OIL SPILL DETECTION BASED ON CBD-NET USING MARINE SAR IMAGE

Yanan Zhang[a], Qiqi Zhu[a], Qingfeng Guan[a]*

[a] School of Geography and Information Engineering, China University of Geosciences, Wuhan, China
*Corresponding author E-mail: zhuqq@cug.edu.com

IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium | 978-1-6654-0369-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/IGARSS47720.2021.9553884

ABSTRACT

Oil spill is one of the most widespread, frequent and harmful marine pollution. Oil spill detection based on synthetic aperture radar (SAR) images detects the oil film by identifying dark spots in the images. Dark spots detection can be achieved using image segmentation techniques. However, natural phenomena such as waves and currents can also cause dark spots, resulting in consistently uneven intensity, high noise and blurred boundaries in oil spill images. In addition, existing oil spill detection models often perform well for large targets, but have poor detection accuracy for small targets, causing part of the oil spill to be ignored. To solve the above problems, oil contextual and boundary-supervised detection network (CBD-Net) is proposed to extract refined oil spill regions by fusing multi-scale features, where the scSE attention block is used to model the global context to improve the internal consistency of oil spill regions. In CBD-Net, boundary details are enhanced with optimized edge supervision. In addition, a manually labeled dataset is proposed, Deep-SAR Oil Spill (SOS) dataset, aiming to solve the problem of insufficient existing oil spill detection dataset. Experimental results demonstrate that the proposed model outperforms other comparative models, and is able to extract robust and accurate oil spill regions from complex SAR images.

Index Terms— Oil spill, deep learning, attention block, edge supervision, new oil spill dataset

1. INTRODUCTION

Large-scale oil spill events and illegal dumping of waste oil have caused serious damage to the marine ecological environment [1]. Therefore, timely and accurate monitoring of oil spills is of great importance to improve the efficiency and effectiveness of maritime law enforcement departments. With the advantages of wide monitoring range and high acquisition efficiency, remote sensing monitoring technology has gradually become one of the main oil spill detection methods. Compared with other remote sensing detection methods, synthetic aperture radar (SAR) has the imaging capability of all-weather and all-time work, and the characteristics of variable viewpoint, multi-polarization [2].

The SAR-based method detects oil film by identifying the dark spots area in the image. Dark spots detection is achieved by using image segmentation techniques. Oil spill detection methods can be classified as threshold-based, edge information-based, and neural network-based segmentation methods [3]. Threshold-based segmentation methods divide grayscale intervals by setting a threshold, and counts the pixels in the same grayscale interval as the same category. Edge information-based methods performs image segmentation by concatenating discrete grayscale pixels to obtain region contours. These algorithms have high requirements for the grayscale distribution of the image. But SAR images have coherent spot noise, and the gray values in the same target area in the image are highly variable [4].

Neural network-based segmentation methods for oil film identification can effectively address the limitations of traditional segmentation methods. Various traditional neural networks and machine learning methods have been applied to detect oil leaks from satellite images, including support vector machines, decision trees, random forests, and artificial neural networks. However, these approaches require manual intervention and a long processing time. In recent years, deep learning models have achieved great success in extracting high-level feature representations of images. Krestenitis *et al.* [5] used several semantic segmentation models and found that the DeepLabv3+ model achieved the best performance in oil spill detection. Temitope Yekeen et al. [6] developed a new deep learning oil spill detection model based on the Mask-Region-based Convolutional Neural Network (Mask R-CNN) model, which is able to segment objects with high accuracy even when overlapping other objects.

Although deep learning methods have performed well in many oil spill detections, they still have the following limitations. First, there are few publicly available SAR oil spill datasets, and deep learning requires large-scale data training. In addition, oil films on the ocean surface exhibit very irregular shapes with highly complex boundaries. Existing neural network-based methods do not target oil spill shapes with extremely irregular shapes to design networks. The segmentation effect is not good for oil spill targets with fuzzy boundaries.

This work was supported by National Natural Science Foundation of China under Grant No. 41901306, Grant from State Key Laboratory of Resources and Environmental Information System.

To solve the above issues, a new oil spill detection framework, oil contextual and boundary-supervised detection network (CBD-Net) is proposed to detect oil spill in SAR images. CBD-Net uses encoder-decoder structure for capturing multi-scale oil spill by gradually recovering the spatial information. The main contributions of the paper are as follows:

(1) To improve the feature representation of complex oil spill in SAR images, CBD-Net is proposed. The multi-parallel dilated convolution module is inserted in the middle to fuse multi-scale feature fusion. The attention mechanism module is introduced to improve the internal consistency of oil spill targets by modeling the global context.

(2) The joint loss function is proposed for the blurring and noise occurring inside the object in SAR images. Following the integral method of computing boundary changes, the edge loss term is designed to weight the binary loss function. This improves the learning ability of edge features in the CBD-Net framework.

