# Oil Spill Detection on Landsat-8/9 Images Based on Deep Learning Methods

Olga Schmidt
*German Remote Sensing Data Center (DFD)*
*German Aerospace Center (DLR)*
Neustrelitz, Germany
Olga.Schmidt@dlr.de

Egbert Schwarz
*German Remote Sensing Data Center (DFD)*
*German Aerospace Center (DLR)*
Neustrelitz, Germany
Egbert.Schwarz@dlr.de

Detmar Krause
*German Remote Sensing Data Center (DFD)*
*German Aerospace Center (DLR)*
Neustrelitz, Germany
Detmar.Krause@dlr.de

***Abstract*—Remote sensing can be used for oil spill detection. To minimize the impact of oil pollution on the ecosystems, it is imperative that oil spills are detected at the earliest possible stage in order that the relevant monitoring frameworks can be put in place and appropriate response measures initiated.**

**This paper presents two different approaches for oil spill detection on optical satellite imagery from the Landsat-8 and Landsat-9 satellites using deep learning techniques. This comprises the application of a (fully connected) deep neural network (DNN) and a convolutional neural network (CNN) in the type of a U-Net architecture. The models were developed to recognise and classify patterns of oil spills against the complex background of marine and coastal environment. Consequently, the performance of the models is evaluated and their efficiency demonstrated on different datasets. The experimental results indicate usability of the analysed methods. This study is based on a limited amount of manually labelled training data and serves to validate the potential of deep learning based oil spill detection on optical satellite remote sensing images.**

***Index terms*—Oil Spill Detection, Optical Remote Sensing, Deep Learning, DNN, CNN.**

## I. Introduction

Oil spills represent a serious threat to the marine and coastal environment and pose a danger to human health. The utilisation of remote sensing technology enables a timely and accurate detection of oil spills, due to its capacity to monitor large marine regions. This can help to prevent pollution spread and support clean-up operations after accidents or deliberate discharges to minimize the negative impacts on the environment as well as to identify the polluter. Satellite images captured by Synthetic Aperture Radar (SAR) sensors are widely used for oil spill detection [1]. Nevertheless, the incorporation of optical observation capabilities enables the reduction in the time required for the repeated monitoring of the same sea area, thereby enhancing temporal resolution. In comparison to optical sensors, active SAR sensors are independent of the daytime and the existence of clouds, dust or smoke. The presence of oil on the surface of the water results in a smoother surface, which appears darker than water on SAR images. Algal blooms and regions with very low wind speeds appear darker than water as well. In this particular case it is already a challenge to distinguish between oil spills and these so called "look-alikes". Optical sensors acquire the reflected radiation from the surface in several spectral bands. Oil changes the spectral characteristics of water. The spectral analysis not only provides some information about the properties of oil spills for each spectral band but can also provide more information, such as oil thickness or type of oil. Nevertheless, similar reflectance like oil may be shown by thin clouds, cloud shadows, dust, suspended sediments or areas characterised by shallow water. On optical images oil spills can appear darker (negative contrast) or brighter (positive contrast) than the surrounding water. The contrast variation is depending on the position of the sun and the sensor. Additionally, there is a high variability of the contrast intensity which is influenced by the optical properties of the oil (oil type, oil thickness), the scattering of the sea water, the sea state (depending mostly by the wind patterns), the depth of the sea and the bathymetry. Furthermore, the reflectance of oil spills and water is similar across all spectral bands. However, the visibility of oil spills can be enhanced through the application of image processing techniques [2].

The objective of this study is to identify the optimal deep learning approach for the automatic oil spill detection on images captured by the Landsat-8 and Landsat-9 optical satellites. Therefore, a deep neural network (DNN) and a convolutional neural network (CNN) in the type of a U-Net architecture are trained. The two networks under consideration utilize different techniques for handling training data. A DNN is trained pixel by pixel ignoring the spatial component, unlike the U-Net, but its implementation is quite simple. The U-Net network is able to use the localization and context information at the same time. Several published studies, for example [3] and [4], have shown that the performance of the U-Net architectures is convincing even if only a sparse training dataset is available. In this study, the training was done by using three independent indices calculated from specific spectral bands. The following sections give an overview of the methods and data used and some results obtained.

