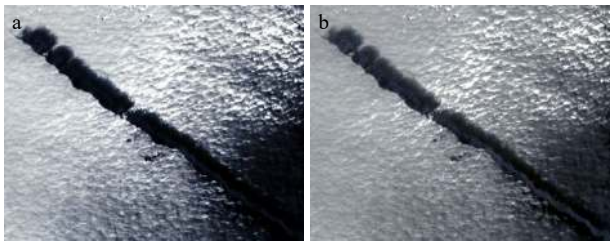


**Fig. 16.** The logarithmic transformation of the image. a. Origin image; b. transformed image.

$$s = c \times \ln(1 + r), \tag{15}$$

where $r$ is the input image pixel value; $s$ is the output image pixel value; and $c$ is a constant. In our experiment, $c$ is set to 1.

As shown in Fig. 17, we selected 14 695 pixels as the training data and 25 572 pixels as the test data. These data include two categories: oil slicks and background, in which the background includes seawater and black background. The detailed information of training data and test data are listed in Table 6.

To evaluate the effectiveness of the CBF-CNN oil detection model, it was experimentally compared with DNN, SVM, and RF.

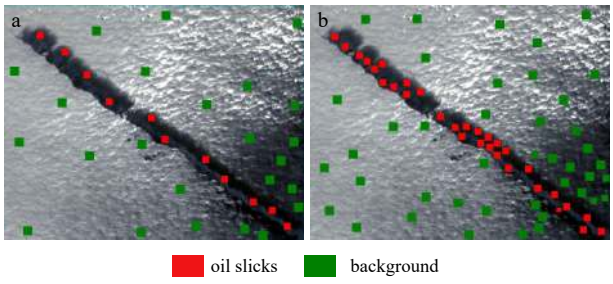


**Fig. 17.** Training (a) and test data (b) of study area 3.

**Table 6.** Statistics of train data and test data

| Class | Oil slicks | Background | Total |
|---|---|---|---|
| Training | 4 111 | 10 584 | 14 695 |
| Test | 7 790 | 17 782 | 25 572 |
| Total | 11 901 | 28 366 | 40 267 |

**Table 7.** Parameters of different models

| Models | Parameters |
|---|---|
| CBF-CNN | Epoch: 100; Batch size: 450; Optimizer: Adam; Loss function: CBF; Learning rate: 0.001; Spatial neighborhood-scale: 11×11 |
| DNN | Hidden layer sizes: (100, 70); Activation: ReLU; Optimizer: Adam; Learning rate: 0.2; Epoch: 200 |
| SVM | Kernel: RBF; C: 800; Gamma: 0.35; Degree: 3 |
| RF | The number of the trees: 80; The minimum number of samples on leaf: 20; The maximum number of elements: 3 |

The parameters of different models are shown in Table 7. From Fig. 18, we can see that the detection result of oil slicks by the CBF-CNN model is closer to the human visual sense, while other methods have more detection errors. The CBF-CNN oil spill detection model has a stronger anti-interference ability and higher applicability.

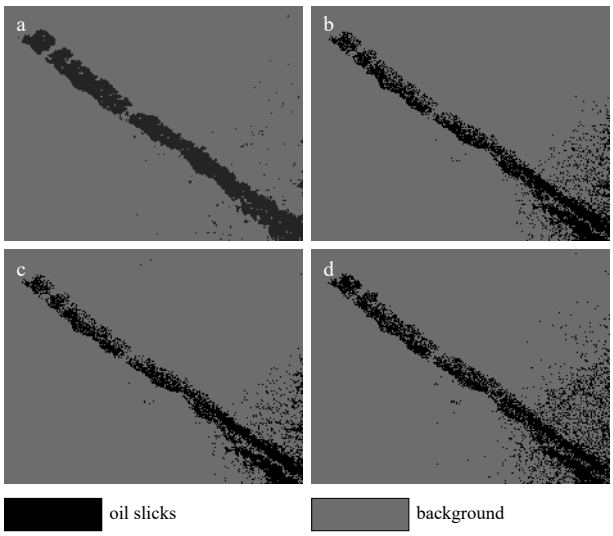


**Fig. 18.** The results of different classification methods in study area 3. a. Classification result of the CBF-CNN; b. classification result of the DNN; c. classification result of the SVM; d. classification result of the RF.

