

A DEEP NEURAL NETWORK FOR OIL SPILL SEMANTIC SEGMENTATION IN SAR IMAGES

Georgios Orfanidis, Konstantinos Ioannidis, Konstantinos Avgerinakis, Stefanos Vrochidis, Ioannis Kompatsiaris

Centre for Research and Technology Hellas (CERTH)-Information Technologies Institute (ITI)

ABSTRACT

Oil spills pose a major threat of the oceanic and coastal environments, hence, an automatic detection and a continuous monitoring system comprises an appealing option for minimizing the response time of any relevant operation. Numerous efforts have been conducted towards such solutions by exploiting a variety of sensing systems. Previous studies, including neural networks, have shown that the use of satellite Synthetic Aperture Radar (SAR) can effectively identify oil spills over sea surfaces in any environmental conditions and operational time. Moreover, in recent years, deep Convolutional Neural Networks (CNN) have presented some remarkable abilities to surpass previous state-of-the-Art performances in a great diversity of fields including identification tasks. This paper describes the development of an approach that combines the merits of a deep CNN with SAR imagery in order to provide a fully automated oil spill detection system. The deployed CNN was trained using multiple SAR images acquired from the sentinel-1 satellite provided by ESA and based on EMSA records for maritime pollution events. Experiments on such challenging benchmark datasets demonstrate that the algorithm can accurately identify oil spills leading to an effective detection solution.

Index Terms— Oil pollution, Synthetic aperture radar, Convolutional neural networks.

1. INTRODUCTION

Oil spill pollution is tightly connected not only with the ocean ecosystem but also with the increase of maritime commerce and activities. Since early measures in such cases are of major importance, numerous algorithms have been presented to accurately and automatically identify such pollution spots. The vast majority of relevant methods exploits data acquired from satellites which are equipped with Synthetic Aperture Radar (SAR) capabilities due to the advantages they display. More specific, such satellites can cover large areas of interest without the necessity of deploying extra equipment and vehicles,

while on the other hand SAR imagery is indispensable because it manifests light and weather condition invariability.

A typical process for oil spill detection can be concluded into four separate stages [1]. The first stage includes the detection of the dark formations in SAR images while in the second process, features for such formations are extracted. Sequentially, the extracted features are compared with some predefined values and a decision making model follows for labeling each formation. The approach poses several disadvantages, mainly due to the necessity of extracting a number of features, the non-unanimous agreement over their nature and the lack of research over their effectiveness. In addition, the limitation of providing a single label to each of the input image poses constraints regarding the identified objects.

The most common approaches involve a two-class classification process, one class that includes only oil spills and a second, more diverse class that corresponds to dark formations of all similar phenomena that resembles oil spills. The latter class can be further divided into subclasses like current shear, internal waves, algae blooms, shoals, floating vegetation and grease ice [2]. Also, contextual information, like existence of similar formations, ship routes etc around the detected instances may affect significantly the final characterization of these "dark spots".

In contrast to the above approach, the presented work aims at deploying a deep Convolutional Neural Network (CNN) in the field of oil spill detection in order to alleviate the corresponding shortcomings. While there have been some efforts dealing with oil spill detection using conventional neural networks [1, 3], spatial features are required to be initially computed. Also these algorithms were limited to image classification, i.e. labeling the whole image and not semantic segmented territories inside the image. The problem can be considered as a rather complicated task since it requires specialized knowledge of the various phenomena related to the marine environments while optical means often fail to provide adequate solutions. Thus, the final pollution affirmation includes the mobilization of the relevant national/regional authorities and in situ identification. The proposed approach comprises a promising solution that could accurately discriminate oil spills from similar instances without the prerequisite of extracting additional features leading to a completely automatic detection system.

This work was supported by ROBORDER and EOPEN projects partially funded by the European Commission under grant agreements No 740593 and No 776019, respectively.

2. RELATED WORK

One of the first attempts for oil spill detection was relied on the use of visible spectrum images. Various approaches were proposed such as utilizing polarized lenses [4] and hyper-spectral imaging [5] among others. Although all relevant researches have proven that there is no wider distinction between oil and water in this spectrum, the field is active and the research is ongoing.

