BEYOND SEGMENTATION: AN OIL SPILL CHANGE DETECTION FRAMEWORK USING SYNTHETIC SAR IMAGERY

A.Chenyang Lai, B. Shuaiyu Chen, C.Zeyu Fu

University of Exeter A-C
Department of Computer Science A-C
Devon,UK A-C

ABSTRACT

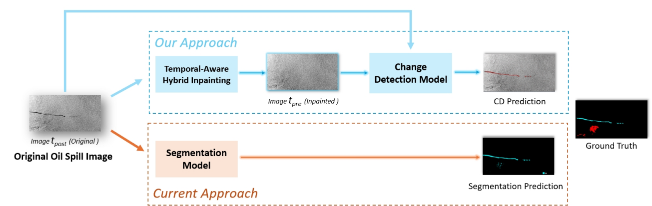
Single-image oil-spill detection in Synthetic Aperture Radar (SAR) imagery is prone to false alarms from look-alike dark formations such as low wind, biogenic films, or algal blooms, and suffers from scarce, imbalanced data. We reformulate the task as **oil-spill change detection (OSCD)**, which highlights only new slick pixels after a spill, suppressing persistent distractors. Since aligned pre-spill SAR images are rarely available, we propose **Temporal-Aware Hybrid Inpainting (TAHI)** to synthesize oil-free counterparts for polluted scenes. TAHI combines: (1) *dynamic foreground simulation*, perturbing vessel positions to mimic plausible temporal changes; (2) *high-fidelity hybrid inpainting*, which removes oil via PatchMatch and restores structure with a partial-convolution U-Net; and (3) *temporal realism enhancement*, aligning radiometry and injecting Gamma-distributed speckle with low-frequency sea-state drift, so that oil becomes the sole systematic difference. Using TAHI we construct the **OSCD dataset**, providing synthetic clean/real polluted SAR pairs and pixel-level change annotations. Image-quality assessment shows realistic inpainting, while benchmarks confirm that change-detection models trained on OSCD significantly surpass differencing and state-of-the-art segmentation. To our knowledge, this is the first framework dedicated to CD-based oil-spill monitoring, establishing a reliable and temporally aware solution for maritime surveillance.

***Index Terms*—** Oil spill change detection, Synthetic Aperture Radar, foundation models, Remote Sensing

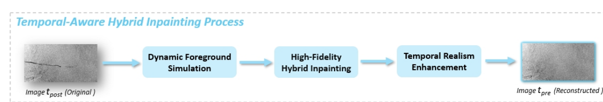
1. INTRODUCTION

Marine oil spills threaten ecosystems and economies, demanding rapid and reliable detection. Synthetic Aperture Radar (SAR), with its all-weather and day–night capability, is widely used for maritime monitoring [1, 2]. Yet traditional approaches—thresholding, texture descriptors, or statistical classifiers—struggle with robustness and often confuse dark ocean features with real slicks [3, 4].

Deep learning–based segmentation methods [5, 6, 7, 8, 9, 10, 11] improve generalization but still analyse each SAR scene independently. Without modelling temporal evolution, they remain vulnerable to look-alikes such as low wind or biogenic films. Moreover, limited annotated data and extreme foreground–background imbalance further hinder robust training [12].



(a) Segmentation vs. OSCD (ours).



(b) TAHI pipeline.

Fig. 1. (a) Our OSCD formulation suppresses look-alikes. (b) TAHI synthesizes oil-free pre-spill SAR from post-spill imagery.

To address these challenges, we introduce **oil spill change detection (OSCD)** (Fig. 1a), a bi-temporal formulation that isolates newly formed slicks. The main difficulty is the scarcity of co-registered pre-spill SAR images, since ocean dynamics and acquisition gaps rarely provide suitable pairs [13, 14].

We therefore develop **Temporal-Aware Hybrid Inpainting (TAHI)** (Fig. 1b) to synthesize realistic pre-spill counterparts directly from annotated post-spill imagery. TAHI combines three steps: Dynamic Foreground Simulation (DFS) perturbs vessel positions to emulate plausible temporal changes; High-Fidelity Hybrid Inpainting (HFHI) removes slicks with PatchMatch and restores texture via a partial-convolution U-Net; and Temporal Realism Enhancement (TRE) aligns radiometry while injecting Gamma speckle and low-frequency sea-state drift. The resulting pairs enable OSCD training and evaluation without real pre-event acquisitions.

