



OPEN

DATA DESCRIPTOR

A dataset of drone-captured, segmented images for oil spill detection in port environments

Thomas De Kerf^{1,3} , Seppe Sels^{1,3}, Svetlana Samsonova^{1,2} & Steve Vanlanduit¹

The high incidence of oil spills in port areas poses a serious threat to the environment, prompting the need for efficient detection mechanisms. Utilizing automated drones for this purpose can significantly improve the speed and accuracy of oil spill detection. Such advancements not only expedite cleanup operations, reducing environmental harm but also enhance polluter accountability, potentially deterring future incidents. Currently, there's a scarcity of datasets employing RGB images for oil spill detection in maritime settings. This paper presents a unique, annotated dataset aimed at addressing this gap, leveraging a neural network for analysis on both desktop and edge computing platforms. The dataset, captured via drone, comprises 1268 images categorized into oil, water, and other, with a convolutional neural network trained using an Unet model architecture achieving an F1 score of 0.71 for oil detection. This underscores the dataset's practicality for real-world applications, offering crucial resources for environmental conservation in port environments.

Background & Summary

The topic of oil spills stands out in the discourse on environmental disasters, highlighted by extensive analyses in various seminal works^{1,2}. These oil spills have brought widespread devastation to ecosystems. While such large-scale spills often dominate media narratives, it is important to note that most oil spills are actually smaller in scale (under 700 tonnes) and tend to occur in or near ports. Recent data reveals that about 66% of oil spills are of medium size, ranging from 7 to 700 tonnes, and more than half of these incidents are reported within the confines of port areas, as indicated by the latest oil tanker spill statistics³. This prevalence of medium-sized spills in such critical areas underscores the need for focused research and improved response strategies in port environments.

Presently, oil spill detection in ports relies largely on accidental discovery by port authority inspectors, leading to delays in reporting and clean-up. This delay is critical, as oil slicks in ports can move at speeds of 0.4 to 0.75 cm/s, potentially travelling over 2 km within an hour, making swift identification (within 30 minutes) crucial for the following benefits⁴. Due to the extended area that a port area can take up, using fixed cameras, is not a feasible option. Therefore drone based measurements are much more suitable as they can traverse a large area, autonomously and signal if there is an oil spill detected. There are implementations. Additionally, the added detection rate has the following advantages:

- Reduced environmental harm due to quicker clean-up.
- More efficient containment and clean-up of the oil.
- Minimal disruption to port operations, reducing economic losses.
- In some cases, clearer identification of the responsible polluter.

Past studies have predominantly concentrated on monitoring large-scale oil spills through satellite imagery and Synthetic Aperture Radar (SAR)^{5,6} techniques. While these methods are effective in open sea conditions, their effectiveness is compromised within port environments due to several limitations. SAR technology, in particular, is less efficient in ports where wave action is minimal and oil spills typically manifest as thin layers, not providing the significant dampening effect needed for SAR detection.

¹InViLab Research Group, Faculty of Applied Engineering, University of Antwerp, Antwerpen, Belgium. ²Port of Antwerp-Bruges, Antwerp, Belgium. ³These authors contributed equally: Thomas De Kerf, Seppe Sels. ✉e-mail: thomas.dekerf@uantwerpen.be

Furthermore, the use of satellite and high-altitude aerial imagery presents challenges in port settings. These methods, while beneficial for broad surveillance, do not offer the fine spatial resolution necessary for identifying smaller spills common in the intricate and bustling nature of port areas. Additionally, the financial implications of satellite usage are substantial, with the high costs becoming a barrier to the routine and detailed monitoring required for managing the complex dynamics of ports.

The integration of RGB and hyperspectral imaging technologies^{7–9} has also been explored, yet these approaches introduce constraints in low-budget commercial scenarios. Given these considerations, there is a pressing demand for alternative strategies better tailored to the specific demands of port environments. These strategies must balance cost-efficiency with the capability to provide accurate, high-resolution monitoring for effective oil spill detection and management.