On the contrary, microwave sensors including radars are widely utilized for such applications in order to overcome the constraints that the optical sensors pose (weather and operation time dependency). For radar imaging, Synthetic Aperture Radar (SAR) is predominantly used [6] as they have been proven to be at a large scale invariant to light condition changes and clouds/fog occurrences [7].

Capillary waves produce "bright" image regions known as sea clutters which in the case of oil spills are depressed and appear as dark formations. This type of detection is not exclusively observed for oils spills but also include among others wind slicks, wave shadows behind land, algae blooms, land territories etc. [7, 2]. Since the identification of all the above instances requires the definition of separate classes, an acceptable simplification regarding the desired oil spill identification could include the reduction of the problem into a two class problem, i.e. oil spills vs the rest of the phenomena. In addition, until recently, oil spill detection approaches were based on the initial extraction of features that represent and/or simulate the physics of the oil dispersion. Such features include geometrical, physical or textural distinctive marks [1, 8] based on which the models are trained to identify the oil spills.

Contrary to the aforementioned approaches, the proposed method introduces a new deep CNN which does not require the extraction of any feature and can semantically annotate multiple regions in SAR images. In addition, the deployed model was properly modified aiming at reducing the total computational cost and thus, the operational time.

3. METHODOLOGY

The proposed approach for oil spill detection aims at semantically segmenting the input images and highlighting the included objects instead of just specifying a simple label to the entire representation. The assignment of multiple labels/tags in each image [9] or the extraction of bounding boxes with the use of object detection techniques [10] could be potential alternatives for the presented problem. Nonetheless, since oil spills display a large variety of irregular shapes and could be heavily intersected with look-alike objects, semantic segmentation could be considered as the most effective solution. The advantage of the approach relies on manipulating images that can contain multiple objects of different nature without the prerequisite of splitting the image into multiple image patches manipulating oil spills, look-alikes etc.

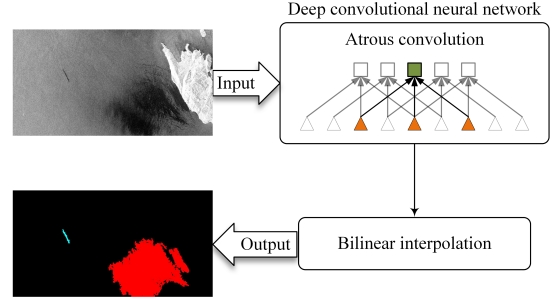


Fig. 1. High-level representation of the altered "DeepLab" model.

The model was inspired from the "DeepLab", initially proposed in [11], which has been proven to be sufficiently effective in multi-class segmentation. In the referred work, multiple experiments were conducted on a variety of methods and network models, including *VGG-16* [12] and *ResNet-101* [13] with the latter giving the best performance.

Similar to the "DeepLab", the proposed oil spill detector includes the use of a deep convolutional neural network trained in the task of semantic segmentation and the convolution with upsampled filters, which originally developed in [14] and utilized in the DCNN context by [15]. To efficiently extract the required dense features and widen the *field-of-view* of filters, atrous convolution was utilized as well as an Atrous Spatial Pyramid Pooling (ASPP) to employ parallel filters with various rates. The resulted maps are enlarged with bilinear interpolation to restore their original resolution. A higher level representation is provided in Fig. 1.

More specific, the proposed model for the required application initially uses a DCNN in a fully convolutional fashion and relies on a *ResNet-101* network model. The *ResNet-101* network was selected due to the highest detection rates it results in image semantic segmentation objectives. The model was redefined so as to fine-tuning it in our case. However, the repeated combination of max-pooling and striding at subsequent layers decreases the final resolution of the extracted feature maps and significantly increases the overall computational cost. Thus, atrous convolution was applied to explicitly control the resolution at which feature responses are extracted within DCNN. In the context of DCNNs, atrous convolution can be utilized in a chain of layers to effectively compute the networks' responses at a randomly high resolution.

For example in one-dimensional signals, the output $y[i]$ of atrous convolution of a 1- D input signal $x[i]$ with a filter $w[k]$ of length K is defined as:

$$y[i] = \sum_{k=1}^K x[i + r \cdot k]w[k] \quad (1)$$

The rate parameter r corresponds to the stride with which we sample the input signal. Basic convolution is a special case for $r = 1$.

