Oil Spill detection and Fingerprinting Using Semantic Segmentation and Data-driven Modeling

Saeed Hashemi Halvaei

A Thesis

in

The Department

of

Building, Civil and Environmental Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science (Civil Engineering) at

Concordia University

Montreal, Quebec, Canada

August 2024

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Saeed Hashemi Halvaei

Entitled: Oil Spill Detection and Fingerprinting Using Semantic Segmentation and Data-Driven Modeling

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Civil Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____Chair
Dr. C. An

_____Examiner
Dr. C. An

_____Examiner
Dr. J. Hwang

_____Supervisor
Dr. Zhi Chen

Approved by:

_____

Dr. Mohammed Ouf, Graduate Program Director

August 29th 2024:

_____

Dr. Mourad Debbabi, Dean of Gina Cody School of Engineering and Computer Science

Abstract

Oil Spill Detection and Fingerprinting Using Semantic Segmentation and Data-Driven Modeling

Saeed Hashemi Halvaei

Oil spills significantly threaten marine environment, damaging ecosystem, wildlife, and coastal communities. This thesis addresses these challenges by employing both advanced machine learning techniques and satellite imagery analysis technologies to enhance the accuracy and efficiency of oil spill detection and or source identification. By utilizing Synthetic Aperture Radar (SAR) images and examining semantic segmentation models, the research aims to accurately detect oil spills based on satellite images. Additionally, oil fingerprinting techniques, involving unsupervised classification are used to identify the sources of marine oil spills, providing a comprehensive framework for oil spill monitoring and management. The methodology involves the use of three distinct datasets: a multi-class dataset for detecting oil spills using satellite images, a binary dataset focusing on oil spill incidents in the Gulf of Suez from 2017 to 2021 as a case study for oil spill detection and a dataset for oil fingerprinting based on samples from the MV Manolis L shipwreck. For oil spill detection, semantic segmentation models were trained and evaluated using these datasets. Performance metrics such as Intersection over Union (IoU) were used to assess the modeling accuracy. Secondly for oil fingerprinting, PCA and HCA were applied to analyze the chemical composition data of the MV Manolis L. oil samples to identify their similarities and differences for oil source classification.

The results indicate that DeepLabv3+ and UNet++ models achieved the highest mean Intersection over Union (mIoU) scores for multi-class and binary segmentation tasks, respectively, demonstrating their robustness in detecting oil spills. Specifically, DeepLabv3+ achieved a mIoU of 68.3% in the multi-class dataset, excelling in complex categories like oil spills and look-alikes. UNet++ achieved a mIoU of 87.5% in the binary dataset, highlighting its effectiveness in distinguishing oil from non-oil regions. For oil fingerprinting, the Support Vector Classifier (SVC) model exhibited the highest accuracy, particularly in predicting the composition of n-alkanes, PAHs, and TPH, with F-scores of 1.0, 0.987, and 0.975, respectively. These findings underscore the effectiveness of coupling advanced machine learning models with established chemical

analysis techniques, offering a reliable approach for oil spill detection and the subsequent effective cleanup.

ACKNOWLEDGEMENTS

I would first like to thank my thesis advisor, Dr. Zhi Chen. The door to Dr. Chen's office was always open whenever I encountered a trouble spot or had a question about my research or writing. He consistently allowed this thesis to be my own work but steered me in the right direction whenever he thought I needed it. I would also like to thank my co-supervisor, Dr. Zeyu Yang from Environment and Climate Change Canada (ECCC), for her invaluable insights and support throughout my research.

Special thanks to my friends and colleagues at Concordia University for their support and camaraderie: Chudi Wu, Afzal Ahmed Dar, Zhaoyang Yang, Kamyar Soleymani, and many others who have contributed in various ways to my research.

Finally, I must express my very profound gratitude to my partner, Bahar, for her unwavering support and love. This accomplishment would not have been possible without her encouragement and constant companionship. My deepest thanks also go to my parents for their unfailing support and continuous encouragement throughout my years of study and the process of researching and writing this thesis. This achievement would not have been possible without them.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

TP	Correct positive predictions.
FP	Incorrect positive predictions.
FN	Incorrect negative predictions.
IoU	Accuracy of object detection, overlap over union.
mIoU	Average IoU across classes.
Precision	Accuracy of positive predictions.
Recall	Completeness of positive predictions.
F1 Score	Balance of precision and recall.

List of Abbreviations

APAH	Alkylated Polycyclic Aromatic Hydrocarbons
ASPP	Atrous Spatial Pyramid Pooling
CNN	Convolutional Neural Network
CDO	Chemically Dispersed Oil
COCO	Common Objects in Context
CP-ANN	Counter Propagation-Artificial Neural Network
DCNN	Deep Convolutional Neural Networks
EVC	Ensemble Vote Classifier
FCN	Fully Convolutional Network
FID	Flame Ionization Detector
FPN	Feature Pyramid Network
GC	Gas Chromatography
GC×GC	Comprehensive Two-dimensional Gas Chromatography
GC/FID	Gas Chromatography-Flame Ionization Detection
GC/MS	Gas Chromatography-Mass Spectrometry
HCA	Hierarchical Cluster Analysis
IRMS	Isotopic Resolution Mass Spectrometry
KNN	K-Nearest Neighbor

LR	Logistic Regression
LRC	Logistic Regression Classifier
ML	Machine Learning
PCA	Principal Component Analysis
PAH	Polycyclic Aromatic Hydrocarbons
PSPNet	Pyramid Scene Parsing Network
RFC	Random Forest Classifier
SAR	Synthetic Aperture Radar
SVC	Support Vector Classifier
TPH	Total Petroleum Hydrocarbons
YOLO	You Only Look Once