(3) A new SAR oil spill detection dataset, Deep-SAR Oil Spill (SOS) dataset, is constructed to fully advance the oil spill detection task. SOS dataset provides a benchmark resource for the development of state-of-the-art algorithms for SAR image oil spill detection or other related tasks.

2. SOS: DEEP-SAR OIL SPILL DATASET

2.1. Study Area 1: Oil Spill Area in the Gulf of Mexico

On April 20, 2010, the Deepwater Horizon rig in the United States Gulf of Mexico exploded, resulting in a major oil spill. The oil slick covered an area of about 160 km in length and 72 km at its widest point. PALSAR images of the Gulf of Mexico region were acquired between May and August 2010.

2.2. Study Area 2: Oil Spill Area in the Persian Gulf

The Persian Gulf accounts for more than 50% of the world's oil reserves. Saudi Arabia, Iran, Kuwait, Iraq and United Arab Emirates along the Persian Gulf coast are important oil-producing countries. The Al-Khafji oil field is an important oil exploration site jointly operated by Kuwait and Saudi Arabia. Sentinel-1 Interferometric Wide Swath Level 1 products are used in this research.

2.3. Production of SOS dataset

Existing oil spill detection datasets are difficult to obtain. In this paper, an oil spill monitoring dataset is created, namely the Deep-SAR Oil Spill Dataset. 19 fully polarized SAR oil spill images are collected. Data enhancement techniques such as random cropping, rotation and scaling are used to extend the original dataset. Relevant sub-regions were sampled by GIS experts. The samples with smaller percentage of oil spill area were filtered. In the end, a total of 6708 images from the Mexican oil spill area were used for training and 1678 images for testing. There are 6202 images in the Persian Gulf oil spill area for training and 1552 images for testing. Each image has a size of 256 × 256 pixels.

2.4. Characteristics of SOS dataset

First, SOS dataset has rich image variability. Tolerance for image variation is an important requirement for oil spill monitoring systems. The images in this dataset are carefully selected under a wide range of weather, seasons, lighting conditions, imaging conditions and scales. Second, the images of SOS dataset are high in class diversity and inter-class similarity. Many performance methods built on deep neural networks have reached saturation in classification accuracy in most existing datasets due to their simplicity, or rather lack of variation and diversity. For this purpose, images were obtained under different imaging conditions. Its images were obtained from multiple remote sensing satellite images with different spatial resolution and spectral coverage.

3. SAR OIL SPILL DETECTION BASED ON THE PROPOSED CBD-NET FRAMEWORK

The network structure of CBD-Net is shown in Fig. 1. The input image is first input to the residual module in the encoder to obtain a reduced dimensional feature map. Then the feature map is fed to the attention block. To preserve more oil spill edge details, a multi-parallel expanding convolution module is added between the encoder and the decoder. Then, the multiscale contextual feature is input into the decoder. Finally, the fused features are fed into the softmax layer to predict the final segmentation result. Furthermore, the joint loss function is designed to enhance the edge detection of oil spill images.

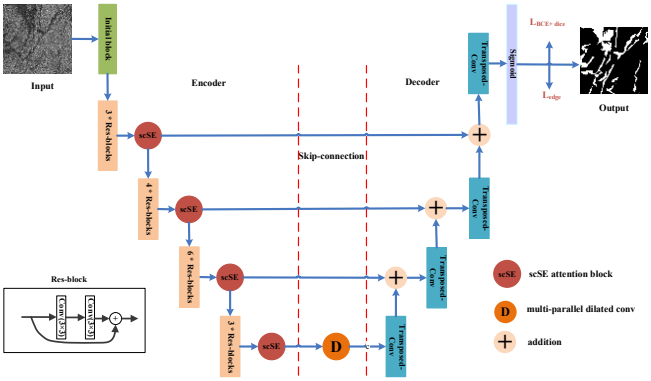

Fig. 1. The overall flowchart of the proposed framework for oil spill detection.

3.1. Encoder with Contextual and Detailed Information

As shown in Fig. 1, the pre-trained ResNet34 is initialized to

enhance the generalization of CBD-Net. This can greatly training. The encoder has 4 down-sampling modules. Therefore, the size of the feature map output by the encoder is reduced to 1/16 of the original image. To address the high noise and internal inconsistency of the oil spill target, each residual block is followed by an attention module. Parallel spatial and channel squeeze excitation (scSE) blocks [7] is utilized to model the interdependencies between channel and space.

scSE attention block: As shown in Fig. 2, the feature map $V=[v_1,v_2,v_3,\cdots,v_c]\in R^{C\times H\times W}$ is input to the attention block.