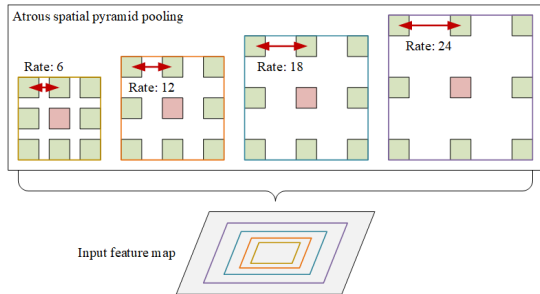


Fig. 2. Atrous Spatial Pyramid Pooling (ASPP).

Despite their ability to sufficiently represent scale by trained on multi-resolution images, DCNNs' competence for object scale can still be improved to detect both large and small objects. In applications that involve satellite image processing where operational heights vary, and oil spill detection where the size and shape of the object display extreme diversity, the scale problem can significantly affect the detection results. Therefore, to handle the scale variability, the deployed model adopted an approach which was based on an R-CNN spatial pyramid pooling method initially proposed in [16] where regions of a random scale can be efficiently classified by resampling features extracted at a single scale. These features are extracted for each sampling rate and further processed and fused to compute the final result. The deployed ASPP is depicted in Fig. 2. The final processing step involves the feature map resolution increment to restore the original resolution by applying the basic bilinear interpolation as in the "DeepLab" system. It should be highlighted that the final CRF module of the "DeppLab" system was excluded from the deployed model since it mostly used for refining the segmentation results. For the oil spill detection, the instances in SAR display vague optical limits, hence, the CRF will not significantly improve the segmented regions and thus, pointless computational overhead will be applied.

4. EXPERIMENTAL RESULTS

4.1. DATASET

One of the main challenges that the research community has to face is the lack of a publicly available dataset for such applications. Previous works [1, 17, 8] confronted this problem by utilizing a manually created dataset. Nonetheless, the comparison with other related works using typical standards is limited. These restrictions in utilizing public benchmark datasets also limited our options into collecting SAR data from European Space Agency (ESA) databases, the Copernicus Open Access Hub¹. The downloaded SAR images were acquired using the Sentinel-1 European Satellite. The required geographic coordinates and time of the confirmed

¹<https://scihub.copernicus.eu/>

Intersection-over-union (IoU)			
mIoU	Oil spills	Look-alikes	Background
0.6098	0.4130	0.4564	0.9599

Table 1. Segmentation results using mIoU/IoU.

oil spills were provided by the European Maritime Safety Agency (EMSA) based on the CleanSeaNet service and its records covering a period from 28/09/2015 up to 31/10/2017.

SAR raw data were properly preprocessed using fundamental remote sensing preprocessing algorithms including:

1. All potential oil spills were localized.
2. Cropped regions from the initial SAR data with the oil spills were extracted. All resulted images were rescaled to have the same resolution of 1252x609 pixels.
3. A radiometric calibration of the image was applied to project the images into the same plane.
4. A speckle filtering process followed to mitigate the effects of the sensor noise. A basic median filter of mask 7x7 was applied since speckle noise in remote sensing is similar to the salt-n-pepper in image processing.
5. A linear transformation from db to actual luminosity values was finally applied.

Numerous SAR images were processed in order to create a convenient database with a sufficient number of images which include confirmed oil spills, look-alikes and other geographical regions. The annotation of the images was based on information provided by EMSA and human identification (manually annotation). This process produced image masks where every desired object was marked with a distinct color (2 foreground+1 background). The processed images were randomly divided into a training and a testing set comprised of 571 and 106 images, respectively. Finally, it must be highlighted that the database is updated continuously and will be publicly available after proper confirmations to the community in order to provide a common basis as a benchmark.

4.2. RESULTS

For our experiments, two foreground classes were defined for the classification process, for oil spills and for look-alikes as well as one class for background pixels. The performance of the deployed model was initially measured in terms of pixel intersection-over-union (IoU) averaged across all classes (mIoU). In addition, the resulted IoU for each class is also provided so as to clarify the individual performance of the model and its effectiveness in each class. Table 1 includes the results from the initial experiments.