Chapter 1 Introduction

1.1. Background and Motivation

Oceanic and maritime oil pollution presents a significant and ongoing challenge, as shown by numerous studies (International Tanker Owners Pollution Federation (ITOPF), 2018; Liu et al., 2017; Z. Wang et al., 1999). Both intentional and accidental discharges of petroleum hydrocarbons into waterbodies have led to numerous ecological catastrophes, adversely impacting marine ecosystems, reducing the productivity and quality of marine habitats. Given that oceans cover nearly two-thirds of the Earth's surface and play a vital role in human life and economic well-being, preserving marine environments is crucial for sustainability in the short and long term(Bayindir et al., 2018). In marine environments, oil spills are particularly harmful and destructive compared to those occurring on land. They can quickly spread across vast distances and form a thin layer of oil that blankets the shorelines. The process of detecting and monitoring these spills is both time-consuming and expensive. Nevertheless, establishing immediate response mechanisms is essential to mitigate their devastating impacts (Raeisi et al., 2018). Effectively reducing the environmental consequences of such pollution hinges on consistent monitoring of the marine areas. This allows for precise measurements of the spread of oil, facilitating swift actions for mitigation and recovery. Over recent decades, the detection of marine oil spills has garnered significant attention due to their threats to human health and the profound environmental and economic damages they inflict on marine life, fisheries, wildlife, coastal communities, mangrove forests, and other socioeconomic aspects (Dabboor et al., 2018).

Offshore oil platforms are a major source of marine oil pollution. These platforms, used to explore, extract, store, and process oil and natural gas, significantly increase the risk of oil spills with catastrophic impacts on the marine environment. The North Sea and the Gulf of Mexico in the United States are particularly vulnerable to such spills due to their high concentrations of drilling rigs, numbering 184 and 175 respectively (Fazeres-Ferradosa et al., 2019) .The Persian Gulf, far East Asia, and Southeast Asia also face similar threats with their substantial numbers of oil platforms. Marine pollution often results from various factors including accidents involving

ships or oil rigs, breakdowns of old and damaged infrastructures, human errors, and conflicts (International Tanker Owners Pollution Federation (ITOPF), 2018) . These incidents lead to significant contamination of ocean waters with liquid petroleum hydrocarbons, causing long-term damage to marine ecosystems (Hoffman & Jennings, 2011). In extreme cases, massive oil spills release millions of gallons of oil, leading to widespread environmental degradation, affecting wildlife and resulting in economic losses, particularly in tourism. Historical examples of such major oil spills have been summarized in **Table 1.1**.

Table 1.1 Major oil spill disasters in the world history ranked by the amount of spill size (Jafarzadeh et al., 2021).

No.	Spill/Tanker	Location	Date	Amount Spilled (million gallons)
1	Gulf War oil spill	Persian Gulf, Kuwait	19 January 1991	380–520
2	Deepwater Horizon	Macondo Prospect, Central Gulf of Mexico	22 April 2010	206
3	Ixtoc-I Oil Spill	Bay of Campeche off Ciudad del Carmen, Mexico	3 June 1979	140
4	Atlantic Empress Oil Spill	Off the coast of Trinidad and Tobago	19 July 1979	90
5	Kolva River Oil Spill	Kolva River, Russia	6 August 1983	84
6	Nowruz Oil Field Spill	Persian Gulf, Iran	10 February 1980	80
7	Castillo de Bellver Oil Spill	Off Saldanha Bay, South Africa	6 August 1983	79
8	Amoco Cadiz Oil Spill	Portsall, France	16 March 1978	69
9	ABT Summer Oil Spill	About 700 nautical miles off the coast of Angola	28 May 1991	51–81
10	M/T Haven Tanker Oil Spill	Genoa, Italy	11 April 1991	45
11	Odyssey Oil Spill	Off the coast of Nova Scotia, Canada	10 November 1988	40.7

No.	Spill/Tanker	Location	Date	Amount Spilled (million gallons)
12	The Sea Star Oil Spill	Gulf of Oman	19 December 1972	35.3
13	The Torrey Canyon Oil Spill	Scilly Isles, U.K.	18 March 1967	25–36
14	Sanchi	Off Shanghai, China	6 January 2018	34
15	Irenes Serenade	Navarino Bay, Greece	23 February 1980	30
16	Urquiola	La Coruna, Spain	12 May 1976	30
17	Hawaiian Patriot	300 nautical miles off Honolulu	23 February 1977	30
18	Independenta	Bosphorus, Turkey	15 November 1979	28.9
19	Jakob Maersk	Oporto, Portugal	25 January 1975	26.4
20	Braer	Shetland Islands, UK	5 January 1993	25.5

In the event of an oil spill, effective detection and control measures are crucial. Oil spills composed of mineral oil on the ocean's surface can be identified using imaging radar technology. These spills suppress the minor surface waves that are essential for radar backscatter, causing them to appear as dark regions on synthetic aperture radar (SAR) imagery, as the darkness correlates to the normalized radar cross section (NRCS), indicating the intensity of the reflected radar signals (Hasimoto-Beltran et al., 2023a) . Yet not every dark area on SAR imagery is attributable to mineral oil spills. Such appearances might also result from natural surface films created by marine organisms like plankton or fish, calm wind conditions found behind geographical barriers like islands or mountains, sandbanks exposed during low tide, water disturbances from ship wakes, turbulence induced by rainfall on the water's surface, or wastewater released from terrestrial industrial or urban sources (Alpers et al., 2017). These alternative sources generate radar signatures similar to those of oil spills, known as “look-alikes,” to oil spills. Identifying true oil spills amidst these look-alikes in SAR imagery is thus highly challenging as shown in Fig 1.1.

