Title:
Automating the Detection of Oil Spills from Sentinel-1 SAR Imagery Using Deep Learning

Author(s): *Mohamed Al Hashmi*

Name of Academic Institution: Newcastle University

Level of study or work: Bachelor thesis
(Bachelor thesis, master, research, project, etc.)

Information about you (and your team): Mohamed is a final year undergraduate student studying BEng Geospatial Surveying and Mapping at Newcastle University.

Area of interest

(Identifying the problem, explain why it is important and the current relevance of the topic, up to 250 words)

Oil spills are described as the discharge of liquid petroleum hydrocarbons into the marine environment because of human activities (Li, 2016). The rising global demand for crude oil has led to the expansion of marine oil transportation, increasing the risk of oil spills. It causes severe harm to the marine ecosystem and nearby coastal areas, leading to significant environmental and economic consequences (De Kerf et al., 2020). Oil spills contaminate water sources, affect the availability of clean water, cause desalination plants to shut, and damage mangroves, beaches and aquatic animals (Anselain et al., 2022). It could happen in freshwater bodies or oceans. Authorities desperately need reliable and automated oil spill monitoring systems to enable effective response and mitigation efforts, as well as monitoring of regulatory compliance by ships (De Kerf et al., 2020). Moreover, minimising the damage caused by oil spills will contribute to achieving the United Nations Sustainable Development Goal 6, introduced to improve water quality, and Goal 14, designed to protect marine life.

Manually extracting oil spills from Synthetic Aperture Radar (SAR) imagery can be time-consuming because the process involves visual inspection of large amounts of data. Machine learning is vital for oil spill detection because it can improve the speed and efficiency of detecting oil spills. Deep learning (DL) offers advantages over traditional machine learning methods in that it can automatically learn the features to extract to solve a task from the training data itself. It can also handle large amounts of data more efficiently and accurately than manual methods. This project aimed to develop an automated system for oil spill detection using deep learning and Sentinel-1 SAR imagery to enable faster, more accurate and reliable detection and response to oil spills.

Approach to the problem

(Describe your methodology or technology and how it will solve the problem you identified, up to 300 words)

To build the automated oil spill detection system, the first step was to develop the DL model. This included the following: (a) collecting, pre-processing and labelling of Sentinel-1 oil spill images and (b) training and evaluating the DL models. The second step was to write a script in Python to automatically download Sentinel-1 images, preprocess the images, pass the pre-processed images into the DL models for interference, post-process the predictions of the models and plot the output on an interactive map.

The training images were chosen to include oil spills from multiple geographical areas, as shown in Figure 1, and oil spills with various sizes. Furthermore, many images of lookalikes were included, such as low wind areas, biogenic slicks, and images with rain cells, to assist the model in distinguishing between oil spills and these common lookalikes in SAR images, as illustrated in Figure 2. The pre-processing involved the following operations: extracting VV polarisation band, updating satellite orbits, removing speckle noise, masking land areas, radiometrically calibrating the images, converting the backscatter values to decibel scale, downsampling and normalising the images.

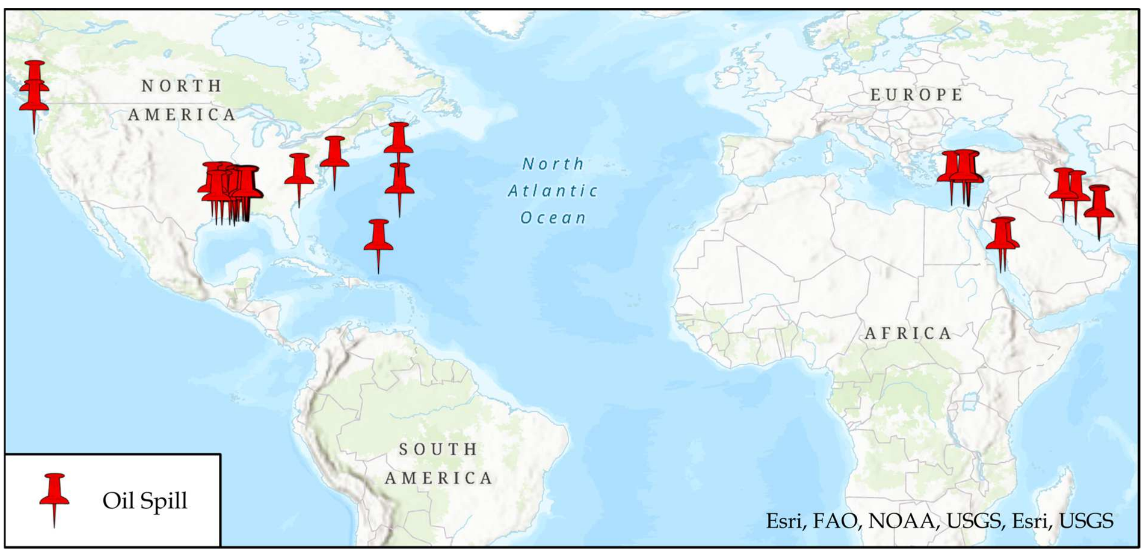

Figure 1. The location of oil spill scenes used in this study.

