

A New Technique for Segmentation of the Oil Spills From Synthetic-Aperture Radar Images Using Convolutional Neural Network

Fatemeh Mahmoudi Ghara, Shahriar Baradaran Shokouhi , *Member, IEEE*, and Gholamreza Akbarizadeh 

Abstract—Oil spills have proven to have detrimental effects on the marine-based environment and economy. Thus, it is necessary to identify oil spills and classify them in the sea to reduce oil-induced pollution in seas and oceans. Synthetic-aperture radar (SAR) imaging is a good option for rapid oil detection, as it covers a wide area, collects data at short intervals, and allows taking images in all weather conditions throughout the day. The reason for using deep neural networks is that training several images enhances segmentation accuracy significantly. This article intended to separate the oil spills of SAR images using U-NET and DeeplabV3 neural networks, separately with the lowest number of images and the highest accuracy possible. Each of these neural networks carries out image segmentation with different architectures independently and thus we could not combine these two networks for oil spill segmentation. We managed to find two accurate convolutional neural networks (CNNs) for oil spill segmentation because we did not have access to sufficient hardware facilities, such as GPU, to train dozens of neural networks. The two networks we used in the article are among the most well-known and widely used networks. Our purpose was to figure out which network was the best in SAR oil spill detection. Given the limited number of SAR oil spill images and as the input of CNNs needs many images for training, we increased the number of input images to 9801 using the augmentation technique. Then, we carefully identified oil spills with 300 epoch and a batch size of 5 using the Python programming language on the GoogleColab server. The oil spill detection accuracy was 78.8% in the U-NET network and 54% in the DeepLabV3 network. Accordingly, we conclude that the most accurate identification of SAR oil spills in images belong to the U-NET network.

Index Terms—Convolutional neural networks (CNNs), DEEPLABV3, feature extraction, oil spill, segmentation, synthetic aperture radar (SAR), U-NET.

I. INTRODUCTION

OIL spills near the coast can result from ship accidents, broken pipelines, explosion of oil rig platforms, and

Manuscript received 22 May 2022; revised 1 September 2022 and 25 September 2022; accepted 4 October 2022. Date of publication 11 October 2022; date of current version 20 October 2022. This work was supported in part by the Shahid Chamran University of Ahvaz (SCU) and Iran University of Science and Technology (IUST) and in part by the Research Council of Shahid Chamran University of Ahvaz under Grant SCU.EE1401.269. (*Corresponding author: Gholamreza Akbarizadeh.*)

Fatemeh Mahmoudi Ghara and Shahriar Baradaran Shokouhi are with the Department of Electronic Engineering, Faculty of Electrical Engineering, Iran University of Science and Technology, Tehran 16846-13114, Iran (e-mail: mahmoudi_fatemeh65@yahoo.com; bshokouhi@iust.ac.ir).

Gholamreza Akbarizadeh is with the Department of Electrical Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz 61357-83151, Iran (e-mail: g.akbari@scu.ac.ir).

Digital Object Identifier 10.1109/JSTARS.2022.3213768

deliberate discharge of tank-cleaning wastewater from ships [1], [2], [3], [4], [5]. This is caused by the frequent tanker accidents and oil spills in waters as the main cause of oil leaks in oceans and seas. By the leakage of oil into water bodies, a thin layer is rapidly formed by spreading over the water surface, known as an oil spill. Due to several environmental factors, marine oil spills are hazardous and can rapidly spread over an extensive area.

One of the most significant origins of marine contamination is oil spills with severe economic and environmental effects on the coastal zone and ocean [1]. Oil spills caused by unintentional or intentional releases into coastal or oceanic waters present a primary threat to marine ecosystems. Hence, the adverse effects of oil spills on these ecosystems are the subject of significant environmental, political, and scientific concerns [6]. The NEREIDS program made the first serious attempt with the support of the European Commission to utilize metocean, shipping, and geological data to characterize oil spills in one of the key oil exploration areas in the world to hinder any major oil spill accidents. These data revealed that oil spill models were generated to simulate trajectories, develop oil spills, assess the susceptibility of the coastal zone, and find suitable measures to alleviate its environmental effects [7]. Hence, identifying and classifying oil spills are essential in preventing water contamination [8]. Nonetheless, some natural phenomena (e.g., waves, ocean currents, and human factors) can alter light intensity over the sea's surface, which leads to non-uniform intensity or high noise from oil spills or lookalikes, sometimes making it very difficult to segment oil spills automatically. Hence, the accurate segmentation technique crucially contributes to oil spill control [9]. Synthetic-aperture radar (SAR) is a cohesive imaging technology capable of producing high-resolution, large-scale images of the earth and targets. SAR can function at both night and day and in adverse conditions to overcome the limitations of optical and infrared systems [10]. Thus, oil spills are detected as "dark" areas in SAR images [11]. Traditionally, SAR imaging is conducted by a moving aircraft or spacecraft [12]. Oil spills emerge in SAR images in form of dark patches, because they produce low-backscatter responses in comparison to nearby clean sea regions [13]. An image of oil spills appears as dark gray pixels in SAR images. It is not simple to transfer image segmentation approaches to SAR images. These images comprise speckle arising with scatterers' complex summation within a resolution cell from coherent signals. Through such

noise, together with SAR sensor representation of geometry, it becomes challenging to consider edge information in the segmentation of the SAR image [14]. Nevertheless, SAR images are normally polluted by multiplication noise or speckle caused by the destructive and constructive interference scattered from coherent returns scattered by small reflectors within each resolution cell. Computer vision systems and human explainers encounter difficulties in interpretation and processing due to the presence of speckle noise in SAR images. Therefore, removing speckles from SAR images is essential to enhance the performance of various computer vision algorithms, such as recognition, segmentation, and detection [2]. It is still difficult to extract information from SAR images, even if manual labels are based on accurate auxiliary information. Data augmentation is a technique used to enhance the generalizability of the network by increasing the number of samples in the training set. However, it has many ambiguities in separating dark spots. Indeed, there is a hybrid area in the feature space for oil spills and lookalikes. For instance, SAR image properties of biological oil layers are very similar to those of oil spills. Without other auxiliary information like environmental information and remote sensing images, even experts cannot make a clear judgment about samples in a hybrid area of a feature space [15]. Over the recent years, convolutional neural networks (CNNs) have been extensively developed in radar imaging and semantic segmentation [16]. A CNN is an in-depth learning method specifically designed for image recognition and classification. This network is designed to be similar to multi-layered neural networks. CNNs are the same as biological neural networks employed for speech recognition and visual and image processing. CNNs can extract spatial information effectively and share weights among nodes for reducing the number of parameters [17]. A CNN works well in recognizing two-dimensional shapes and learning structural features and presents a way to distinguish features from points automatically. Short training time makes it easier to use artificial neural networks and improves diagnostic accuracy [18]. A CNN also works well to automatically learn features from raw data, particularly for structural features [19]. CNN was successfully used in image classification and recognition, indicating the satisfactory results of CNN in hand-written recognition [20]. CNN is also used in image segmentation. A study in 2005 reviewed oil spill classification and segmentation methods [21]. According to Solberg et al., there are three feature extraction steps for oil spill detection in SAR images, including oil pollution detection, dark spot detection, and look-alikes. The oil extraction location is affected by the accuracy of dark spot detection. In this article, different methods were reviewed for satellite-based oil spill detection in the marine environment. Manual and automatic methods were considered based on various satellite sensors and oil spills under different circumstances to differentiate between oil spills and look in terms of pattern recognition [21]. In another study, two neural networks were implemented by Duan et al. [19] for segmenting dark objects and separating oil spills from similar ones. The presented technique was very auspicious in detecting dark formations and oil spills from similar cases because dark formations are recognized with an overall accuracy of 94%. Moreover, 89% of the examined items were correctly recognized. The framework was used in different other assessments