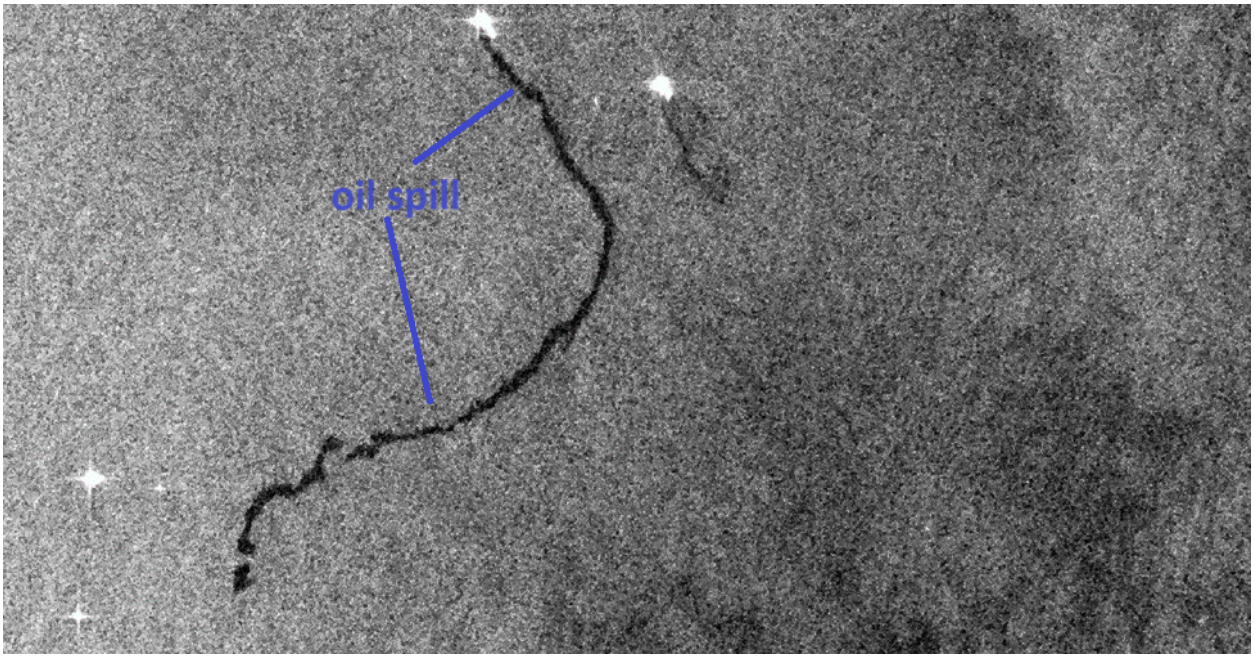


Fig. 1.1 A sample of Sentinel 1 SAR images showing oil spill (Krestenitis et al., 2019b).

Oil spill SAR images, which are captured using various satellite sensors, exhibit distinct characteristics, such as spatial resolution and waveband. These differences can make it challenging to accurately segment different types of oil spill images. Additionally, having a substantial amount of high-quality oil spill SAR image data is advantageous for training deep learning models that segment oil spills effectively. However, because oil spill incidents are infrequent, there is a limited supply of such image data available for research purposes. This scarcity of data can hinder the full training of deep neural networks for segmentation, consequently impairing the effectiveness of oil spill image analysis (Hasimoto-Beltran et al., 2023b) .

In the realm of object detection for oil spill monitoring, the integration of advanced models has led to significant developments. Primarily, machine learning-based methods utilize algorithms trained on SAR data to identify oil spills by recognizing patterns specific to oil characteristics. This approach often involves complex computational models that can distinguish oil spills from natural sea surface features, providing high accuracy and efficiency in detection. Secondly, object detection methods focus on identifying and delineating the physical boundaries of oil spills within the imagery. Techniques such as convolutional neural networks are used to segment spills, allowing for precise quantification and monitoring over time. Lastly, semantic and instance segmentation involves classifying each pixel in an image to determine whether it is part of a spill

or not, with instance segmentation further separating individual instances of oil spills. This method is crucial for detailed analysis and is used in comprehensive environmental impact assessments, facilitating targeted cleanup and mitigation strategies.

The exploration of machine learning-based methods for oil spill detection has been advanced by several significant studies. A notable study by Tong et al. (2019) demonstrated the use of a self-similarity parameter in Radarsat-2 and UAVSAR data to differentiate genuine oil slicks from look-alikes, achieving accuracies of 92.99% and 82.25%, respectively. This approach was notably effective in low wind conditions, providing a robust tool for environmental monitoring. Similarly, Conceição et al. (2021) advanced through the employment of two random forest classifiers to differentiate between various types of oil spills using SAR imagery, achieving a remarkable detection accuracy of 90% through enhanced feature application. In the realm of object detection for oil spill monitoring, the integration of advanced models has led to significant developments. Zhu et al. (2022) enhanced the YOLOX-S model with a truncated linear stretch module and score loss, improving its ability to detect marine oil spills. Concurrently, Huang et al. (2022) utilized the Faster R-CNN to rapidly and accurately detect oil spills with high precision. These advancements underscore the potential of deep learning techniques in real-time environmental monitoring and operational maritime activities, offering a rapid response to ecological threats. Semantic and instance segmentation technologies have also seen notable applications in oil spill detection. The use of deep convolutional neural networks (DCNNs), as highlighted in studies like those by Krestenitis et al. (2019a) and Ma et al. (2022) has improved the accuracy of oil spill segmentation by training on diverse and complex image data. These methods, especially the super pixel segmentation technique by Ma et al., allow for adaptable and efficient processing of SAR images, providing detailed and reliable detection of oil spills even under challenging conditions. This integration of advanced segmentation techniques showcases the evolving landscape of oil spill detection technologies, pushing the boundaries of environmental preservation efforts.

