

Article

Two-Stage Oil Spill Detection in SAR Using a Domain-Adapted Segment Anything Model

George Giannopoulos ^{1,*}, Maria Kremezi ¹, Vasilia Karathanassi ¹, Vassilis Andronis ¹, Dimitris Bliziotis ², Katerina Kikaki ², Ana Sofia Oliveira ³ and Ariane Mütting ³

¹ School of Rural Surveying and Geoinformatics Engineering, National Technical, University of Athens, 157 72 Athens, Greece; mkremezi@central.ntua.gr (M.K.); karathan@survey.ntua.gr (V.K.); andronis@central.ntua.gr (V.A.)

² Hellenic Space Center, Ministry of Digital Governance, 153 42 Agia Paraskevi, Greece; dimitris.bliziotis@hsc.gov.gr (D.B.); akikakh@hsc.gov.gr (K.K.)

³ ESTEC—European Space Research and Technology Centre, Keplerlaan 1, 2200 AG Noordwijk, The Netherlands; anasofia.oliveira@esa.int (A.S.O.); ariane.mutting@esa.int (A.M.)

* Correspondence: g_giannopoulos@mail.ntua.gr

Highlights

What are the main findings?

- OSDA-SAM, a domain-adapted SAM segmentor, improves mask quality over CNN baselines with only ~2.6M trainable parameters.
- OSDA-SAM benefits from incorporating GLCM-derived texture information alongside VV backscatter

What are the implications of the main findings?

- Foundation segmentation models can be adapted to SAR effectively, using parameter-efficient tuning with limited labels, for oil spill segmentation.
- Feature integration may improve the transferability of SAM-based approaches to other SAR segmentation tasks.

Abstract

Synthetic Aperture Radar (SAR) is widely used for marine oil spill surveillance due to its all-weather capabilities and sensitivity to sea surface roughness. However, oil slicks often appear as dark formations that can be confounded with visually similar “look-alikes”, making automated detection and boundary delineation challenging. This study proposes a two-stage deep learning framework for oil spill mapping in Sentinel-1 SAR imagery. First, a ConvNeXt-T classifier screens image patches for likely slick presence, reducing the search space for dense prediction. Second, spill boundaries are extracted with a domain-adapted Segment Anything Model (SAM) configured for prompt-free, single-shot segmentation. The input representation is enhanced by combining preprocessed Sentinel-1 VV backscatter with Gray-Level Co-occurrence Matrix (GLCM) texture measures (homogeneity and variance) to better separate oil from heterogeneous background sea at the segmentation level. Quantitative evaluation against established segmentation baselines demonstrates that our adapted SAM achieves the highest overall accuracy, reaching an F1-score of 0.86. This outperforms traditional models such as UNet and CBDNet (0.83), as well as DeepLabV3, SegNeXt, and OFCNet (all at 0.82). Furthermore, an analysis of the wind speed on the test set shows that wind speed affects detectability but does not by itself determine segmentation quality. The results indicate that combining transformer-based screening with efficient foundation-model adaptation can provide accurate and scalable oil spill mapping for operational SAR monitoring.



Academic Editor: Merv Fingas

Received: 8 May 2026

Revised: 4 June 2026

Accepted: 9 June 2026

Published: 12 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

Keywords: oil spill detection; SAR; deep learning; semantic segmentation; Segment Anything Model (SAM); domain adaptation; GLCM texture features

1. Introduction

Sea surface oil spill pollution is a major environmental threat that has adverse effects to the marine environment and the living organisms that inhabit it [1–3]. Even though most of the oil that ends up at sea comes from natural underwater seepages, a significant amount is anthropogenic. Most of this anthropogenic oil does not originate from accidents, but rather from illegal routine sea operations. Thus, oil spills have a high correlation with ship routes, oil platforms, pipelines and other offshore installations [4–6]. Due to its high kinematic viscosity (higher than the surrounding sea water), oil has a dampening effect on the short gravity/capillary waves, thereby reducing the backscatter coefficient and appearing as dark patches in SAR images. This effect varies spatially across the scene, depending on the type of oil, thickness of the spill and the weather conditions present [7,8].

SAR images have been successfully used for oil spill identification and delineation for decades [9–11]. Their ability to monitor the Earth regardless of time of day and weather conditions, combined with the minimal effect of the atmosphere on the microwave region of the spectrum, makes them very useful for observing various marine phenomena. For an oil spill to be observable in an SAR image, it needs to be distinguishable from the surrounding water. This is very prominent in VV polarization and at wind speeds ranging from approximately 3 to 7–10 m/s. At speeds lower than that, the sea has equally low backscatter values, and at higher speeds it mixes completely with the surrounding water (unless the oil is thick enough). Additionally, the backscattering coefficient decreases as the incident angle increases, making oil spills more visible at angles between 20 and 45 degrees.

However, oil spills are not the only phenomenon that can cause a damping effect. Other phenomena, termed look-alikes, appear very similar to an actual spill. Most commonly, look-alikes can be produced by low winds in the lee of islands, natural surface films produced by fish and plankton and algal blooms. It has been observed that oil spills produce a dampening effect in the range of 0.6–13.0 dB while look-alikes produce one in the range of 0.8 dB to 11.3 dB. This overlap in values makes their discrimination difficult. One way to separate oil spills from look-alikes is to consider their shape. Oil spills are more linear and have clearer boundaries, especially on the downwind side. They may also exhibit feathering due to the wind [12,13].

Deep learning is the current state of the art for pixel-wise image segmentation. It has largely replaced traditional methods that require handcrafted feature extraction. Encoder–decoder designs such as U-Net [14] have become the standard for combining multi-scale context via skip connections, which is especially effective for thin, fragmented targets and precise boundary delineation. Many modern models strengthen this encoder–decoder architecture with residual backbones, most commonly ResNet [15], to improve optimization stability and representation. DeepLabV3 [16] uses atrous (dilated) convolutions and multi-rate context aggregation to improve robustness to scale variation, while SegNeXt [17] revisits efficient convolutional attention as a practical alternative to global self-attention. Swin-T [18] and ViT [19] provide strong hierarchical and token-based representations for dense prediction, while ConvNeXt [20] shows that modernized convolutional backbones can achieve comparable segmentation performance to transformer-based architectures. More recently, foundation models have reshaped segmentation when ground truth is scarce. The Segment Anything Model (SAM) performs promptable segmentation, producing masks conditioned on points, boxes, or text to enable strong zero-shot transfer

across diverse image distributions [21]. In remote sensing, however, systematic domain shifts caused by sensor physics, spatial resolution and spatial and temporal variations motivate dedicated evaluation and adaptation. SAM has been studied as a baseline and annotation tool for point and box-driven workflows [22], and it has also been used to build large-scale remote-sensing segmentation resources that support pretraining when pixel labels are expensive [23]. Several works propose adaptations tailored to remote-sensing characteristics: RS-SAM [24] integrates multi-scale modeling, SAM-RSIS [25] uses box-prompted progressive adaptation for instance segmentation, and SAM-assisted semantic segmentation injects object/boundary constraints into conventional semantic heads to improve mask quality [26]. Parameter-efficient fine-tuning methods like LoRa provide a pragmatic route to adapting SAM-like encoders to geographically diverse distributions with limited trainable parameters and potentially limited labels [27]. Related training-free or low-supervision approaches further demonstrate that SAM's representations can be repurposed beyond natural images when domain gaps are explicitly addressed [28].

In SAR oil spill mapping, deep learning is increasingly used to perform semantic segmentation on data that are noisy and hard to interpret. Speckle, variations in incidence angle, and many look-alike features can appear similar to oil slicks, so accurate boundary delineation and low false alarm rates are key goals in practice [29–31]. To cope with strong class imbalance and limited labeled data, ref. [29] proposed a two-stage approach that first classifies patches and then applies a U-Net-style segmentation model with an imbalance-aware loss. OSCNet [30] showed that CNN features can improve separation of oil from look-alikes in classification tasks that are often used prior to segmentation. Benchmark datasets and evaluations, such as those introduced by [31], have supported more systematic comparison of segmentation models and highlight that distinguishing oil from look-alikes remains difficult at the pixel level.

Recent studies have focused on making models scalable and improving boundary quality. Ref. [32] introduced a fully convolutional model (OFCN/OFCNet) designed to handle variable-sized SAR images and run efficiently at a large scale (e.g., using sliding-window inference). The study reported performance comparable to that of human operators in large-area detection and categorization when training data were sufficiently diverse and inference was carefully engineered. Other work has targeted boundary errors more directly: CBD-Net combines multi-scale features with edge-focused supervision to sharpen spill outlines, supported by the manually labeled SOS dataset that helps with more consistent benchmarking [33]. Attention-based encoder–decoder variants, such as dual-attention U-Net models, also aim to reduce missed spill regions and unclear boundaries in noisy SAR scenes [34]. Reliability can further improve by adding extra inputs to reduce look-alike confusion, for example, by using SAR intensity together with derived statistics (e.g., variance) and environmental/context information (e.g., wind) rather than relying on a single channel [35]. More recently, hybrid CNN–Transformer models have been explored to capture both local texture and wider context, while multi-task approaches (including GAN-based frameworks) attempt to learn oil and look-alike discrimination and pixel-level segmentation jointly under limited labeled data [36,37]. In [38], a modified version of SegNeXt was introduced for segmenting different kinds of marine phenomena and polluting agents in Sentinel 2 images including marine debris and oil spills.

A common limitation in SAR oil spill segmentation is the lack of labeled data and the difficulty in training models that generalize well. To reduce this dependence on fixed datasets, self-evolving training approaches iteratively create new training samples and update the model as additional SAR scenes become available, which can partially overcome the limits of static labeled corpora [39]. At the same time, detector-based methods trained on large Sentinel-1 collections support scalable monitoring, but accurate spill

extent detection still depends on reliable downstream segmentation [40,41]. Other studies, including CNN-based segmentation, adversarial learning methods, and dual-stream U-Net variants, show that improving generalization with limited supervision remains a key open problem [42–44].

Recent studies have also begun to explore the Segment Anything Model (SAM) and its variants for oil spill detection. SAM-OIL [45] first introduced SAM into SAR-based oil spill detection by combining YOLOv8-generated bounding boxes, an adapted SAM, and ordered mask fusion to produce class-aware segmentation masks. Subsequent works extended this idea through stronger domain adaptation and feature fusion strategies. CoRemoteSAM-Oil [46] and HSRD-Net [47] combine RemoteSAM with ResNet-based visual features to improve generalization and suppress false alarms, while DADS-SAM [48] applies parameter-efficient SAM fine-tuning to complex UAV-based port oil spill scenes. More recently, OilSAM2 [49] introduced a memory-augmented SAM2 framework to reuse multi-scale information across SAR image collections. Overall, these works show that SAM-based foundation models are promising for oil spill segmentation but also highlight the need for SAR-specific adaptation, efficient fine-tuning, and robust handling of look-alikes, scarce labels, and variable sea-state conditions.

A complementary line of radar-polarimetry research shows that scattering behavior is not fixed but changes with polarization basis, time, frequency, and observation geometry. The General Polarimetric Correlation Pattern (GPCP) [50] formalizes this idea by visualizing and characterizing target scattering diversity across multiple domains. PolSAR and dual-polarimetric SAR data provide extra scattering information that can reduce the ambiguity of dark-spot interpretation. Several studies have therefore incorporated polarimetric information in segmentation networks. Ma et al. [51] used Sentinel-1 dual-polarimetric amplitude, phase, and Cloude decomposition parameters in an improved DeepLabv3+ model. Wang et al. [52] proposed BO-DRNet using quad-polarimetric RADARSAT-2 features and Bayesian hyperparameter optimization. Another Wang et al. study [53] introduced a Cloude–Pottier-based relative polarimetric feature. Liao et al. [54] used PolSAR and deep learning for coastal oil spill risk monitoring in Jiaozhou Bay. More recent work combines polarimetric inputs with stronger segmentation designs. OSDTAU-Net [55] uses dual-polarimetric SAR with Transformer and attention mechanisms. A scene-adaptive PolSAR network [56] uses dynamic convolution and boundary constraints. Xiang et al. [57] combine composite polarimetric scattering power entropy with multi-scale hybrid feature fusion to improve detection of small and elongated slicks. CDANet [58] uses a multi-year, multi-region, multi-polarimetric SAR dataset to improve generalization. Finally, PBITU-Net [59] combines dual-polarimetric features with oil–seawater boundary information.