[6], [22]. According to Krestenitis et al., semantic segmentation with deep CNN (DCNNs) deployment can be used successfully in oil spill detection. The established DCNN segmentation model was trained and evaluated using a common database on a case-by-case basis. Generally, the best performance was reported by the DeepLabv3 model with the highest accuracy based on greater inference time. The mentioned study is advantageous since the complexity of the problem was extensively discussed, along with relative figures. The superiority of the DeepLabv3 model was confirmed by further comparison of various DeepLab architectures with previous methods. It is not possible to use this model for oil spills and similar classes since some oil spill pixels are similarly classified. Though, it is expected that this deficiency was caused by the small number of specimens for the small object size and the training method [23]. A superpixel segmentation method was introduced by Zhang et al., to hypersegment SAR images. A statistical dissimilarity measurement technique was proposed in the superpixel integration section to transform soft superpixels into a self-connected weight graph. Moreover, phase integration and superpixel generation were run under a unified deep network. This method is advantageous since its shape is alternately adjusted based on the segmentation results and boundaries during training to achieve the desired segmentation results. The segmentation method is efficient computationally due to the simple network structure and includes advanced performance and good generalization [24]. A boundary clustering method was designed by Ma et al. to assess the specific tasks of superpixels. A soft graph convolutional network was proposed in the segmentation section taking the connectivity map as input and subdividing superpixels intelligently. The superpixel and graph complexity parts can be trained under a unified framework until obtaining optimal parameters of the two parts, which is the advantage of this technique. Furthermore, the superpixel shape can be adjusted by the network gradually in terms of segmentation. The results revealed that the boundary information could be preserved by the suggested superpixel generation model, which has good resistance to stains. This technique was efficient computationally and performed consistently with generalizability [25]. A benchmark solution was presented by Hang et al., by establishing a general framework for deep learning of multiple MDL components. Generally, the MDL framework is used to classify pixel tasks and model spatial information with artificial neural networks. A general framework was presented comprising two subnetworks, i.e., ex-framework and FU-net, to present a basic solution for classification tasks of pixel-level images using MULTIMODAL data. Various hybrid strategies have been presented in this regard. In networks, the three “what,” “how,” and “how” were focused along with two feature extraction approaches associated with FC-NET and CNN to classify pixels and spatio-spectral classes, respectively [26]. The CNN network can record spatial-spectral properties to classify hyperspectral images. Recently, GCNNs have been used in data analysis and visualization despite the sampling restrictions. Hong et al., assessed the complete HS image arrangement with GCNN and CNN. MINIGCNN can infer the output data without altering the network and enhancing the classification performance. It also facilitates the training of large-scale graph networks through the MINIBATCH technique. MINIBATCH allows the joint use of

GCNN and CNN to extract more distinct and diverse properties for the classification of HS images [27]. Usually, HS images are combined into a data cube using spatio-spectral data. Generally, it can be considered a sequence of data along with the spectral dimension. Inverters can describe available sequential attributes as a powerful architecture. An inverter-based backbone network was proposed by Hong et al., by further focusing on the extraction of spectral information from classifying these images without using it. This method can be improved by investigating self-organizing learning and creating a weighted network in terms of transformers to reduce the network complexity while maintaining more performance [28].

In this article, we needed many images to train convolutional networks. For this case, we increased 99 images of oil spills received from the Sentinel-1 and EnviSat satellites from the desired sites to 9801 images using the augmentation technique. Thus, we removed the oil spill 2 from the image and placed it on the oil spill 1. Several images are required by the CNNs for training. Only 99 images were used in the article as the database. The method as the main creativity and innovation of the article was used to increase the number of images of each oil spill on the remaining 98 oil spills. Thus, the images with various oil spills were incremented to 9801. Furthermore, the method was used to train CNNs well and perform image segmentation with higher accuracy. The main background of the new image was the oil spill 1. By this innovation, CNNs are trained without overfitting and perform segmentation with considerable accuracy. The remaining sections of the article were arranged as follows: The collection procedure for images required for training neural networks was introduced in Section II. Section III described the proposed method. Section IV explained the analysis results of segmented images of the oil spill output from CNNs in detail with their accuracy tables. Section V presented conclusions and implications.

II. DATASETS

This section deals with the way SAR radar oil spill images were collected for the database to train CNNs. SARs send information to Earth in Earth's orbit in any weather conditions and at any time of day or night. Images sent by the Sentinel-1 and EnviSat radars at alternating time periods per month can be downloaded from the relevant websites. This dataset shelters major types of oil spill candidates detected under various sea conditions [29]. In oil spill images, it is necessary to consider satellite type, frequency band, resolution, polarization, and SAR or POLSAR type. For instance, on the ESA SciHUB site, images sent from the Sentinel-1 radar can be received and converted to desired formats, like JPG, using SNAP software. A high resolution (up to MB) SAR radar is needed to train the neural network of convulsive images of oil spills. In this article, about 99 original images of oil spills were collected from different SARs. SAR images are needed in addition to annotated images of oil spills to train CNNs for oil spill segmentation. An annotated image of an oil spill made using Supervisely software is presented in Fig. 1. The mask image shows all the borders of oil spills accurately.

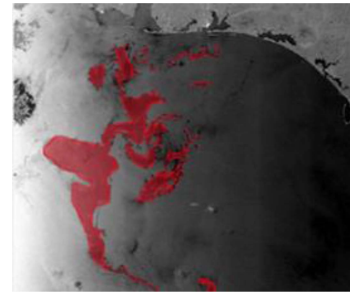


Fig. 1. Annotated sample of oil spill image from oil bank.

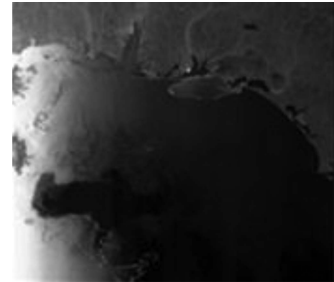


Fig. 2. Image of oil spills dataset.

A. Creating a Database of SAR Oil Spills Images in Google Drive

We put each of the oil spills images on the other oil spills to increase the images for training CNNs. An example of an oil spills shown on an oil spills is shown in Fig. 2.

For instance, we removed the oil spill in the first image. To do so, we multiplied the mask in the same image, and the resulting matrix returned the image values in the spots and was zero in spots without oil spills. Using matrix techniques, we placed the spot image in Fig. 2. The oil spill of the first image may overlap with the second oil spill, which does not affect the output of our work in this case. Using this method, we can produce many new images. The block diagram of the image data augmentation technique is shown in Fig. 3. We put all the 9801 images of SAR oil spills along with annotated images in a folder in the Google drive to provide a better access and established the link for addressing the folder in the program code. Each folder in the Google drive had two original images and a mask. The original image was the image received from SARs, and the annotated image was the image obtained from the software for training CNNs. The database of the images is available here and online. The original and annotated images were preserved in two separate folders.