This research employing state-of-the-art deep learning techniques. The use of advanced models such as U-Net, LinkNet, and DeepLabv3+ allows for effective handling of complex textures and small-scale features in satellite images. These models are designed to capture fine-grained details and can distinguish between oil spills and other similar-looking features with high

accuracy. Techniques such as data augmentation and transfer learning are employed to overcome the issue of data scarcity (Li et al., 2023). Data augmentation artificially increases the size of the multi-class dataset by creating modified versions of existing data, while transfer learning leverages pre-trained models on similar tasks to improve performance on the target task. Models like DeepLabv3+ and Pyramid Scene Parsing Network (PSPNet) incorporate mechanisms to capture multi-scale features, which are crucial for maintaining high accuracy across different environmental conditions. Additionally, this research uses two case studies to evaluate the effectiveness of the latest semantic models on binary segmentation and multi-class segmentation in oil spill detection, ensuring reliable oil spill detection and monitoring (Fustes et al., 2014). Despite these advancements, gaps remain in the field of oil spill detection using semantic segmentation. One significant challenge is the lack of binary segmentation datasets specifically designed for distinguishing between oil and non-oil spills. This gap limits the ability of models to effectively learn and differentiate between these two critical categories. Additionally, the accuracy of oil spill class predictions in current models is not consistently high enough to ensure reliable detection and monitoring in all scenarios. These limitations highlight the need for more comprehensive datasets and further refinement of segmentation models to enhance their performance.

While detecting oil spills using satellite imagery is crucial, the next vital step in this research is identifying the sources of these spills. Oil fingerprinting is an essential technique in environmental forensics used to pinpoint the origin of oil spills and understand their impact on marine ecosystems. However, traditional oil fingerprinting methods face several significant challenges (Bayona et al., 2015a). These methods often rely on manual analysis and basic statistical techniques, which can be both time-consuming and prone to human error, leading to limited accuracy and speed. Additionally, the chemical composition of oil is complex, comprising various hydrocarbons and other compounds, making it difficult for traditional methods to accurately identify and differentiate between different oil sources. This complexity complicates tracking the environmental behavior and degradation of oil over time. Furthermore, oil's chemical composition can change due to various environmental factors such as weathering, biodegradation, and photo-oxidation, which traditional methods may not adequately account for, resulting in inaccurate fingerprinting and source identification (Wang et al., 2006a). In this context, integrating advanced deep learning models and robust analytical techniques into oil fingerprinting can

significantly enhance the accuracy and efficiency of identifying oil spill sources. This approach not only aids in precise source identification but also helps in understanding the environmental impact and behavior of spilled oil over time, thereby supporting effective response and remediation efforts.

Petroleum is composed of diverse compounds from saturates, aromatics, resins, and asphaltenes. Semi-volatile components belonging to saturates and aromatics that can be detected by gas chromatography (GC) usually range from C8 to C50 (Fig. 1.2). Specified hydrocarbon groups, including aliphatic hydrocarbons like n-alkanes, iso-alkanes, cycloalkanes, cyclic and aromatic petroleum biomarkers, as well as polycyclic aromatic hydrocarbons (PAHs) and their alkylated derivatives, have been identified by gas chromatography-mass spectrometry (GC/MS) effectively. Among these individual hydrocarbon groups, petroleum biomarkers derived from the original organic material reflect the depositional environment and geological history. They include sesquiterpenes, steranes, and terpenes, which are crucial for oil forensic identification. Additionally, polar compounds such as resins, primarily heterocyclic compounds containing elements like sulfur, nitrogen, and oxygen, are also present (Bayona et al., 2015b). High molecular weight compounds, often exceeding 1000 atomic mass units and known as asphaltenes, are dispersed as colloids in the heavier fractions of petroleum. Oil fingerprinting by gas chromatography (GC), an essential technique for oil spill analysis, relies on specific and stable chemical parameters unaffected by environmental changes like evaporation, dissolution, photo-oxidation, and biodegradation. The MS profiles reveal the presence of hopane (m/z 217), sterane (m/z 231), and specific poly aromatic hydrocarbons (PAHs) (m/z 216), each with characteristic retention times and intensities. These biomarkers are stable chemical parameters unaffected by environmental changes like evaporation, dissolution, photo-oxidation, and biodegradation, making them essential for oil fingerprinting. Oil fingerprinting, an essential technique for oil spill analysis, relies on these stable markers. The process is tiered: Tier 1 involves assessing the gas chromatography-flame ionization detector (GC/FID) profile to determine the type of oil **(Fig 1.2(A)).** Tier 2 utilizes GC/MS to identify specific petroleum biomarkers and hydrocarbons (**Fig 1.2 (B**)). Tier 3 involves calculating diagnostic ratios between certain compounds to compare potential sources with spilled samples. These methodologies allow for the precise identification and comparison of oil samples, aiding in environmental forensic investigations and the management of oil spill incidents.