Despite the strong performance reported by the aforementioned studies, with F1-scores reaching as high as 0.98, their experimental validation is typically based on small image collections and limited test sets, averaging approximately three test images. Consequently, their reported performance may not fully reflect robustness across diverse oil spill appearances, acquisition conditions, and sea states.

In this study, we aim to exploit the deep features of the foundational model SAM through OSDA-SAM, a novel architecture for segmenting oil spills. Unlike standard SAM, which was trained primarily on natural RGB images and typically requires prompts, OSDA-SAM operates in a prompt-free, single-shot segmentation setting while keeping the SAM backbone frozen. The SAR-to-SAM domain gap is addressed through three lightweight adaptation components. First, the Input Domain Adaptation Block (IDAB) maps SAR-derived inputs into a SAM-compatible representation using a residual convolutional adapter and learnable channel-wise normalization. This helps the model make the SAR input cleaner and more suitable for SAM. Second, LoRA introduces trainable low-rank

updates into selected encoder projections, allowing the internal SAM representations to adapt to SAR oil spill features without full fine-tuning. This preserves the general segmentation knowledge of SAM while reducing trainable parameters and improving stability under limited labeled data. Third, the Residual Feature Space Adapter (RFSA) is applied after the frozen encoder to refine the generated image embeddings through lightweight pointwise convolutions and residual feature reweighting. This provides an additional feature-level correction.

The proposed OSDA-SAM differs from existing SAM-based oil spill segmentation approaches by avoiding manual prompts, memory modules, or substantial modifications to the SAM architecture. It also differs from approaches that either apply SAM directly in a zero-shot manner or mainly focus on improving multi-scale encoder features. Instead, OSDA-SAM keeps the SAM backbone largely unchanged and introduces only lightweight adaptation modules at selected stages. Because marine SAR images differ substantially from optical images, we place particular emphasis on adapting the input before it enters the encoder. This makes the SAR data more compatible with the type of input SAM was originally trained on, while keeping the overall architecture simple and efficient.

Our approach employs a two-stage methodology similar to [29]. In the first stage, a ConvNeXt classification model is trained to distinguish oil spill patches from look-alikes. In the second stage, the proposed OSDA-SAM segmentation model is used to delineate the oil spill boundaries. Since the SAM image encoder is constrained to three-channel inputs, our approach uses VV backscatter together with selected texture descriptors rather than incorporating the full set of dual-polarimetric features, which would require modifying the pretrained encoder and reduce the benefit of using SAM as a foundation model. By evaluating this architecture on a broader dataset, this study aims to provide a more robust assessment of oil spill segmentation performance under diverse real-world conditions. The main contributions of this study are as follows:

- (1) We propose OSDA-SAM, a novel network architecture for segmenting images containing oil slicks. The architecture is based on the foundational model SAM and is efficiently adapted to the task by LoRa linear layers, taking advantage of the rich knowledge of the model.
- (2) A Domain Adaption Block is designed to map SAR-derived inputs to a SAM-compatible appearance, which consists of a residual convolutional adapter and a learnable channel-wise normalization.
- (3) The effect of incorporating GLCM-derived statistics alongside VV backscatter as an input to deep learning models for oil spill detection and segmentation is analyzed.
- (4) A statistical analysis of the effect of wind speed on the performance of our deep learning model is conducted.

The paper is organized as follows: Section 2 provides a description of the entire methodology and data used in the study; Section 3 provides the results of the proposed method compared to the current state of the art; and Section 4 discusses the effectiveness of the algorithm and Section 5 concludes the study.

2. Materials and Methods

A two-stage pipeline is employed for the detection of oil spills on Sentinel-1 SAR images. Preprocessed Sentinel-1 VV image patches (2048×2048) are combined with Gray-Level Co-occurrence Matrix (GLCM) [60] texture features (homogeneity and variance) to enable better separation of oil–water boundaries. Although the radar sensor captures dual-polarization data, the VH channel was intentionally excluded as an input to the network. This decision was motivated by the physical interaction of cross-polarized signals with the marine surface. While the VV channel is sensitive to surface scattering from the

sea surface and oil slicks, the VH cross-polarization channel tends to produce stronger responses from volume scatterers, such as ships and oil platforms. Consequently, raw VH backscatter over both clean water and oil-covered areas remains consistently low, providing limited discriminatory information for oil spill segmentation, even though it may return a stronger signal for the oil source (in cases where it is present). Since SAM expects a three-channel input, textural features derived from the VV signal were used instead to enhance the information content of the input data. In the first stage, VV backscatter image patches are resized to 512×512 and fed into a ConvNeXt-T classifier to identify candidate patches that may contain oil. In the second stage, the positives are fed into the segmentation pipeline: the VV backscatter is combined with the GLCM texture features, and oil slick segments are obtained using the proposed OSDA-SAM. OSDA-SAM is a domain-adapted Segment Anything Model that operates prompt-free in a single-shot fashion. The SAM foundation backbone is frozen, and the lightweight adaptation components, which include an input-domain mapping block and LoRA [61] updates, are applied to adapt the model to the SAR image domain, which is different from the RGB-like inputs that the SAM was originally designed for.

2.1. Data Collection

2.1.1. Data Annotation

To train the models, a dataset containing patches of Sentinel-1 VV backscatter images of size 2048×2048 was created. The patches were extracted from 660 images containing oil spills and 200 images not containing oil spills but look-alikes and sea (background patches). The background scenes were chosen from various regions and periods and with as much visual variety as possible. The absence of any oil spills was evaluated by photointerpretation and cross-checking with known oil spill alerts. The dataset was enhanced with textural information, specifically the homogeneity and variance measures calculated from the Gray-Level Co-occurrence Matrix (GLCM). In total, 2168 negative and 660 positive patches were created. Out of the negative patches, approximately 930 depict scenes that contain look-alikes.

All Sentinel-1 images were orbit-corrected, and thermal and border noise was removed. The images were then radiometrically calibrated to backscatter values and terrain-corrected. Although the images depict marine environments, where terrain does not affect backscatter values, terrain correction is necessary to project the images to ground geometry so that they can be correctly combined with other geospatial data. Additionally, speckle noise was suppressed using the Lee speckle filter. The Lee filter was applied with a window size of 7×7 , a target window size of 3×3 , and a sigma value of 0.9. All preprocessing was performed using ESA's SNAP software, version 12.0.

The dataset was annotated in a task-specific manner, using image-level labels for the classification model and pixel-wise masks for the segmentation model, depending on the task. Information on the oil spills and their spatial extents was acquired from the Office of Satellite and Product Operations (OSPO, NOAA/NESDIS; Suitland, MD, USA) in vector geocoded format [62], while the Sentinel-1 images were obtained through the European Space Agency's (ESA; Paris, France) Copernicus Data Space Ecosystem [63].

OSPO is part of the National Oceanic and Atmospheric Administration (NOAA) and is responsible for managing NOAA's operational satellite systems and ensuring the distribution of satellite data. The NESDIS Satellite Analysis Branch (SAB) of the office is responsible for identifying marine anomalies, including oil spills, from satellite imagery. Oil slicks are mapped based on visual inspection and an automated oil spill mapping tool. This product, called the Marine Pollution Surveillance Report (MPSR), is publicly available and offers mapped oil spill areas in geocoded format. From 2016 to present,

thousands of oil spill events have been detected and mapped using optical and radar satellite imagery, mostly along the coasts of the USA and the Gulf of Mexico, including some in international waters.

In the MPSRs, oil spills are categorized based on the analyst's confidence that the anomaly spotted is indeed an oil spill event. An event can have High, Medium-High, Medium and Low Confidence. An event has High Confidence when there is a report that clearly confirms the presence of an oil slick, and the spill is distinguishable from natural phenomena. Medium-High confidence applies in the same situation as high confidence, but when the event exhibits moderate to low contrast to the background. An event is assigned Medium Confidence when there is no additional confirmation that it is indeed an oil spill but is well defined in the SAR image. Lastly, an event is assigned Low Confidence when it is considered an oil spill by the analyst but there is no additional evidence that supports the claim and, additionally, there are natural phenomena nearby. In the collected data, none of the events have Low Confidence, while most of them are Medium-High and above. Figure 1 shows the locations of these events.

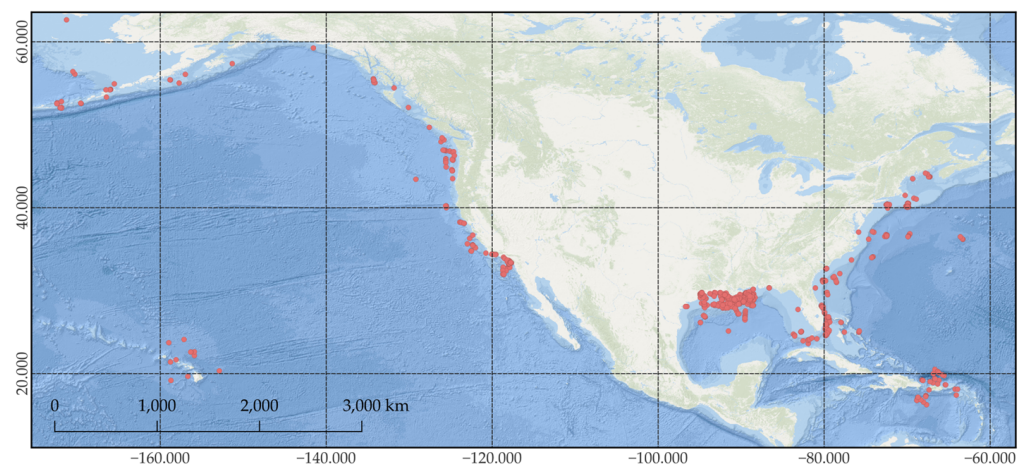


Figure 1. Mapped oil spill events from Sentinel 1 imagery by OSPO, from 2016 to 2025. Basemap: © OpenStreetMap contributors (ODbL).