In addition to the visual results, we through three evaluation indexes, namely, the precision values, recall values, and F1-Score. The statistical results of detection performance by different methods are given in Table 8.

Results in Table 8 indicate that the CBF-CNN model accurately detected oil slicks. Also, the oil slicks evaluation measures of the classification result of the CBF-CNN model outperformed other methods with precision (0.94), recall (0.97), F1-Score (0.96).

The above experiments indicated that the CBF-CNN oil spill detection model has strong applicability, which can accurately identify oil spills under the interference of clouds, shadows, and other look-alikes.

### 3.6 *Prospect of oil spill monitoring*

During the operation of HY-1C, the CZI detected multiple oil spill events, two of which are studied in this paper. Figure 19 shows the color composite images of some oil spill events occurring in the Bohai Sea and the South China Sea. With the construction of ground receiving stations and the development of data transmission technology, the HY-1C/D satellite constellation will soon achieve a true 1.5 days global coverage, and it is expected that more oil spill accidents will be detected. The CZI sensors of HY-1C/D have a wide observation range and high temporal resolution, so they can effectively monitor oil spills in offshore China waters and have great potential for global oil spill monitoring. It is predicted that HY-1C/D CZIs will play an important role in sea oil spill monitoring missions in the future.

**Table 8.** Detection performance of different methods for oil slicks

| Methods | Precision | Recall | F1-Score |
|---|---|---|---|
| CBF-CNN | 0.94 | 0.97 | 0.96 |
| DNN | 0.69 | 0.84 | 0.76 |
| SVM | 0.66 | 0.86 | 0.75 |
| RF | 0.70 | 0.76 | 0.73 |

## 4 Conclusions

For a wide range of remote sensing images, the number of oil spill samples is usually far less than the number of seawater samples, and this limits the performance of CNN models. Moreover, the marine environment tends to be heterogeneous in large area images because of various sea conditions and sunglint, which makes the identification of oil spills more difficult. To address these problems, we proposed the CBF loss function, which comprehensively considers both the precision and recall rates and rebalances the loss according to the effective sample numbers of various types. We built a CBF-CNN model using the CBF loss function and set out a series of experiments using the images obtained by the HY-1C CZI in the Andaman Sea and the Karimata Strait.

By analyzing the sunglint intensity, we found that the three study areas all belong to the negative contrast area with weak sunglint. In this area, the reflectivity of oil emulsions is higher than that of seawater, while the reflectivity of oil slicks is lower than that of seawater. Therefore, the categories of oil spill variety that we detected in the experiments were oil emulsions and oil slicks. Through a series of experiments, we explored the influence of the size of the spatial neighborhood on the accuracy of oil spill detection and finally determined that the optimal size of the spatial neighborhood is 11×11. Through experiments, we found that the data augmentation method combining flipping and adding Gaussian noise can effectively improve the classification accuracy of the CBF-CNN. The influence of parameters $\alpha$ and $\beta$ on the performance of the CBF loss function is explored. It was found that when $\alpha = 1.1$ and $\beta = 0.99$, the CBF loss function can achieve optimal performance. We compared the CBF loss function with other commonly used loss functions. The experimental results show that the F1-Score of the detection result of oil emulsions is 0.87, which is 0.03–0.07 higher than other loss functions, and the F1-Score of the detection result of oil slicks is 0.94, which is 0.01–0.09 higher than other loss functions. Then, the CBF-CNN was compared with other classical oil spills detection methods