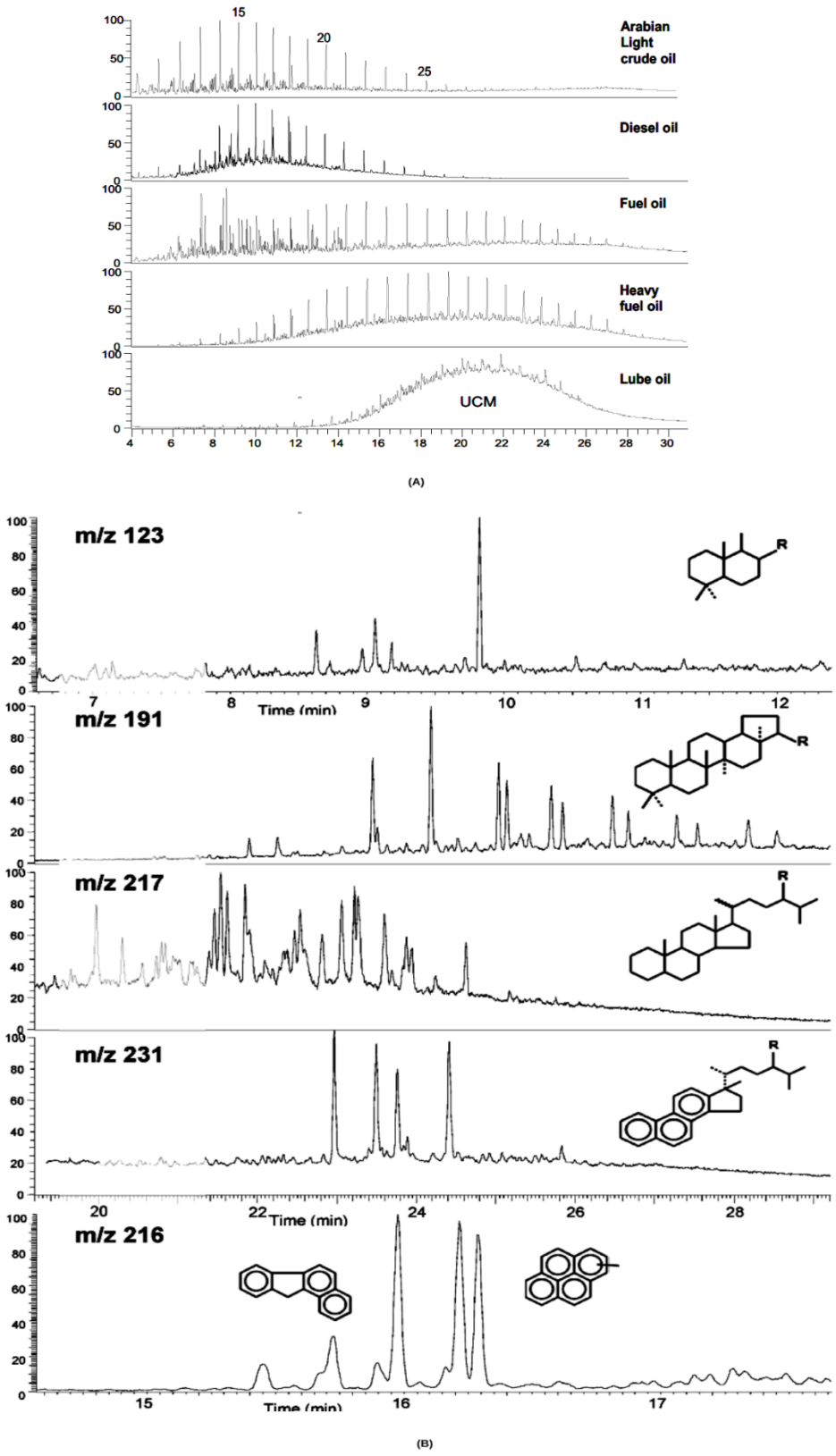


Fig. 1.2 (A) GC–FID chromatograms of possible oil spill sources (B) GC–MS trace of selected molecular markers or biomarkers (Bayona et al., 2015b)

A research project examined the changes in the composition of a fuel oil spill from 2000 in mangrove sediment cores in Brazil over four years using GC/FID and GC/MS techniques (Farias et al., 2008). The study differentiated between highly affected, less affected, and unaffected areas. Diagnostic ratios involving hopanes and steranes confirmed the presence of the oil, while analyses of total PAHs and specific PAH ratios like fluoranthene/pyrene and alkylated PAHs indicated the level of contamination. In the most affected regions, sediments showed evidence of oil migration into deeper layers. The challenging environmental conditions and ongoing contamination complicated the relationship between observed PAH levels and the spill, making it difficult to distinctly trace the oil's origin in the presence of other residues. The uneven distribution of oil in the mangrove sediments highlighted the difficulties of determining the contamination state and studying oil's environmental impact.

Advanced techniques such as comprehensive two-dimensional gas chromatography (GC) have proven essential for addressing these challenges, providing detailed insights into oil weathering processes and enhancing the understanding of how oil behaves in complex environmental contexts. GC has significantly advanced the analysis of spilled oil samples, addressing many of the previous limitations. This technique is especially useful for assessing oil weathering, as it enables researchers to measure retention indices that predict liquid-vapor pressures, solubilities in water, and partition coefficients between octanol-water and air-water, as well as vaporization enthalpies. Gaines et al. (1999) compared a spilled diesel sample against a database to pinpoint its origin. The two-dimensional chromatograms used flame ionization detection to assess the spill and potential sources, focusing on the presence and intensity of peaks in key regions including naphthalenes, anthracenes, phenanthrenes, alkanes, and cycloalkanes. This comparison, visualized through bar graphs of peak ratios, helped quantify similarities and differences, facilitating a better understanding of the spill's characteristics and origin.

After preprocessing GC chromatograms, it's crucial to apply statistical methods or linear algebra to derive meaningful chemical insights from the dataset. Chemometrics enhances the extraction of information from multidimensional GC data by applying mathematical models through multivariate statistics, enabling the establishment of valid statistical correlations. Techniques such as Principal Component Analysis (PCA) for exploratory analysis, Partial Least

Squares (PLS) for multivariate regression calibration, and Partial Least Squares Discriminant Analysis (PLSDA) for distinguishing between known sample groups are used. Effective preprocessing, careful selection of parameters, and rigorous validation of chemometric models are essential for determining the performance and robustness of these models and for understanding the main sources of variance within the data.

Despite the advancements in environmental forensics, there remain significant gaps that our research aims to address, particularly in the application and integration of cutting-edge technologies and methodologies.

1.Integration of Machine Learning: Traditional oil fingerprinting techniques, while effective, often lack the integration of advanced computational methods that can significantly enhance the precision and efficiency of analyses. Our research leverages machine learning to process complex, multivariate datasets, enabling a more precise classification of oil types. This approach not only increases the accuracy of identifying oil spill sources but also improves the speed of these determinations, which is essential for timely environmental response and remediation efforts.