2.1.2. Wind Speed and Incidence Angle

For our dataset, we examined the wind speed and incidence angle of every scene. Figures 2 and 3 show the kernel density estimates of the per-patch mean incidence angle and wind speed for each class and data split. For incidence angle, the full Sentinel-1 IW range (30–46°) is represented, with the oil spill class slightly skewed toward higher angles. Regarding wind speed, background patches exhibit higher values, although all distributions peak at similar values. This is desirable because it allows the dataset to include dark areas produced under stronger wind conditions. In contrast, oil spills are typically detectable at wind speeds of roughly 2–3 up to 10 m/s, which is also reflected in the statistics. Wind speed was calculated from the 10 m eastward (u) and northward (v) wind components, provided by the Copernicus Climate Change Service (C3S), and implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF; Reading, United Kingdom) [64]. It was calculated as follows:

$$\text{wind speed} = \sqrt{u^2 + v^2} \quad (1)$$

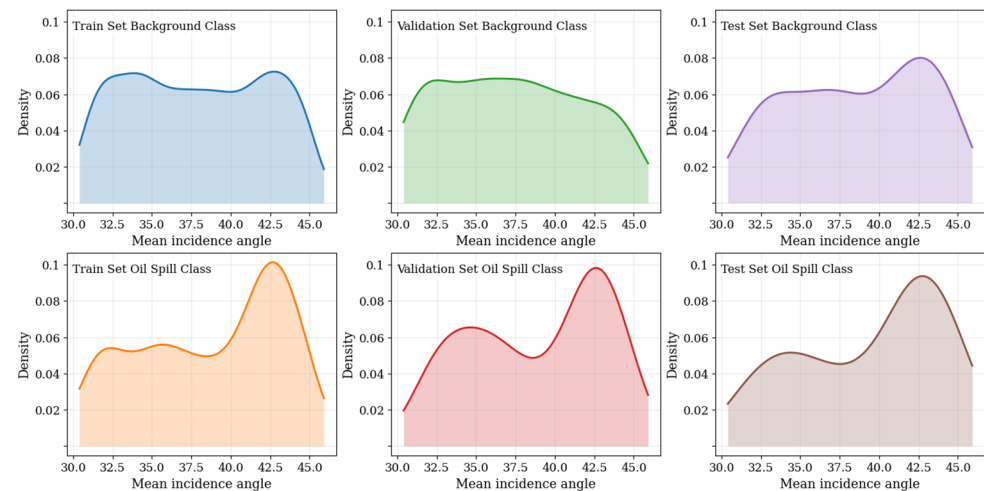


Figure 2. Kernel density estimates of the patch average SAR incidence angle for every dataset split (columns) and for each class (rows).

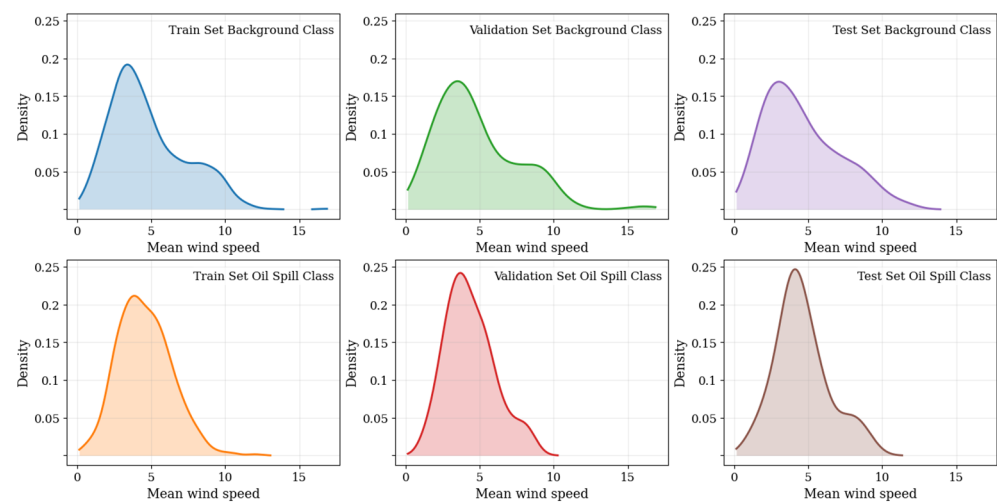


Figure 3. Kernel density estimates of the patch average wind speed for every dataset split (columns) and for each class (rows).

2.2. Texture Analysis

The GLCM is a statistical method commonly used for second-order image texture analysis. It is a normalized symmetrical matrix, that represents the probability of a pair of gray intensity values to occur in the image, given a small offset between the two pixels that constitute the pair. Two texture measures were calculated from the GLCM of the preprocessed VV backscatter values: homogeneity and variance. They are defined as follows:

$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} \frac{p_{ij}}{1 + (i - j)^2} \quad (2)$$

$$\text{Variance} = \sum_{i,j=0}^{N-1} p_{ij}(i - \mu)^2 \quad (3)$$

where p_{ij} is the element of the GLCM at row i and column j , μ is the GLCM mean and N stands for the number of gray levels in the image. Homogeneity is a measure of similarity between neighboring pixels with higher values indicating less variation in the intensity values of the image. Variance measures the spread of these intensity values. Oil spills are expected to have higher homogeneity values and lower variance values than the sea.

This information enhances the model's ability to differentiate boundaries between the oil spills and the sea during the segmentation process, as well as limiting false positives. Figure 4 highlights the Sentinel 1 VV backscatter image patches of the oil spills. They are accompanied by their computed texture measures and their corresponding RGB composites. When looked at as a color composite (RGB: VV, homogeneity, variance) all dark areas appear green, while bright areas appear purple. Furthermore, most oil spills, in contrast to most look-alike areas, have clear defining boundaries with the background. The first and third rows show scenes with higher wind speeds and more turbulent water than the other two. For the calculation of the two GLCM features, a window size of 9×9 was used. Larger window sizes may lead to loss of information while smaller may lead to loss of broader textures. The images were binned into 32 gray levels since the sea surface is not expected to present huge value variations.

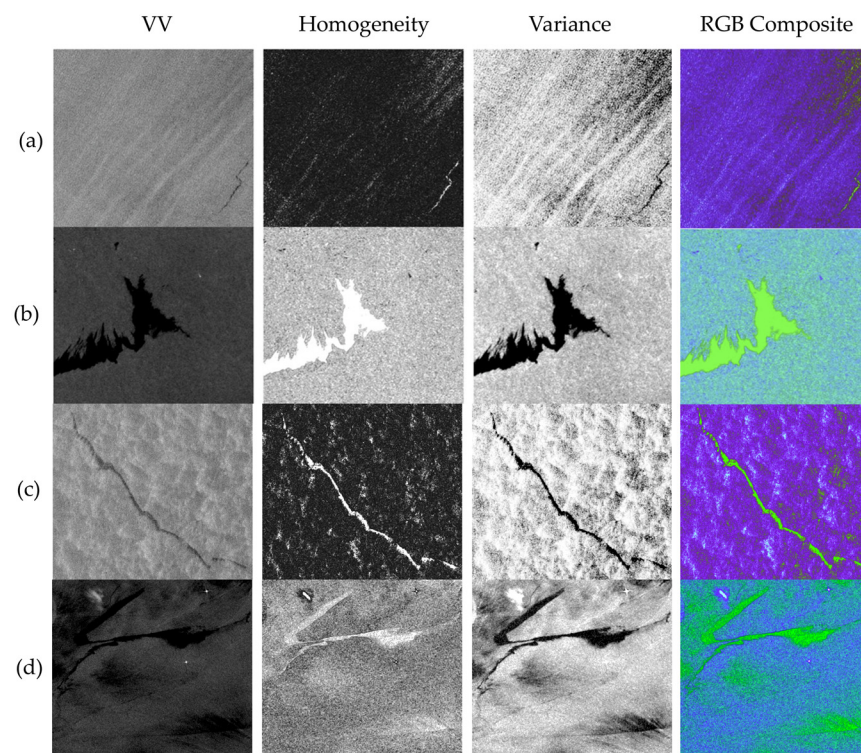


Figure 4. Images containing oil spills. The preprocessed VV backscatter alongside the two GLCM homogeneity and variance texture features and their RGB visualization is displayed. The rows illustrate: (a) a small, elongated oil spill; (b) a large oil spill area with clear boundary separation; (c) an elongated oil spill in a high-wind area; and (d) an oil spill in a scene containing look-alikes.

2.3. Methodology

2.3.1. Classifier

ConvNeXt-T (Tiny) is a hierarchical convolutional backbone that updates the classic ResNet-style design with several Transformer-inspired refinements, including large kernel depthwise convolutions and LayerNorm-based blocks. It is organized into four stages that progressively increase feature dimensionality (96, 192, 384, and 768 channels) while reducing spatial resolution through downsampling. The network produces multi-scale feature representations and, for classification, applies global average pooling followed by a linear prediction head. In this work, the final head is replaced to match the target classes, and pretrained weights are used to improve initialization and training stability.

2.3.2. OSDA-SAM

The Segment Anything Model (SAM) is a promptable segmentation framework designed to produce object masks from images given simple user inputs such as points, boxes, or coarse masks. It couples a large, general-purpose image encoder with a prompt encoder that converts user guidance into embeddings, and a lightweight mask decoder that combines image and prompt embeddings to predict one or more segmentation masks. Trained at scale for broad segmentation generalization, SAM can transfer effectively to many domains and tasks, either directly in a zero-shot setting or with lightweight adaptation modules when the input distribution differs substantially from natural images. OSDA-SAM is an adaptation of SAM to the SAR domain for oil spill segmentation. SAM is kept frozen through the entire training process, while a set of small and lightweight trainable components reconcile the input domain shift at the pixel level, adjust normalization statistics and modify SAM's encoder representation through low-rank updates. The overall framework is described in Figure 5.

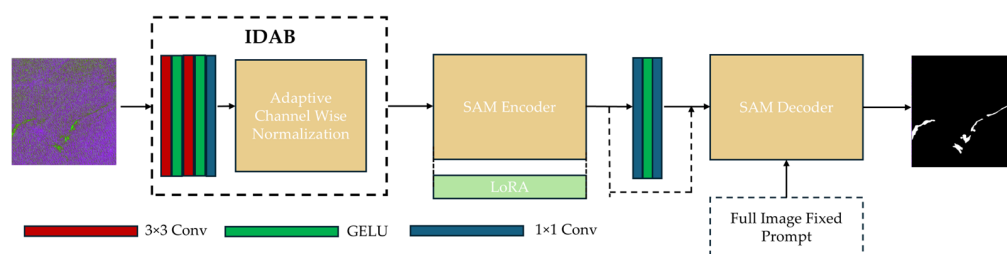


Figure 5. OSDA-SAM network architecture.

First, the Input Domain Adaptation Block (IDAB) is implemented in pixel space to map SAR-derived inputs to a SAM-compatible appearance. IDAB consists of a shallow residual convolutional adapter followed by adaptive, learnable channel-wise normalization. These two parts address two distinct aspects of the domain shift: local structure and global statistics. The adapter is made of two 3×3 convolutions with GELU activations [65], followed by a final 1×1 convolution that maps the output back to three channels. It can learn content-dependent, spatial corrections (e.g., suppressing speckle-like texture, reshaping local edge/contrast cues), while the residual connection ensures it can stay close to the identity when little correction is needed. The subsequent adaptive normalization aligns channel statistics via learnable mean and scale parameters:

$$\text{AdaNorm}(x) = \frac{x - \mu}{\sigma + \varepsilon}, \sigma = \exp(s) \quad (4)$$

with $\mu, s \in \mathbb{R}^3$ learned per channel. In other words, the pixel adapter adjusts the local representation of the input, and the normalization module aligns the overall per-channel statistics (shifts and scales). Together, these convert the SAR input into a form that the frozen SAM backbone can process effectively, without changing its weights.

Second, we incorporate low-rank adaptation (LoRA) to efficiently adapt the frozen SAM image encoder by augmenting selected linear transformations with trainable low-rank updates. For an input feature vector x , each adapted linear layer calculates

$$y = Wx + b + \frac{\alpha}{r}BAx \quad (5)$$

where W and b are the original (frozen) weight and bias of the linear layers, $A \in \mathbb{R}^{r \times d_{in}}$ and $B \in \mathbb{R}^{d_{out} \times r}$ are trainable matrices, r is the LoRA rank and α is a scaling factor. This parameterization constrains the update to rank r while adding only $r(d_{in} + d_{out})$ trainable

parameters, compared to the $d_{in}d_{out}$ parameters of a full linear layer, enabling efficient fine-tuning. To preserve the pretrained behavior at initialization and stabilize optimization, the low-rank branch is initialized such that the induced update is initially zero. Thus, the layer starts as an exact copy of the pretrained mapping and learns only a gradual correction during training. LoRA modules are inserted into the attention and feed-forward projections of the image encoder (e.g., query-key-value projections, output projections, and MLP linear layers), enabling the internal representations of the encoder to accommodate a SAR oil spill domain shift. Following the original author's recommendations [46], the LoRA rank and scaling factors were kept identical. Both parameters were set to 8, which is a relatively small value. Increasing this value resulted in a slight decrease in performance, characterized by a rise in false positives.