Our contributions are threefold: 1) We formalize **OSCD** as a new task and release the first open bi-temporal dataset for oil-spill CD. 2) We propose **TAHI**, a hybrid inpainting framework producing temporally realistic pre-spill SAR imagery. 3) We conduct extensive quality and benchmark studies, showing that OSCD with TAHI outperforms segmentation baselines and supports reliable maritime monitoring.

2. RELATED WORK

2.1. SAR-based Oil Spill Detection

Early approaches used thresholds, texture measures, or statistical classifiers to identify dark areas in SAR imagery [1, 3], but these lacked robustness and required expert validation. Deep learning has since advanced segmentation: from early CNNs [15] to FCNs, attention U-Nets, and transformer-based models [8, 4, 6]. Recent designs such as YOLO-based detectors [11], DeepLabv3+ with speckle modelling [9], and state-space models like OSDMamba [7] further improve accuracy. However, all rely on single images, leaving them prone to false alarms from persistent look-alikes and unable to capture temporal dynamics.

2.2. Change Detection in Remote Sensing

Change detection (CD) compares multi-temporal images to highlight new or disappeared structures [16, 17]. Modern CD employs CNNs, transformers, and pseudo-labelling [18, 19], achieving strong results in urban or land-use monitoring. Yet CD has rarely been applied to oil spills, largely due to the lack of co-registered pre-event SAR imagery, which is hard to obtain given revisit gaps and ocean variability [14, 13].

2.3. Synthetic Data for Change Detection

Synthetic generation of bi-temporal SAR pairs offers a solution to scarcity and imbalance. Exemplar-based inpainting [20] and partial-conv networks [21] create plausible pre-event scenes, while recent datasets such as anomaly-aware CD [22] and PRISM [23] improve realism. However, synthetic frameworks have not been tailored for oil-spill CD, leaving the temporal anomaly signature of spills underexplored.

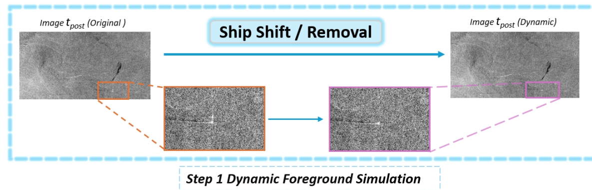
3. METHODOLOGY

3.1. Problem Formulation

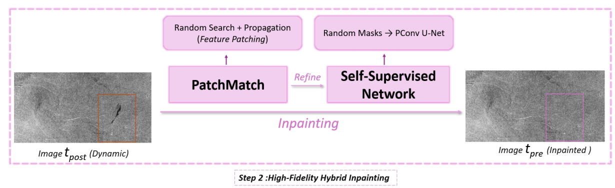
We define **oil spill change detection (OSCD)** as identifying spill-specific changes between bi-temporal SAR observations while suppressing persistent look-alikes. In practice, matched pre-spill imagery is rarely available due to revisit gaps and ocean variability. To overcome this, we generate a synthetic oil-free reference from each post-spill image I_t and mask M, denoted $\hat{I}_{t-1}$. The pair $(\hat{I}_{t-1}, I_t)$ is then fed into a change detection (CD) network to predict a binary change map C.

3.2. Temporal-Aware Hybrid Inpainting (TAHI)

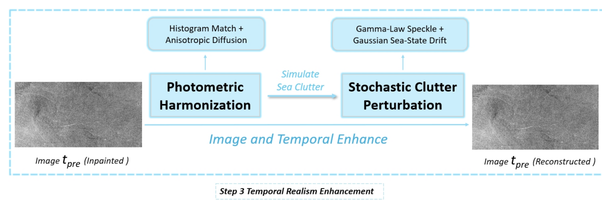
Our **TAHI** pipeline synthesizes realistic pre-spill images in three stages:



(a) Dynamic foreground simulation.



(b) High-fidelity hybrid inpainting.



(c) Temporal realism enhancement.

Fig. 2. Overview of the three TAHI stages for synthetic pre-spill generation.

1) Dynamic Foreground Simulation (DFS). Ship regions annotated in auxiliary maps are randomly shifted or removed, mimicking vessel drift or departure before the spill. The perturbed scene serves as input for subsequent inpainting.

2) High-Fidelity Hybrid Inpainting (HFHI). We first apply PatchMatch [20] to fill oil and removed-ship regions with similar background patches, then refine seams with a U-Net inpainting model using partial convolutions [21]. This combination balances global structural continuity and local speckle restoration.

