

Deep Learning Framework for Edge Detecting and Classification in SAR Images of Oil Spills using CNN

Dr. C. Pabitha¹, Dr. B. Vanathi², Ms. G. Sangeetha³ and Ms. S. Suma⁴

¹⁻⁴SRM Valliammai Engineering College, Department of Computer Science and Engineering, Tamil Nadu, India
Email: pabithac.cse@srmvalliammai.ac.in, vanathib.cse@srmvalliammai.ac.in, sangeethag.cse@srmvalliammai.ac.in, sumas.cse@srmvalliammai.ac.in

Abstract— Oil spills are a type of pollution that occurs when small or large quantities of crude oil or distilled oils, such as gasoline or diesel, pollute bodies of water and coastal lands. The spill contains toxic chemicals that can have serious short- and long-term health consequences for oil workers, nearby residents, marine life, animals, and the environment. There are numerous types of oil spills, the majority of which are minor, such as when oil spills from a ship while it is being refueled. Serious incidents, on the other hand, have long-term consequences. The most of major oil spills occur when a pipeline ruptures, a tanker sinks or runs aground, or when a drilling operation goes wrong. An oil spill is an untidy and risky occurrence that can harm people and animals in the surrounding area. The proposed work uses a Deep Learning model named MASK RCNN algorithm (MASK Region Based Convolutional Neural Network) to automatically detect the area or size of the oil spill to overcome the hazardous situation. Convolutional neural networks, a subset of artificial intelligence (AI), are trained on a visual dataset of oil spill images from various altitudes and geographical locations in this study. Experiments show that the proposed algorithm has a good ability to detect oil spills at a higher detection rate, with an identification rate greater than 80% and a false alarm rate less than 12%. The evaluation results show that the proposed algorithm performs well in detecting oil spills, with an overall detection rate of more than 80%.

Index Terms— Oil Spill, MASK RCNN, Deep Learning, Resnet 101, SAR images.

I. INTRODUCTION

An oil spill is an environmental hazard that harms bodies of water, land, and air. Additionally, it can pollute sea surfaces and harm bird species, fish, and other aquatic creatures. They are primarily caused by oil tanker, ship, and pipeline accidents in which crude oil, gasoline, fuel, and oil by-products are released into the water. Oil slick removal is critical for maintaining a safe and clean environment and protecting aquatic life. If an oil spill occurs in an area populated by wildlife, the consequences can be severe. Oil degrades the insulating ability of mammalian fur and reduces the water repellency of bird feathers; without insulation or water repellency, mammals and birds can die from hypothermia. Dolphins and whales can inhale oil, which affects their immune systems and reproduction. While fish and shellfish are not immediately affected, because oil floats on water, fish can experience impacted growth, enlarged livers, fin erosion, and a reduction in reproductive capabilities as the oil mixes and sinks. The impact on fish and shellfish can also be lethal; even if not lethal, they are frequently no longer safe for human consumption. Due to the sheer high oil consumption by various sectors, ocean exploration

and transportation of oil have increased significantly over the last three decades. Oil spills, as one of the primary polluters of the ocean, have negative consequences for coastal and deep-ocean environments [1]. For example, OS causes seafood contamination and the extinction of marine life [2]. As a result, developing accurate and timely OS detection (OSD) algorithms is critical for preventing OS spread and minimizing its negative effects on the ocean and the ecological ecosystem [3]. Remote sensing (RS) systems provide valuable datasets in this regard [4]. Satellites, for example, collect data over large and remote ocean areas at a low cost and in a timely manner [5]. The rest of this article is structured as follows. A review of literature work is presented in Section 2, which is followed by a description of the dataset used and the suggested approach in Section 3. In Section 4, the experimental study's findings are discussed. Finally, Section 5 represents the conclusion and future ideas.