III. PROPOSED ALGORITHM

In this article, we first masked all the oil spill images with supervisely software. Deep learning models become hungry of data, particularly with large architectures comprising numerous trainable parameters. For learning general classification rules and features and not overfitting the training data, the model must

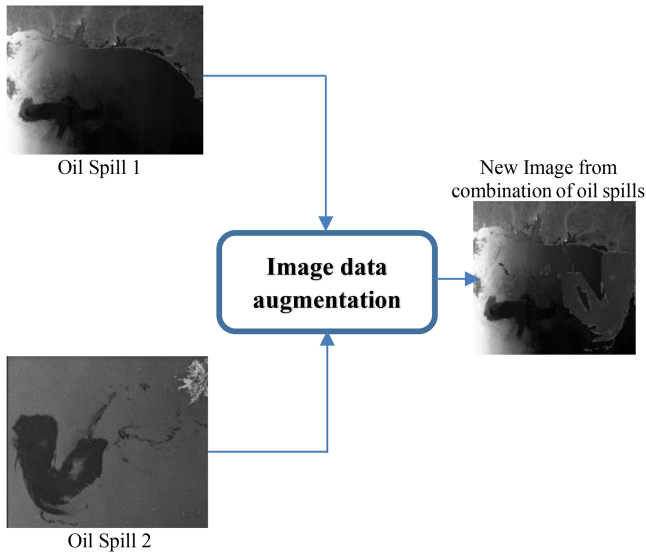


Fig. 3. Image data augmentation technique.

be exposed to numerous input-output pairs that are segmentation masks and SAR images for oil spill detection. Although unlabeled data are inexpensive and accessible in large degree, there are normally scarce labels expensive for achievement [30]. By augmenting the dataset during training, overfitting the training data was avoided in the model, and the generalization ability of the model was enhanced over invisible instances. The generalization performance can be improved by randomized data augmentation in various computer vision tasks, such as applications on remote sensing [29]. Then, using the image data augmentation technique, we increased the SAR oil spill images to 9801 images. We needed many images to train convolutional networks to avoid overfitting. Next, we trained the images to U-NET [31] and DEEPLABV3 networks separately, and at the end, we gave some images to the network as a test sample. The oil spill segmentation images were also displayed. The block diagram of the proposed flowchart is shown in Fig. 4.

The encoder-decoder networks have been applied successfully to several computer vision tasks, such as human pose estimation [32], object detection [20], [33], [34], and semantic segmentation [35], [36], [37], [38], [39], [40], [41], [42]. Characteristically, encoder-decoder networks comprise: an encoder module gradually reducing feature maps and capturing higher semantic information and a decoder module gradually recovering spatial information. Hence, we proposed to utilize DeepLabv3 as the encoder module and added an effective yet simple decoder module to achieve sharper segmentations [43].

IV. EXPERIMENTAL RESULTS

Although unlabeled data are inexpensive and accessible in large degree, there are normally scarce labels expensive for achievement [30]. By augmenting the dataset during training, overfitting the training data was avoided in the model, and the generalization ability of the model was enhanced over invisible instances. The generalization performance can be improved by

randomized data augmentation in various computer vision tasks, such as applications on remote sensing [29].

Then, using the image data augmentation technique, we increased the SAR oil spill images to 9801 images. We needed many images to train convolutional networks to avoid overfitting. Next, we trained the images to U-NET [31] and DEEPLABV3 networks separately, and at the end, we gave some images to the network as a test sample. The oil spill segmentation images were also displayed. The block diagram of the proposed flowchart is shown in Fig. 4.

The encoder-decoder networks have been applied successfully to several computer vision tasks, such as human pose estimation [32], object detection [20], [33], [34], and semantic segmentation [35], [36], [37], [38], [39], [40], [41], [42]. Characteristically, encoder-decoder networks comprise: an encoder module gradually reducing feature maps and capturing higher semantic information and a decoder module gradually recovering spatial information. Hence, we proposed to utilize DeepLabv3 as the encoder module and added an effective yet simple decoder module to achieve sharper segmentations [43].

A. SAR Oil Spill Detection Results Using U-NET Network

We considered the size of all the oil spill images at 1024×1024 pixels. We programmed the database address as/content/drive/My Drive/DataSets/OilSpillGen.

Multiplying the original images by the annotated images gave us original fragmented images. As seen, an oil spill was augmented to the image, and in the annotated images, both parts of the mask oil spill were removed. This algorithm was used in the U-NET network application code. As observed in Fig. 5, the final oil spills had a black background and the spots were quite clear. The network was trained with 90% of the input images. Here, we considered 90% of the original images for training and 10% for testing. Then, after the image augmentation technique, we used 10% of this educational data for validation. Identifying oil spills means that areas containing oil spills surrounding the SAR image are entirely isolated. The U-NET network training results in Table I are given here. Then, images with 300 epoch and a batch size of 5 were considered for network training. As observed, in the epoch $i = 22$, the values of the parameters were $\text{mean_iou} = 0.9058$ and $\text{val_mean_iou} = 0.7888$, which had the closest and highest values. The proximity of these two parameters' values showed the good training of the network, and the value of val_mean_iou showed the oil spill detection accuracy by the network, which was among the highest values (0.788 or 78.8%). Furthermore, we saved weights obtained from the network training in the program algorithm so that no time was spent on re-testing the network training, and we could use the weights. In Table I, the loss showed the network error value, and val_loss showed the validation error value. Additionally, the mean_iou parameter indicated the oil spill detection accuracy, and the val_mean_iou grid showed the network accuracy for validation data. Diagrams of network training parameters are shown below.

We observed the loss diagram based on the number of epochs in the U-NET network. It calculated the output or prediction with