3) Temporal Realism Enhancement (TRE). To reduce artifacts, we align intensities between inpainted and surrounding regions, smooth seams with anisotropic diffusion, and apply perturbations: Gamma-distributed speckle and a low-frequency drift field. These ensure that the only systematic difference between $(\hat{I}_{t-1}, I_t)$ is the oil spill itself.

3.3. Change Detection Module

For the CD backbone, we adopt existing Siamese encoder–decoder networks such as FC-Conc [24] or ChangeStar [25]. Training uses binary cross-entropy plus Dice loss against the oil mask M. Our focus is not on architectural novelty, but on demonstrating that synthetic pairs generated by TAHI enable CD models to robustly distinguish new oil slicks from persistent look-alikes.

4. DATASET CONSTRUCTION AND QUALITY ASSESSMENT

4.1. Data

We build our dataset from the M4D collection [4], which provides Sentinel-1 SAR VV imagery and masks. All confirmed oil-spill samples are selected as post-event inputs, while other labels (ships, coastlines, low wind) are treated as background or distractors. Using real SAR ensures realistic spill signatures and diverse sea context. Each polluted image is paired with a synthetic oil-free counterpart generated by TAHI.

4.2. Quality Metrics

To assess realism of pre-spill synthesis we compute: (i) ENL (Equivalent Number of Looks) for speckle smoothness, (ii) CNR for global contrast, (iii) ISLR/PSLR for side-lobe leakage, and (iv) a Dice score from a Unet++ spill segmenter (near-zero implies residual oil is undetectable).

Method	ENL↑	CNR↓	Dice↓	ISLR	PSLR
Baseline (Orig.)	10.93	10.52	0.495	8.29	0.295
PatchMatch	11.11	10.14	0.106	8.29	0.299
U-Net	11.98	8.27	0.011	7.75	0.316
PatchMatch+U-Net	**12.69**	**7.91**	**0.003**	8.29	0.296

Table 1. Quantitative evaluation of inpainting methods.

4.3. Results

Fig. 3 compares restoration strategies. PatchMatch leaves bright artifacts, U-Net oversmooths textures, while our hybrid approach removes spills and preserves sea clutter. Table 1 confirms this: ENL is highest, CNR lowest, and Dice nearly zero, showing effective removal of oil signatures.

4.4. OSCD Dataset

The final **OSCD** dataset contains 879 bi-temporal pairs (791 train / 88 test), each with a synthetic pre-spill and real post-spill SAR image plus binary change mask. Oil pixels form only about 1% of the total, making it highly imbalanced (Table 2).

In summary, OSCD augments real SAR spills with synthetic pre-event views, providing the first benchmark dedicated to oil-spill CD and enabling development of temporally aware monitoring algorithms.

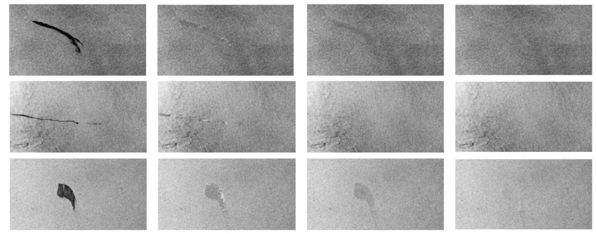


Fig. 3. Restoration comparison: Original, PatchMatch, U-Net, and PatchMatch+U-Net (ours).

Table 2. Pixel-level statistics of post-spill masks.

Subset	Pairs	Oil Pixels	Ratio
Train	791	8.16M	1.02%
Test	88	0.96M	1.07%

5. OSCD EVALUATION

5.1. Experimental Setup

We benchmark both segmentation and change detection (CD) models on OSCD. As a non-learning baseline, **Diff-Otsu** applies absolute differencing with Otsu thresholding [26]. For segmentation, we report results of **YOLOv8-SAM**, **DeepLabv3+**, **TransOilSeg**, and **OSDMamba**, all trained on single post-spill frames. For CD, we evaluate four representative networks: **CGNet** (encoder–decoder fusion), **ChangeStar** (Siamese ResNet with star-shaped head), **FC-Conc** (feature concatenation), and **DDLNet** (cross-feature interaction). All CD models share the same data splits and loss (BCE + Dice). Evaluation uses Precision, Recall, F_1, and oil-specific IoU (OS IoU). Relative improvement over Diff-Otsu is computed as Δ_{rel}.