Third, a Residual Feature Space Adapter module (RFSA) is implemented on the image embeddings generated by SAM's frozen encoder. It employs two 1×1 pointwise convolutions with a GELU activation between them and returns the output back to the original embeddings through a residual connection. Since the convolutions are pointwise, it retains the spatial grid while learning a versatile cross-channel mixing and reweighting of encoder features at each local area. This residual design ensures that the update is conservative but allows the model to re-align the embedding space towards oil spill cues (e.g., enhancing the separation between slick textures and nearby sea). Practically, this allows additional adaptation to be added after the frozen backbone with limited parameter costs and can elevate downstream mask quality.

To enable training at a lower computational cost, SAM's positional information is resized to a 512×512 grid, allowing the frozen image encoder to operate on a smaller square input. The absolute positional embeddings defined on a 2D token grid are resized by interpolation to the new grid, and the decomposed 1D relative positional embeddings used in attention are resized accordingly to remain consistent with the new token geometry. In parallel, the method runs SAM in a prompt-free setting by using a fixed implicit prompt that spans the full image, allowing automatic single-shot segmentation without interactive inputs. The final output is upsampled to the original resolution.

2.4. Training Implementation

All models were trained for 100 epochs using the Adam optimizer [66] with a learning rate of $1e-4$, a batch size of 4, and a cosine annealing schedule [67] with warm restarts. The initial five epochs served as warm-up epochs. For classification, the binary cross-entropy loss function was used, weighted by the fraction of negative samples in the dataset. For segmentation, the dice loss was used [68]. Prior to inference, each patch was bilinearly resized to 512×512 pixels. To improve robustness, we applied random rotations during training and included native-resolution patches of real oil spills, ensuring the models could not rely solely on the relative area of dark regions (oil spills are typically small). During validation and testing, test-time augmentations were applied. The classification and segmentation models were trained separately.

The classification model was trained on 70% of the patches without texture information and the segmentation model only on 70% of the patches that contained oil spills due to strong class imbalance. Texture information was incorporated into the segmentation model only. In both cases, the remaining 20% of the patches were used for validation and 10% for testing purposes (Table 1).

Table 1. Train–test split for the datasets. Only patches with oil spills were used for training the segmentation models.

	Oil Spills	Background
Train	462	1519
Validation	132	433
Test	66	216

2.5. Evaluation Details

For the classification problem, four metrics were used: recall, precision, specificity, and F1-score. For the pixel-wise segmentation problem, recall, precision, F1-score, and Intersection over Union (IoU) were used. They are defined as follows:

$$\text{Precision (P)} = \frac{TP}{TP + FP} \quad (6)$$

$$\text{Recall (R)} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{Specificity (S)} = \frac{TN}{TN + FP} \quad (8)$$

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (9)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (10)$$

where TP represents the true positives, FP the false positives, TN the true negatives and FN the false negatives.

3. Results

3.1. Classification Results

The first stage of the proposed pipeline performs patch-level classification to discriminate true oil spills from visually similar dark look-alike phenomena, reducing the amount of imagery that must be processed by the second-stage segmentor.

The summary of the quantitative comparison of the six backbones is shown in Table 2. In general, all models have high specificity (0.94–0.98), which means that most of the non-oil patches are correctly rejected, and this is a desirable characteristic since there is a large imbalance in the number of background/look-alikes compared to oil spills. However, the models vary in terms of how they balance the number of missed spills and false alarms.

Table 2. Quantitative classification performance on the test set for all compared models reported in terms of recall, precision, specificity and F1-score. Best values are highlighted in bold.

Model	Recall	Precision	Specificity	F1-Score
ResNet-18	0.90	0.93	0.98	0.92
ResNet-50	0.92	0.93	0.98	0.93
VGG-16	0.97	0.91	0.97	0.94
SWIN-T	0.98	0.86	0.94	0.93
ViT	0.98	0.91	0.97	0.95
ConvNeXt-T	0.98	0.94	0.98	0.96

Among the tested architectures, ConvNeXt-T offers the best overall trade-off, with 0.98 recall, 0.94 precision, 0.98 specificity, and the best F1-score of 0.96. These properties make it particularly well-suited for a first-stage filter: high recall ensures that actual spills

are rarely missed before segmentation, while its improved precision over other models with high recall rates prevents unnecessary subsequent segmentation of background patches. Transformer-based architectures are also very competitive. ViT has the same best recall of 0.98 and an F1-score of 0.95 with 0.91 precision and 0.97 specificity, suggesting a slightly higher false positive rate than ConvNeXt-T but with comparable sensitivity. Swin-T has the same high recall of 0.98 but with lower precision of 0.86 and specificity of 0.94, suggesting a higher rate of misattribution with look-alike structures, which would result in more patches being sent to stage two.

The CNN baselines remain strong but show a more conservative detection profile. ResNet-18 and ResNet-50 maintain high precision (0.93) and specificity (0.98), but recall is lower (0.90–0.92), meaning more missed oil spill patches compared to the best-performing models. VGG-16 improves sensitivity (0.97 recall) while keeping good specificity (0.97), but precision (0.91) and F1 (0.94) remain below ViT and ConvNeXt-T. In practice, these differences matter because false negatives at this stage cannot be recovered by the segmentor, whereas false positives primarily increase computational cost and may introduce occasional spurious masks.

Qualitative examples for ConvNeXt-T are shown in Figure 6. In the figure, the Sentinel-1 VV backscatter patches are visualized alongside the target labels T and prediction outputs P (Oil Spill or Background). Classification errors occur in hard cases under challenging conditions, where the contrast between oil spills and the sea is very low (last example of Figure 6d), or where look-alikes have very similar shape to those of slicks. Wind speed and incidence angle values for the false positive and false negative cases overlap with those of correctly classified patches, reinforcing the idea that the models base their decisions primarily on shape. For patches that contain both look-alikes and oil slicks (Figure 6c), ConvNeXt-T has no trouble identifying the slick. Taken together, the results justify selecting ConvNeXt-T as the screening model for the two-stage framework due to its best overall F1 and its favorable high-recall/high-specificity operating point.

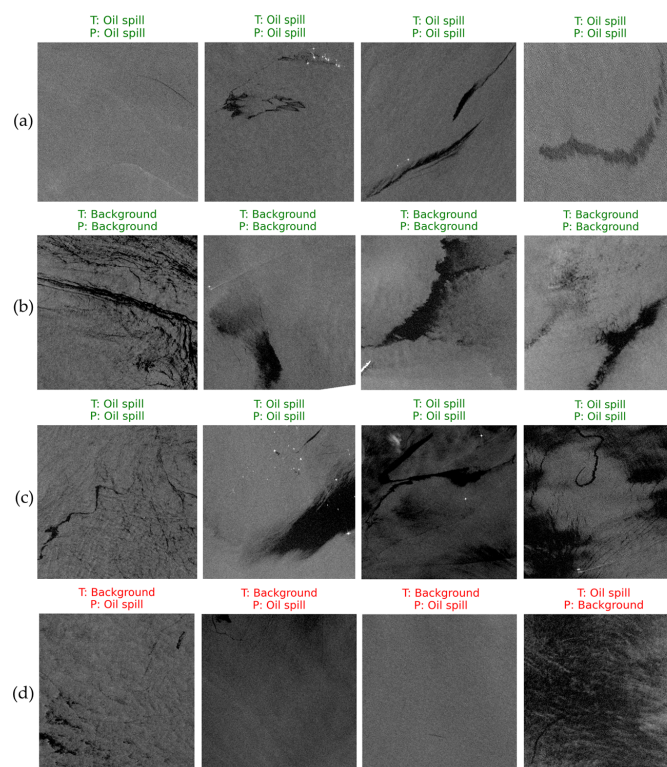


Figure 6. Example classification outputs of the ConvNeXt-T oil spill detector on randomly selected test-set patches. For every presented patch, T (Target) denotes the ground truth and P (Prediction)

denotes the model output. Green and red are used to denote correct and false predictions respectively. The rows illustrate: (a) successfully detected oil spills; (b) look-alikes correctly classified as background; (c) successfully detected oil spills in scenes containing look-alikes; and (d) examples of false-positive and false-negative cases.

3.2. Segmentation Results

The second stage of the proposed pipeline focuses on the pixel-wise delineation of oil slicks in Sentinel-1 SAR patches. All segmentation models were evaluated on the test set using the metrics defined earlier (recall, precision, F1-score, and IoU). Table 3 summarizes the quantitative comparison between representative encoder–decoder and modern segmentation baselines (U-Net, DeepLabV3, SegNeXt, OFCNet, CBDNet) and the proposed SAM-adapted segmentor (OSDA-SAM).

Table 3. Quantitative segmentation performance on the test set for all compared models and OSDA-SAM, reported in terms of recall, precision, F1-score, and Intersection over Union (IoU). Best values are highlighted in bold.

Model	Recall	Precision	F1-Score	IoU	PR-AUC
UNet	0.80	0.88	0.83	0.71	0.9003
DeepLabV3	0.80	0.84	0.82	0.70	0.8889
SegNext	0.77	0.86	0.82	0.69	0.8622
OFCNet	0.77	0.86	0.82	0.69	0.8863
CBDNet	0.80	0.87	0.83	0.72	0.9046
OSDA-SAM	0.86	0.86	0.86	0.75	0.9259

Overall, the results show that the SAM-based adaptation provides the most accurate and balanced masks. OSDA-SAM achieves the highest F1-score (0.86) and IoU (0.75), indicating improved overlap with ground truth and stronger boundary consistency. In addition, it attains the best recall (0.86), demonstrating that it is less prone to missing thin or low-contrast slick regions compared to the CNN baselines. The strongest competing method is CBDNet (IoU 0.72, F1 0.83), while U-Net reaches a similar F1-score (0.83) with slightly lower IoU (0.71). Among the remaining baselines, DeepLabV3 produces moderate overlap (IoU 0.70, F1 0.82), and SegNeXt/OFCNet show the lowest IoU (0.69) and lower recall (0.77), suggesting a tendency toward under-segmentation on subtle slick pixels.

U-Net attains the highest precision (0.88), implying fewer false positive pixels, but this comes with lower recall (0.80), consistent with conservative masks that may omit faint spill regions. In contrast, OSDA-SAM balances recall and precision at 0.86/0.86, yielding the best overall F1 and IoU.

In Figure 7a, the upper portion of the slick lies in a low-contrast region. OSDA-SAM is the only method that segments it correctly, with UNet delivering the second-best result and SegNeXt performing the worst. For highly elongated spills such as the example in Figure 7b, the proposed network successfully delineates the full length of the slick, unlike the competing approaches. Figure 7c shows a very thin spill against a bright background. Most models perform well, particularly CBDNet and UNet, but OSDA-SAM produces the thinnest mask. Finally, Figure 7d depicts a challenging scenario with a very dark sea, resulting in low contrast between the slick and the water. Here, OSDA-SAM captures most of both the small and large slicks, although all methods exhibit substantial misclassifications. In this case, CBDNet performs very well, while UNet fails to detect the small region entirely.