In our previous work¹⁰, we demonstrated the potential benefits of combining RGB and infrared imagery for oil spill detection. However, the practical application of this approach has been hindered by the high minimum requirements for thermal cameras, rendering it economically unviable at present. Addressing this challenge, this paper introduces a pioneering solution: the first annotated dataset of RGB images specifically captured in a port environment.

This dataset represents a significant advancement in the field of oil spill detection in a port environment. The dataset consists of 1268 images that are fully annotated into three categories: oil, water, and other. Offering an invaluable resource for training and validating detection models. While there exists a body of research and datasets^{11,12}, that focuses on the segmentation of ships and port infrastructure, these studies notably omit the consideration of oil spills within the port environment.

Utilizing this dataset, we have developed a segmentation model that showcases the practical viability of drone-based RGB imaging for oil spill detection. The model demonstrates an f1 score of 0.72, 0.91 and 0.75 for the oil, water and other category respectively. This marks an overall effective classification, underscoring the feasibility of this approach. Our findings reveal that drone-based RGB imaging, without the need for thermal imaging, is not only feasible but also efficient for oil spill detection in port environments.

This research opens up new avenues for cost-effective and accurate oil spill monitoring, leveraging the agility and extensive reach of drone technology combined with the simplicity and effectiveness of RGB imaging. The results from this study are poised to make a significant impact on environmental protection strategies in port areas, offering a ready-to-deploy solution for timely and reliable oil spill detection.

Methods

Data collection. The dataset for this study was meticulously gathered using a drone (Dronematrix YACOB and DJI Mavic2) equipped with high-resolution cameras (4000*2250 and 3840*2160 pixels). The drone flights, conducted from 09/2021 to 09/2023, captured a wide array of images under different environmental conditions. This strategy guaranteed a varied and exhaustive dataset, mirroring a multitude of situations that might arise within a port context. The deliberate selection of the periods for capturing images, encompassing different months and times throughout the day. This ensured the inclusion of a wide spectrum of lighting scenarios, weather changes, and the array of operational tasks characteristic of a port setting.

During the capturing of the images, the height was varied between 30 and 70 meters. Also the camera angle was varied, from a top down approach to a slight incline. This was done on purpose to be able to generalize the dataset.

To align with privacy and confidentiality guidelines, we implemented specific alterations to the raw data, a vital measure to facilitate the dataset's broad utilization without compromising privacy. Critical adjustments involved identifying and processing text¹³, people, and logos. All elements that could reveal individual identities and specific locations within the images. These identified areas were subsequently inpainted¹⁴ to generate anonymized images, ensuring they remain suitable for image processing, without impacting its efficacy.

Furthermore, specific identifiable markings on ships and dockside were also blurred. This measure was taken to ensure that no sensitive information about the port's operations or the identities of the vessels within the port was disclosed. Such modifications are essential in maintaining the integrity and confidentiality of the operations within the port, as well as respecting the proprietary rights of shipping companies.

These modifications, while necessary for privacy and ethical considerations, were implemented in a manner that did not compromise the quality of the data for its intended use in oil spill detection. The alterations were carefully executed to ensure that the key elements necessary for identifying and analyzing oil spills - such as the texture, spread, and coloration of the oil against the water - remained intact and discernible in the images.

Data annotation. The annotation process for our dataset was meticulously carried out using the software platform CVAT (Computer Vision Annotation Tool), accessible at <https://cvat.ai>. To ensure comprehensive coverage and accuracy, the images were divided into several subsets. Each subset was then assigned to different members of the InViLab research group for initial annotation. This division of labour allowed for a more focused and detailed approach to the annotation process, as each researcher could concentrate on a manageable portion of the dataset, thereby enhancing the overall quality and consistency of the annotations.

After the initial annotation phase, a rigorous quality control process was implemented. Senior researchers within the group undertook a thorough review of the annotated images. This step was crucial to ensure the highest level of accuracy and reliability of the annotations. The senior researchers meticulously double-checked each annotation, making corrections and adjustments where necessary. This collaborative and multi-tiered approach to annotation and quality assurance ensured the quality of the dataset, with annotations that accurately represented the various aspects and characteristics of oil spills in a port environment.