II. RELATED WORKS

Topouzelis et al. deployed a two artificial neural networks were shown for detecting oil spills [6]. Dark formation is accurately detected by the first network with a 1-3-1 topology 94% of the time. A further set of ten features was recovered from the dark formation, and when a fully connected multilayer perceptron with a topology of 10-51-2 was applied to the features, it was able to distinguish oil spills from similar-looking objects with an accuracy of 89%. Proposed combining two neural networks to detect oil leaks. Singha et al. proposed a second neural network with 14-14-5-1 topology is used to classify a pixel as an oil spill or lookalike, while the first network with 3-6-2 topology is used to segment the SAR images to a dark formation or backdrop. 91.6 percent of the oil spill pixels and 98.3 percent of the look-alike pixels were correctly categorized by the suggested network [7]. On a chosen set of totally polarimetric features taken from SAR pictures [8], Song et al. implemented an optimized wavelet neural network with an overall accuracy of 96.55 percent and 97.67 percent on two separate datasets. In order to detect instances of oil spills, Chen et al. proposed Stacked Autoencoder (SAE) and deep belief networks using polarimetric SAR features [9]. With a testing error below 1.4 percent when the SAE was utilized, the usage of the aforementioned deep-learning methodologies was found to be more promising than Support Vector Machine (SVM) and conventional artificial neural networks. The classification method used by the image segmentation algorithm is not uniform, and the choice of segmentation algorithm is heavily influenced by the shape of the image to be segmented, the characteristics of the pixel distribution, and other factors. These segmentation algorithms are primarily classified as threshold segmentation, clustering segmentation, region growth, and so on.

The OTSU algorithm was used by Xu et al. to monitor oil spills [10]. In order to extract oil spill dark patches from SAR pictures, Jin et al. used FCM [11]. To retrieve information about oil spills, Zou et al. employed the SVM supervised classification approach [12]. Japanese researcher OTSU proposed the OTSU algorithm, commonly referred to as the maximum interclass variance approach, in 1979. From the perspective of properties like geometry, grayscale, and texture in SAR images, an oil spill can be located [13]. Additionally, the features of distinct SAR remote sensing images are also highly different. Finding appropriate characteristics that can recognize oil spills is challenging. The practice of manually looking for features can be avoided by using convolutional neural networks. It reduces the need to spend time manually looking for suitable features and can increase the precision of detection. Coherent speckle noise in SAR remote sensing photos is a natural result of the imaging technology used for SAR, which is coherent microwave imaging. Coherent speckle noise makes it extremely challenging to analyze SAR images [14]. It is challenging to categorize each pixel in the SAR remote sensing image using semantic segmentation because coherent speckle noise exists. Therefore, the bigger picture is divided into numerous smaller pictures to lessen the influence of coherent speckle noise, and then the convolutional neural network is used to categorize the cropped smaller pictures in place of categorising each pixel individually in semantic segmentation. In essence, it conceptually segments the original image using the classification of small images. This technique increases the detection accuracy while also significantly reducing the interference of speckle noise on SAR remote sensing images. Because the input image is so small, the requirements have been met by employing a shallow network. The traditional network model AlexNet [15] is the one utilized in this study. The model is modified to fit the input smaller image given that the input image is a smaller image. Jiao et al. introduced the Otsu algorithm to reduce the false positive rate and the Maximally Stable External Regions (MSER) algorithm to locate the oil spill by creating a detection box [16]. They also introduced a pre-trained deep convolutional neural network based on VGG-16 for classifying oil spill instances. The VGG-16 scored an f-measure of 98.5 percent and a recall of 99.5%. While the oil spill was tagged using the MSER approach at the right threshold level, the Otsu algorithm was used to increase precision from 97.7 percent to 98.3 percent. For the purpose of classifying the thickness of an oil coating on hyperspectral remote sensing photos, Zhu et al. tested with SVM, fully connected neural networks, SAE, and CNN [17]. When compared to

SAE, CNN performed and delivered accuracy that was over 5% better. A pre-processed dataset of SAR images was created by Krestenitis et al. [18] using data from the Sentinel-1 satellites of the European Space Agency (ESA). Additionally, the dataset was used to train the networks UNet, LinkNet, PSPNet, DeepLabV2, and DeepLab3+, using MobileNetV2 as the foundation network. These networks generated modest Jaccard scores for the oil spill.

III. DATASETS

The oil spill dataset, oil slicks dataset, or just oil, a common unbalanced machine learning dataset is used for this work. Machine Learning for the Detection of Oil Spills in Satellite Radar Images was the title of a 1998 publication by Miroslav Kubat, et al. that first mentioned the dataset. Robert Holte, a co-author of the publication, is frequently given credit for the dataset. Starting with satellite photographs of the ocean, some of which show an oil leak and some of which do not, the dataset was created. Images were segmented into patches using computer vision methods, and the contents of each patch were then described by a vector of features. The test samples and training samples were not duplicated, and a total of 1,100 photos representing 10 percent were used as test samples. Before the image is input, the convolution kernel size is set to 3, and the median filtering method is used to remove the salt and pepper noise in the image in order to eliminate interference brought on by noise.