## II. Methods and Data

### A. Workflow

*Figure 1* gives an overview of the implemented workflow employed for training the DNN and the U-Net for oil spill detection. The Landsat-8/9 data, described in section II.B, were

subjected to manual examination to ascertain their oil content. Data, on which oil could be identified, was indexed in a data cube (Open Data Cube, ODC) [5] and, afterwards, pre-processed to be incorporated into the model training process. The pre-processing consists of the generation of the train images and the preparation of corresponding oil masks, which together form the training dataset. A comprehensive overview is provided in section II.C. Subsequently, the dataset was split into three independent datasets: the training data (80 %), the validation data (10 %) and the test data (10 %). During model training process, the training data were used to train two models while the validation data was used to fine-tune the hyperparameters. Due to the occurrence of numerous false positive detections, especially caused by cloud shadows, three masks (no-data mask, cloud / cloud shadow mask and land mask), were employed in the post-processing step. Finally, the test data were used to evaluate the performance of the models.

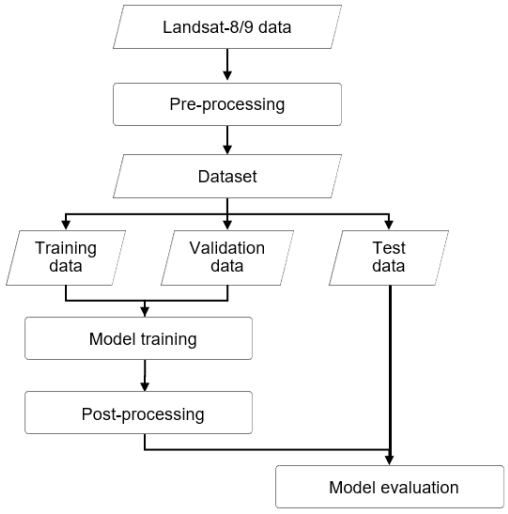


Figure 1 Overview of the implemented workflow for oil spill detection.

### *B. Landsat-8/9 data*

The Landsat program represents a collaboration between National Aeronautics and Space Administration (NASA) and United States Geological Survey (USGS) which has been in operation since 1972. The two currently active satellites Landsat-8 (launched in 2013) and the most recent Landsat-9 (launched in 2021) are equipped with the Operational Land Imager (OLI/OLI-2) instrument. Each satellite acquires data in accordance with its respective Long Term Acquisition Plan (LTAP). OLI/OLI-2 is a multispectral sensor that measures the Earth's surface in nine bands, comprising visible, near-infrared and shortwave-infrared spectra. The spatial resolution is 30 meters for bands 1-7 and 9 and 15 meters for band 8 (panchromatic) along a 185 km swath. The resolution of each single frame image is ≈ 7 600 x 7 600 pixels. Both satellites cover each Landsat scene area once in 16 days. Due to the offset between the two satellite orbits, a repeat cycle of 8 days is possible [6], [7]. The data can be downloaded free of charge from the EarthExplorer website [8].

### *C. Training data*

In this study, the Landsat-8/9 Collection 2 Level-1 data products, downloaded from the USGS catalogue, are used. These are calibrated scaled digital numbers (DN) that are influenced by the scattering and absorption in the atmosphere. In order to reduce this influence, an atmospheric correction was applied using the ATCOR software [9]. At the beginning of this study, it was discovered that the selection of band combinations based on specific spectral bands is more effective for oil spill detection than the usage of individual bands directly. Based on this conclusion the combination of the Normalised Difference Oil Index (NDOI) (eq. (1)) [3], the Green-Shortwave Infrared Index (G-SWIR) (eq. (2)) (adapted from [10]) and the index CaBGS (eq. (3)) (adopted to Landsat images from [11]) were chosen to create the training dataset. The NDOI is defined as the ratio of the difference and the sum of the surface reflectance values for the green band ($\rho_\lambda$(green)) and the near infrared band ($\rho_\lambda$(NIR)). The index G-SWIR is based on the ratio of the difference and the sum of the surface reflectance values for the green band ($\rho_\lambda$(green)) and the shortwave infrared band ($\rho_\lambda$(SWIR2)). Both indices have shown to enhance the visual differences between oil slicks and the surrounding waters. CaBGS is a combination of the surface reflectance values of the spectral bands coastal aerosol ($\rho_\lambda$(coastal_aerosol)), blue ($\rho_\lambda$(blue)), green ($\rho_\lambda$(green)) and SWIR2 ($\rho_\lambda$(SWIR2)). This value shows an ability to better distinguish oil slicks from look-alikes in shallow water areas. Figure 2 shows the result of each of these three indices individually. The contrast between oil and the surrounding water is quite low on the coastal aerosol band, the oil spill is difficult to identify. In contrast, the visibility of the oil spill is very high on each chosen index.