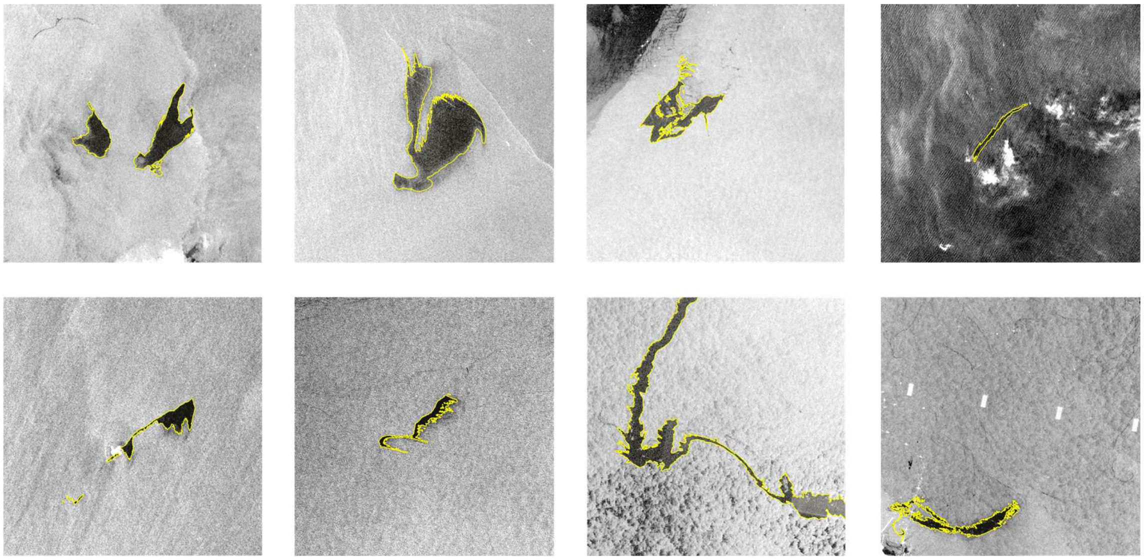

Figure 2. Examples of training images. The oil spills are yellow outlined.

Two DL models, which are U-Net and DeepLabV3+, were trained. The post-processing of the output of the models included reducing the prediction noise by removing small regions of oil spill pixels. Finally, the coordinate reference system and transformation parameters were copied from the original input image and used to transform the predicted oil spill pixels to an appropriately georeferenced polygon, which can be plotted on a map. The developed system will enable automated detect oil spills as soon as new Sentinel-1 acquisitions become available.

Results, conclusions and next steps

(Present your research results and conclusions of your study, up to 250 words)

Quantitative results

The U-Net model achieved an Intersection over Union (IoU) score of 78.43% on validation images, whereas DeeplabV3+ achieved 75.59% (Figure 3). IoU is the ratio of the area of overlap between the predicted and ground truth segmentation to the area of union between the two, as defined in Equation 4 (Jadon, 2020). The highest IoU score was 78.82%, obtained by combining the prediction of U-Net and DeepLabV3+.