Each pixel is categorised into one of either these categories: oil, water or other. The other class is a combination of ships, quays, buildings and air. The choice of such an encompassing super category is made because the

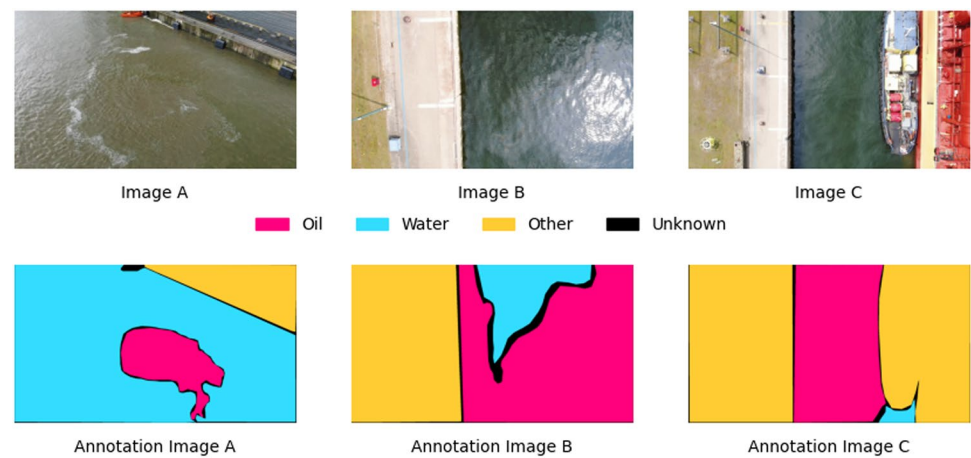


Fig. 1 Top row: Three randomly chosen images from the dataset. Bottom row: the annotated masks of the respective images.

	Category		
	Oil	Water	Other
# Images	994	929	1106
# Annotated Pixels	527361085	835851102	939502502
Relative Annotation (%)	22.9	36.3	40.8

Table 1. Dataset Overview by Category. This table provides a comprehensive breakdown of the dataset across three categories: Oil, Water, and Other. The first row indicates the total number of images featuring each respective category. The second row details the count of annotated pixels labelled under each category, reflecting the extent of detailed annotations provided. The final row presents the percentage of relative annotations per category, thereby illustrating their proportional representation in the entire dataset.

goal of the dataset is to distinguish between oil and water, not to identify the different other elements in a port environment. An overview of three randomly selected images, together with the annotated regions can be found in Fig. 1:

Data Records

The dataset is publicly available and has been contributed to the open-source community, accessible via Zenodo¹⁵. It comes pre-partitioned into training, testing, and validation subsets, following a 70/15/15 percentage split, respectively, to facilitate immediate use in machine learning workflows. Annotation data is provided in three formats to ensure compatibility with various training paradigms: the CamVid format¹⁶, which was also employed during the network’s training phase, as well as the widely recognized COCO¹⁷ and ImageNet¹⁸ formats. These multiple annotation formats enhance the dataset’s versatility, allowing for seamless integration with different models and frameworks.

Dataset statistics. In Table 1, we present essential statistics of our dataset, segmented into three key categories: Oil, Water, and other. This comprehensive overview is critical for understanding the scope and distribution of our data.

The first row of the table indicates the total number of images in which each category appears. The count of 994 images for Oil, 929 for Water, and 1106 for other reflects the dataset’s diversity and the varying frequency of these elements within our collected imagery.

Moving to the second row, we detail the number of annotated pixels for each category. With 527,361,085 annotated pixels for Oil, 835,851,102 for Water, and 939,502,502 for other, these figures demonstrate the extensive level of detail that has been captured in our annotations. This granularity is pivotal for accurate analysis and application in related fields, such as machine learning and environmental monitoring.