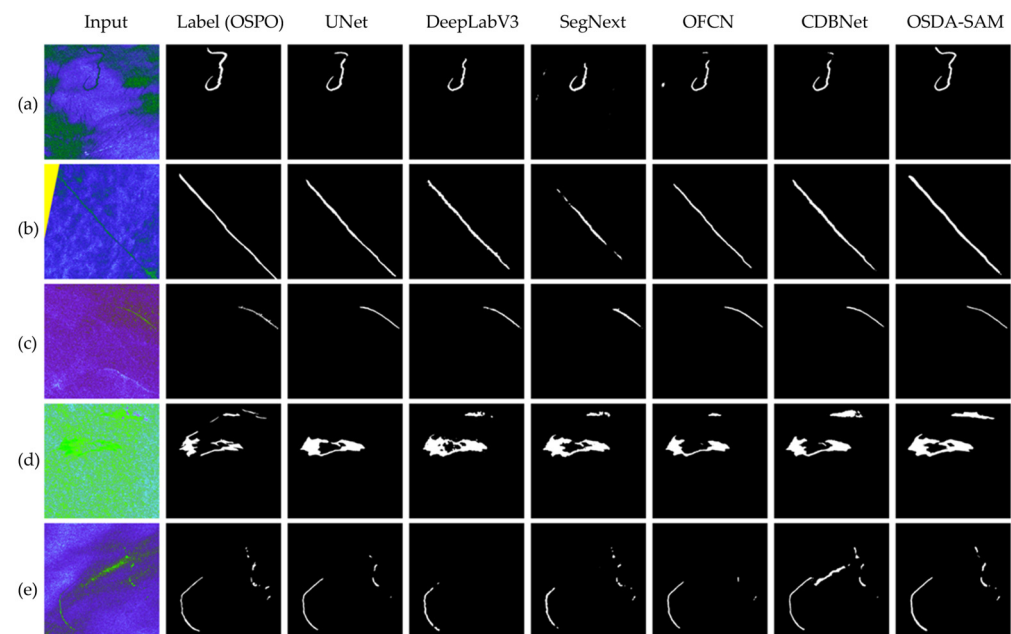


Figure 7. Qualitative comparison of oil spill segmentation on representative Sentinel-1 SAR patches (a–e). From left to right: input patch, ground-truth label mask, and predicted masks from U-Net, DeepLabV3, SegNeXt, OFCNet, CDBNet, and the proposed OSDA-SAM. White pixels denote oil slick; black pixels denote background.

For scenes that contain both look-alikes and oil spills, OSDA-SAM can segment the slick regions more effectively. For instance, Figure 7e presents an oil slick that, according to OSPO, is in the same area as an elongated look-alike feature. The results indicate that the proposed model can distinguish between the two and accurately delineate the oil spill region. Therefore, OSDA-SAM helps separate look-alike features from oil spills in many scenes where both are present, adding an additional layer of robustness.

Examining the precision–recall (PR) curves on the test dataset further reinforces the previous findings. Figure 8 shows the calculated curves. The PR curves provide an informative view of model performance under class imbalance, since they focus directly on the retrieval quality of the positive class. The proposed model achieves the highest PR-AUC (0.9259), indicating the best overall trade-off between precision and recall across decision thresholds. Its curve remains above the competing methods over most of the recall range, showing that it preserves higher precision while maintaining strong sensitivity. In practical terms, this suggests that the proposed model is more effective at identifying positive samples without incurring the same level of false positives as the alternative approaches.

Taken together, these findings indicate that adapting a foundation segmentation model to SAR imagery can improve oil slick delineation relative to conventional CNN segmentors, particularly by preserving more complete slick structures without sacrificing precision.

Table 4 reports the computational overhead of OSDA-SAM compared with the baseline segmentation models. The comparison was conducted on a system equipped with an NVIDIA RTX 4070 SUPER GPU (NVIDIA Corporation, Santa Clara, CA, USA), an Intel Core i5-13600 CPU (Intel Corporation, Santa Clara, CA, USA), and 32 GB of DDR5 RAM. Runtime was averaged over 10 runs using the same input resolution and batch size of 1 for all models. As expected, OSDA-SAM has a substantially larger total parameter count because it is built on the SAMasoundation model. However, only a small fraction of these parameters is trainable, since the SAM backbone remains frozen and adaptation is performed through lightweight modules. Specifically, OSDA-SAM has 639.94M total parameters but only 2.66M trainable parameters, which is lower than all baseline models.

This confirms the parameter-efficient nature of the proposed adaptation strategy. In terms of inference speed, OSDA-SAM is slower than the CNN-based baselines, with an inference time of 134.74 ms per batch and a throughput of 7.42 images/s. This additional cost is mainly due to the use of the large SAM image encoder. Although OSDA-SAM is slower, its inference time remains within a usable range for patch-based SAR analysis, especially since the first classification stage reduces the number of patches that need to be passed to the segmentation model.

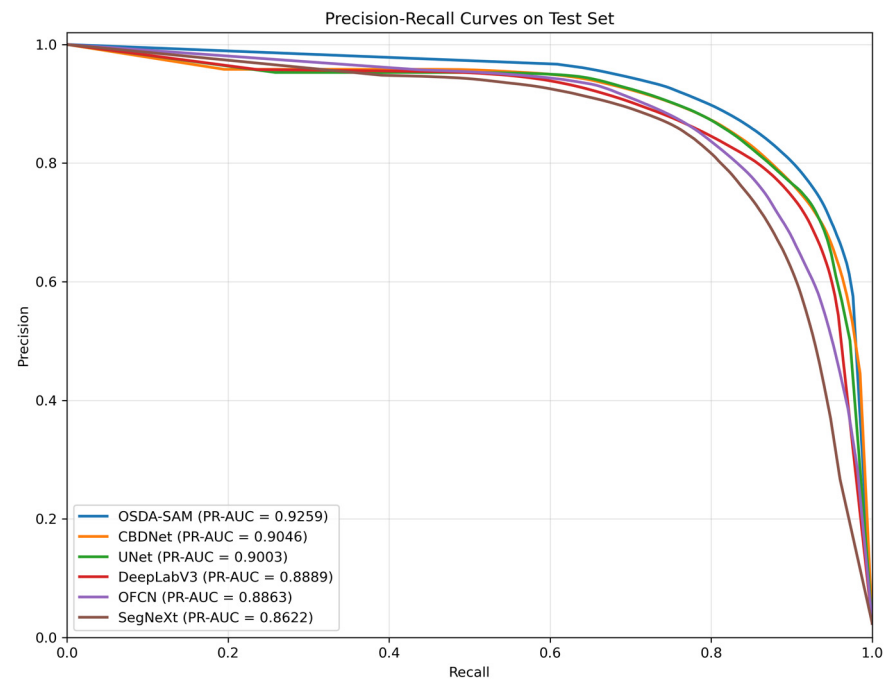


Figure 8. The calculated precision–recall (PR) curves on the test datasets for the segmentation models.

Table 4. Computational efficiency comparison between OSDA-SAM and baseline segmentation models.

Model	Total Params (M)	Trainable Params (M)	Inference Time/Batch (ms)	Throughput (Images/s)
UNet	24.44	24.436369	4.27	234.07
DeeplabV3	11.69	11.680185	5.64	177.28
SegNeXt	17.66	17.655745	17.52	57.08
OFCN	7.89	7.894303	7.70	129.94
CBDNet	31.47	31.471141	4.58	218.36
OSDA-SAM	639.94	2.657353	134.74	7.42

3.3. Effect of Wind on Segmentation Performance

To examine whether wind conditions influence segmentation performance, we compared patch-level F1-scores ($N = 49$ patches) from the OSDA-SAM test set with the mean wind speed of each patch derived from collocated ERA5 data (Figure 9). The resulting scatter does not support a simple linear relationship; instead it suggests a “wind-window” effect: low F1-scores occur across a wide range of wind speeds, whereas higher F1-scores—particularly those above the operational “good” threshold ($F1 = 0.8$)—are predominantly observed within 1–7 m/s, a range commonly reported to enhance oil spill visibility in SAR imagery. To quantify this observation, we performed a two-stage analysis: (i) we tested whether wind speed dispersion differs across F1 bins using the median-centered Brown–Forsythe (Levene) test and (ii) we tested whether the probability of achieving good performance ($F1 \geq 0.8$) is higher inside the 1–7 m/s wind window than outside it using Fisher’s exact test.

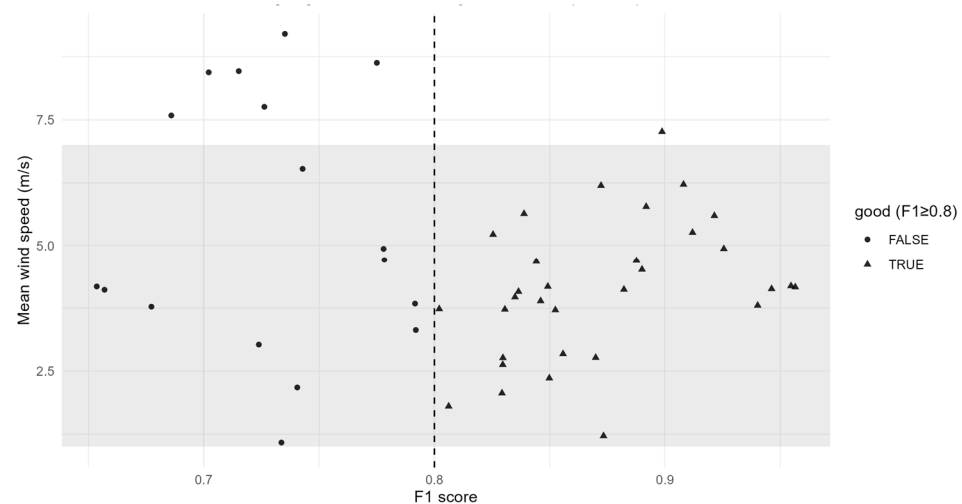


Figure 9. Patch-level F1-score versus mean wind speed. The shaded band highlights the 1–7 m/s wind window, and the vertical dashed line marks the good-performance threshold ($F1 = 0.8$).

The Brown–Forsythe test [69] is a robust variant of Levene’s test for equality of variances. It evaluates whether groups have the same variability, but it centers each group on the median rather than the mean, making it less sensitive to outliers and non-normal distributions. We applied the test after splitting the patches into three equal F1 tertiles (Figure 10). The test yielded $p = 0.003348$, indicating that wind speed variance differs significantly across tertiles. The ratio of the interquartile ranges (IQRs) between the lowest and highest tertile was 2.6, suggesting that wind speed exhibits $\sim 2.6\times$ larger typical dispersion in the low-F1 tertile than in the high-F1 tertile. However, a bootstrap 95% confidence interval for the IQR ratio was wide (0.729–7.58) and included values < 1 , implying that the exact effect size cannot be estimated precisely from these data. Overall, the results provide strong evidence of heteroscedasticity (different wind dispersion across F1 levels), while the magnitude of the dispersion difference remains uncertain due to the limited sample size.

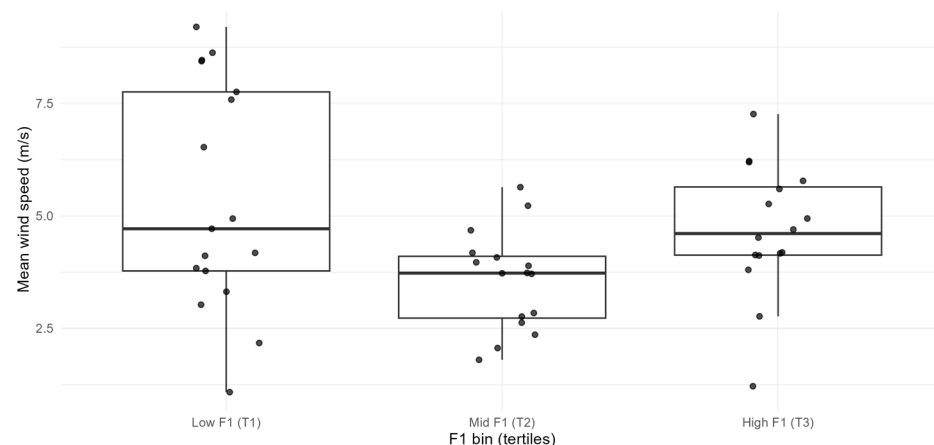


Figure 10. Wind dispersion across the three equal F1 tertiles. As F1 increases, wind values tend to cluster within a narrower range.