$$\text{NDOI} = \frac{\rho_\lambda(\text{green}) - \rho_\lambda(\text{NIR})}{\rho_\lambda(\text{green}) + \rho_\lambda(\text{NIR})} \quad (1)$$

$$\text{G-SWIR} = \frac{\rho_\lambda(\text{green}) - \rho_\lambda(\text{SWIR2})}{\rho_\lambda(\text{green}) + \rho_\lambda(\text{SWIR2})} \quad (2)$$

$$\text{CaBGS} = \frac{\rho_\lambda(\text{coastal\_aerosol}) + \rho_\lambda(\text{blue})}{\rho_\lambda(\text{green}) + \rho_\lambda(\text{SWIR2})} \quad (3)$$

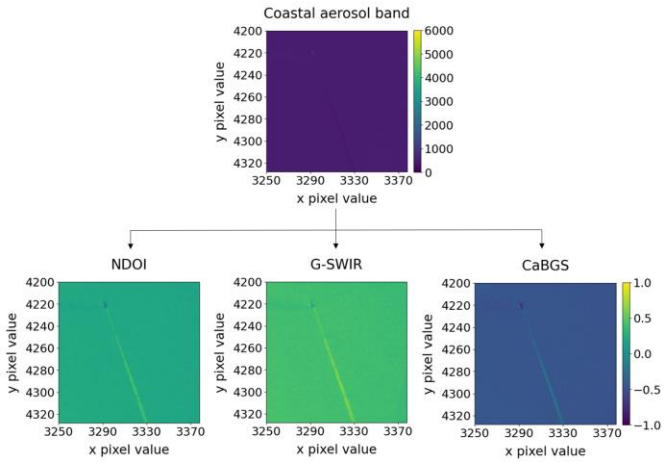


Figure 2. Coastal aerosol band (top) and the individual indices NDOI, G-SWIR and CaBGS (lower row) showing a long and narrow oil spill behind a ship, presented in patches of the size of 128 x 128 pixels.

In the current study, the dataset used to train the model includes 26 Landsat-8/9 images containing oil spills. In order to test whether it is possible to train a model that is applicable to different geographical areas no spatial restriction was made. This is the reason why the selected images were acquired over different regions of the North Sea, Baltic Sea, Mediterranean Sea, Red Sea, South China Sea and in the Gulf of Guinea.

Binary oil masks were generated for the selected Landsat images to be used in the model training. Initially, the masking was done manually using the VisualAnalyst software [12] by drawing a polygon around each individual oil slick. Subsequently, a segmentation method based on the NDOI, G-SWIR and the pixel values of the coastal aerosol band was implemented. As mentioned above, these two indices were chosen because of their ability to enhance the visual distinction between oil spills and the surrounding water. Furthermore, the coastal aerosol band was used to remove very dark and very bright pixel values and to enhance the contrast between oil slicks and water. To distinguish the oil pixels from the surrounding water pixels, separate thresholds were introduced for each of these three values. Both methods depend on the image quality and the contrast between oil slicks and the surrounding water. Especially different illumination around the oil spills is challenging issue as it has a significant influence on the pixel values. This meant that individual thresholds had to be defined for each Landsat image, or even for each separate oil slick. Due to these challenges, the process of oil mask generation took a lot of time.

The precise creation of the oil masks is crucial for the model training, as any pixel not marked as an oil pixel (pixel value 1) is automatically considered as a non-oil pixel (pixel value 0). This includes pixels of the land surface, clouds, cloud shadows and no-data values (black border outside the Landsat image). The no-data areas all have the same pixel value, while the cloud pixel values vary in the high range and the land surface values can be low or high depending on the surface characteristics. This would lead to a high variability of pixel values for this label and negatively influence the model training. To exclude these areas from training, three additional masks were applied in addition to the oil mask: the no-data mask, the cloud/cloud shadow mask, and the land mask. The no-data pixels were removed using the blue band, the land areas were excluded using OpenStreetMap [13] and clouds and cloud shadows were masked using a combination of the Landsat QA band and the UKIS Cloud Shadow MASK (ukis-csmask) [14].

The handling of these additional masks was different depending on the deep learning method used. In the case of the DNN model training, these pixels were excluded from the training. Moreover, an extra buffer was applied to make sure that, for example, small clouds or cloud shadows, that were not excluded by the masking algorithm, but were close to detected ones, were excluded from this label too. Since the training of a DNN model is done pixel by pixel, the applied masking allowed to clean the training dataset with respect to the areas to be excluded. From the remaining dataset, 130 411 pixels were chosen randomly. This number corresponds to the number of oil pixels available in the binary oil masks and was chosen to obtain a balanced training dataset.