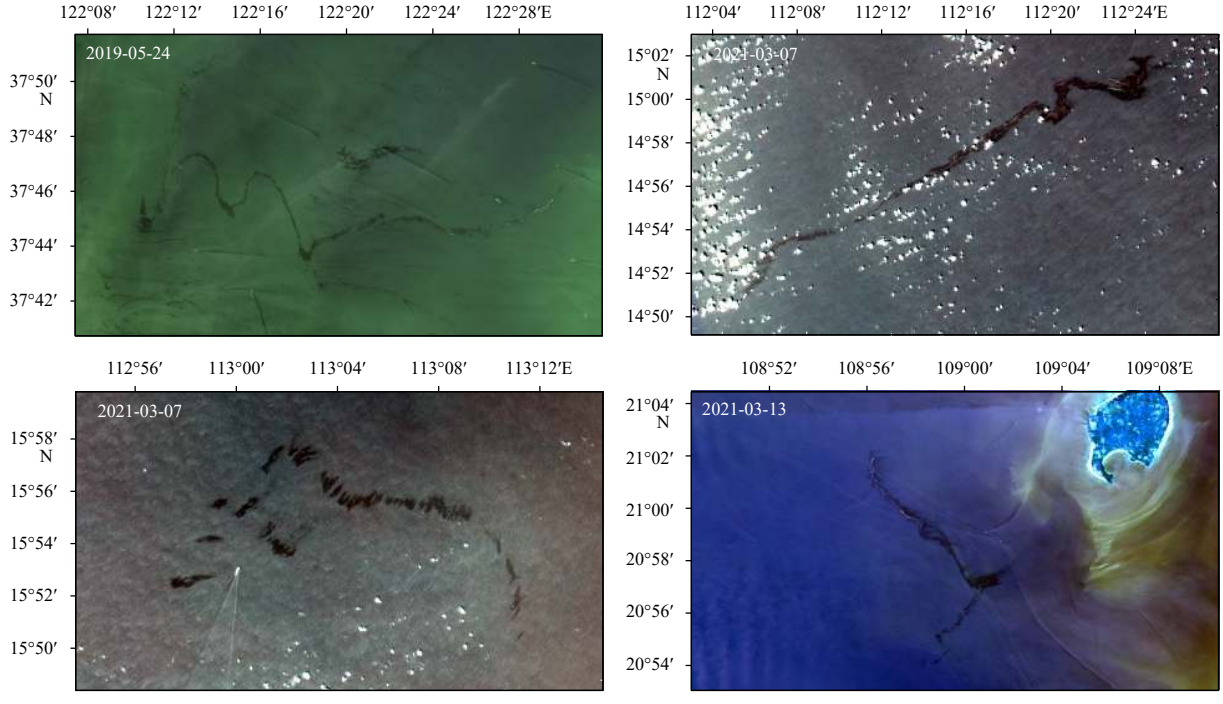


**Fig. 19.** Various oil spills could be identified from HY-1C CZI RGB images.

such as DNN, SVM, and RF models. The results show that the F1-Score of the CBF-CNN model for the detection result of oil emulsions is 0.05–0.12 higher than that of other methods, and the F1-Score for the detection result of oil slicks is 0.15–0.22 higher than that of other methods.

In order to verify the applicability of the CBF-CNN model in different observation scenes, we used the images obtained by HY-1C CZI in the Karimata Strait to carry out experiments, which include two study areas: study area 2 and study area 3. In the experiment of study area 2, the F1-Score of CBF-CNN model for the detection result of oil emulsions is 0.88, which is 0.16–0.24 higher than that of other methods, and the F1-Score for the detection result of oil slicks is 0.97, which is 0.06–0.07 higher than that of other methods. In study area 3, because no oil emulsion can be observed, we conducted a two-class experiment of oil spill and background. In the experiment, the F1-Score of CBF-CNN for oil slicks detection is 0.96, which is 0.20–0.23 higher than other methods.

Sum up, the following conclusions are drawn:

(1) The CBF loss function can restrain the influence of oil spill and seawater sample imbalance on oil spill detection of CNN model thus improving the detection accuracy of oil spills.

(2) The CBF-CNN model is suitable for the detection of oil slicks and oil emulsions in the negative contrast area of CZI images with weak sunglint.

(3) The CBF-CNN oil spill detection model can be applied to different scenarios of CZI images.

Due to the high time resolution of the HY1-C/D, the CZI has acquired many images of oil spills, and the images are still accumulating. In the future, we could use oil spill big data to train the CNN model to realize the effective remote sensing detection of oil spills.

***Acknowledgements***

We thank NSOAS (http://www.nsoas.org.cn/) for providing HY-1C/D data.

**References**

Abbriano R M, Carranza M M, Hogle S L, et al. 2011. Deepwater horizon oil spill: a review of the planktonic response. Oceanography, 24(3): 294–301, doi: 10.5670/oceanog.2011.80

Adamo M, De Carolis G, De Pasquale V, et al. 2009. Detection and tracking of oil slicks on sun-glittered visible and near infrared satellite imagery. International Journal of Remote Sensing, 30(24): 6403–6427, doi: 10.1080/01431160902865772

Breiman L. 2001. Random forests. Machine Learning, 45(1): 5–32, doi: 10.1023/A:1010933404324