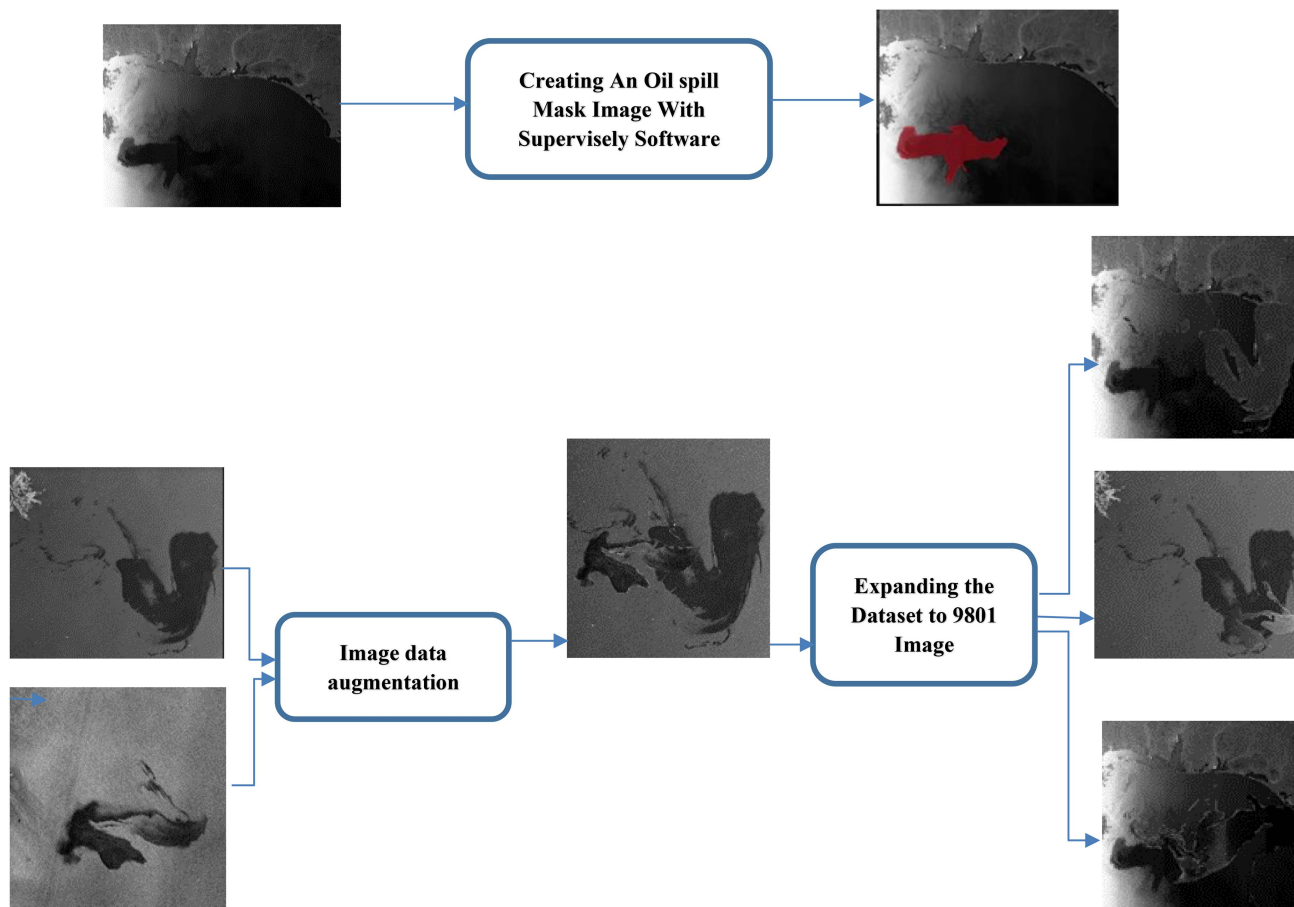


Fig. 4. Proposed flowchart.

TABLE I
DATA FROM U-NET NETWORK TRAINING WITH OIL SPILL IMAGES

I=epoch	loss	Val_loss	Mean_iou	Val_mean_iou
1	0.08196484907353759	0.10172418751146482	0.788085547260455	0.6894875977350318
2	0.07191911635357448	0.1707098355759745	0.8154206526175185	0.5331065761006397
3	0.0636149129737078	0.10124399935745675	0.8348091069764599	0.659610287780347
4	0.05951163240059472	0.10711092430612315	0.8466555221609054	0.6737405160199041
5	0.05413024892634101	0.11803870252940966	0.858266018512066	0.6455114641915197
6	0.051569041453910114	0.11522038967713066	0.8650610864857543	0.6533996372119241
7	0.048948323097083614	0.10665370880261711	0.8713654678231657	0.6717446394588636
8	0.04688558819096452	0.08674332700345827	0.8755788798466599	0.7051838105139525
9	0.04763690522800411	0.07664671430121298	0.8775860735000743	0.7374000139858412
10	0.04267840763414595	0.08835555198399916	0.8869562840091012	0.7218527955205544
11	0.043157789057918956	0.11547743166270463	0.8863751553469659	0.6410029398358387
12	0.04242155753778415	0.07814529988428821	0.8885653570974997	0.7000519340452941
13	0.041096594093224456	0.08591998384698578	0.8921406724941858	0.695846459399099
14	0.040715936151404654	0.06834082980842694	0.8928562622260305	0.7477461291396099
15	0.038037953149813314	0.08196463218849638	0.8989827165450717	0.6863521249397941
16	0.0383271293911949	0.057220512741933696	0.9003753101976318	0.7802569477454476
17	0.037384332215596276	0.07280522163795389	0.9018070927745978	0.7369055685789688
18	0.038210178180566046	0.07637368942732396	0.9011538861676023	0.7157978314420451
19	0.036782321111960704	0.06423269145838592	0.9048507175362145	0.7511574522308682
20	0.036091085260895064	0.10402997411463571	0.9066292671350974	0.6730545497458914
21	0.03591657073333845	0.07320086964122627	0.907800073700118	0.7305297841196475
22	0.036507004616733195	0.05187258883986784	0.9058043515369426	0.7888550820557968

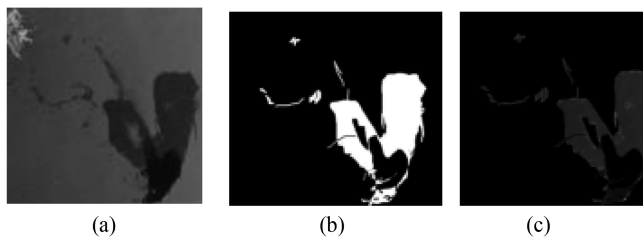


Fig. 5. (a) Original oil spill image. (b) Oil spill annotated image. (c) Multiplying 2 side images with original fragmented images and U-NET network.

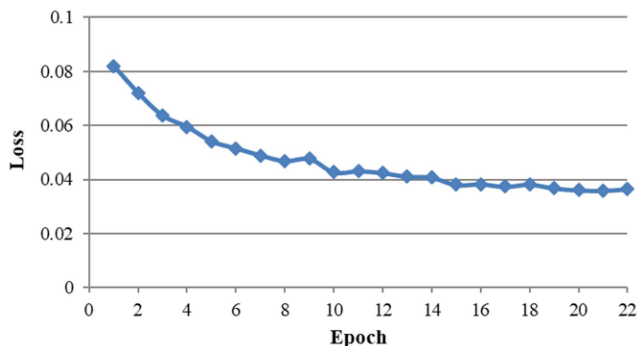


Fig. 6. U-NET network Loss diagram.

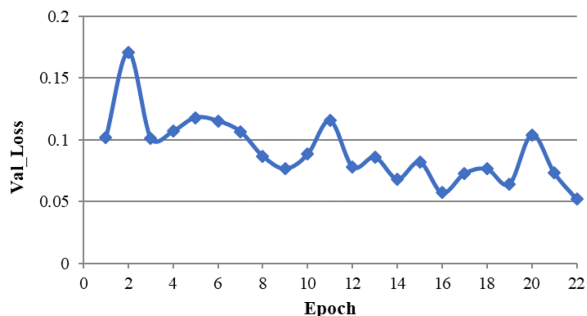


Fig. 7. U-NET network Val_Loss diagram.

the actual value according to the formula and obtained the error, and the network reduced this value during training by backpropagation and weight adjustment. In Fig. 6, we observed that as the number of epoch iterations increased, the error decreased and reached a constant value.

We observed the error value in the validation data. As the epochs increased in Fig. 7, the error reached a fixed state of 0.05.

The mean_iou parameter is used in detection and segmentation studies. Iou means that we had the main ground prediction range. The prediction detection range was also introduced. We intended to measure the extent to which these two areas were fit. The parameter was more fit if it was closer to one. The mean_iou parameter means that the entire test data will be trained. The mean of the data is the mean_iou value. This diagram showed that the mean_iou value in the U-NET network increased with the increase in epoch and reached a constant value of about 0.9, as shown in Fig. 8.

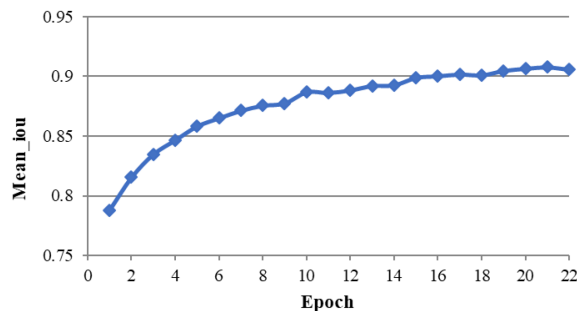


Fig. 8. U-NET network Mean_iou diagram.