IV. SAR IMAGES

Synthetic-aperture radar, or SAR image, is a type of image that can provide both two- and three-dimensional reconstructions. To create a precise 3D reconstruction, the radar antenna is moved over the target area. UAVs with fixed SARs and sophisticated surround-sensing radars are frequently deployed. The complete synthetic aperture is produced by the distance that these microwaves travel in the time it takes for the radar pulses to return to the receiver. In general, the resolution of the image will be higher the larger the aperture, Oil spill image represented in Figure 1.



Figure 1. Oil Spill image

In order to “irradiate” the target area and create a SAR image, continuous microwave beams are propagated; the echo of each pulse is then captured and processed. Using just one beam-forming antenna, the pulses are sent out and the echoes are picked up. The relative location of the SAR antenna with regard to the target area changes over time when a SAR device-loaded spacecraft moves. The integration of the received data is made possible by processing these continuous received pulse echoes. The SAR aperture is created as a result of this technique. These photos can be produced using a variety of satellites, including RADARSAT-2 and RISAT-1, with Sentinel-1A being the most recent to be used.

V. PREPROCESSING

Speckle filtering will perform better at identifying various target situations and making image segmentation simple. The speckle noise is taken into account when using various analysis techniques. With only a little amount of information loss, an excellent speckle reduction filter can be applied. The filter can keep the edge and radiometric information in homogeneous areas. When it comes to textured areas, the filter can preserve the radiometric and spatial signal variability that make up the texture information. Based on mean and median, the non-adaptive filters can be further minimized. While moving the window around the sample, the intensity of each individual sample interval in the image is modified to the mean of the pixel values. This is the fundamental idea behind non-adaptive filters. The benefit of a mean filter is that while it blurs the texture region, it can maintain radiometry information. Better texture information can be preserved with this filter, however the homogenous area is altered. The rapid Fourier transform filter is an illustration of a non-adaptive filter. When it comes to minimized speckle noise, adaptive filters perform better than non-adaptive filters. The advantage of

using adaptive filters is that they take into account the local terrain backscatter and can adapt to local property changes.

In satellite photos, medical images, and synthetic aperture radar (SAR) images, the speckle noises are generally dispersed. A method of adaptive filtering based on bidimensional empirical mode decomposition (BEMD)[19] for despeckling SAR images. The initial BIMF level is filtered because it contains the high-frequency noise component after the noisy SAR picture is divided into several bidimensional intrinsic mode function (BIMF) levels using BEMD.

VI. EDGE DETECTION

A. Laplacian of Gaussian (LoG)

A mathematical model was created by Marr and Hildreth [1980] using their knowledge of biological vision. The most crucial elements are local averaging carried out by applying a filter to smooth the image and looking for extreme values for an intensity change (representing an edge). The Gaussian filter is the most widely used smoothing filter. The function being convolved with the image is the two-dimensional Gaussian function.

$$G_{\sigma}(x, y) = e^{-\frac{(x^2 + y^2)}{2\sigma^2}} \quad (1)$$

where represents the corresponding Gaussian probability distribution's standard deviation [Parker, 2011]. Convolution of the image is followed by application of the Laplacian operator: ∇

$$\nabla^2 = \frac{\delta^2}{\delta x^2} + \frac{\delta^2}{\delta y^2} \quad (2)$$

The Laplacian has the benefit of being rotation-invariant, thus it responds to changes in intensity similarly independent of the orientation of the mask [Gonzalez and Woods, 2010]. The Laplacian operator and the Gaussian filter can both be combined to form the Laplacian of Gaussian because the order is irrelevant (LoG)

$$\nabla^2 G_{\sigma} = \left(\frac{r^2 - 2\sigma^2}{\sigma^4} \right) e^{\left(\frac{-r^2}{2\sigma^2} \right)} \quad \text{where } r = \sqrt{x^2 + y^2} \quad (3)$$

B. Canny edge detection

The signal-to-noise ratio should be as high as possible, the distance between the calculated edge pixels and the edge should be as small as possible, and the edge detector should not identify multiple edge pixels when there is only one single edge. Canny [1986] took this approach to improve current methods by defining these three performance criteria. The best filter was identified by the author as one that meets all three requirements. The first derivative of a Gaussian function serves as an accurate approximation for those requirements.