5.2. Results

As shown in Table 3, all CD models outperform Diff-Otsu, with gains up to +259% (DDLNet). FC-Conc also achieves strong balance of precision and recall, while ChangeStar provides a lightweight option. Even CGNet, though simple, exceeds the unsupervised baseline. Segmentation models trained on single frames remain competitive, reflecting domain-specific architectures and pretraining, but lack temporal cues and thus misclassify look-alikes.

Qualitative examples (Fig. 4) confirm that CD models suppress false positives on biogenic slicks or wind shadows, which segmentation often over-detects. These results show that synthetic pre-spill imagery enables robust temporal discrimination, with further improvements hinging on advanced fusion designs.

Table 3. Performance on OSCD test set.

Method	OS IoU↑	P↑	R↑	F1↑	Δ_{rel}
Diff-Otsu	23.12	31.1	47.4	37.6	-
YOLOv8-SAM	41.8	-	-	-	+81
DeepLabv3+	53.4	-	-	-	+131
TransOilSeg	62.4	-	-	-	+170
OSDMamba	65.6	-	-	-	+184
CGNet	43.5	71.5	91.9	59.9	+88
ChangeStar	67.8	75.1	87.4	80.7	+193
FC-Conc	79.6	87.3	90.0	88.6	+244
DDLNet	**82.9**	**88.4**	**93.1**	**90.7**	**+259**

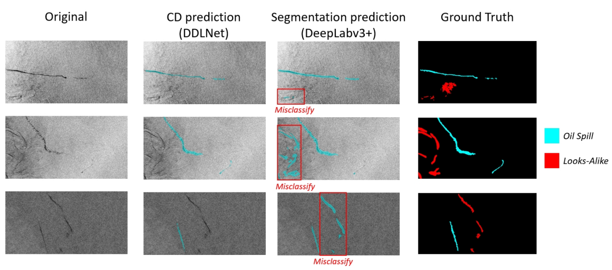


Fig. 4. Comparison of CD and segmentation outputs on OSCD. CD avoids false positives from look-alikes.

6. CONCLUSION

We present the first **OSCD** dataset and the **TAHI** pipeline for synthetic pre-spill generation. TAHI integrates dynamic vessel simulation, hybrid inpainting, and realism enhancement to create oil-free references for bi-temporal CD. Benchmarks demonstrate that CD models trained on OSCD substantially surpass both differencing and strong single-frame segmenters, highlighting the benefit of temporal supervision. Future extensions may incorporate human-in-the-loop refinement and transformer-based fusion to further enhance maritime oil-spill monitoring.

7. REFERENCES

[1] Camilla Brekke and Anne H. S. Solberg, "Oil spill detection by satellite remote sensing," *Remote Sensing of Environment*, vol. 95, no. 1, pp. 1–13, 2005.

[2] Rami Al-Ruzouq, Mohamed Barakat A. Gibril, Abdallah Shanableh, Abubakir Kais, Osman Hamed, Saeed Al-Mansoori, and Mohamad Ali Khalil, "Sensors, features, and machine learning for oil spill detection and monitoring: A review," *Remote Sensing*, vol. 12, no. 20, 2020.

[3] Konstantinos N. Topouzelis, "Oil spill detection by SAR images: Dark formation detection, feature extraction and classification algorithms," *Sensors*, vol. 8, no. 10, pp. 6642–6659, 2008.

[4] Marios Krestenitis, Georgios Orfanidis, Konstantinos Ioannidis, Konstantinos Avgerinakis, Stefanos Vrochidis, and Ioannis Kompatsiaris, "Oil spill identification from satellite images using deep neural networks," *Remote Sensing*, vol. 11, no. 15, pp. 1762, 2019.

[5] Abdulla Alhashmi, Maitham Al Shanqiti, Marwan Darweesh, and Mohammed A. Abed, "Deeplabv3+ with transfer learning for oil spill detection from synthetic aperture radar images," *Scientific Reports*, vol. 15, no. 1, pp. 10119, 2025.

[6] Yu Chai, Xinhai Han, Yiqi Wang, Dan Luo, Jingsong Yang, Peng Chen, and Gang Zheng, "Transoilseg: A novel sar oil spill detection method addressing data limitations and look-alike confusions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 63, pp. 1–16, 2025.

[7] Shuaiyu Chen, Fu Wang, Peng Ren, Chunbo Luo, and Zeyu Fu, "Osdmamba: Enhancing oil spill detection from remote sensing images using selective state space model," *IEEE Geoscience and Remote Sensing Letters*, pp. 1–1, 2025.