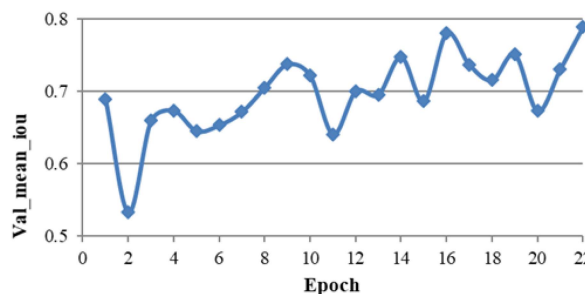


Fig. 9. U-NET network Val_mean_iou diagram.

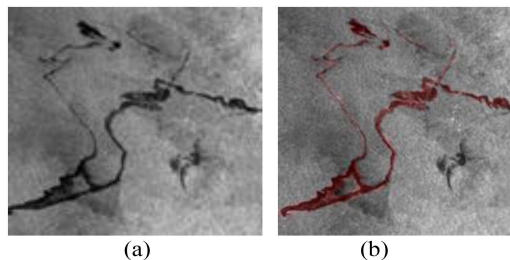


Fig. 10. (a) Original oil spill 1. (b) Oil spill detection 1 with U-NET network.

The val_mean_iou parameter calculated the men_iou value in the validation data, reaching around 0.8 with the increase of epoch, as shown in Fig. 9. Network testing was performed on three images for which the network was not yet trained. Parts of the oil spills were highlighted in red.

The target image was developed at 723×543 pixels on November 20 at ESRIN, Italy. As shown in Fig. 10, the scattering oil spill was not collected in one part, and the U-NET network managed to identify areas containing oil and even narrow oil borders.

In Fig. 11, the oil spill accumulated in one part and was not dispersed. The U-NET network identified oil spills well in the tested image and was marked in red with full details.

Fig. 12 shows the detection on November 20 at ESRIN, Italy. As seen, the oil spill had a closed-loop shape, and there were clear areas inside the loop with no oil spill. The U-NET network recognized well around the thick borders of the spot, and the borders were well displayed.

Deeplabv3+, an encoding-decoding DCNN, extended Deeplabv3 by adding a simple yet effective decoder module to

TABLE II
DATA FROM DEEPLABV3 NETWORK TRAINING WITH OIL SPILLS IMAGES

i =epoch	loss	Val loss	Mean_iou	Val_mean_iou
1	0.8032314038	0.3932342112	0.454296371	0.4567289326
2	0.4542275932	0.5280497518	0.4583567742	0.4579200217
3	0.3700664638	0.8705132067	0.4583983907	0.4619714612
4	0.3621124886	0.4383109769	0.4650602353	0.4664458355d
5	0.3507502279	0.4973744184	0.4677395403	0.4702208513
6	0.3431110072	0.4065105887	0.7068734789	0.4733036327
7	0.3100349888	0.4726145835	0.4744875705	0.4761986917
8	0.290338994	0.638282671	0.4780522391	0.4808789197
9	0.2753489763	0.2469013497	0.4834528753	0.4857358956
10	0.2685007556	0.4828741257	0.487895129	0.4899188921
11	0.2919004406	1.153913648	0.4916976067	0.4945708647
12	0.2714103538	0.5199286111	0.497454083	0.4994683668d
13	0.2536790152	1.299947506	0.5010962749	0.5031436205
14	0.2511523475	0.3945304415	0.5049613637	0.5063102776
15	0.2474990679	0.5756871331	0.5076857179	0.5091737151
16	0.2415248938	0.5984489825	0.5105202883	0.5123457861
17	0.2695244965	0.8850448307	0.5142025834	0.5138342136
18	0.2964571683	0.3173567598	0.5133885455	0.512697407
19	0.2525779346	0.898651104	0.5124000138	0.5132769257
20	0.2408069737	0.4714079872	0.5140178508	0.5145885998
49	0.1589323011	0.5833733169	0.5413329709	0.541946324

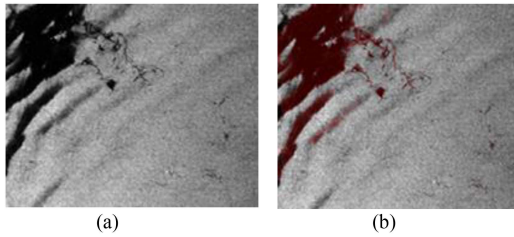


Fig. 11. (a) Original oil spill 2. (b) Identification of oil spill 2 with U-NET network.

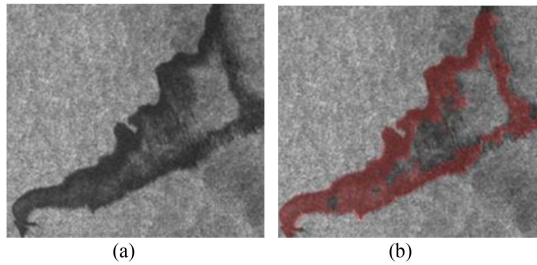


Fig. 12. (a) Original oil spill 3. (c) Oil spill 3 detection with U-NET network.

refine the segmentation results, especially along object boundaries, which vastly improved semantic image segmentation. Due to the unique radar imaging mechanism of SAR, the SAR image structure is complex, and the content is extremely rich, making the semantic segmentation of SAR imagery more difficult than that of optical imagery [44]. We considered the size of all the oil spill images at 512×512 pixels. We provided the address of the database in the program.

B. SAR Oil Spill Detection Results Using DeepLabV3 Network

The DeeplabV3 network tutorial results are given in Table II. Here, images with 300 epochs and a batch size of 4 were

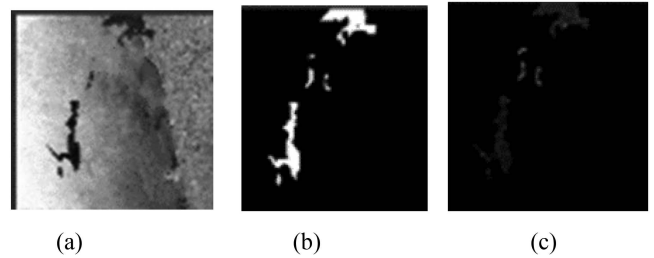


Fig. 13. (a) Original oil spill image. (b) Oil spill annotated image. (c) Multiplying 2 side images with original fragmented images and DeepLabv3 network.

considered for network training. As observed, in the epoch $i = 49$, the values of the parameters were $\text{mean_iou} = 0.5413$ and $\text{val_mean_iou} = 0.5419$, with the closest and highest values. The proximity of these two parameters' values indicated the good training of the network, and the val_mean_iou value showed the oil spill detection accuracy by the network, which was among the highest values (0.54 or 54%).

Multiplying the original images by the annotated images gave us original fragmented images. As shown, an oil spill was augmented to the image, and in the annotated images, both parts of the mask oil spill were removed. The same algorithm was used in the DeepLabV3 network application code. As shown in Fig. 13, the final oil spills had a black background, and the oil spills were quite clear. The network was trained with 90% of the overview images. Here, we considered 90% of the original images for training and 10% for testing. Then, after the augmentation technique, we used 10% of this educational data for validation. Identifying oil spills means that areas containing oil spills surrounding the SAR image are entirely isolated.

As shown in Fig. 14, the Deeplabv3 network error value decreased with an increase in the number of epochs and reached a constant value.

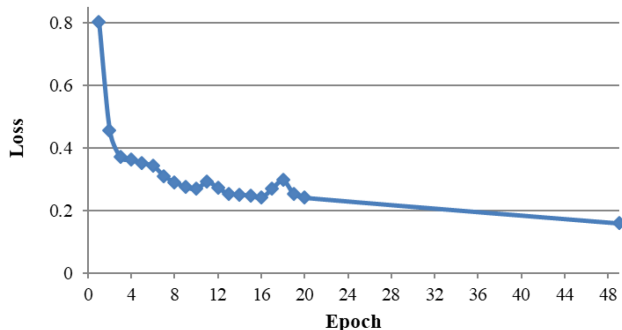


Fig. 14. DeepLabV3 network Loss diagram.