Chen Weitao, Li Xianju, He Haixia, et al. 2018. A review of fine-scale land use and land cover classification in open-pit mining areas by remote sensing techniques. Remote Sensing, 10(1): 15

Corucci L, Nardelli F, Cococcioni M. 2010. Oil spill classification from multi-spectral satellite images: exploring different machine learning techniques. In: Proceedings of SPIE 7825, Remote Sensing of the Ocean, Sea Ice, and Large Water Regions 2010. Toulouse: SPIE, 782509

Cui Yin, Jia Menglin, Lin T Y, et al. 2019. Class-balanced loss based on effective number of samples. In: Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA: IEEE

Esbaugh A J, Mager E M, Stieglitz J D. et al 2016. The effects of weathering and chemical dispersion on Deepwater Horizon crude oil toxicity to mahi-mahi (*Coryphaena hippurus*) early life stages. Science of The Total Environment, 543: 644–651, doi: 10.1016/j.scitotenv.2015.11.068

Feng Lian, Hou Xuejiao, Li Junsheng, et al. 2018. Exploring the potential of Rayleigh-corrected reflectance in coastal and inland water applications: a simple aerosol correction method and its merits. ISPRS Journal of Photogrammetry and Remote Sensing, 146: 52–64, doi: 10.1016/j.isprsjprs.2018.08.020

Hu Chuanmin. 2009. A novel ocean color index to detect floating algae in the global oceans. Remote Sensing of Environment, 113(10): 2118–2129, doi: 10.1016/j.rse.2009.05.012

Hu Chuanmin, Li Xiaofeng, Pichel W G, et al. 2009. Detection of natural oil slicks in the NW Gulf of Mexico using MODIS imagery. Geophysical Research Letters, 36(1): L01604

Hu Chuanmin, Lu Yingcheng, Sun Shaojie, et al. 2021. Optical remote sensing of oil spills in the ocean: what is really possible?. Journal of Remote Sensing, 2021: 9141902

Ioffe S, Szegedy C. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 32nd International Conference on Machine Learning. Lille: ACM, 448–456

Jiang Zongchen, Ma Yi. 2020. Accurate extraction of offshore raft aquaculture areas based on a 3D-CNN model. International Journal of Remote Sensing, 41(14): 5457–5458, doi: 10.1080/01431161.2020.1737340

Jiang Zongchen, Ma Yi, Yang Junfang. 2020. Inversion of the thickness of crude oil film based on an OG-CNN Model. Journal of Marine Science and Engineering, 8(9): 653, doi: 10.3390/jmse8090653

Kingma D, Ba J. 2015. Adam: a method for stochastic optimization. In: Proceedings of the 3rd International Conference on Learning Representations. San Diego, CA: arXiv.org

Kolokoussis P, Karathanassi V. 2018. Oil spill detection and mapping using sentinel 2 imagery. Journal of Marine Science and Engineering, 6(1): 4, doi: 10.3390/jmse6010004

LeCun Y, Bengio Y. 1995. Convolutional networks for images, speech, and time series. In: Arbib M A, ed. The Handbook of Brain Theory and Neural Networks. Cambridge: MIT Press

Lin T Y, Goyal P, Girshick R, et al. 2017. Focal loss for dense object detection. In: Proceedings of 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2999–3007

Lu Yingcheng, Li Xiang, Tian Qingjiu, et al. 2013. Progress in marine oil spill optical remote sensing: detected targets, spectral response characteristics, and theories. Marine Geodesy, 36(3): 334–346, doi: 10.1080/01490419.2013.793633

Lu Yingcheng, Shi Jing, Hu Chuanmin, et al. 2020. Optical interpretation of oil emulsions in the ocean–Part II: Applications to multiband coarse-resolution imagery. Remote Sensing of Environment, 242: 111778, doi: 10.1016/j.rse.2020.111778

Lu Yingcheng, Sun Shaojie, Zhang Minwei, et al. 2016. Refinement of the critical angle calculation for the contrast reversal of oil slicks under sunglint. Journal of Geophysical Research: Oceans, 121(1): 148–161, doi: 10.1002/2015JC011001

Lu Jinshu, Xu Zhenfeng, Xu Song, et al. 2015. Experimental and numerical investigations on reliability of air barrier on oil containment in flowing water. Marine Pollution Bulletin, 95(1): 200–206, doi: 10.1016/j.marpolbul.2015.04.020