To test whether the probability of obtaining a “good” F1-score ($F1 \geq 0.8$) is higher when wind speed falls within the “good window” than when it falls outside it, Fisher’s exact test [70] was performed on a 2×2 contingency table (good vs. not-good performance; inside vs. outside the window). Fisher’s exact test is appropriate here because the number of patches outside the window is small ($n = 7$). The test returned $p = 0.004837$, providing strong

evidence of an association between wind-window membership and good performance. In our test set, 31/42 patches inside the window achieved $F1 \geq 0.8$ ($p = 0.738$), compared with 1/7 patches outside the window ($p = 0.143$). The estimated odds ratio was 15.84 (95% confidence interval: 1.65–798), indicating substantially higher odds of good performance inside the window; however, the confidence interval is wide because of the small number of samples outside the window and only one “good” case outside. For interpretability, the corresponding risk ratio was 5.167 and the risk difference was 0.595; i.e., being inside the window is associated with an approximately 60-percentage-point-higher probability of good performance. Note that in this dataset, “outside the window” corresponds almost entirely to wind speeds > 7 m/s (no samples < 1 m/s), so the result primarily reflects reduced performance under higher winds.

Overall, the 1–7 m/s wind window appears to be a favorable condition for achieving good patch-level segmentation performance ($F1 \geq 0.80$). Under moderate wind conditions that favor clearer slick contrast in SAR imagery, performance is more stable, whereas low F1-scores occur across a broader range of wind speeds. This suggests that wind speed influences detectability but does not, by itself, determine segmentation quality. As a robustness check, aggregating patches by acquisition (one observation per scene ID) yielded consistent results (Fisher $p = 0.004702$; Wilcoxon $p = 0.009338$).

3.4. Ablation on GLCM Texture Features

To measure the effect of the proposed texture-enhanced input representation, an ablation study was conducted in which the two GLCM statistics, homogeneity and variance, were introduced together with the preprocessed Sentinel-1 VV backscatter and compared against the VV-only input representation. The reasoning behind this is that oil slicks are expected to have higher homogeneity and lower variance than the surrounding sea surface, and this information can improve the boundary separability and reduce false positives.

3.4.1. Effect on Classification (Patch-Level)

Across the six classification backbones, adding texture channels produces no consistent improvement in recall, precision, specificity, or F1-score. In most cases, the changes are marginal (near zero), and, in several cases, performance slightly decreases (Table 5). This indicates that, for the classification stage, the texture descriptors do not add discriminative information beyond what the networks already learn from the intensity channel.

Table 5. Effect of the two added GLCM features on the performance of all the classification models on the test set. Negative numbers mean worse performance with texture features, and positive numbers mean better performance.

Model	ΔR	ΔP	ΔS	$\Delta F1$
ResNet-18	0.02	−0.01	−0.01	−0.01
ResNet-50	−0.02	0.00	0.00	0.00
VGG-16	−0.02	0.01	0.01	0.01
SWIN-T	−0.07	0.00	0.01	0.01
ViT	−0.03	−0.04	−0.02	−0.02
ConvNeXt-T	−0.03	0.00	0.00	0.00

This result is expected for two main reasons. First, classification is context-driven, and the context is global rather than local. The classifier makes a prediction for the whole patch based on a single label. In such a scenario, the classification decision may be dependent on mesoscale information, such as the geometric arrangement of dark patches, their form and connectivity, and the relationship between slick patterns and the sea background. While texture maps may provide useful local information, they do not directly represent the global

geometry that may be required to distinguish between actual slicks and their look-alike patterns. Hence, the inclusion of GLCM features is redundant, at least for architectures that are good at capturing multi-scale context information. Second, a practical factor is performance saturation. The strongest backbone already achieves very high scores with baseline input, leaving limited headroom for handcrafted channels to help.

3.4.2. Effect on Segmentation (Pixel-Level)

To measure the contribution of texture to segmentation performance, we remove the two channels of the GLCM and compare the VV-only inputs to the texture-aided scenario (Table 6). In the SAR domain, slicks are generally more locally homogeneous and smoother than the surrounding ocean, but intensity contrast may be low, and speckle noise may obscure or break up boundaries, leading to increased uncertainty in boundary regions. Homogeneity and variance offer a direct characterization of local structure that, together with VV, can alleviate such ambiguity in low-contrast boundary areas, potentially aiding completeness. This is most apparent in the case of OSDA-SAM, where the addition of texture results in the most pronounced gains ($\Delta\text{Recall} +0.04$, $\Delta\text{F1} +0.02$, $\Delta\text{IoU} +0.03$ with $\Delta\text{Precision} \approx 0.00$), suggesting that texture is mainly beneficial for segmenting noisy and difficult cases. In the case of fully trainable CNN segmentors, the impact is more modest and less robust since many such models can learn filters that resemble texture from intensity images. For example, U-Net makes only a slight improvement ($\Delta\text{Recall} +0.01$, $\Delta\text{Precision} +0.01$, $\Delta\text{IoU} +0.01$), while other models display near-zero or mixed changes.

Table 6. Effect of the two added GLCM features on the performance of all the segmentation models on the test set. Negative numbers mean worse performance with texture features and positive numbers mean better performance.

Model	ΔR	ΔP	ΔF1	ΔIoU
UNet	0.01	0.01	0.00	0.01
DeepLabV3	−0.02	0.00	−0.01	−0.01
SegNext	−0.03	0.01	0.00	−0.01
OFCNet	−0.02	0.04	0.02	0.02
CBDNet	−0.02	0.03	0.00	0.01
OSDA-SAM	0.04	0.00	0.02	0.03

Figure 11 presents the feature maps produced by the SAM encoder of the OSDA-SAM architecture for representative oil spill patch inputs. Since the feature maps have dimensions of $256 \times 32 \times 32$, they are averaged across channels and resampled to the input resolution for visualization. The VV backscatter of each patch is shown alongside the aggregated feature maps from the model trained using VV only, the model trained with texture features, and the corresponding ground-truth information. Figure 11a illustrates a case where the texture-based model exhibits stronger activation over small oil slicks. Figure 11c shows that the model assigns reduced relevance to ocean objects that are not associated with oil slicks. The remaining examples further demonstrate that incorporating texture information suppresses activation in dark areas unrelated to oil spills, making them less prominent in the resulting feature maps.

The benefit to OSDA-SAM is much stronger because of its domain adaptation architecture. The SAM backbone is kept mostly frozen, and the SAR-to-SAM gap is filled with lightweight adaptation (pixel-space mapping, adaptive normalization, and low-rank updates). In this setting, the texture channels are remarkably useful because they bring in a SAR-relevant second-order structure that the frozen backbone is not quite relearning from scratch, assisting in the separation of slick from sea under speckle and low contrast. When the adaptation modules align these channels into a SAM-compatible space, the model can

leverage the new structure to better include thin slick extensions and faint boundary pixels, leading to the recall-driven improvement seen in the ablation.

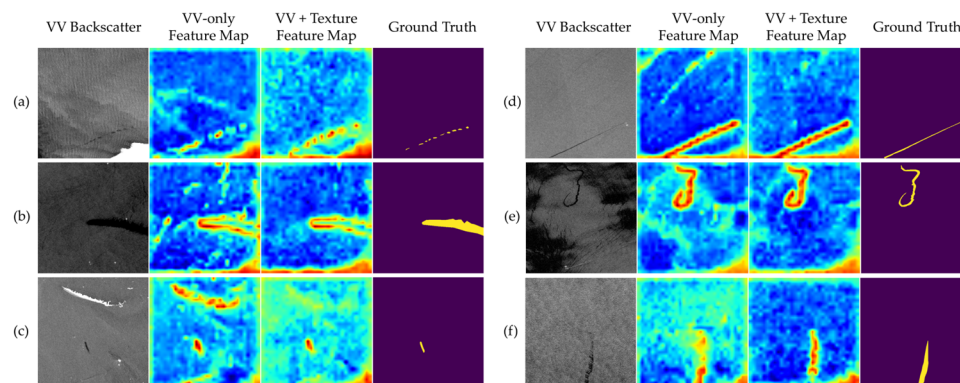


Figure 11. Aggregated feature map (from the SAM encoder) comparison between the model trained with VV only and the model trained with the addition of texture information for the OSDA-SAM architecture. The VV backscatter is visualized alongside the aggregated feature maps and ground-truth information.

3.5. Ablation on OSDA-SAM Components

To further investigate the significance of every adaptation module in the OSDA-SAM architecture, an ablation study was conducted by iteratively deactivating one component at the time and retraining the model on the train dataset. The results on the test dataset are presented in Table 7. The most significant module is LoRA, where its omission drops the F1-score by 5 percentage points to 0.81, followed by IDAB to 0.83 and RFSA to 0.85. This is expected as LoRA is adapting the weights of the model itself, while IDAB transforms the inputs and RFSA simply offers some extra guidance inside the feature representation after the SAM encoder.

Table 7. Ablation over the different OSDA-SAM modules. Every row is a different run, with the tick symbol signifying the modules that are activated for that run. The first row shows the performance of SAM while the last row shows the performance of OSDA-SAM.

LoRA	IDAB	RFSA	F1-Score
			0.37
	✓	✓	0.81
✓		✓	0.83
✓	✓		0.85
✓	✓	✓	0.86

4. Discussion

This work proposed a two-step deep learning approach for operational oil spill mapping in Sentinel-1 SAR images, targeting the challenge of distinguishing real oil slicks from look-alike regions and accurately segmenting slick boundaries. The first stage uses a ConvNeXt-T patch classifier to filter SAR images for potential patches containing oil slicks, and the second step carries out prompt-free, single-shot image segmentation using the proposed OSDA-SAM, a domain-adapted version of the Segment Anything Model. OSDA-SAM bridges the SAR-to-RGB domain gap with an Input Domain Adaptation Block that includes residual pixel-space correction and channel-wise normalization, along with low-rank updates in the frozen SAM image encoder and a lightweight residual adapter on top of the image embeddings.

Quantitative analysis on the test set indicates that the screening step reaches high sensitivity with a low false alarm rate. The best trade-off of the competing backbones is provided by ConvNeXt-T, with a recall of 0.98, precision of 0.94, specificity of 0.98, and F1-score of 0.96. This makes it a good choice for shrinking the search space in dense prediction tasks while rarely eliminating actual spills. For boundary delineation, OSDA-SAM provides the strongest and most balanced masks, with a recall of 0.86 and precision of 0.86, and the highest overlap with the ground truth (F1-score of 0.86 and IoU of 0.75). These results outperform the state-of-the-art CNN-based baselines, such as CBDNet (IoU of 0.72, F1-score of 0.83) and U-Net (IoU of 0.71, F1-score of 0.83). This indicates that after aligning the input distribution, the segmentation priors captured by a foundation model can be leveraged for more coherent and complete oil-slick masks.

The classification-stage results compare favorably with those reported in prior work. For instance, ref. [29] reports 99% accuracy, 84% precision, and an F1-score of 80% in a Sentinel-1 two-stage framework. Recall is comparable, while precision and F1 are higher in the present study. For segmentation, ref. [32] reports an F1-score of 0.892, which is slightly higher, likely due to the substantially larger training dataset used. Ref. [36] reports an F1-score of 78.48 on a large and varied dataset, like ours, which is lower than the value achieved here, further indicating strong robustness under diverse operating conditions. Studies [30,34,38] were trained and tested on smaller datasets with less scene variation and therefore we cannot draw direct comparisons. In [31], multiclass oil spill segmentation yields a markedly lower F1-score of 53.79, consistent with the increased difficulty of multiclass discrimination. Finally, study [35] reports very high performance, but a direct comparison is not appropriate due to differences in sensor resolution and evaluation on a small set of large scenes.