The U-Net utilises a spatial component for training. For this reason, the pixels corresponding to the three masks described were summarised in a new training class, defined with a pixel value of 2, and excluded from the model training. Before training the U-Net, the Landsat images were cropped into 128 x 128 pixels patches. In order to make the model more robust, Keras [15] ImageDataGenerator class was used. Horizontal and vertical flips and rotation were applied on each training image on the fly while the model was being trained. This class creates new variations of the images at each epoch. The original images are replaced by the transformed ones. For the purpose of improving the balancing of oil pixels/non-oil pixels in the training dataset, patches that contain less than 5 % of oil pixels were removed from the dataset. However, the resulting dataset is still unbalanced: 91 962 oil pixels and 521 677 non-oil pixels. Here, the lower number of available oil pixels in comparison to the number used for DNN indicates that 38 449 oil pixels were removed because they were less than 5 % of the patch size. Table 1 summarizes the available pixel numbers used for training.

Table 1
AVAILABLE NUMBER OF OIL AND NON-OIL PIXELS FOR THE DNN AND THE U-NET

| **Pixel** | **Number of pixels, DNN** | **Number of pixels, U-Net (128 x 128px)** |
|---|---|---|
| Oil | 130 411 | 91 962 |
| Non-Oil | 130 411 | 521 677 |
| | → balanced | → imbalanced |

Finally, each chosen pixel (DNN) and each pixel from the cropped patch (U-Net) of this edited oil mask, respectively, are assigned to the corresponding calculated three indices NDOI, G-SWIR and CaBGS. These data were then used to train the different neural networks, which architectures are described in the following two sections. Both networks were implemented using Keras, which is a high-level API for TensorFlow and one of the most widely used libraries for deep learning in Python.

### *D. Deep Neural Network (DNN)*

A fully connected DNN consists of an input layer, several hidden layers and an output layer. The layers contain nodes (neurons) and edges. The edges describe the linear and non-linear relationships between each node of one layer and all nodes of the following layer. The linear relationships are based on weights and biases. An activation function is applied after each hidden layer, which allows very complex non-linear problems to be handled. The DNN includes an iterative update of the weights according to a backpropagation, where the gradient descent optimization method is used to minimize the loss using a loss function [16], [17]. The aim of this study is to find the optimal combination of weights and biases in the

hidden layers of the DNN that allows a reliable discrimination between the oil pixels and the non-oil pixels based on the three indices used. As shown in Figure 3, the architecture of the network implemented for this study consists of three nodes in the input layer corresponding to the selected indices NDOI, G-SWIR and CaBGS, 80 nodes in each of the six hidden layers and two nodes in the output layer (binary oil mask). The selection of these values in the hidden layers section is the result of several tests to find the highest accuracy. The network was defined as a sequential model of the Keras library. In order to optimize the model, the adaptive moment estimation (Adam) algorithm was chosen. This is the most widely used optimizer. The errors were minimized using the binary cross-entropy as it is a binary classification. The rectified liner unit (ReLU) activation function was taken for the hidden layers. The final classification was done using the sigmoid activation function because the output is either 0 or 1 [18]. The best model performance could be reached after 300 epochs. The summary of the selected hyperparameters is given in Table 2.

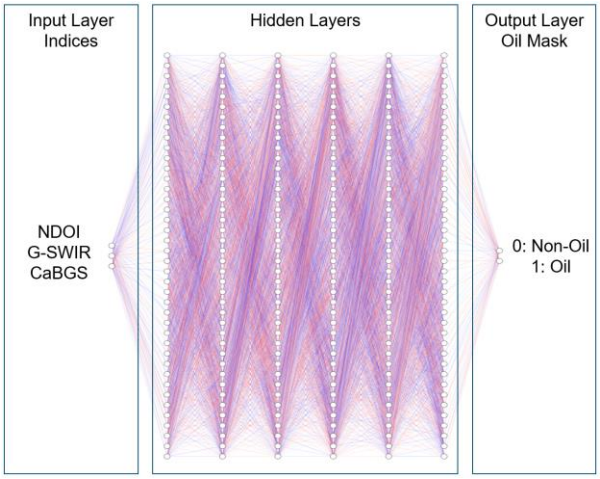


Figure 3. Model architecture of the used Deep Neural Network (DNN).