The final row presents the percentage of relative annotations for each category, illustrating their proportional representation in the entire dataset. Here, Oil annotations constitute 22.9%, Water 36.3%, and Other 40.8%. These percentages are indicative of the emphasis placed on each category within our dataset and underscore the balanced yet distinct focus on each element, providing a multifaceted perspective on the environmental subjects at hand.

Neural network architecture. We utilized a modified U-Net architecture¹⁹ featuring an EfficientNet encoder for image segmentation tasks. This combination harnesses EfficientNet’s efficient feature extraction capabilities and U-Net’s precision in segmentation application²⁰.

Metric	Category		
	Oil	Water	Other
F1 Score	0.72	0.91	0.75
Precision	0.77	0.91	0.85
Recall	0.79	0.98	0.77
IoU	0.69	0.86	0.62

Table 2. This table provides a detailed breakdown of key performance indicators for the used multiclass image segmentation model. For each category, the table displays the F1 Score, Precision, Recall, and Intersection over Union (IoU), offering a comprehensive view of the model’s effectiveness in accurately identifying and classifying each specific category within the images.

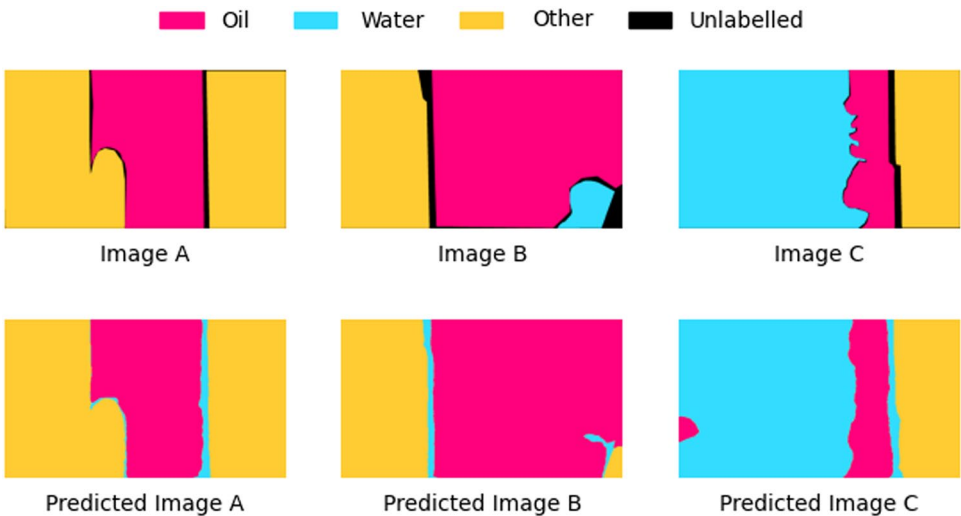


Fig. 2 Predictions on three images from the dataset. The top row represents the annotated mask, and the bottom row is the predicted mask by the neural network.

Encoder. The encoder is based on EfficientNet b4²¹, starting with a convolutional layer for initial feature extraction and followed by batch normalization. The core consists of Mobile Inverted Bottleneck Convolution Blocks with depthwise separable convolutions, incorporating Squeeze-and-Excitation layers for channel-wise feature recalibration and the MemoryEfficientSwish activation function for balance between performance and memory efficiency. The encoder was pretrained on the ImageNet Database¹⁸.

Decoder. The U-Net decoder includes multiple DecoderBlocks with convolutional layers and ReLU activation, designed to upsample the feature maps to the original image size. Attention mechanisms are integrated to focus on salient features.

Segmentation head. The final part of the network, the Segmentation Head, transforms decoded features into a segmentation map through a convolutional layer, followed by a softmax activation for multiclass classification.

Segmentation results. Table 2 offers a detailed evaluation of the multiclass image segmentation model used in our study. For each category - Oil, Water, and Other - the model’s performance is assessed using four critical metrics: F1 Score, Precision, Recall, and Intersection over Union (IoU). These metrics collectively provide a multifaceted view of the model’s accuracy and efficiency in correctly identifying and classifying each category within the dataset.