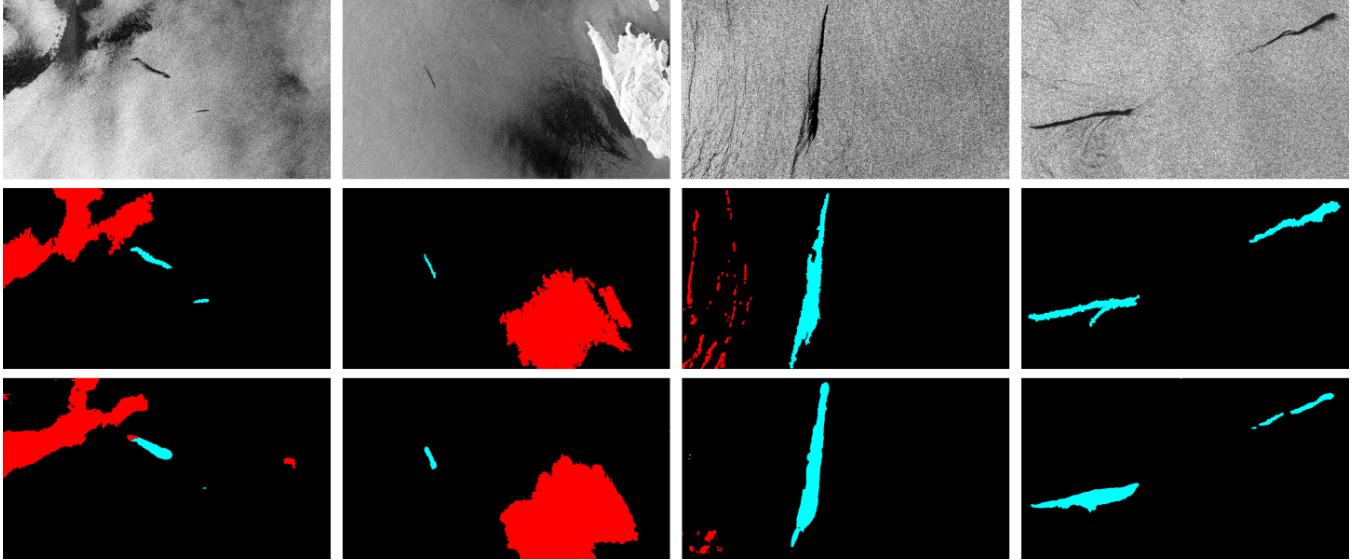


Fig. 3. Example of 4 testing images (from top to bottom): SAR images, ground truth masks and resulted detection masks.

For comparison reasons, we additionally utilized the accuracy so that the model could be somehow comparable to pure image classification models. Every image pair (ground truth and resulted detection mask) were cropped automatically in a predefined number of overlapping patches. In order to acquire a dataset which complies with the rules of image representation, some constraints were imposed:

1. A minimum number of pixels belonging to either of the two classes should be present in the patch. A threshold equal to 2% was applied meaning that the number of pixels of the largest class compared to the number of pixels of the background class should be at least 2%.
2. One of the basic classes should be dominant to label the image patch. The applied threshold was set at most 50% of the pixels of the non dominant class in relation to the pixels of the dominant class.
3. Patches recalcitrant to the above rules were discarded in accuracy calculation.

The results of patch image classification are provided in Table 2. The values are dependent of the number of patches cropped from each sample. Two different pair values for the horizontal and vertical number of patches, respectively, are presented. The results provided in Tables 1 and 2 are somehow dissimilar due to the fact that for the second metric a single label per patch is evaluated and not on every pixel.

The comparison with relevant approaches could be considered as invalid due to the lack of a common image dataset, nonetheless, some results are compared. The neural network based method in [3] reported a 91.6% and 98.3% accuracy for oil spills and look-alikes, respectively, with much higher number of look-alikes. The method in [1] using a decision

Image patch classification precision results					
Number of patches: 3,3			Number of patches: 5,3		
Overall	Oil spills	Look-alikes	Overall	Oil spills	Look-alikes
0.8063	0.8621	0.7588	0.8166	0.8932	0.7540

Table 2. Segmentation results with accuracy.

tree forest resulted an 85.0% accuracy as the highest value. Also, relevant methods like the probabilistic based in [18] reported results equal to 78% for oil spill and 99% for look-alikes. Thus, the initial results of the deep CNN based approach are similar to the corresponding state-of-the-art methods nonetheless without the need of extracting relevant features and with the merit of semantically annotated regions.