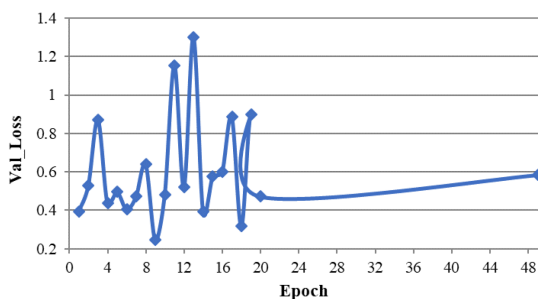


Fig. 15. DeepLabV3 network Val_Loss diagram.

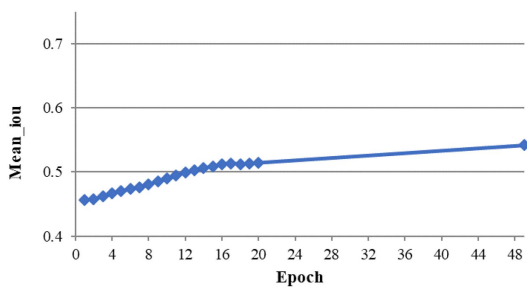


Fig. 16. DeepLabV3 network Mean_iou diagram.

The Deeplabv3 network validation data error value reached a value between 0.5 and 1 after a series of fluctuations around the value of 1. There were more errors in this network compared to U-NET. The loss val diagram in the DeepLabV3 network is shown in Fig. 15.

The mean_iou diagram in the Deeplabv3 network reached a fixed value of 0.5 after increasing epochs compared to that in the U-NET network. Accordingly, the network was weaker in the detecting oil spill segmentation. The mean_iou diagram in the DeepLabV3 network is shown in Fig. 16.

As the diagram in Fig. 17 shows, with the increase of epochs, the mean_iou value in the validation data increased to 0.55, which was lower than the U-NET value, and segmentation in the validation data in Deeplabv3 had a more unsatisfactory performance. Network testing was performed on three images for which the network was not yet trained. Parts of the oil spills were highlighted in red.

The target image was produced at 723×543 pixels on November 20 at ESRIN, Italy. As the image shows, the oil spill

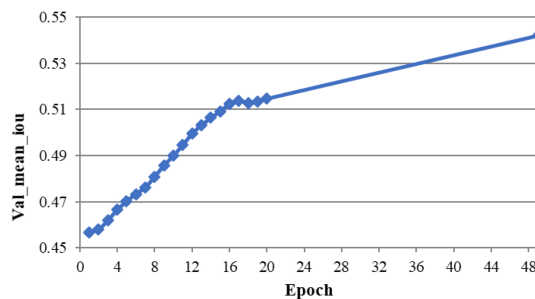


Fig. 17. DeepLabV3 network Val_mean_iou diagram.

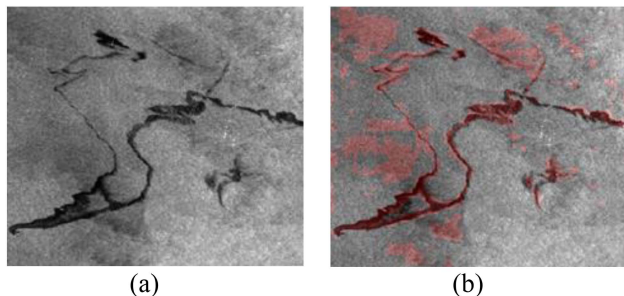


Fig. 18. (a) Original oil spill 1. (b) Oil spill 1 detection with DeepLabV3.

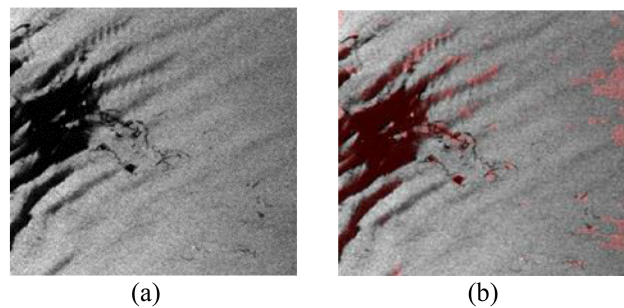


Fig. 19. (a) Original oil spill 2. (c) Oil spill 2 detection with DeepLabV3.

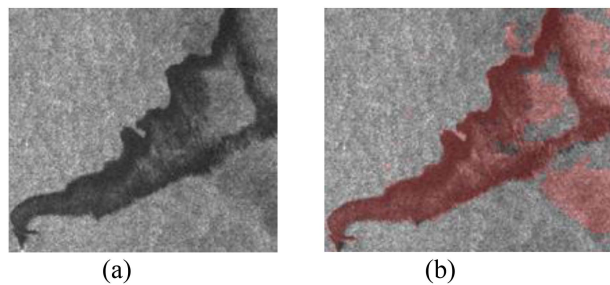


Fig. 20. (a) Original oil spill 3. (d) Oil spill 3 detection with DeepLabV3.

was diffused and did not shrink in one part, and the DeepLab3 network had moderately identified areas with oil spills and narrow oil borders. As shown in Fig. 18, the left part of the image had marked parts with a pale red in addition to the oil spill, which was considered a problem.

In Fig. 19, the oil spill accumulated in one part and was not dispersed.

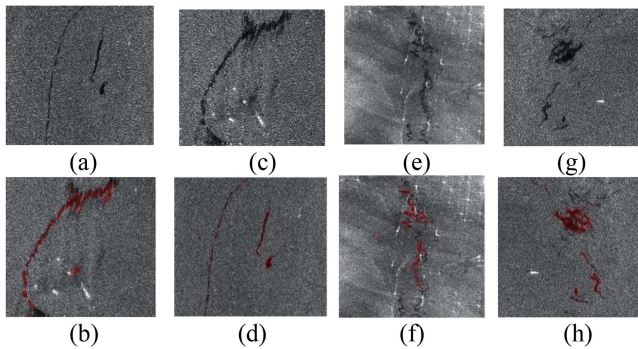


Fig. 21. Segmentation results of oil spill SAR images with noise. (a) Original SAR oil spill-1. (b) Segmented oil spill-1 image with U-NET. (c) Original SAR oil spill-2. (d) Segmented oil spill-2 image with U-NET. (e) Original SAR oil spill-3. (f) Segmented oil spill-3 image with U-NET. (g) Original SAR oil spill-4. (h) Segmented oil spill-4 image with U-NET.

Fig. 20 shows the detection on November 20 at ESRIN, Italy. As observed, the oil spill had a closed-loop shape, and there were clear areas inside the loop without the oil spill. The U-NET network recognized well around the thick borders of the spot, and the borders were well shown. However, the network detected excess oil spills inside the loop, on the border, and outside the loop.

In this article, the following hyperparameters were used in U-NET and DeepLabv3 networks: optimizer; activation function; learning rate; epochs; and batch size. In DeepLabv3 and U-NET networks, the activation function parameter is equal to ReLU. In U-NET, the optimizer parameter is equal to RMSprop, and DeepLabv3 networks are equal to Adam. The learning rate parameter in DeepLabv3 and U-NET is equal to 0.001. In U-NET, the batch size parameter is 5, while in DeepLabv3, it is 4. The epoch parameter in U-NET and DeepLabv3 equals 300. In other words, during the test, the hyperparameters did not change.