Michel J, Owens E H, Zengel S, et al. 2013. Extent and degree of shoreline oiling: *Deepwater horizon* oil spill, Gulf of Mexico, USA. PLoS ONE, 8(6): e65087, doi: 10.1371/journal.pone.0065087

Niclòs R, Doña C, Valor E, et al. 2013. Thermal-infrared spectral and angular characterization of crude oil and seawater emissivities for oil slick identification. IEEE Transactions on Geoscience and Remote Sensing, 52(9): 5387–5395

Serra-Sogas N, O'Hara P D, Canessa R, et al. 2008. Visualization of spatial patterns and temporal trends for aerial surveillance of illegal oil discharges in western Canadian marine waters. Marine Pollution Bulletin, 56(5): 825–833, doi: 10.1016/j.marpolbul.2008.02.005

Shen Yafeng, Liu Jianqiang, Ding Jing, et al. 2020. HY-1C COCTS and CZI observation of marine oil spills in the South China Sea. Journal of Remote Sensing, 24(8): 933–944

Sun Shaojie, Lu Yingcheng, Liu Yongxue, et al. 2018. Tracking an oil tanker collision and spilled oils in the East China Sea using multisensor day and night satellite imagery. Geophysical Re-

search Letters, 45(7): 3212–3220, doi: 10.1002/2018GL077433

Tong Cheng, Mu Bing, Liu Rongjie, et al. 2019. Atmospheric correction algorithm for HY-1C CZI over turbid waters. Journal of Coastal Research, 90(SI): 156–163

Wen Yansha, Wang Mengqiu, Lu Yingcheng, et al. 2018. An alternative approach to determine critical angle of contrast reversal and surface roughness of oil slicks under sunglint. International Journal of Digital Earth, 11(9): 972–979, doi: 10.1080/17538947.2018.1470687

Yang Junfang, Wan Jianhua, Ma Yi, et al. 2019. Oil spill hyperspectral remote sensing detection based on DCNN with multi-scale features. Journal of Coastal Research, 90(SI): 332–339

Yekeen S T, Balogun A L, Yusof K B W. 2020. A novel deep learning instance segmentation model for automated marine oil spill detection. ISPRS Journal of Photogrammetry and Remote Sensing, 167: 190–200, doi: 10.1016/j.isprsjprs.2020.07.011

Yin Liping, Zhang Min, Zhang Yuanling, et al. 2018. The long-term prediction of the oil-contaminated water from the *Sanchi* collision in the East China Sea. Acta Oceanologica Sinica, 37(3): 69–72, doi: 10.1007/s13131-018-1193-5

Zhu Xueyuan, Li Ying, Zhang Qiang, et al. 2019. Oil film classification using deep learning-based hyperspectral remote sensing technology. ISPRS International Journal of Geo-Information, 8(4): 181, doi: 10.3390/ijgi8040181

# Appendix: Corresponding machine learning glossary

**1. Rectified linear unit (ReLU)**

This is a layer of neurons that applies the non-saturating activation function $f(x) = \max(0, x)$. It increases the nonlinear properties of the decision function and the overall network without affecting the receptive fields of the convolution layer.

**2. Adaptive Moment Estimation (Adam)**

Adam optimization is a stochastic gradient descent method that is based on the adaptive estimation of first-order and second-order moments.

**3. Batch normalization (BN)**

The BN layer was proposed by Google in 2015. It is a technique for deep neural network training that not only speeds up the convergence of models but also alleviates the "disappearing gradient" problem in deep networks to some extent, thus making the training of deep network models easier and more stable.

**4. Softmax**

The softmax function is a function that turns a vector of K real values into a vector of K real values that sum to 1. The input values can be positive, negative, zero, or greater than one, but the softmax transforms them into values between 0 and 1 so that they can be interpreted as probabilities. If one of the inputs is small or negative, the softmax turns it into a small probability, and if the input is large, then it turns it into a large probability, but it will always remain between 0 and 1.

**5. Batch size**

The batch size is the number of samples that will be passed through to the network at one time. Given that a single epoch is one single pass of all the data through the network, it will take batches to make up a full epoch.

**6. Epoch**

An epoch in machine learning means one complete pass of the training dataset through the algorithm. This epochs number is an important hyperparameter for the algorithm. It specifies the number of epochs or complete passes of the entire training dataset passing through the training or learning process of the algorithm. With each epoch, the dataset's internal model parameters are updated.