[8] Yulong Chen and Zhi Wang, "Marine oil spill detection from SAR images based on attention u-net model using polarimetric and wind speed information," *International Journal of Environmental Research and Public Health*, vol. 19, no. 19, pp. 12315, 2022.

[9] Chenglei Li, Duk jin Kim, Soyeon Park, Junwoo Kim, and Juyoung Song, "A self-evolving deep learning algorithm for automatic oil spill detection in sentinel-1 sar images," *Remote Sensing of Environment*, vol. 299, pp. 113872, 2023.

[10] Lingfei Shi, Xianhu Wei, Kun Yang, and Gong Chen, "A few-shot oil spill segmentation network guided by

multi-scale feature similarity modeling," *Frontiers in Marine Science*, vol. Volume 11 - 2024, 2024.

[11] Wenhui Wu, Man Sing Wong, Xinyu Yu, Guoqiang Shi, Coco Yin Tung Kwok, and Kang Zou, "Compositional oil spill detection based on object detector and adapted segment anything model from sar images," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1–5, 2024.

[12] Faezeh Mahmoudi Ghara, Seyed Behzad Shokouhi, and Ghassem Akbarizadeh, "A new technique for segmentation of the oil spills from synthetic-aperture radar images using convolutional neural network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 8834–8844, 2022.

[13] Xiaofeng Li, Chunyan Li, Zizang Yang, and William Pichel, "Sar imaging of ocean surface oil seep trajectories induced by near inertial oscillation," *Remote Sensing of Environment*, vol. 130, pp. 182–187, 2013.

[14] Achille Carlo Ciappa, "Oil trajectory analysis for oil spill surveillance by sar in the mediterranean sea," *Marine Pollution Bulletin*, vol. 190, pp. 114825, 2023.

[15] F. Del Frate, A. Petrocchi, J. Lichtenegger, and G. Calabresi, "Neural networks for oil spill detection using ers-SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 5, pp. 2282–2287, 2000.

[16] Dengsheng Lu, Paul Mausel, Eduardo Brondízio, and Emilio Moran, "Change detection techniques," *International Journal of Remote Sensing*, vol. 25, no. 12, pp. 2365–2407, 2004.

[17] Ting Bai, Lichao Wang, Daoliang Yin, Kaichang Sun, Yifen Chen, Wei Li, and Deren Li, "Deep learning for change detection in remote sensing: a review," *Geospatial Information Science*, vol. 26, no. 3, pp. 262–288, 2023.

[18] Weiming Li, Lihui Xue, Xueqian Wang, and Gang Li, "Mctnet: A multi-scale cnn-transformer network for change detection in optical remote sensing images," in *2023 26th International Conference on Information Fusion (FUSION)*. June 2023, p. 1–5, IEEE.

[19] Gianluca Murdaca, Federico Ricciuti, Alessio Rucci, Bertrand Le Saux, Alfio Fumagalli, and Claudio Prati, "A semi-supervised deep learning framework for change detection in open-pit mines using sar imagery," *Remote Sensing*, vol. 15, no. 24, 2023.

[20] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B. Goldman, "PatchMatch: A randomized correspondence algorithm for structural image editing," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 28, no. 3, pp. 24:1–24:11, 2009.

[21] Minyoung Huh, Andrew Liu, Andrew Owens, and Alexei A. Efros, "Fighting fake news: Image splice detection via learned self-consistency," in *Computer Vision – ECCV 2018*, Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, and Yair Weiss, Eds., Cham, 2018, pp. 106–124, Springer International Publishing.

[22] Jingtao Li, Qian Zhu, Xinyu Wang, Hengwei Zhao, and Yanfei Zhong, "Anomalycd: A benchmark for earth anomaly change detection with high-resolution and time-series observations," 2024.

[23] Enki Cho, Soyoun Won, Oh Sung Choo, and Seong Tae Kim, "Prism: Pseudo-labeling and region-based inpainting for synthetic change detection modeling," *IEEE Geoscience and Remote Sensing Letters*, vol. 22, pp. 1–5, 2025.

[24] Rodrigo Caye Daudt, Bertr Le Saux, and Alexandre Boulch, "Fully convolutional siamese networks for change detection," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 4063–4067.

[25] Zhuo Zheng, Ailong Ma, Liangpei Zhang, and Yanfei Zhong, "Change is everywhere: Single-temporal supervised object change detection in remote sensing imagery," 2023.

[26] Nobuyuki Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.