2.Enhanced Analytical Methods: Further bridging the gaps in traditional environmental forensics, our study incorporates sophisticated analytical methods such as PCA and hierarchical clustering (HCA). These techniques allow for a more comprehensive analysis of the geochemical characteristics of hydrocarbons. By doing so, we can categorize crude oil more effectively based on its physicochemical properties, providing deeper insights into its behavior and environmental impact. This advancement addresses a crucial need for more detailed and nuanced analysis within the field of oil spill forensics.

3.Application in Complex Scenarios: The practical utility of our methods is demonstrated through their application to complex real-world scenarios, such as the MV Manolis L shipwreck. This case study exemplifies how our enhanced methods can be applied successfully where traditional techniques may fall short, offering robust tools that stand up to the challenges posed by complicated environmental disasters.

4.Data-Driven Decision Making: By employing a data-driven approach, our research supports more informed decision-making in environmental management and spill response

strategies. This aligns with the contemporary needs for rapid, accurate environmental assessments, enabling authorities and stakeholders to act more decisively and effectively in mitigating the impacts of oil spills.

Through these enhancements, our research not only fills existing gaps in environmental forensics but also sets a new standard for the field, pushing the boundaries of what can be achieved in the study and mitigation of environmental pollutants. This progression is crucial for developing more resilient ecosystems and for ensuring the health of our planet in the face of industrial accidents and natural disasters.

1.2. Thesis Objective

The thesis objectives are proposed as follows:

1. Developing and Testing Advanced Segmentation Models: The primary goal is to create and implement advanced semantic segmentation models—such as UNet, LinkNet, UNet++, FPN, DeepLabv3+, and PSPNet—to accurately detect and segment oil spills from SAR imagery and evaluating on real case datasets. The study will assess these models' ability to differentiate oil spills from other similar-looking features, like algae blooms or ship wakes, under various environmental conditions. The objective is to identify the most effective model for real-world applications in both multi-class and binary segmentation tasks, ensuring prompt and accurate oil spill detection.

2. Optimizing Model Training and Performance: This research aims to enhance the training and validation of the selected segmentation models by addressing data scarcity and variability issues in SAR imagery. Transfer learning will be employed to improve model accuracy and reduce training time, enabling the models to generalize effectively across different scenarios. The focus is on optimizing model architectures and applying transfer learning to boost the efficiency and reliability of oil spill detection in dynamic marine environments. This used to lower training epochs by using pretrained models on massive dataset like Imagnet which will be evaluated on models training process.

3. Enhancing Oil Spill Fingerprinting Using Analytical Methods: The research seeks to improve oil spill classification and source identification by integrating advanced analytical techniques like PCA and HCA with machine learning algorithms. The objective is to automate the

analysis of complex datasets, overcoming the limitations of traditional fingerprinting methods, and ensuring accurate identification of oil sources despite environmental changes. This will support timely and effective environmental response efforts.

4. Validating Models Through Real-World Case Studies: The practical utility of the developed techniques will be demonstrated through real-world case studies, such as the MV Manolis L shipwreck and canal Suez oil spill detection case. These case studies will test the models' effectiveness and robustness in actual spill scenarios, providing insights into their strengths and areas for improvement. The goal is to confirm the models' applicability to real-world incidents and contribute to better decision-making in environmental management and spill response strategies.

By achieving these objectives, the research will advance the accuracy, efficiency, and reliability of tools for detecting and fingerprinting oil spills, ultimately aiding in the protection of marine environments.

1.3. Thesis Outline

This thesis delves into the application of machine learning and semantic segmentation techniques within the realm of environmental forensics, with a particular focus on identifying the sources of marine oil spills and detecting these spills through satellite imagery. The structure of this thesis is organized into five comprehensive chapters, each addressing a critical component of the research. Below is an overview of the thesis structure, accompanied by a simple explanation of each chapter.

Chapter 1 discusses the significance of oceanic and maritime oil pollution, highlighting the detrimental impacts of oil spills on marine ecosystems and the importance of preserving marine environments. It emphasizes the challenges in detecting and monitoring oil spills and the need for immediate response mechanisms. The chapter also reviews major sources of marine oil pollution, including offshore oil platforms, and outlines the environmental and economic impacts of significant oil spill incidents. It also discusses the objectives of the thesis, which include enhancing oil spill detection accuracy using advanced semantic segmentation models, optimizing model training and performance validation through transfer learning, developing robust oil fingerprinting

techniques using advanced analytical methods, and evaluating model performance in real-world scenarios through case studies.

Chapter 2 provides a comprehensive review of methodologies for marine oil spill segmentation and oil fingerprinting, focusing on the application of deep learning algorithms in environmental monitoring. The chapter first discusses semantic segmentation methods, including the use of self-similarity parameters and random forest classifiers, as well as advanced object detection techniques like enhanced YOLOX-S and Faster R-CNN models. It highlights the effectiveness of deep convolutional neural networks (DCNNs), U-NET, DeepLabV3, and Mask R-CNN in accurately detecting oil spills in SAR images. Additionally, the chapter explores oil fingerprinting for marine oil spill identification, reviewing traditional methods such as gas chromatography-mass spectrometry (GC/MS) and modern data-driven models employing machine learning algorithms. By integrating advanced analytical methods with traditional techniques, the chapter underscores the improved classification and source identification of oil spills. Case studies, including the MV Manolis L shipwreck, are used to demonstrate the practical utility of these approaches.

Chapter 3 provides a detailed description of the methodology used for oil spill detection and oil fingerprinting. It outlines the datasets, data preparation processes, and the models and techniques applied to achieve accurate detection and classification of oil spills. The comprehensive methodology ensures a systematic approach, integrating advanced imaging techniques, thorough data processing steps, and sophisticated machine learning models. The chapter begins with an explanation of the framework, highlighting the use of SAR images and semantic segmentation models like U-Net, LinkNet, UNet++, FPN, DeepLabv3+, and PSPNet, trained on high-quality datasets. It also details the preprocessing steps undertaken to ensure high-quality input data for the segmentation models. Following detection, oil fingerprinting analyzes the chemical composition of detected oil samples using GC/MS and GC/FID data, with unsupervised classification techniques like PCA and k-means clustering, and further validated by HCA. The chapter also covers the development and optimization of machine learning models such as KNN, SVC, RFC, DTC, LRC, and EVC for classifying oil samples, with a focus on hyperparameter optimization and avoiding overfitting. Finally, it describes the performance evaluation methods using cross-

validation and the F-score metric, ensuring a robust and reliable classification process for real-world applications in oil spill detection and fingerprinting.