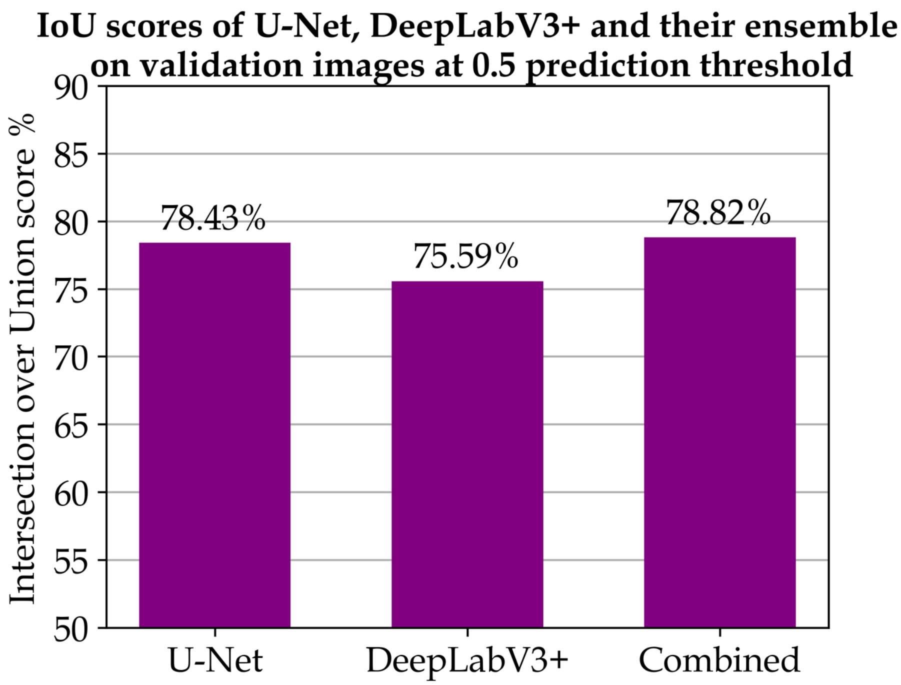

Figure 3. Comparison between DeepLabV3+ and U-Net.

Qualitative results

The results of running the developed system on a test satellite image is shown in Figure 4. The system succeeded in capturing most of the oil spill, with minor omissions and without false positives. Moreover, noise was greatly reduced by combining the predictions of U-Net and DeepLabV3+ and removing small regions. The extracted oil spill polygon is illustrated in Figure 5.

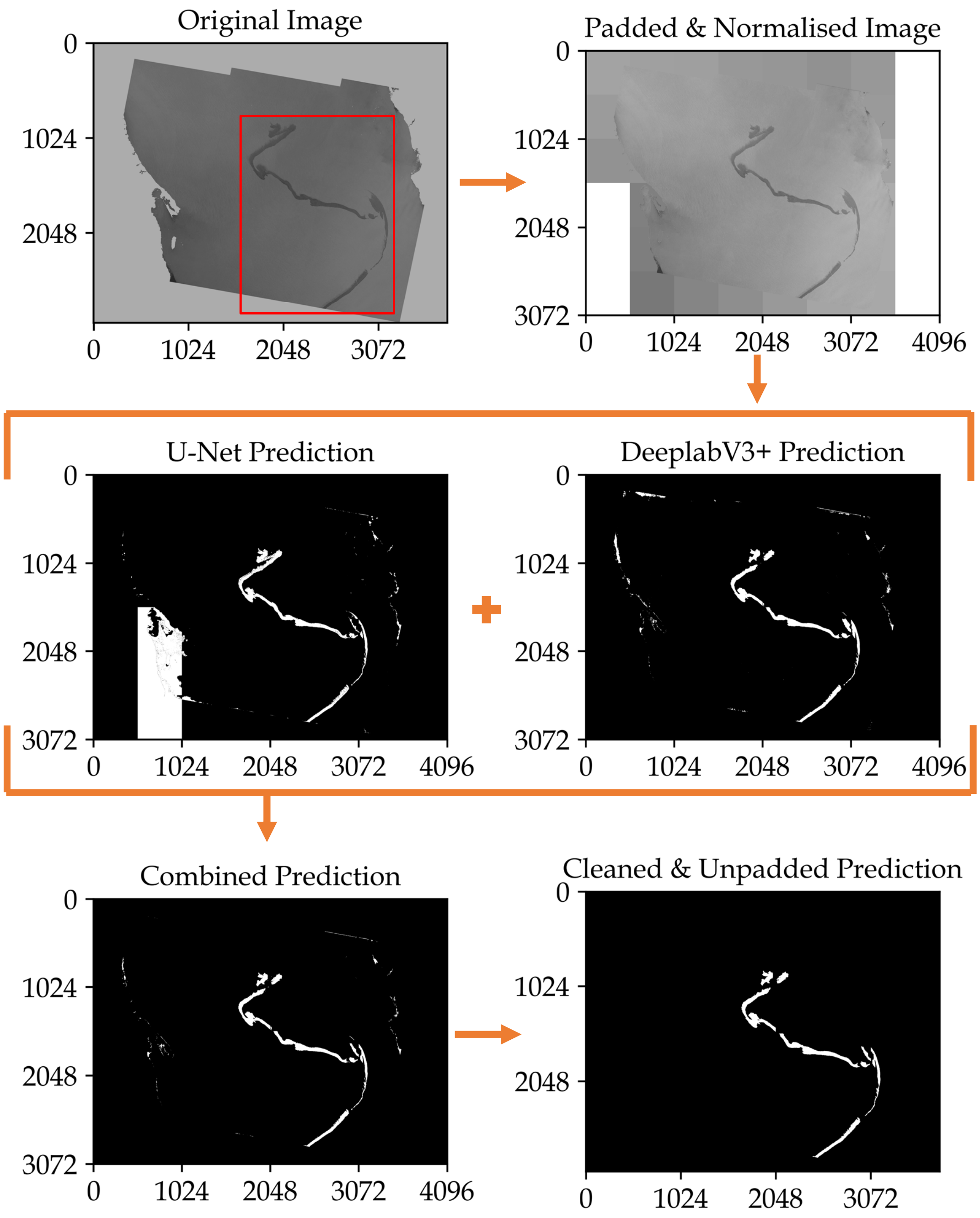

Figure 4. The results of running the system on the test image. The red bounding box indicates the oil spill.