As observed, the model demonstrates strong performance across all categories, with particularly high scores in Precision and Recall for the ‘Water’ category. The F1 Score, which balances Precision and Recall, is notably high for ‘Water’ and ‘Oil’, indicating a robust capability of the model in handling these categories. The Intersection over Union (IoU) metric, which assesses the overlap between predicted and ground truth areas, also shows commendable results, especially for ‘Water’.

While the model achieves a good overall accuracy, a closer inspection of its predictions, as shown in Fig. 2, reveals notable limitations in distinguishing between specific classes. A significant challenge is observed in differentiating ‘oil’ from ‘water,’ especially in areas where these substances coexist or are in close proximity. The model frequently struggles to accurately delineate the boundaries between oil and water, resulting in a high rate of misclassifications in these regions. This difficulty is compounded by the nature of the transitions between oil and water, which often appear as gradual rather than sharp, well-defined borders. Such transitions are particularly problematic for the model in scenarios where the boundary is subtle or where oil has partially dispersed

on the water surface, creating a blended zone that does not distinctly represent either category. This lack of clear demarcation poses a significant challenge to the model's ability to accurately classify these overlapping or closely related regions.

Overall, while the model's general accuracy is commendable, these observed discrepancies underscore the need for further refinement. Enhancing the model could involve adjusting the model's architecture and hyper-parameters to better capture the nuances between different categories.

Technical Validation

Our dataset was carefully curated through the deployment of drones equipped with cameras, ensuring the capture of a wide array of images across varying environmental conditions and locations. These varying environmental conditions are primarily related to weather, particularly the sun's position and brightness. Variations in sunlight intensity, angle, and shadow can alter the appearance of objects in images, affecting contrast, color, and texture. In this dataset, these varying conditions are present, making it a representative for real-life cases.

To uphold privacy while retaining the dataset's efficacy for oil spill detection, specific alterations were implemented. This included the use of advanced inpainting techniques and selective blurring to anonymize identifiable features effectively, ensuring the resulting images remained highly relevant for analytical purposes.

Annotation validation. The annotation process was a collaborative effort led by the InViLab research group, with equitable distribution of images among members for initial annotation. Dr. Sels played a pivotal role in reviewing and refining these annotations to guarantee precision. Post-annotation, the images were anonymized and further processed using inpainting to preserve the integrity of the dataset without compromising network performance. This careful anonymization process was meticulously reviewed by both Dr. Sels and Dr. De Kerf, ensuring the highest standards of data quality and reliability were maintained.

Influence of the inpainting process. To assess the impact of our anonymization technique on model performance, we conducted a comparative analysis. Initially, a model was trained using the original dataset, which had not been subjected to the anonymization process. This model's accuracy was validated against a distinct validation set to establish a performance benchmark. Subsequently, the same model was evaluated using the anonymized dataset, ensuring consistency in testing conditions. Remarkably, the comparative analysis revealed no discernible difference in the validation results between the anonymized and non-anonymized datasets. This finding demonstrates that our inpainting-based anonymization process does not compromise the model's ability to detect oil spills, affirming the integrity of our data processing methodology.

Code availability

The data for this study was generated independently of any coding process. For reproducibility and further research, the complete code base for our analysis, figures, and tables creation is available on GitHub at https://github.com/thomasdekerf/OilSpillDataset_Analysis. This repository also contains the code for training and evaluating our neural network, alongside the optimally trained model saved in a .pth file format, facilitating direct use and exploration by other researchers.

Received: 27 February 2024; Accepted: 8 October 2024;