$$G(x) = e^{-\frac{x^2}{2\sigma^2}} \quad (4)$$

Therefore, the derivative with regard to x is:

$$G'(x) = \left(-\frac{x}{\sigma^2} \right) e^{-\left(\frac{x^2}{2\sigma^2} \right)} \quad (5)$$

A Gaussian in two dimensions is given by

$$G(x, y) = \sigma^2 e^{-\left(\frac{x^2 + y^2}{2\sigma^2} \right)} \quad (6)$$

In addition, derivatives of G exist in both the x and y directions. By convolving the input image with G' , an image E with improved edges will be produced because this is the approximation to Canny's ideal filter for edge detection [Parker, 2011]. The following step is a non-maximum suppression, which suppresses both the magnitude of gradients and all values. By following the edges by a lower and an upper threshold, hysteresis thresholding is used to filter out weak edges that are not coupled to strong edges [Canny, 1986; McIlhagga, 2011].

VII. DEEP LEARNING FRAMEWORK

The preprocessed image is taken and patches of size $64 \times 64 \times 3$ were created from the denoised photos. Additionally, patches are divided into two categories: patches with oil spill pixels that are more than 1% and patches with oil spill pixels that are less than 1%. Nearly 40 pixels out of a total of 4096 pixels with a "1" label can be deemed background patches in patches with a 1 percent oil spill. Additionally, when combined with other patches, the patches with less than 1% oil spill worsen the dataset's imbalance and decrease the model's accuracy. A two-stage deep learning framework is then given patches. An Inception CNN that has 23 convolutional, rectified linear units (ReLU), pooling, fully linked, and SoftMax layers is used in the first stage. The cross-entropy is determined and defined at the SoftMax layer as the expected probability distribution of each output produced by each node of the final two-node fully linked layer:

$$\tilde{P}(c_i) = \frac{e^{c_i}}{\sum_{j=1}^2 e^{c_j}} \quad (7)$$

where $P(c_i)$ is the anticipated cross-entropy of c_i and c_i is the output of node i of the fully connected layer. The decision regarding the input patch is predicted after the cross-entropy for each of the two outputs is compared with respect to some threshold (i.e., significant or insignificant oil spill patch). To update the CNN model's parameters and provide a precise forecast for the patch class, the cross-entropy loss (CEL) is additionally computed and minimized using a backpropagation gradient descent approach. This is how CEL is described:

$$CEL = - \sum_{i=1}^2 P(c_i) \log[\tilde{P}(c_i)] \quad (8)$$

where the expected or desired probability distribution of the output value c_i is denoted by $P(c_i)$.