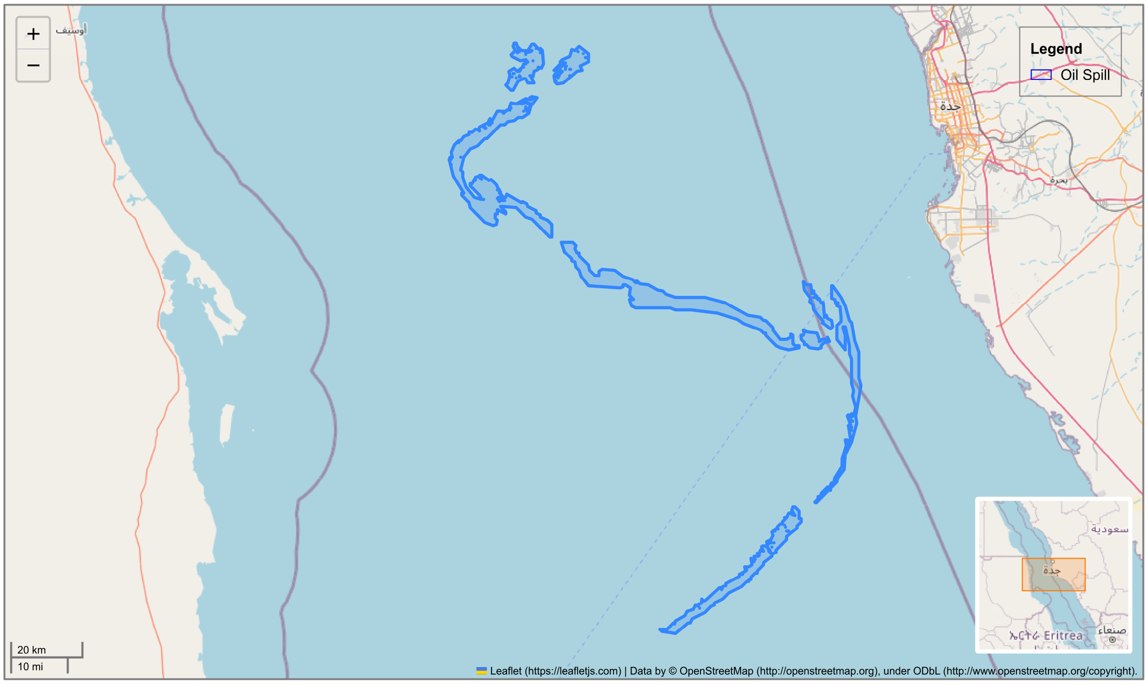

Figure 5. Test image oil spill extent on an interactive leaflet map. The map was generated using folium library.

Conclusions

The developed system showed excellent results for automated oil spill detection. Averaging the output of two models results in better prediction accuracy because they complemented each other. The system has the weakness of confusing certain lookalikes and patches with sudden changes in backscatter coefficient values as oil spill. Nevertheless, this can be easily solved in subsequent iterations of the model by identifying the patches on which the model performed poorly and repassing it back to the model as training data.

Next steps

The next step would be to increase the temporal coverage and reliability of the model. Sentinel-1 satellites has a revisit time of 12 days. This means that there is a chance that oil spills may occur without being detected. By modifying the system to incorporate data from more than one SAR platform and bandwidths, such as X-band, this issue can be solved.

References
(Additional information, publications, or links, up to 200 words)

Anselain, T., Heggy, E., Dobbelaere, T. and Hanert, E. (2023) Qatar Peninsula's vulnerability to oil spills and its implications for the global gas supply, *Nature Sustainability*, pp. 1-11.

De Kerf, T., Gladines, J., Sels, S. and Vanlanduit, S. (2020) Oil spill detection using machine learning and infrared images, *Remote sensing*, 12(24), p. 4090.

Jadon, S. (2020) A survey of loss functions for semantic segmentation. In *2020 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB)*, pp. 1-7.

Li, P., Cai, Q., Lin, W., Chen, B. and Zhang, B. (2016) Offshore oil spill response practices and emerging challenges, *Marine pollution bulletin*, 110(1), pp. 6-27.

END