Table 2
SELECTED HYPERPARAMETERS FOR THE DNN AND THE U-NET

| | **DNN** | **U-Net** |
|---|---|---|
| Activation function | ReLU | ReLU |
| Optimizer | Adam | Adam |
| Loss function | Binary cross-entropy | Sparse categorical cross-entropy |
| Classification rule | Sigmoid | Softmax |
| Number of epochs | 300 | 50 |

### *E. U-Net (Convolutional Neural Network, CNN)*

In addition to the DNN described above, a CNN model in the type of a U-Net architecture [19] was trained in the study to determine which model is more suitable for detecting oil spills. The U-Net architecture, shown in Figure 4, consists of two parts: the contracting path (encoder) and the expanding path (decoder). While the contracting path captures the context of an image (feature extraction), the expanding path enables precise localization of containing object. The contracting path is a typical convolutional neural network that contains a collection of convolutional layers and "max-pooling" layers that gradually reduces the spatial resolution of the input image (down-sampling) and increases the number of features in each layer. The expanding path is the mirrored part of the contracting path: the spatial resolution increases (up-sampling) and the number of features decreases gradually in each layer [20]. In this study, patches of the size of 128 x 128 pixels were used as input. Each patch is a combination of three layers corresponding to the calculated three indices NDOI, G-SWIR and CaBGS. The U-Net was instantiated using the model class (functional model) of the Keras library. The contracting path and the expanding path were built with four encoder blocks and four decoder blocks, respectively. Both paths are connected by skip connections where the features from the contracting path are used to improve the segmentation results in the corresponding expanding path. As already used for the DNN, the Adam algorithm and ReLU were chosen as optimizer and activation function, respectively. To minimize errors, it was decided to take the sparse categorical cross-entropy due to its ability to handle classes which should be excluded from the model training. As mentioned in section II.C, an additional class was created, comprising pixels of the land surface, clouds, cloud shadows and no-data values. This is also the reason for the usage of the softmax activation function as the final segmentation [18]. In this case, the best model performance could be reached after 50 epochs. These selected hyperparameters are summarised in the last column of Table 2. Due to the exclusion of this third class from the model training the output is a binary oil mask.

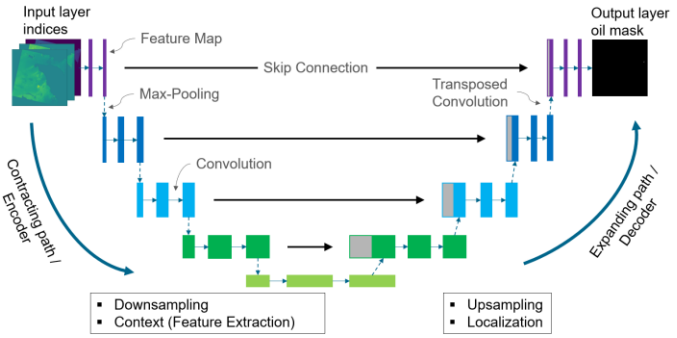


Figure 4. Model architecture of the used U-Net (CNN), adopted from [19].

## III. RESULTS

In this section, some results obtained from the training of the two different deep learning models DNN and U-Net (CNN) for oil spill detection are presented. In order to evaluate the performance and effectiveness of both models a selection of several evaluation metrics [21] has been used. The evaluation

metrics were applied to the test dataset, that was not included in the model training (see section II.A), by comparing the predicted oil spill pixels or patches with the pixels or patches from the manually generated binary oil spill masks (see section II.C) depending on the deep learning method used. The latter are subsequently referred to as the ground truth. In this study, a confusion matrix [22] was generated and precision, recall, F1-score and the intersection over union (IoU) [23] were determined.

Figure 5a and Figure 5b show the confusion matrices determined for the DNN and the U-Net, respectively. The confusion matrix for the DNN shows that 93 % of the predicted oil pixels were also masked as oil in the manually labelled binary oil mask. Only 7 % of the labelled oil pixels were missed and 8 % were incorrectly detected by the model. The prediction of the non-oil pixels matches 92 % with the non-oil pixels containing in the labelled binary oil mask. This indicates that the performance of the DNN is quite good with respect to the test data used. In contrast, the confusion matrix for the U-Net indicates a less favourable performance outcome: While 72 % of the predicted oil pixels were containing in the labelled binary oil mask, 28 % were missed and 8 % were wrongly detected. The prediction of labelled non-oil pixels behaves in similar way to the DNN: there is a match of 92 %. As already mentioned, unlike the DNN, the U-Net contains an additional class with the masked pixels, which are labelled as a third class during pre-processing. Thus, 97 % of these pixels were correctly identified as non-oil, while 3 % were erroneously classified as oil. As this third class was excluded from the model training, no predictions were made in this case. This is shown by values of "0.00" in the right column.