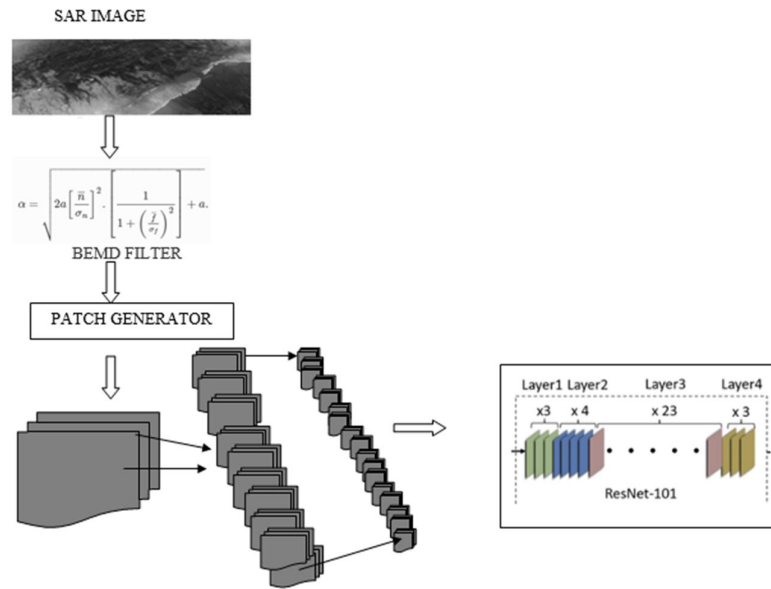


Figure 2. Deep Learning Framework

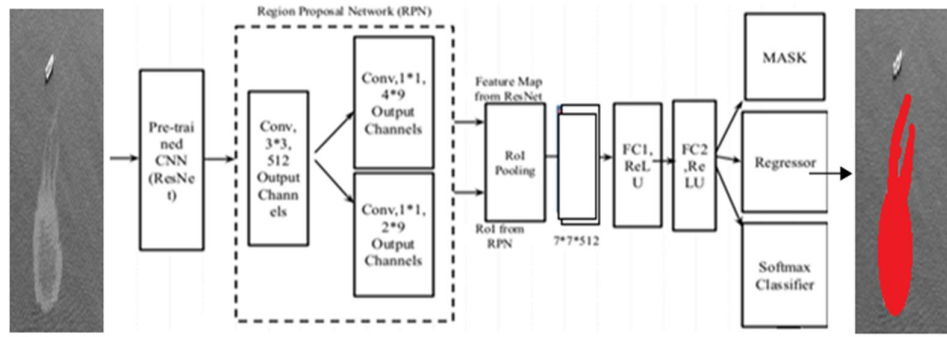


Figure 3. Architecture of MASK RCNN model

A. MASK R CNN Algorithm

Modern instance segmentation deep learning model Mask R-CNN combines pixel-level image classification with object recognition. It is a Faster-RCNN framework enhancement that has the extra capability of doing segmentation of the instance elements (ROIs). The foundation for creating each instance of an interest class in an image with a mask output is an object detection method. This model's superior capacity to segment items with high accuracy, especially when they overlap other objects, over other conventional classifiers and semantic segmentation models is critical for oil spill detection. The Mask R-CNN model typically consists of three stages: feature extraction from the input images by the backbone network; generation of Region of Interests (ROIs) from the Region Proposal Network (RPN) based on the feature maps output from the backbone; and extraction of objects of interest based on the output from the RPN which are sent to the fully convolutional layer (FC) and the fully convolutional network (FCN) for classification and instance segment. The working of Deep learning framework and general architecture of Mask R-CNN is depicted in Fig 2 and Fig.3 respectively.

VIII. VALIDATION AND DISCUSSION

Utilizing precision, accuracy, and recall metrics, the Resnet101 model's effectiveness at identifying photos based on the presence of oil spills is determined. Eq. (9), in which true positive (TP) and true negative (TN) are proportional to the number of correctly predicted values, calculates accuracy as the ratio of the number of images that are correctly classified to the total number of images. Positive predictions that were accurate are denoted by TP, and accurate negative predictions are denoted by TN. On the other hand, false positive (FP) and false negative (FN) refer to the quantity of wrongly predicted values. False positive predictions are denoted by FP, while false negative predictions are denoted by FN. Therefore, the total number of all predictions made by the model is the denominator of Eq. (9). Eqs. (10) and (11) can be used to determine precision and recall values given TP, TN, FP, and FN.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

Overall, the model achieves a 94.89 percent accuracy and a 20.51 percent test loss. The lack of an oil spill is similarly properly identified in 90% of the test photographs, while it is mistakenly identified in 10% test images. The comparison with underlying backbone networks is depicted in Table I.

TABLE I. COMPARISON STUDY

Different Approaches	Accuracy	Recall	Precision	Dice
Inception v2	92.50%	81.40%	80.95%	81.28%
VGG 16	90.35%	77.63%	75.01%	76.30
Resnet 101	94.89%	78.3%	85.1%	79.6%

IX. CONCLUSION

In this study, the categorization of SAR dark patches was done using Resnet 101, a relatively deep DCNN. The implementation of the data augmentation technique for the large data set derived from 1100 SAR images allows Resnet 101, which has up to 16 weight layers, to perform better. When compared to hand-crafted features, the distinguishability of features acquired from the data set by Resnet 101 is substantially superior. As a result, Resnet 101 performs classification substantially better than more conventional techniques. From 92.50 percent, 81.40 percent, and 80.95 percent, the accuracy, recall, and precision increased to 94.2 percent, 78.3 percent, and 85.1 percent, respectively.