An ablation study on the texture-enhanced input representation further helps clarify the importance of domain-specific information. Adding Gray-Level Co-occurrence Matrix statistics (homogeneity and variance) to VV backscatter does not provide any consistent improvement in patch-level classification, with marginal and sometimes negative changes across backbones. This indicates that the screening decision is dominated by mesoscale morphology and context information already captured from intensity images. On the other hand, the same texture augmentation does help pixel-wise segmentation in most models, and the improvement is most significant for OSDA-SAM, with a relative improvement of 0.04, 0.02, and 0.03 in recall, F1-score, and IoU, respectively, compared to VV-only inputs, while precision is left essentially unchanged. This is consistent with the interpretation that homogeneity and variance capture SAR-relevant second-order structural information useful for delineating low-contrast boundaries under speckle noise, especially when the backbone is mostly frozen and only lightly adapted.

At the dataset level, the samples span the full Sentinel-1 IW incidence angle range as well as the expected wind speed range, indicating that the reported results are not limited to a narrow set of viewing geometries or environmental conditions. For the classification stage, the qualitative error analysis shows that most misclassifications occur in inherently difficult cases (e.g., faint slicks or highly linear look-alikes), while the associated wind speed and incidence angle values remain within typical ranges. This suggests that classification performance is driven more by scene appearance and shape-related cues than by acquisition or environmental parameters alone. For the segmentation stage, the analysis suggests that patch level segmentation performance is generally more consistent under moderate wind conditions, where oil slick contrast in SAR is typically clearer, while lower F1-scores appear across a broader range of wind speeds. In turn, this indicates that wind speed affects detectability but does not, by itself, determine segmentation quality.

Several limitations should be considered when interpreting these results. Although manual inspection and correction were performed, the reference labels originate from operational mapping products that combine visual interpretation with automated tools and may therefore contain boundary uncertainty or occasional omissions. Also, oil spills are undetectable in the presence of dense look-alike areas (last example of row d in Figure 7). Furthermore, inaccuracies in land masking and a dense presence of fish farms can heavily affect image statistics, influencing the binning process during the GLCM estimation and ultimately leading to loss in sea texture details. Finally, the two-stage design introduces an unrecoverable error path: spills missed by the classifier are not passed to the segmentation stage. The missed oil spills, which account for approximately 2% of the cases, are typically characterized by low contrast, diffuse or fragmented boundaries, and limited separability from the surrounding sea surface. These characteristics suggest that they are mainly thin or low-volume spills, as well as weathered, aged, or dispersed oil slicks. In some cases, the spills are also mixed with look-alike areas, primarily under low-wind conditions, placing them near the SAR detectability limits. Since the number of missed cases is small and they generally correspond to weak or ambiguous slick signatures, their impact on the overall pollution area assessment is expected to be limited.

These limitations point to several directions for future work. Possible avenues for limiting the effect of misclassifications during the first stage include joint/curriculum training of both stages and active learning to focus annotation efforts on hard cases, as well as test-time augmentations. Even though noise is currently handled on the data preprocessing level (Lee filter, thermal noise) and Sentinel 1 is considered a low-noise SAR system, greater integration of noise robustness in the model design can also be investigated in the future. Furthermore, the fusion of other SAR features such as multiple polarizations and auxiliary environmental information, such as wind speed and sea surface temperature, could be examined. Finally, investigating additional GLCM features like entropy as input could lead to further improvements in the segmentation stage.

5. Conclusions

This work introduced a two-stage deep learning framework for oil spill detection and delineation in Sentinel-1 SAR imagery, combining a ConvNeXt-T patch classifier with OSDA-SAM, a domain-adapted variant of the Segment Anything Model. The proposed framework achieved strong classification and segmentation performance and outperformed conventional CNN-based segmentation baselines, demonstrating that parameter-efficient adaptation of large segmentation models can be effective in the SAR domain.

The results also showed that SAR-specific texture information is more beneficial for boundary delineation than for patch-level screening, highlighting the importance of lightweight domain adaptation and task-specific feature design. Overall, the proposed pipeline provides an accurate and scalable solution for oil spill mapping while requiring only limited task-specific labeled data. At the same time, the identified limitations indicate clear opportunities for further improvement in future work.

Author Contributions: Conceptualization, G.G.; methodology, G.G. and V.A.; software G.G.; validation, G.G.; data curation, G.G.; writing—original draft preparation G.G.; writing—review and editing, M.K., V.K. and V.A.; visualization, G.G. and V.A.; supervision, M.K. and V.K.; project administration V.K., K.K., D.B., A.S.O. and A.M. All authors have read and agreed to the published version of the manuscript.

Funding: The project is part of the National Recovery and Resilience Plan ‘Greece 2.0’, which is funded by the Recovery and Resilience Facility (RRF), core programme of the European Union-NextGenerationEU under contract No. 4000145363/24/NL/FFi.

Data Availability Statement: Sentinel 1 data will be made available after reasonable request.

Acknowledgments: The project is being carried out under an ESA Contract in the frame of the Greek National Satellite Space Project. The Project Small-Satellites (Measure ID 16855) is implemented by the Hellenic Ministry of Digital Governance with European Space Agency (ESA) Assistance in Management and Implementation. The project is part of the National Recovery and Resilience Plan ‘Greece 2.0’, which is funded by the Recovery and Resilience Facility (RRF), core programme of the European Union-NextGenerationEU. Views expressed herein can in no way be taken to reflect the official opinion of the European Union/European Commission/European Space Agency/Greek Ministry of Digital Governance. Views and opinions expressed are those of the author(s) only and the European Union/European Commission/European Space Agency/Greek Ministry of Digital Governance cannot be held responsible for any use which may be made of the information contained therein.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ESA	European Space Agency
HSC	Hellenic Space Centre
SAR	Synthetic Aperture Radar
SAM	Segment Anything Model
CNN	Convolutional Neural Network
FCN	Fully Convolutional Network
GLCM	Gray-Level Co-occurrence Matrix
LoRA	Low Rank Adaptation
OSDA-SAM	Oil Spill Domain Adapted Segment Anything Model
OSPO	Office of Satellite and Product Operations
NOAA	National Oceanic and Atmospheric Administration
SAB	Satellite Analysis Branch
MPSR	Marine Pollution Surveillance Report
IDAB	Input Domain Adaptation Block
IoU	Intersection over Union
IQR	Interquartile Range
RRF	Recovery and Resilience Facility

References

1. Kingston, P.F. Long-Term Environmental Impact of Oil Spills. *Spill Sci. Technol. Bull.* **2002**, *7*, 53–61. [[CrossRef](#)]
2. Saadoun, I.M.K. Impact of Oil Spills on Marine Life. In *Emerging Pollutants in the Environment—Current and Further Implications*; Larramendy, M.L., Soloneski, S., Eds.; InTech: London, UK, 2015. [[CrossRef](#)]
3. Jernelöv, A. The Threats from Oil Spills: Now, Then, and in the Future. *AMBIO* **2010**, *39*, 353–366. [[CrossRef](#)]
4. Chen, J.; Zhang, W.; Wan, Z.; Li, S.; Huang, T.; Fei, Y. Oil Spills from Global Tankers: Status Review and Future Governance. *J. Clean. Prod.* **2019**, *227*, 20–32. [[CrossRef](#)]
5. Zhang, B.; Matchinski, E.J.; Chen, B.; Ye, X.; Jing, L.; Lee, K. Marine Oil Spills—Oil Pollution, Sources and Effects. In *World Seas: An Environmental Evaluation*; Elsevier: Amsterdam, The Netherlands, 2019. [[CrossRef](#)]
6. Dong, Y.; Liu, Y.; Hu, C.; MacDonald, I.R.; Lu, Y. Chronic Oiling in Global Oceans. *Science* **2022**, *376*, 1300–1304. [[CrossRef](#)]
7. Topouzelis, K.N. Oil Spill Detection by SAR Images: Dark Formation Detection, Feature Extraction and Classification Algorithms. *Sensors* **2008**, *8*, 6642–6659. [[CrossRef](#)] [[PubMed](#)]
8. Solberg, A.; Storvik, G.; Solberg, R.; Volden, E. Automatic Detection of Oil Spills in ERS SAR Images. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1916–1924. [[CrossRef](#)]
9. Brekke, C.; Solberg, A.H.S. Oil Spill Detection by Satellite Remote Sensing. *Remote Sens. Environ.* **2005**, *95*, 1–13. [[CrossRef](#)]
10. Alpers, W.; Holt, B.; Zeng, K. Oil Spill Detection by Imaging Radars: Challenges and Pitfalls. *Remote Sens. Environ.* **2017**, *201*, 133–147. [[CrossRef](#)]