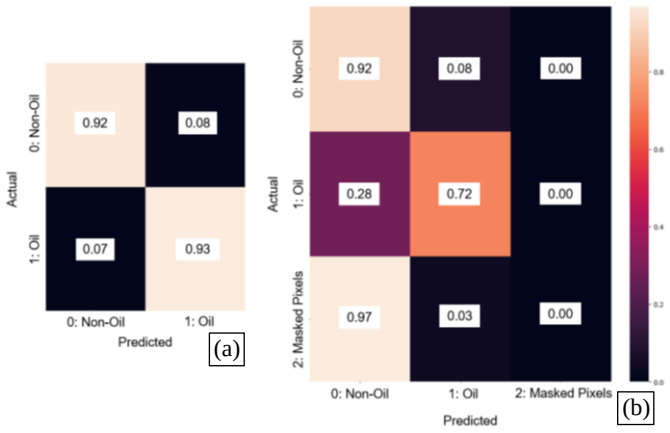


Figure 5: Confusion matrix (a) determined for the used DNN and (b) determined for the used U-Net.

Table 3 shows the values of the selected evaluation metrics precision, recall, F1-score and the IoU. The calculations were performed for the model outputs "oil" (1) and "non-oil" (0) for the DNN as well as for the U-Net. Since the additional third class, which contained the masked pixels, was excluded from the model training using U-Net, the evaluation metrics were not determined. All of these values are very high for the DNN. This confirms the conclusion made from the confusion matrix that the performance of the DNN regarding the test data is very good. The comparatively lower performance of the U-Net as evidenced in the confusion matrix is also reflected in the values of the evaluation metrics. These values are considerably lower except for the recall value of the non-oil pixels.

Table 3
EVALUATION METRICS FOR OIL (1) AND NON-OIL PIXELS (0) FOR THE DNN AND THE U-NET, RESPECTIVELY

| Method | Output | Precision | Recall | F1-score | IoU |
|---|---|---|---|---|---|
| DNN | 0 | 0.93 | 0.92 | 0.93 | 0.86 |
| | 1 | 0.92 | 0.93 | 0.93 | |
| U-Net | 0 | 0.63 | 0.92 | 0.75 | 0.52 |
| | 1 | 0.55 | 0.72 | 0.62 | |

In order to investigate the model performances on entire Landsat images, both deep learning algorithms were applied to several scenes. Figure 6, Figure 7 and Figure 8 show sections of selected examples of three Landsat images visualized as a false colour band combination. The image sections on the left side (a) serve to illustrate the oil spills. The geographical position of the respective image section within the overall Landsat scene is indicated in the top left-hand corner. The same image sections are displayed on the right-hand side (b), overlaid with the results of the deep learning method employed: The results of the DNN are displayed in blue and pink, while those of the U-Net in orange and pink. The pink-coloured areas indicate the regions where the two methods yielded overlapping results.

The examples selected in Figure 6 and Figure 7 show parts of the spilled oil after the sinking of the oil tanker MT Princess Empress at the coast of Mindoro (Philippines) on 28 February 2023, which was carrying 800 thousand liters of industrial fuel.

The Landsat-9 image shown in Figure 6a and Figure 6b was acquired on 12 March 2023. The detections in Figure 6b show that the DNN has demonstrated a tendency to overestimate regions that are supposed to be oil, especially in areas along the coastline. In contrast, the U-Net has demonstrated a tendency to underestimate the extend of the spilled oil. In this particular case, the overlap between DNN and U-Net detections is quite substantial. The majority of the overlap corresponds to the potential oil spill, but there are also some detections in the northern image section which are assumed to be cloud shadows.

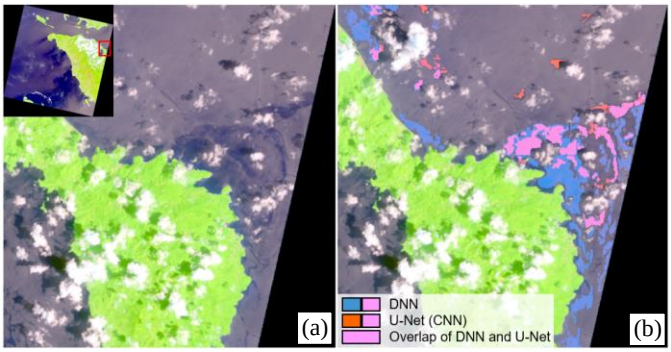


Figure 6. Section of a false-colour Landsat-9 image acquired on 12 March 2023 (a) showing the oil spill with the geographical position of the image section within the entire Landsat image and (b) extending by the oil spill detections of the DNN and the U-Net.