Considering a prominent feature of oil spill segmentation with this technique and networks, it can be concluded that the generality of this method was examined. Hence, noisy images, such as speckle noise, were tested, yielding accurate segmentation results, which are considerably displayed here.

The types of the oil spill segmentation results of SAR radar are presented in Fig. 21, along with the speckle noise of the unit network. Remarkable accuracy was obtained for the segmentation results. It should be noted that for both networks, we reviewed all the assessed output data in 300 epochs. Moreover, after 20 epochs, the recorded epoch had the highest accuracy due to not enough space in the text of the article to include 300 epochs.

Considering the future challenges and motivations, it is worth noting that oil spill segmentation was investigated in this article by using CNN networks. For the next ideas, the number of input images of the networks should increase; for instance, data transmission satellites should be checked, which comprise more images containing oil spills to increment the accuracy of the segmentation. The neural network architecture can be deepened at the same time to obtain better accuracy. Also, the neural

network with optimal settings should be updated; for instance, the optimizer parameter or learning rate should be changed.

There is a problem with spectral variability for hyperspectral images gathered from airborne or satellite sources, inevitably making it difficult to accurately estimate the spectral mixing. An advanced ALMM linear mixing model was introduced by Heng et al., to cope with spectral variability by using a data-based learning strategy for hyperspectral mixing inverse problems. Then, other spectral variables were modeled, such as temperature, local humidity, and atmospheric influence, as well as instrument settings such as noise and nonlinear effects. The ALMM model considered both the main scale factor and other spectral variables by introducing the spectral diversity dictionary to increment the end member dictionary scalability. More importantly, the presented technique can achieve more accurate frequency estimation than other advanced algorithms because it models spectral variables separately as scale and other spectral variables based on distinctive features [45]. Radar images are formed by the coherent interaction of transmitted microwaves with targets, unlike optical images. Therefore, the speckle noise effect is caused by the coherent summation of randomly scattered signals in each pixel. There is more noise in radar images than in optical images. SAR images are degraded inherently due to the coherent nature of the scattering phenomena known as speckle. The utility of the SAR images is reduced by the presence of speckles by reducing the capability to discover ground objects, which has adverse effects on the image quality and hampers the observation of crucial information in the image. In this article, noise-free images of SAR radar oil spills were considered for the training and input of the network. The inclusion of noise in SAR radar oil spill images should be assessed as a new issue. Here, the test data for the U-NET and DeepLabv3 networks were the oil spill images of noisy radar. As seen, good results are obtained for oil spill segmentation with noise by using the U-NET network.

V. CONCLUSION

Deep network training needs a significant sum of data and usually faces the overfitting issue with a small volume of data. In this article on SAR images, our data was limited, and it was shown that with this number of deep segmentation networks, the networks faced the overfitting issue and were not efficient. In overcoming this problem, one approach is to use augmentation techniques, although the desired number of several thousand images cannot be reached with these techniques. The innovative task carried out in this article was to combine the existing images to produce a large number of images. For instance, here, the number of original images was 99, where 98×98 images were produced using this approach, as each oil spill of the image can be placed on 98 other images. With this number of new images, we managed to overcome the overfitting issue. As stated in the previous sections about the training and testing results of SAR oil spill images using the U-NET and DeepLab3 networks, U-NET and DeepLabV3 could detect oil spills with 78.8% and 54% accuracy, respectively, and the U-NET network had the highest detection accuracy. Given the architecture and

structural features of the U-NET network, with 300 epochs and testing of images, the network lasted about 8 h on the Google Colab server and performed oil spill separation and detection acceptably. However, in addition to detecting oil spills, the DeepLabV3 network indicated additional areas as oil spills, which were network errors. The oil spill detection accuracy reached about 78.8% using the U-NET network. Regarding the remaining motivations and challenges, it should be said that this article investigated oil spill segmentation by using CNN networks. For the next ideas, it is better to increase the number of input images of the networks, for instance, to check data transmission satellites containing more images of oil spills and to increase the segmentation accuracy. At the same time, the neural network architecture can be deepened or updated by optimal settings (e.g., changing the optimizer parameter or learning rate) to achieve better accuracy. The following are some approaches to improve the accuracy of oil spill segmentation: In other words, increasing the input data will lead to more training of neural networks, and as a result, the accuracy of oil spot spill segmentation will be improved. Another idea can be mentioned to increase the number of epochs and further training. Again, more CNN training will lead to better accuracy. Here, it can be said that using a deeper and different neural network will definitely increase the accuracy of oil spill segmentation. Finally, updating and optimizing hyperparameters (such as the optimizer and learning rate) will contribute significantly to improving the accuracy of oil spill segmentation.