Published online: 30 October 2024

References

- Beyer, J., Trannum, H. C., Bakke, T., Hodson, P. V. & Collier, T. K. Environmental effects of the deepwater horizon oil spill: A review. *Marine Pollution Bulletin* **110**, 28–51, <https://doi.org/10.1016/j.marpolbul.2016.06.027> (2016).
- Readman, J. W. *et al.* Recovery of the coastal marine environment in the Gulf following the 1991 war-related oil spills. *Marine Pollution Bulletin* **32**, 493–498, [https://doi.org/10.1016/0025-326x\(95\)00227-e](https://doi.org/10.1016/0025-326x(95)00227-e) (1996).
- Itpof. Oil Tanker Spill Statistics 2019. Tech. Rep., International Tanker Owners Pollution Federation (2019).
- Janati, M., Kolahdoozan, M. & Imanian, H. Artificial Neural Network Modeling for the Management of Oil Slick Transport in the Marine Environments. *Pollution* **6**, 399–415, <https://doi.org/10.22059/poll.2020.289549.684> (2020).
- Topouzelis, K. N. Oil spill detection by sar images: Dark formation detection, feature extraction and classification algorithms. *Sensors* **8**, 6642–6659 (2008).
- Fiscella, B., Giancaspro, A., Nirchio, F., Pavese, P. & Trivero, P. Oil spill detection using marine sar images. *International Journal of Remote Sensing* **21**, 3561–3566 (2000).
- Salem, F., Kafatos, M., El-Ghazawi, T., Gomez, R. & Yang, R. Hyperspectral image analysis for oil spill detection. In *Summaries of NASA/JPL Airborne Earth Science Workshop, Pasadena, CA*, 5–9 (2001).
- Khanna, S. *et al.* Comparing the potential of multispectral and hyperspectral data for monitoring oil spill impact. *Sensors* **18**, 558 (2018).
- Koirala, B. *et al.* Study on the potential of oil spill monitoring in a port environment using optical reflectance. *Remote Sensing* **15**, 4950, <https://doi.org/10.3390/rs15204950> (2023).
- De Kerf, T., Gladines, J., Sels, S. & Vanlanduit, S. Oil spill detection using machine learning and infrared images. *Remote sensing* **12**, 4090 (2020).
- Prasad, D. K., Rajan, D., Rachmawati, L., Rajabally, E. & Quek, C. Video processing from electro-optical sensors for object detection and tracking in a maritime environment: A survey. *IEEE Transactions on Intelligent Transportation Systems* **18**, 1993–2016, <https://doi.org/10.1109/TITS.2016.2634580> (2017).
- Shao, Z., Wu, W., Wang, Z., Du, W. & Li, C. Seaships: A large-scale precisely annotated dataset for ship detection. *IEEE Transactions on Multimedia* **20**, 2593–2604, <https://doi.org/10.1109/TMM.2018.2865686> (2018).
- Baek, Y., Lee, B., Han, D., Yun, S. & Lee, H. Character region awareness for text detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 9365–9374 (2019).
- Suvorov, R. *et al.* Resolution-robust Large Mask Inpainting with Fourier Convolutions, 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, pp. 3172–3182, <https://doi.org/10.1109/WACV51458.2022.00323> (2022).

15. De Kerf, T., Sels, S. & Vanlanduit, S. Annotated rgb of oil spills in a port environment. *Zenodo* <https://doi.org/10.5281/zenodo.10555314> (2024).
16. Brostow, G. J., Fauqueur, J. & Cipolla, R. Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters* **30**, 88–97 (2009).
17. Lin, T.-Y. *et al.* Microsoft coco: Common objects in context (2015).
18. Deng, J. *et al.* ImageNet: A Large-Scale Hierarchical Image Database. In *Cvpr09* (2009).
19. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation (2015).
20. Baheti, B., Innani, S., Gajre, S. & Talbar, S. Eff-unet: A novel architecture for semantic segmentation in unstructured environment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 358–359 (2020).
21. Tan, M. & Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, 6105–6114 (PMLR, 2019).

Acknowledgements

The authors wish to thank the contribution of the Port of Antwerp-Bruges for the collaboration and images. This research is made possible thanks to the Belgian SPF Economy ETF:Phairywind Project.

Author contributions

Must include all authors, identified by initials, for example: A.A. conceived the experiment(s), A.A. and B.A. conducted the experiment(s), C.A. and D.A. analysed the results. All authors reviewed the manuscript.

Competing interests

The authors declare that they have no competing interests or personal relationships that could have influenced the work reported in this paper.

Additional information

Correspondence and requests for materials should be addressed to T.D.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024