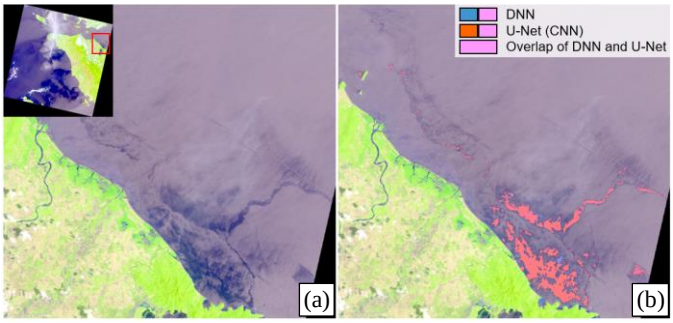


Figure 7. Section of a false-colour Landsat-8 image acquired on 20 March 2023 (a) showing the oil spill with the geographical position of the image section within the entire Landsat image and (b) extending by the oil spill detections of the DNN and the U-Net.

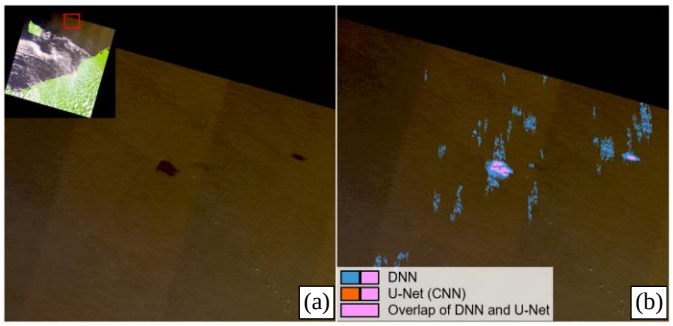


Figure 8. Section of a false-colour Landsat-8 image acquired on 13 July 2017 (a) showing the oil spill with the geographical position of the image section within the entire Landsat image and (b) extending by the oil spill detections of the DNN and the U-Net.

Figure 7a and Figure 7b shows a Landsat-8 image of the same region acquired eight days subsequent to the previous Landsat-9 image on 20 March 2023. It is evident that the oil spill has expanded in a north-west direction, with the spilled area increasing in size. As depicted in Figure 7b, in this particular case, both detection models underestimated the extent of the oil spill. The U-Net model has detected only the darkest regions of the potential oil spill, whereas only few oil slicks were detected by the DNN. The degree of overlap between the two methods in the present case was relatively limited.

The third example showed in Figure 8 was acquired in the Baltic Sea on 13 July 2017. The oil spill detection using the DNN shows an overestimation of the two oil slicks and, additionally, some water pixels were falsely detected as oil. As can also be seen in the two other images, the U-Net model has underestimated the oil spill. Regarding this particular image section, no false detections were made.

## IV. Conclusion

This paper presents an approach to detect oil spills on Landsat-8/9 images using deep learning methods, specifically the DNN and the U-Net (CNN). The study is based on a limited number of labeled training data. The model training was done using the combination of three indices (Normalised Difference Oil Index (NDOI), the Green-Shortwave Infrared Index (G-SWIR) and the index CaBGS), that are based on specific spectral bands.

The current results of the study indicate that the proposed methodology is, in principle, an effective approach. This conclusion is proven by the confusion matrices and the calculated evaluation metrics based on the test dataset.

The trained DNN model demonstrated excellent performance in distinguishing between oil and non-oil samples, as evidenced by the high accuracy rates observed in the confusion matrix and the calculated evaluation metrics. However, the processing of additional Landsat images revealed a tendency for the model to overestimate oil contamination on one side. In some instances, the oil was not correctly detected. The evaluation metrics are determined using the test data. The test data probably did not contain pixels that are similar to oil spill pixels. That could be a reason why the number of false detections is low. Additionally, the training dataset did not contain any such pixels. That is the reason, why the model cannot distinguish properly between oil pixels and pixels of similar values.

Despite the lower performance of the U-Net model in comparison to the DNN model according the confusion matrix and the evaluation metrics its application on additional Landsat images demonstrates superior performance in terms of accuracy. However, it tends to underestimate the extent of potential oil spills. The findings of the study indicate that the performance of the model is enhanced when the oil slicks exhibit a broad shape. In the training dataset, regions containing less than 5 % oil pixels were excluded from the training process, which had a discernible impact on the models' capacity to recognise narrow and small-scale structures. This is postulated to be a potential factor contributing to the inability of the models to recognise these structures.

Both methods show too many false positive detections which are predominantly attributable to dark image areas, cloud shadows or suspended sediments. One potential explanation for this phenomenon is the limited training dataset.

To reduce false detections and improve model performance, a significantly larger training dataset is required. Another approach is to train localized models for different geographical regions. Oil has a high spectral variability depending on several properties of the oil itself but of the environment conditions as well. In particular, the look-alikes caused by cloud shadows could be reduced by implementing an enhanced mask that is more precise in accounting for cloud shadows.