REFERENCES

- [1] A. Raeisi, G. Akbarizadeh, and A. Mahmoudi, "Combined method of an efficient cuckoo search algorithm and nonnegative matrix factorization of different zernike moment features for discrimination between oil spills and lookalikes in SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 11, pp. 4193–4205, Nov. 2018.
- [2] T. M. Alves, E. Kokinou, G. Zodiatis, R. Lardner, C. Panagiotakis, and H. Radhakrishnan, "Modelling of oil spills in confined maritime basins: The case for early response in the eastern Mediterranean sea," *Environ. Pollut.*, vol. 206, pp. 390–399, Nov. 2015.
- [3] R. Lardner and G. Zodiatis, "Modelling oil plumes from subsurface spills," *Mar. Pollut. Bull.*, vol. 124, no. 1, pp. 94–101, Nov. 2017.
- [4] T. Soomere et al., "The potential of current- and wind-driven transport for environmental management of the Baltic sea," *Nat. Center Biotechnol. Inf., Ambio*, vol. 43, no. 1, pp. 94–104, Feb. 2014.
- [5] A. Chrastansky and U. Callies, "Model-based long-term reconstruction of weather-driven variations in chronic oil pollution along the German north sea coast," *Mar. Pollut. Bull.*, vol. 58, no. 7, pp. 967–975, Jul. 2009.
- [6] S. Singha, T. J. Bellerby, and O. Trieschmann, "Satellite oil spill detection using artificial neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 6, pp. 2355–2363, Dec. 2013.
- [7] K. Topouzelis, V. Karathanassi, P. Pavlakakis, and D. Rokos, "Detection and discrimination between oil spills and look-alike phenomena through neural networks," *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 4, pp. 264–270, Sep. 2007.
- [8] J. S. Murgan and V. Parthasarathy, "AETC: Segmentation and classification of the oil spills from SAR imagery," *Environ. Forensics*, vol. 18, no. 4, pp. 258–271, Nov. 2017.
- [9] H. Guo, G. Wei, and J. An, "Dark spot detection in SAR images of oil spill using segnet," *Appl. Sci.*, vol. 8, no. 12, pp. 1–17, Dec. 2018.
- [10] P. Wang, H. Zhang, and V. M. Patel, "SAR image despeckling using a convolutional neural network," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1763–1767, Dec. 2017.
- [11] G. Chen, Y. Li, G. Sun, and Y. Zhang, "Application of deep networks to oil spill detection using polarimetric synthetic aperture radar images," *Appl. Sci.*, vol. 7, no. 10, pp. 1–15, Sep. 2017.
- [12] J. F. Penner and D. G. Long, "Ground-Based 3D radar imaging of trees using a 2D synthetic aperture," *Electronics*, vol. 6, no. 1, pp. 1–13, Jan. 2017.
- [13] F. M. Bianchi, O. Martine, M. Espeseth, and N. Borch, "Large-Scale detection and categorization of oil spills from SAR images with deep learning," *Remote Sens.*, vol. 12, no. 14, Jul. 2020, Art. no. 2260.
- [14] D. Malmgren-Hansen and M. Nobel-Jorgensen, "Convolutional neural networks for SAR image segmentation," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, 2015, pp. 231–236.
- [15] J. Zhang, H. Feng, Q. Luo, Y. Li, J. Wei, and J. Li, "Oil spill detection in quad-polarimetric SAR images using an advanced convolutional neural network based on superpixel model," *Remote Sens.*, vol. 12, no. 6, pp. 1–26, Mar. 2020.
- [16] R. Li et al., "DeepUNet: A deep fully convolutional network for pixel-level sea-land segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 11, pp. 3954–3962, Nov. 2018.
- [17] B. Liu, Y. Li, G. Li, and A. Liu, "A spectral feature based convolutional neural network for classification of sea surface oil spill," *ISPRS, Int. J. Geo-Inf.*, vol. 8, no. 4, pp. 1–14, Mar. 2019.
- [18] H. Guo, D. Wu, and J. An, "Discrimination of oil slicks and lookalikes in polarimetric SAR images using CNN," *Sensors*, vol. 17, no. 8, pp. 1–20, Aug. 2017.
- [19] Y. Duan, F. Liu, L. Jiao, P. Zhao, and L. Zhang, "SAR image segmentation based on convolutional-wavelet neural network and Markov random field," *Pattern Recognit.*, vol. 64, pp. 255–267, Apr. 2017.
- [20] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 936–944.
- [21] C. Brekke and A. H. Solberg, "Oil spill detection by satellite remote sensing," *Remote Sens. Environ.*, vol. 95, no. 1, pp. 1–13, Mar. 2005.
- [22] O. Garcia-Pineda, I. R. MacDonald, X. Li, C. R. Jackson, and W. G. Pichel, "Oil spill mapping and measurement in the gulf of Mexico with textural classifier neural network Algorithm(TCNNA)," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 6, pp. 2517–2525, Dec. 2013.
- [23] M. Krestenitis, G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis, and I. Kompatsiaris, "Oil spill identification from satellite images using deep neural networks," *Remote Sens.*, vol. 11, no. 15, pp. 1–22, Jul. 2020.
- [24] F. Ma, F. Zhang, D. Xiang, Q. Yin, and Y. Zhou, "Fast task-specific region merging for SAR image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5222316.
- [25] F. Ma, F. Zhang, D. Xiang, Q. Yin, and Y. Zhou, "Fast SAR image segmentation with deep task-specific superpixel sampling and soft graph convolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5214116.
- [26] D. Hong et al., "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4340–4354, May 2021.
- [27] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Jul. 2021.
- [28] D. Hong et al., "SpectralFormer: Rethinking hyperspectral image classification with transformers," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–15, Nov. 2021, Art. no. 5518615, doi: [10.1109/TGRS.2021.3130716](https://doi.org/10.1109/TGRS.2021.3130716).
- [29] L. Xu, M. J. Shafiee, A. Wong, F. Li, L. Wang, and D. Clausi, "Oil spill candidate detection from SAR imagery using a thresholding-guided stochastic fully-connected conditional random field model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2015, pp. 79–86.
- [30] Y. Bengio, A. C. Courville, and P. Vincent, "Unsupervised feature learning and deep learning: A review and new perspectives," Jun. 2012, *arXiv:1206.5538v1[cs.LG]*.
- [31] Q. Fan et al., "Ship detection using a fully convolutional network with compact polarimetric SAR images," *Remote Sens.*, vol. 11, no. 18, pp. 1–15, Sep. 2019.
- [32] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," in *Proc. Eur. Conf. Comput. Vis., Comput. Vis.*, Sep. 2016, pp. 483–499.
- [33] A. Shrivastava, R. Sukthankar, J. Malik, and A. Gupta, "Beyond skip connections: Top-down modulation for object detection," Sep. 2017, *arXiv:1612.06851v2[cs.CV]*.
- [34] C. Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, "Dssd: Deconvolutional single shot detector," Jan. 2017, *arXiv:1701.06659v1[cs.CV]*.
- [35] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3431–3440.
- [36] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 2380–2504.

- [37] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
- [38] T. Pohlen, A. Hermans, M. Mathias, and B. Leibe, "Full-resolution residual networks for semantic segmentation in street scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 3309–3318.
- [39] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, "Large kernel matters—improve semantic segmentation by global convolutional network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1743–1751.
- [40] M. A. Islam, M. Rochan, N. D. B. Bruce, and Y. Wang, "Gated feedback refinement network for dense image labeling," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4877–4885.
- [41] J. Fu, J. Liu, Y. Wang, J. Zhou, C. Wang, and H. Lu, "Stacked deconvolutional network for semantic segmentation," *IEEE Trans. Image Process.*, early access, Jan. 2019, doi: [10.1109/TIP.2019.2895460](https://doi.org/10.1109/TIP.2019.2895460).
- [42] Z. Zhang, X. Zhang, C. Peng, X. Xue, and J. Sun, "Exfuse: Enhancing feature fusion for semantic segmentation," in *Proc. 15th Eur. Conf., Comput. Vis.*, Sep. 2018, pp. 273–288.
- [43] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 801–818.
- [44] Y. Kong, Y. Liu, B. Yan, H. Leung, and X. Peng, "A novel deeplabv3+ network for SAR imagery semantic segmentation based on the potential energy loss function of Gibbs distribution," *Remote Sens.*, vol. 13, no. 3, pp. 454–466, Jan. 2021.
- [45] D. Hong, N. Yokova, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.



Fatemeh Mahmoudi Ghara was born in Ghom, Iran, on May 16, 1986. She received the B.S. degree in electronic engineering from Semnan University, Semnan, Iran, in 2009, and the M.S. degree in digital electronic systems engineering from Iran University of Science and Technology, Tehran, Iran, in 2019, respectively.

Her research interests include image processing, machine vision, and machine learning.



Shahriar Baradaran Shokouhi (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from the Iran University of Science and Technology (IUST), Tehran, Iran, in 1986 and 1989, respectively, and the Ph.D. degree in electronic and electrical engineering from the University of Bath, Bath, U.K., in 1999.

He was a Research Associate with the Department of Electrical and Computer Engineering, Western University, London, ON, Canada, from 2012 to 2014.

He is currently a Faculty Member with the Department of Electrical Engineering, IUST. His current research interests include embedded systems, trusted hardware designs, machine vision, and intelligent system designs.

Dr. Shokouhi is a Member of the Circuit and Systems Society.



Gholamreza Akbarizadeh was born in Shiraz, Iran, on July 14, 1981. He received the B.S. degree from the Khajeh-Nassir Tousei University of Technology, Tehran, Iran, in 2003 and the M.S. and Ph.D. degrees from the Iran University of Science and Technology, Tehran, Iran, in 2005 and 2011, respectively, all in electrical and electronics engineering.

He is currently an Associate Professor and the Faculty Dean with the Electrical Engineering Department, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran. He has authored and

co-authored more than 70 papers in different international electrical engineering conferences and journals. His research interests include machine vision, image processing, and remote sensing analysis.

Dr. Akbarizadeh is a Member of some international scientific societies such as Iranian Society of Machine Vision and Image Processing. He is currently an Associate Editor for *IET Image Processing*, *International Journal of Remote Sensing*, and the Reviewer of some international journals, such as *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*, *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, *IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS*, *IEEE TRANSACTIONS ON IMAGE PROCESSING*, *Measurement*, *ISPRS Journal of Photogrammetry and Remote Sensing*, etc.