Chapter 4 is divided into two main sections: oil spill detection results and oil fingerprinting results. In the first section, the performance of different segmentation models is evaluated using quantitative metrics like intersection-over-union (IoU) and mean IoU (mIoU). The qualitative analysis includes visual comparisons of model outputs, highlighting their strengths and weaknesses in detecting and segmenting oil spills. The second section focuses on the application of PCA and HCA for oil fingerprinting. It presents key findings from the chemical composition analysis and evaluates the performance of various machine learning models using accuracy and F-scores.

Chapter 5 synthesizes the key findings of the research, discussing the overall effectiveness and limitations of the methodologies employed. It outlines the advantages of integrating advanced remote sensing techniques with machine learning models for environmental monitoring, highlighting improvements in detection accuracy and operational efficiency. However, the chapter also addresses significant challenges, such as the high computational demands and issues related to data acquisition, including the availability and quality of satellite imagery. It explores the uncertainties related to model predictions, which can arise from environmental variations, such as changing weather conditions and differing geographical features, as well as from incomplete data coverage. These uncertainties can lead to variability in model performance and potential inaccuracies in detecting and analyzing oil spills.

Chapter 6 comprises a summary of the thesis, contributions of the selected methods, and recommended future studies. It presents a comprehensive framework for oil spill detection and fingerprinting, integrating advanced remote sensing techniques with machine learning models. The methodology involves detecting oil spills using SAR images processed through semantic segmentation models and analyzing the chemical composition of detected oil samples. The contributions enhance detection accuracy and reliability, providing valuable insights into chemical composition and source identification. Future research recommendations include expanding datasets, integrating additional data sources, developing explainable AI, refining models, and applying methodologies to other environmental monitoring tasks.

This detailed outline provides a roadmap for understanding the comprehensive and multifaceted approach taken in this thesis to address the challenges of marine oil spill detection and fingerprinting using cutting-edge machine learning techniques.

Chapter 2 Literature Review

2.1. Marine Oil Spill Segmentation with Image Processing Methodologies

Research in oil spill detection techniques can be categorized into two main areas: "machine learning-based techniques for oil spill detection" and "deep learning-based techniques for oil spill detection." Machine learning-based techniques typically involve algorithms such as Support Vector Machines (SVM), Random Forests (RF), and k-Nearest Neighbors (KNN) (Boateng et al., 2020a, 2020b; Cristianini & Shawe-Taylor, 2000). These techniques rely on manually extracted features from images and then use these features to train models for detecting oil spills. Although, these methods often require significant preprocessing and feature engineering, and their performance can be limited by the quality and quantity of the available data. Deep learning algorithms, which consist of sophisticated neural networks designed to extract complex features from large image datasets (Al-Ruzouq et al., 2020), analyze data through multiple layers to identify unique patterns. These advancements have been successfully applied in various fields, consistently outperforming traditional methods (Bhatnagar et al., 2017; He et al., 2016). In environmental monitoring, particularly for oil spill detection, deep learning models have demonstrated substantial effectiveness. The "Deep learning-based techniques for oil spill detection" category is further divided into two principal detection techniques:

Object detection technique identifies and locates oil spills within larger scenes. Object detection models, such as YOLO (You Only Look Once) and Faster R-CNN (Region-based Convolutional Neural Network), generate bounding boxes around detected oil spills (Sudha & Saro Vijendran, 2024; Y. J. Yang et al., 2021). This approach is advantageous for rapidly pinpointing the location and extent of oil spills within vast areas covered by satellite or aerial imagery. It provides a high-level overview of the spill's presence and its approximate boundaries, making it useful for quick assessments and directing cleanup efforts. Instance and semantic segmentation technique goes beyond merely locating oil spills by categorizing each object and pixel in an image. Instance segmentation distinguishes and labels individual objects within an

image, ensuring that each detected oil spill is treated as a separate entity, even if they overlap. Models like Mask R-CNN are commonly used for this purpose (J. Zhang et al., 2023). Semantic Segmentation, on the other hand, classifies each pixel of an image as part of an oil spill or the surrounding environment, providing a detailed pixel-level map of the spill. Models like U-Net and DeepLabV3+ are effective for semantic segmentation. This approach is crucial for understanding the precise boundaries and area of the oil spill, facilitating detailed analysis and accurate measurement of the spill's impact.

These advanced techniques enhance the accuracy and reliability of oil spill detection, providing critical tools for environmental protection and disaster response (Basit et al., 2022). By combining object detection with instance and semantic segmentation, researchers can achieve both broad situational awareness and detailed analytical insights, ensuring comprehensive monitoring and effective intervention strategies**.**

2.1.1. Machine-learning-based methods

In the article by Tong et al. (2019) introduced a novel approach that utilizes a self-similarity parameter sensitive to the randomness of scattering targets, aiding in distinguishing genuine oil slicks from look-alikes. The study evaluates the effectiveness of this method using datasets from Radarsat-2 and UAVSAR, achieving oil spill detection accuracies of 92.99% and 82.25%, respectively. Additionally, the self-similarity parameter proves particularly effective in low wind conditions (2–3 m/s) and less so in higher wind conditions (9–12 m/s). This research marks a significant advance in the remote sensing of marine oil spills, offering more reliable tools for environmental monitoring and protection. The research conducted by Conceição et al. (2021) introduced a set of open-source methodologies adept at addressing oil-like spills through the deployment of two random forest classifiers. The first classifier leverages ocean Synthetic Aperture Radar (SAR) imagery to categorize various inputs such as biofilm and multi-substance oil spills. The second classifier, referred to as the Radar Image Oil Spill Seeker (RIOSS), targets oil spill detection on marine surfaces using Sentinel-1 SAR images. RIOSS enhances the feature application to the random forest algorithm, significantly boosting the accuracy of the detection system by 90%. This enhancement aids in achieving more precise and reliable detection outcomes in the challenging marine environment.