## References

[1] J. Fan, F. Zhang, D. Zhao and J. Wang, "Oil Spill Monitoring Based on SAR Remote Sensing Imagery," *Aquatic Procedia*, vol. 3, pp. 112-118, 2015.

[2] F. A. M. Alawadi, “Detection and Classification of Oil Spills in MODIS Satellite Imagery,” Faculty of Engineering, Science and Mathematics, PhD Thesis, University of Southampton, 2011.

[3] L. Nava, M. Cuevas, S. Meena, C. F. and O. Monserrat, “Artisanal and Small-Scale Mine Detection in Semi-Desertic Areas by Improved U-Net,” in *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.

[4] R. Azad, M. Asadi-Aghbolaghi, M. Fathy and S. Escalera, “Bi-Directional ConvLSTM U-Net with Densley Connected Convolutions,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.

[5] B. Killough, “Overview of the Open Data Cube Initiative,” in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, Spain, 2018.

[6] V. Ihlen, Landsat 8 (L8) Data Users Handbook, LSDS-1574, Version 5.0, Department of the Interior U.S. Geological Survey, 2019.

[7] K. Sayler, Landsat 9 Data Users Handbook, LSDS-2082, Version 1.0, Department of the Interior U.S. Geological Survey, 2022.

[8] “EarthExplorer,” 2024. [Online]. Available: https://earthexplorer.usgs.gov.

[9] R. Richter and D. Schläpfer, “Atmospheric / Topographic Correction for Satellite Imagery,” ATCOR-2/3 User Guide, Version 9.5, 2023.

[10] J. Coffelt and R. Livingston, “Second U.S. geological survey wildland fire workshop: los Alamos, New Mexico, October 31-November 3, 2000,” U.S. Geological Survey, 10.3133/ofr0211, 2002.

[11] M. Zakzouk, I. A. El-Magd, E. M. Ali, A. M. Abdulaziz, A. Rehman and T. Saba, “Novel oil spill indices for sentinel-2 imagery: A case study of natural seepage in Qaruh Island, Kuwait,” *MethodsX*, Volume 12, https://doi.org/10.1016/j.mex.2023.102520, 2024.

[12] S. Voinov, “Deep Learning-based Vessel Detection from Very High and Medium Resolution Optical Satellite Images as Component of Maritime Surveillance Systems,” PhD Thesis, University of Rostock, 2020.

[13] OpenStreetMap, 2022. [Online]. Available: https://www.openstreetmap.org.

[14] M. Wieland, F. Fichtner and S. Martinis, “UKIS-CSMASK: A Python package for multi-sensor cloud and cloud shadow segmentation,” *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B3-2022, 217–222,* https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-217-2022, 2022.

[15] A. Gulli and S. Pal, Deep Learning with Keras, Implementing deep learning models and neural networks with the power of Python, Packt Publishing Ltd., 2017.

[16] G. Montavon, W. Samek and K.-R. Müller, "Methods for interpreting and understanding deep neural networks," *Digital Signal Processing*, vol. 73, pp. 1-15, 2018.

[17] J. Feng, X. He, Q. Teng, C. Ren, H. Chen and Y. Li, "Reconstruction of porous media from extremely limited information using conditional generative adversarial networks," *PHYSICAL REVIEW E*, vol. 100, 2019.

[18] D. Sumathi and K. Alluri, “Deploying Deep Learning Models for Various Real-Time Applications Using Keras,” in *Prakash, K.B., Kannan, R., Alexander, S., Kanagachidambaresan, G.R. (eds) Advanced Deep Learning for Engineers and Scientists. EAI/Springer Innovations in Communication and Computing*, Springer, Cham, https://doi.org/10.1007/978-3-030-66519-7_5, 2021.

[19] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, https://doi.org/10.1007/978-3-319-24574-4_28, 2015.

[20] I. Ahmed, M. Ahmad and G. Jeon, "A real-time efficient object segmentation system based on U-Net using aerial drone images," *J Real-Time Image Proc*, vol. 18, p. 1745–1758, 2021.

[21] G. Naidu, T. Zuva and E. Sibanda, “A Review of Evaluation Metrics in Machine Learning Algorithms,” in *Silhavy, R., Silhavy, P. (eds) Artificial Intelligence Application in Networks and Systems. CSOC 2023. Lecture Notes in Networks and Systems*, vol. 724, Springer, Cham, https://doi.org/10.1007/978-3-031-35314-7_2, 2023.

[22] J. Liang, “Confusion Matrix: Machine Learning,” *POGIL Activity Clearinghouse*, Vols. 3 (4), https://pac.pogil.org/index.php/pac/article/view/304, 2022.

[23] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid and S. Savarese, “Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression,” pp. 658-666, 2019.