11. Fingas, M.; Brown, C.E. A Review of Oil Spill Remote Sensing. *Sensors* **2017**, *18*, 91. [\[CrossRef\]](#)
12. Topouzelis, K.; Karathanassi, V.; Pavlakis, P.; Rokos, D. Detection and Discrimination between Oil Spills and Look-Alike Phenomena through Neural Networks. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 264–270. [\[CrossRef\]](#)
13. de Araújo Carvalho, G.; Minnett, P.J.; Ebecken, N.F.F.; Landau, L. Oil Spills or Look-Alikes? Classification Rank of Surface Ocean Slick Signatures in Satellite Data. *Remote Sens.* **2021**, *13*, 3466. [\[CrossRef\]](#)
14. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*; Springer: Cham, Switzerland, 2015. [\[CrossRef\]](#)
15. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016. [\[CrossRef\]](#)
16. Chen, L.-C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv* **2017**, arXiv:1706.05587. [\[CrossRef\]](#)
17. Cheng, M.-M.; Guo, M.-H.; Hou, Q.; Hu, S.-M.; Liu, Z.; Lu, C.-Z. SegNeXt: Rethinking Convolutional Attention Design for Semantic Segmentation. *arXiv* **2022**, arXiv:2209.08575.
18. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 10–17 October 2021. [\[CrossRef\]](#)
19. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image Is Worth 16 × 16 Words: Transformers for Image Recognition at Scale. *arXiv* **2020**, arXiv:2010.11929. [\[CrossRef\]](#)
20. Liu, Z.; Mao, H.; Wu, C.-Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A ConvNet for the 2020s. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022. [\[CrossRef\]](#)
21. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.-Y.; et al. Segment Anything. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 1–6 October 2023.
22. Osco, L.P.; Wu, Q.; de Lemos, E.L.; Gonçalves, W.N.; Ramos, A.P.M.; Li, J.; Marcato, J. The Segment Anything Model (SAM) for Remote Sensing Applications: From Zero to One Shot. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *124*, 103540. [\[CrossRef\]](#)
23. Du, B.; Liu, L.; Tao, D.; Wang, D.; Xu, M.; Zhang, J.; Zhang, L. SAMRS: Scaling-up Remote Sensing Segmentation Dataset with Segment Anything Model. *arXiv* **2023**, arXiv:2305.02034. [\[CrossRef\]](#)
24. Zhang, E.; Liu, J.; Cao, A.; Sun, Z.; Zhang, H.; Wang, H.; Sun, L.; Song, M. RS-SAM: Integrating Multi-scale Information for Enhanced Remote Sensing Image Segmentation. In *Asian Conference on Computer Vision (ACCV)*; Springer: Singapore, 2024. [\[CrossRef\]](#)
25. Luo, M.; Zhang, T.; Wei, S.; Ji, S. SAM-RSIS: Progressively Adapting SAM With Box Prompting to Remote Sensing Image Instance Segmentation. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 4413814. [\[CrossRef\]](#)
26. Ma, X.; Wu, Q.; Zhao, X.; Zhang, X.; Pun, M.-O.; Huang, B. SAM-Assisted Remote Sensing Imagery Semantic Segmentation With Object and Boundary Constraints. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5636916. [\[CrossRef\]](#)
27. Lu, X.; Weng, Q. Multi-LoRA Fine-Tuned Segment Anything Model for Urban Man-Made Object Extraction. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5637519. [\[CrossRef\]](#)
28. Zheng, Z.; Zhong, Y.; Zhang, L.; Ermon, S. Segment Any Change (AnyChange): Training-Free Adaptation of SAM for Bitemporal Remote Sensing Change Detection. *arXiv* **2024**, arXiv:2402.01188. [\[CrossRef\]](#)
29. Shaban, M.; Salim, R.; Abu Khalifeh, H.; Khelifi, A.; Shalaby, A.; El-Mashad, S.; Mahmoud, A.; Ghazal, M.; El-Baz, A. A Deep-Learning Framework for the Detection of Oil Spills from SAR Data. *Sensors* **2021**, *21*, 2351. [\[CrossRef\]](#)
30. Zeng, K.; Wang, Y. A Deep Convolutional Neural Network for Oil Spill Detection from Spaceborne SAR Images. *Remote Sens.* **2020**, *12*, 1015. [\[CrossRef\]](#)
31. Krestenitis, M.; Orfanidis, G.; Ioannidis, K.; Avgerinakis, K.; Vrochidis, S.; Kompatsiaris, I. Oil Spill Identification from Satellite Images Using Deep Neural Networks. *Remote Sens.* **2019**, *11*, 1762. [\[CrossRef\]](#)
32. Bianchi, F.M.; Espeseth, M.M.; Borch, N. Large-Scale Detection and Categorization of Oil Spills from SAR Images with Deep Learning. *Remote Sens.* **2020**, *12*, 2260. [\[CrossRef\]](#)
33. Zhu, Q.; Zhang, Y.; Li, Z.; Yan, X.; Guan, Q.; Zhong, Y.; Zhang, L.; Li, D. Oil Spill Contextual and Boundary-Supervised Detection Network Based on Marine SAR Images (CBD-Net). *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5213910. [\[CrossRef\]](#)
34. Mahmoud, A.S.; Mohamed, S.A.; El-Khoriby, R.A.; AbdelSalam, H.M.; El-Khodary, I.A. Oil Spill Identification Based on Dual Attention UNet Model Using Synthetic Aperture Radar Images. *J. Indian Soc. Remote Sens.* **2023**, *51*, 121–133. [\[CrossRef\]](#)
35. Hasimoto-Beltran, R.; Canul-Ku, M.; Díaz Méndez, G.M.; Ocampo-Torres, F.J.; Esquivel-Trava, B. Ocean Oil Spill Detection from SAR Images Based on Multi-Channel Deep Learning Semantic Segmentation. *Mar. Pollut. Bull.* **2023**, *188*, 114651. [\[CrossRef\]](#)
36. Dehghani-Dehcheshmeh, S.; Akhoondzadeh, M.; Homayouni, S. Oil Spills Detection from SAR Earth Observations Based on a Hybrid CNN–Transformer Network. *Mar. Pollut. Bull.* **2023**, *190*, 114834. [\[CrossRef\]](#) [\[PubMed\]](#)

37. Fan, J.; Liu, C. Multi-Task GANs for Oil Spill Classification and Semantic Segmentation Based on SAR Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 2532–2546. [\[CrossRef\]](#)
38. Kikaki, K.; Kakogeorgiou, I.; Hoteit, I.; Karantzas, K. Detecting Marine Pollutants and Sea Surface Features with Deep Learning in Sentinel-2 Imagery. *ISPRS J. Photogramm. Remote Sens.* **2024**, *210*, 39–54. [\[CrossRef\]](#)
39. Li, C.; Kim, D.-J.; Park, S.; Kim, J.; Song, J. A Self-Evolving Deep Learning Algorithm for Automatic Oil Spill Detection in Sentinel-1 SAR Images. *Remote Sens. Environ.* **2023**, *299*, 113872. [\[CrossRef\]](#)
40. Yang, Y.-J.; Singha, S.; Mayerle, R. A Deep Learning Based Oil Spill Detector Using Sentinel-1 SAR Imagery. *Int. J. Remote Sens.* **2022**, *43*, 4287–4314. [\[CrossRef\]](#)
41. Huang, X.; Zhang, B.; Perrie, W.; Lu, Y.; Wang, C. A Novel Deep Learning Method for Marine Oil Spill Detection from Satellite Synthetic Aperture Radar Imagery. *Mar. Pollut. Bull.* **2022**, *179*, 113666. [\[CrossRef\]](#)
42. Ghara, F.M.; Shokouhi, S.B.; Akbarizadeh, G. A New Technique for Segmentation of the Oil Spills From Synthetic-Aperture Radar Images Using Convolutional Neural Network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 8834–8844. [\[CrossRef\]](#)
43. Yu, X.; Zhang, H.; Luo, C.; Qi, H.; Ren, P. Oil Spill Segmentation via Adversarial f-Divergence Learning. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4973–4988. [\[CrossRef\]](#)
44. Li, C.; Wang, M.; Yang, X.; Chu, D. DS-UNet: Dual-Stream U-Net for Oil Spill Detection of SAR Image. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 4014905. [\[CrossRef\]](#)
45. Wu, W.; Wong, M.S.; Yu, X.; Shi, G.; Kwok, C.Y.T.; Zou, K. Compositional Oil Spill Detection Based on Object Detector and Adapted Segment Anything Model From SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 4007505. [\[CrossRef\]](#)
46. Guo, S.; Li, Y.; Cheng, L.; Yang, Q. CoRemoteSAM-Oil: A Method for Marine Oil Spill Pollution Detection through Cross-Domain Knowledge Transfer and Integration. *Measurement* **2026**, *279*, 121659. [\[CrossRef\]](#)
47. Guo, S.; Li, Y.; Zhang, Z.; Wu, P.; Wang, Z. HSRD-Net: A Hybrid RemoteSAM-ResNet with Dynamic Weight Assignment for SAR Oil Spill Detection. *Expert Syst. Appl.* **2026**, *322*, 132384. [\[CrossRef\]](#)
48. Guo, S.; Li, Y.; Shang, J.; Wang, Z.; Yuan, J. DADS-SAM: DALoRA and DSCA Enhanced Fine-Tuning for Port Oil Spill Detection Using UAV Imagery. *IEEE Trans. Geosci. Remote Sens.* **2026**, *64*, 4200620. [\[CrossRef\]](#)
49. Chen, S.; Yin, M.; Ren, P.; Luo, C.; Fu, Z. OilSAM2: Memory-Augmented SAM2 for Scalable SAR Oil Spill Detection. In *ICASSP 2026—2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*; IEEE: New York, NY, USA, 2026; pp. 11382–11385. [\[CrossRef\]](#)
50. Li, H.-L.; Chen, S.-W. General Polarimetric Correlation Pattern: A Visualization and Characterization Tool for Target Joint-Domain Scattering Mechanisms Investigation. *IEEE Trans. Geosci. Remote Sens.* **2026**, *64*, 5200417. [\[CrossRef\]](#)
51. Ma, X.; Xu, J.; Wu, P.; Kong, P. Oil Spill Detection Based on Deep Convolutional Neural Networks Using Polarimetric Scattering Information From Sentinel-1 SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4204713. [\[CrossRef\]](#)
52. Wang, D.; Wan, J.; Liu, S.; Chen, Y.; Yasir, M.; Xu, M.; Ren, P. BO-DRNet: An Improved Deep Learning Model for Oil Spill Detection by Polarimetric Features from SAR Images. *Remote Sens.* **2022**, *14*, 264. [\[CrossRef\]](#)
53. Wang, D.; Song, S.; Yang, J.; Xu, M.; Song, D.; Guo, J.; Wan, J.; Liu, S. Marine Oil Spill Detection Using Improved Polarimetric Feature Based on Polarization SAR Image. *Int. J. Remote Sens.* **2024**, *45*, 911–929. [\[CrossRef\]](#)
54. Liao, L.; Zhao, Q.; Song, W. Monitoring of Oil Spill Risk in Coastal Areas Based on Polarimetric SAR Satellite Images and Deep Learning Theory. *Sustainability* **2023**, *15*, 14504. [\[CrossRef\]](#)
55. Song, W.; Ma, X.; Song, W. Automatic Detection of Marine Oil Spills from Polarimetric SAR Images Using Deep Convolutional Neural Network Model. *Ecol. Indic.* **2024**, *169*, 112934. [\[CrossRef\]](#)
56. Song, D.; Huang, Q.; Gao, H.; Wang, B.; Zhang, J.; Chen, W. Adaptive Oil Spill Detection Network for Scene-Based PolSAR Data Using Dynamic Convolution and Boundary Constraints. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *130*, 103914. [\[CrossRef\]](#)
57. Xiang, D.; Lu, Y.; Guan, D.; Li, G.; Cheng, J.; Li, B. Oil Spill Detection in PolSAR Imagery Using Composite Scattering Power Entropy and Multiscale Hybrid Feature Fusion Network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 13388–13407. [\[CrossRef\]](#)
58. Wang, Z.; Zhang, H.; Guo, Z.; Fan, Y.; Xu, L.; Li, X.; Wang, C. CDANet: A Context-Detail Aware Network for Marine Oil Spill Detection in SAR Imagery. *J. Hazard. Mater.* **2026**, *501*, 140683. [\[CrossRef\]](#)
59. Cui, Y.; Ma, X.; Chen, T.; Shen, H.; Wang, Z.; Hu, Z. Oil Spill Detection on Sea Surface with Dual-Polarimetric SAR Imagery Integrating Polarization Features and Oil-Seawater Boundary Information. *J. Hazard. Mater.* **2026**, *511*, 142159. [\[CrossRef\]](#)
60. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621. [\[CrossRef\]](#)
61. Hu, E.J.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; Chen, W. LoRA: Low-Rank Adaptation of Large Language Models. *arXiv* **2017**, arXiv:2106.09685.
62. National Oceanic and Atmospheric Administration (NOAA), Office of Satellite and Product Operations (OSPO). OSPO | Office of Satellite and Product Operations. Available online: <https://www.ospo.noaa.gov/> (accessed on 17 February 2026).

63. Copernicus Data Space Ecosystem. “Copernicus Browser.” Copernicus Data Space Ecosystem (Ecosystem Services Registry). Available online: <https://dataspace.copernicus.eu/ecosystem/services/copernicus-browser> (accessed on 17 February 2026).
64. Copernicus Climate Change Service (C3S). *ERA5 Hourly Data on Single Levels from 1940 to Present*; Copernicus Climate Data Store (CDS): Berkshire, UK, 2026. [[CrossRef](#)]
65. Hendrycks, D.; Gimpel, K. Gaussian Error Linear Units (GELUs). *arXiv* **2016**, arXiv:1606.08415.
66. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
67. Loshchilov, I.; Hutter, F. SGDR: Stochastic Gradient Descent with Warm Restarts. *arXiv* **2017**, arXiv:1608.03983. [[CrossRef](#)]
68. Milletari, F.; Navab, N.; Ahmadi, S.-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *arXiv* **2016**, arXiv:1606.04797. [[CrossRef](#)]
69. Brown, M.B.; Forsythe, A.B. Robust Tests for the Equality of Variances. *J. Am. Stat. Assoc.* **1974**, *69*, 364–367. [[CrossRef](#)]
70. Fisher, R.A. On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P. *J. R. Stat. Soc.* **1922**, *85*, 87–94. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.