REFERENCES

- [1] D. Song et al., "A novel marine oil spillage identification scheme based on convolution neural network feature extraction from fully polarimetric SAR imagery", *IEEE Access*, vol. 8, pp. 59801-59820, 2020.
- [2] B. G. Gautama, N. Longép , R. Fablet and G. Mercier, "Assimilative 2-D Lagrangian transport model for the estimation of oil leakage parameters from SAR images: Application to the Montara oil spill", *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 11, pp. 4962-4969, Nov. 2016.
- [3] Y. Cheng et al., "Monitoring of oil spill trajectories with COSMO-SkyMed X-band SAR images and model simulation", *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 7, pp. 2895-2901, Jul. 2014.
- [4] S. T. Seydi, M. Hasanlou and M. Amani, "A new end-to-end multi-dimensional CNN framework for land cover/land use change detection in multi-source remote sensing datasets", *Remote Sens.*, vol. 12, no. 12, 2020.
- [5] S. Sun and C. Hu, "The challenges of interpreting oil–water spatial and spectral contrasts for the estimation of oil thickness: Examples from satellite and airborne measurements of the deepwater horizon oil spill", *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2643-2658, May 2018.
- [6] Topouzelis K., Karathanassi V., Pavlakis P., Rokos D. Detection and discrimination between oil spills and look-alike phenomena through neural networks. *ISPRS J. Photogramm. Remote Sens.* 2007; 62:264–270. doi: 10.1016/j.isprsjprs.2007.05.003.
- [7] Singha S., Bellerby T.J., Trieschmann O. Satellite Oil Spill Detection Using Artificial Neural Networks. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2013;6:2355–2363. doi: 10.1109/JSTARS.2013.2251864.
- [8] Song D., Ding Y., Li X., Zhang B., Xu M. Ocean Oil Spill Classification with RADARSAT-2 SAR Based on an Optimized Wavelet Neural Network. *Remote Sens.* 2017;9:799. doi: 10.3390/rs9080799.
- [9] Chen G., Li Y., Sun G., Zhang Y. Application of Deep Networks to Oil Spill Detection Using Polarimetric Synthetic Aperture Radar Images. *Appl. Sci.* 2017;7:968. doi: 10.3390/app7100968.
- [10] J. Xu, B. Li, C. Cui, P. Liu, and X. Y. Zhu, "Research on marine radar oil spill monitoring technology," *Marine Environmental Science*, vol. 37, no. 1, pp. 16–20, 2018.
- [11] J. Jin, Y. N. Wu, and Z. L. Kang, "Feature extraction of oil spill dark spot based on multi-feature in SAR image," *Geomatics & Spatial Information Technology*, vol. 41, no. 2, pp. 53–56, 2018.
- [12] Y. R. Zou, C. Liang, and T. Zeng, "Oil spill identification using SVM based on polarization parameters," *Journal of Marine Sciences*, vol. 31, no. 3, pp. 71–75, 2013.
- [13] C. Shu and S. S. Sha, "Characteristics analysis of the oil spill and looks-alike for sea surface oil spill automatic monitoring," *Ship Ocean Engineering*, vol. 49, no. 2, pp. 64–67, 2020.
- [14] Z. L. Lu, X. Jia, W. G. Zhu, and C. Z. Zeng, "Study on SAR image despeckling algorithm," *Journal of Sichuan Ordnance*, vol. 38, no. 6, pp. 104–108, 2018.
- [15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 60, no. 6, pp. 84–90, 2017.
- [16] Jiao Z., Jia G., Cai Y. A new approach to oil spill detection that combines deep learning with unmanned aerial vehicles. *Comput. Ind. Eng.* 2019;135:1300–1311. doi: 10.1016/j.cie.2018.11.008.
- [17] Zhu X., Li Y., Zhang Q., Liu B. Oil Film Classification Using Deep Learning-Based Hyperspectral Remote Sensing Technology. *ISPRS Int. J. Geo-Inf.* 2019;8:181. doi: 10.3390/ijgi8040181.
- [18] Krestenitis M., Orfanidis G., Ioannidis K., Avgerinakis K., Vrochidis S., Kompatsiaris I. Oil Spill Identification from Satellite Images Using Deep Neural Networks. *Remote Sens.* 2019;11:1762. doi: 10.3390/rs11151762.
- [19] Ranjith Kumar Painam and M. Suchetha, Despeckling of SAR images using BEMD-Based Adaptive Frost Filter, *Journal of the Indian Society of Remote Sensing*, 2022, <https://doi.org/10.1007/s12524-022-01495-x>.