The global average pooling is used to compress the spatial dimension on the feature V to generate the feature map $q\in R^{1\times 1\times C}$.

$$q_k=\frac{1}{H\times W}\sum_i^H\sum_j^W v_k(i,j) \tag{1}$$

$W_1\in R^{\frac{C}{2}\times 1\times 1}$, $W_2\in R^{C\times 1\times 1}$ are used to explicitly model the correlation of feature channels is explicitly modeled.

$$\bar{q}=W_1(\delta(W_2 q)) \tag{2}$$

where $\delta(\cdot)$ denotes rectified linear unit (RELU). The sigmoid function is then used for normalization to obtain the corresponding mask. The formulation of cSE module is defined as:

$$\bar{V}_{cSE}=\left[\sigma(\bar{q}_1)v_1,\sigma(\bar{q}_2)v_2,\cdots,\sigma(\bar{q}_C)v_C\right] \tag{3}$$

Given the feature map of the input tensor is $V=\left[v^{1,1},v^{1,2},\cdots v^{i,j},\cdots,v^{H,W}\right]$. The spatial attention map is then activated using the sigmoid function. The formulation of sSE module is defined as:

$$\bar{V}_{sSE}=\left[\sigma(u_{1,1})v^{1,1},\cdots,\sigma(u_{i,j})v^{i,j},\cdots,\sigma(u_{H,W})v^{H,W}\right] \tag{4}$$

Finally, the above two modules are connected in parallel and added together to form the scSE attention block.

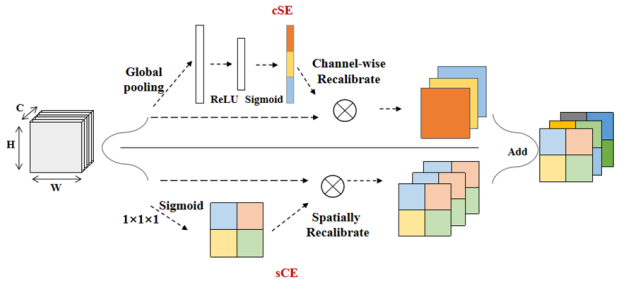

Fig. 2. The overall of the scSE attention block.

3.2. Multi-parallel Dilated Convolution Block

To perform a finer segmentation of the oil spill target, multiple dilated convolutions with different receptive fields are merged in parallel. The dilation rates are 1, 2, 4, 8. The feature maps will obtain different receptive fields improve the convergence speed of the network during to integrate the multi-scale oil spill targets.

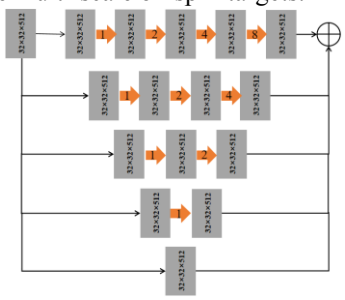

Fig. 3. The overall of the multi-parallel dilated convolution block.

3.3. Decoder

The output of the multi-parallel dilation convolution module is fed to the decoder. Transposed convolution is used to expand the spatial dimension of the feature map. The decoder has four decoder blocks. Skip-connection is used to fuse low-level information with high-level information. This is important for fine-grained segmentation. The size of the feature mapping gradually increases and the number of channels gradually decreases through four layers of repeated decoder blocks. Finally, the prediction results are output using the Sigmoid function.

3.4. Joint Loss Function for Edge Preservation

Instead of only supervising the final segmentation map , the joint loss function is proposed to consist of three components, including L_{edge}, L_{BCE} and L_{dice}. L_{edge} is a binary representation of the contours of oil spill targets in the image. L_{BCE} and L_{dice} denote the BCE (binary cross entropy) and dice coefficient loss, respectively. The total loss L is computed as:

$$L=L_{edge}(L_{BCE}+L_{dice}) \tag{5}$$

4. EXPERIMENTS AND ANALYSIS

4.1. Implementation Details

Adam was adopted as optimizer with the batch size of 4. Learning rate was initially set to be 2e-4 and reduced by a factor of 5 for 3 times while observing the training loss decreasing slowly. All models were trained on 4 NVIDIA RTX2080 GPUs.

4.1. Results and Analysis

The proposed method is compared with U-Net [8], Deeplabv3 [9] and D-LinkNet [10] on SOS dataset. The visual evaluation of these methods is shown in Fig. 4.

Table 1. Comparison of CBD-Net with other oil spill segmentation technologies.

	mIoU (%)		Fscore (%)		Recall (%)		Precision (%)	
	PALSAR	Sentinel-1	PALSAR	Sentinel-1	PALSAR	Sentinel-1	PALSAR	Sentinel-1
U-Net	80.38	80.02	96.36	92.55	95.40	94.02	95.35	89.32
Deeplabv3	82.54	82.87	96.58	93.39	96.15	94.25	97.02	92.54
D-Linknet	81.75	79.74	96.60	92.16	97.29	93.86	95.91	90.52
Ours	**83.23**	**83.40**	**97.69**	**94.08**	**98.62**	**94.14**	**97.56**	**93.02**

Compared with other methods, CBD-Net extracts less noise and smoother edges for the oil spill targets.