5. CONCLUSIONS

In this paper, we introduced a new approach for oil spill detection using SAR images for maritime applications. With the adaptation of accurate DCNNs, oil spill detection can be further automated and be incorporated into a larger detection pipeline. The extracted results are similar to the state-of-the-art results for general classification problems. For the oil spill detection problem, the use of similar deep learning techniques may improve further the identification of such pollution spots. Potential improvements could also be achieved utilizing a more advanced dataset which may include images acquired with improved SAR sensors and larger training sets. Though it is a preliminary work, the initial results could be considered as promising for further exploitation of deep learning algorithms in the oil spill detection field.

6. REFERENCES

- [1] Konstantinos Topouzelis and Apostolos Psyllos, "Oil spill feature selection and classification using decision tree forest on sar image data," *ISPRS journal of photogrammetry and remote sensing*, vol. 68, pp. 135–143, 2012.
- [2] A Yu Ivanov and Victoria V Zatyagalova, "A gis approach to mapping oil spills in a marine environment," *International Journal of Remote Sensing*, vol. 29, no. 21, pp. 6297–6313, 2008.
- [3] Suman Singha, Tim J Bellerby, and Olaf Trieschmann, "Satellite oil spill detection using artificial neural networks," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 6, pp. 2355–2363, 2013.
- [4] Hui-yan Shen, Pu-cheng Zhou, and Shao-ru Feng, "Research on multi-angle near infrared spectral-polarimetric characteristic for polluted water by spilled oil," in *International Symposium on Photoelectronic Detection and Imaging 2011: Advances in Infrared Imaging and Applications*. International Society for Optics and Photonics, 2011, vol. 8193, p. 81930M.
- [5] Carlos Gonzalez, Sergio Sánchez, Abel Paz, Javier Resano, Daniel Mozos, and Antonio Plaza, "Use of fpga or gpu-based architectures for remotely sensed hyperspectral image processing," *INTEGRATION, the VLSI journal*, vol. 46, no. 2, pp. 89–103, 2013.
- [6] G Aalet Mastin, JJ Manson, JD Bradley, RM Axline, and GL Hover, "A comparative evaluation of sar and slar," Tech. Rep., Sandia National Labs., Albuquerque, NM (United States), 1993.
- [7] Merv Fingas and Carl Brown, "Review of oil spill remote sensing," *Marine pollution bulletin*, vol. 83, no. 1, pp. 9–23, 2014.
- [8] Marta Konik and Katarzyna Bradtke, "Object-oriented approach to oil spill detection using envisat asar images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 118, pp. 37–52, 2016.
- [9] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan, "Show and tell: A neural image caption generator," in *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*. IEEE, 2015, pp. 3156–3164.
- [10] Andrej Karpathy and Li Fei-Fei, "Deep visual-semantic alignments for generating image descriptions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3128–3137.
- [11] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *arXiv preprint arXiv:1606.00915*, 2016.
- [12] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [14] Matthias Holschneider, Richard Kronland-Martinet, Jean Morlet, and Ph Tchamitchian, "A real-time algorithm for signal analysis with the help of the wavelet transform," in *Wavelets*, pp. 286–297. Springer, 1990.
- [15] Alessandro Giusti, Dan C Cirean, Jonathan Masci, Luca M Gambardella, and Jurgen Schmidhuber, "Fast image scanning with deep max-pooling convolutional neural networks," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*. IEEE, 2013, pp. 4034–4038.
- [16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *European conference on computer vision*. Springer, 2014, pp. 346–361.
- [17] Marco Cococcioni, Linda Corucci, Andrea Masini, and Fabio Nardelli, "Svme: an ensemble of support vector machines for detecting oil spills from full resolution modis images," *Ocean Dynamics*, vol. 62, no. 3, pp. 449–467, 2012.
- [18] Anne HS Solberg, Camilla Brekke, and Per Ove Husoy, "Oil spill detection in radarsat and envisat sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 746–755, 2007.