As shown in Table 1, the proposed CBD-Net achieves a maximum IoU of 83.23% and an F1 score of 97.69% on the Mexican study area images obtained by PALSAR. The method achieves the best performance in all metrics compared to other methods. In detail, the IoU and F1 scores of CBD-Net are improved by 2.85% and 1.33%, respectively, compared to U-Net, which indicates the effectiveness of multiscale contextual aggregation. Compared with D-LinkNet, CBD-Net, which uses global contextual information modeling, increases by 1.48% and 1.09% in IoU, respectively.

Table 1 shows the accurate results of the proposed CBD-Net on the Persian study area images obtained by Sentinel-1. The proposed CBD-Net achieved the best performance with 3.38%, 1.53%, 0.12%, 3.7% and 2.05% increase in mIoU, F1 score, Recall, Precision and Accuracy, respectively. These indicate the importance of capturing remote spatial relationships to detect oil spill objects, as well as the supervision by enhancing edge features.

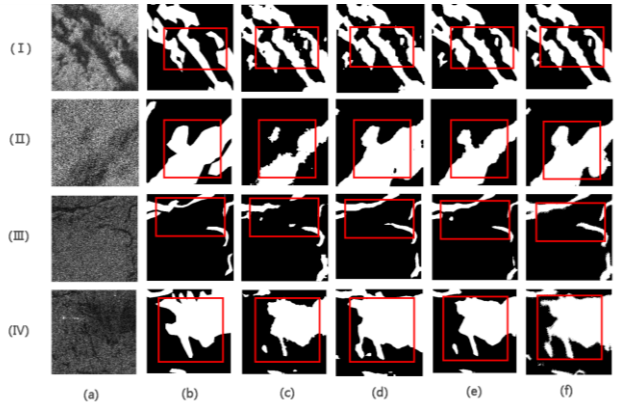

Fig. 4. Oil spill detection results using the SOS dataset. (I)-(II) the Mexican study area images. (III)-(IV) the Persian study area images. (a) SAR image. (b) Ground truth. (c) U-Net. (d) Deeplabv3. (e) D-LinkNet. (f) Ours.

5. CONCLUSION

In this paper, CBD-Net is designed to extract pixel-level oil spill regions from SAR images. Fine-grained oil spill segmentation is performed by modeling the global context and fusing multi-scale features. Enhanced edge feature is obtained by an improved joint loss function. To alleviate the problem of lack of training samples, SOS dataset consisting of SAR images is proposed. The experimental results on the proposed dataset show that the framework performs better in oil spill detection of SAR images. In the future, multimodal learning methods will be considered for oil spill monitoring.

6. REFERENCES

[1] C. Brekke and A. H. S. Solberg, "Oil spill detection by satellite remote sensing," *Remote Sensing of Environment*, vol. 95, no. 1, Art. no. 1, Mar. 2005.

[2] M. Konik and K. Bradtke, "Object-oriented approach to oil spill detection using ENVISAT ASAR images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 118, pp. 37–52, Aug. 2016.

[3] R. Al-Ruzouq *et al.*, "Sensors, Features, and Machine Learning for Oil Spill Detection and Monitoring: A Review," *Remote Sensing*, vol. 12, no. 20, p. 3338, 2020.

[4] W. Wang, H. Sheng, S. Liu, Y. Chen, J. Wan, and J. Mao, "An Edge-Preserving Active Contour Model with Bilateral Filter based on Hyperspectral Image Spectral Information for Oil Spill Segmentation," in *2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Sep. 2019, pp. 1–5.

[5] M. Krestenitis, G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis, and I. Kompatsiaris, "Oil spill identification from satellite images using deep neural networks," *Remote Sensing*, vol. 11, no. 15, p. 1762, 2019.

[6] S. Temitope Yekeen, A. Balogun, and K. B. Wan Yusof, "A novel deep learning instance segmentation model for automated marine oil spill detection," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 167, pp. 190–200, Sep. 2020.

[7] A. G. Roy, N. Navab, and C. Wachinger, "Concurrent spatial and channel 'squeeze & excitation'in fully convolutional networks," in *International conference on medical image computing and computer-assisted intervention*, 2018, pp. 421–429.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234–241.

[9] L. C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587*, 2017.

[10] L. Zhou, C. Zhang, and M. Wu, "D-LinkNet: LinkNet With Pretrained Encoder and Dilated Convolution for High Resolution Satellite Imagery Road Extraction.," in *CVPR Workshops*, 2018, pp. 182–186.