Magri et al. (2021) utilized a SVM classifier to determine the optimal features for the detection and classification of oil spills from satellite imagery**.** This application was specifically used to assess an environmental spill caused by a ship collision**.** The study highlights the SVM's robust classification capabilities, especially when only a limited number of training samples are available**.** The ability to precisely identify oil spills is crucial for initiating timely early warnings and providing essential data for quick remediation actions and emergency responses. Mdakane & Kleynhans (2022) explored the classification of oil spills using images from shipborne radar. The methodology encompassed three stages: First, image preprocessing was conducted to remove any interference and speckles from the original gray-scale images. Following this, wave patterns were classified using a SVM. Subsequent stages involved the selection of areas effective for monitoring and the extraction of oil spills utilizing a local adaptive threshold technique. The study revealed that the SVM effectively extracts relevant wave information from the radar images, and the local adaptive threshold method is versatile in its application for segmenting oil films.

2.1.2. Object detection

Zhang et al. (2022) incorporated a truncated linear stretch module and score loss into the original YOLOX-S model, enhancing its ability for marine oil spill detection. Their research confirms the effectiveness of the modified linear stretching module and score loss in boosting detection accuracy. Future studies aim to test the revised model across diverse SAR remote sensing images to evaluate its performance in identifying oil spills at sea. In the study by Y. J. Yang et al. (2021), two initial experiments were conducted using satellite SAR imagery to detect oil spills, utilizing the YOLOv4 object detection system. The model faced challenges in differentiating between objects, prompting the recommendation to adjust pixel thresholds to 28 in the first test. The average precisions (APs) for the validation and test sets were 67.80% and 65.37%, respectively, indicating that the model did not experience overfitting. Results from the second test revealed that the use of various data enhancements did not affect the study's outcomes. However, the introduction of rotational data augmentation posed a risk of overfitting due to the prevalence of small, nearly circular objects typical of oil spills.

The novel deep learning approach for detecting marine oil spills using satellite SAR imagery, developed by Huang et al. (2022), utilizes the Faster R-CNN to address limitations of existing methods. The faster R-CNN demonstrated robust performance, achieving precision and

recall rates of 89.23% and 89.14%, respectively, with an average precision of 92.56%. The model efficiently processes each full SAR image in less than 0.05 seconds using an NVIDIA GeForce RTX 3090 GPU, showcasing its capability for rapid and accurate oil spill detection under various environmental conditions. This method shows promise for real-time applications in marine conservation and operational maritime monitoring, making it a significant advancement in environmental remote sensing. In a comprehensive study, Sudha & Saro Vijendran (2024)outline an innovative approach to detecting and managing oil spills through advanced computer vision techniques, employing a multi-stage strategy to improve the accuracy and efficiency of oil spill identification. The researchers utilize Contrast Limited Adaptive Histogram Equalization to enhance image quality and reduce data noise, followed by a Fused UNet Segmentation model for precise delineation of contaminated areas. Key features are then extracted using a CNN based on the AlexNet architecture, which significantly improves the model's discriminative ability. The core of their method integrates Faster R-CNN with Enhanced MobileNetV2 architecture, enabling real-time processing and high-performance object recognition.

In the innovative study by Huang et al. (2022), a novel framework named SAM-OIL is introduced to improve the detection of oil spills in SAR imagery. This framework integrates several advanced components: YOLOv8 for initial object detection, an adapted Segment Anything Model (SAM) for generating category-agnostic masks, and an Ordered Mask Fusion (OMF) module for merging these elements into a coherent output. The SAM-OIL framework marks the first use of SAM in the context of oil spill detection, significantly advancing the field by addressing the limitations of previous semantic segmentation-based methods that required extensive finely annotated data. By leveraging the combined capabilities of these components, SAM-OIL achieves a mean Intersection over Union (mIoU) of 69.52%, showcasing its effectiveness and efficiency over existing methods. The study highlights the adaptability and accuracy improvements brought by the Adapter and OMF modules within the SAM-OIL framework.

2.1.3. Instance and semantic segmentation

Krestenitis et al. (2019b) demonstrated that deep convolutional neural networks (DCNNs) are effective for detecting oil spills through semantic segmentation. The DCNN models were specifically trained and evaluated on a standardized dataset on an individual basis. Among these, the DeepLabv3 model achieved the highest accuracy, attributed to its longer inference time.

However, the DeepLabv3 model, despite its advantages, has limitations in differentiating oil spill classes due to similar pixel classifications, which might be a result of the limited sample size and the particularities of the training approach used in the study. The adoption of deep neural networks, such as U-NET and DeepLabV3, has improved the accuracy of oil spill segmentation due to their ability to train on multiple images. These networks function with distinct architectures, making their integration challenging. The study by Ghara et al. (2022) utilized U-NET and DeepLabV3 to analyze SAR images, with U-NET achieving a higher detection accuracy of 78.8% compared to 54% by DeepLabV3.

Ma et al. (2022) developed a super pixel segmentation method designed to enhance the processing of SAR images. One of the key benefits of this approach is its adaptability during training, as the algorithm adjusts shapes based on segmentation outcomes and edge definitions to meet targeted results. Furthermore, the simplicity of the network structure contributes to computational efficiency, while still maintaining high performance and strong generalization capabilities. Ronciet al. (2020) introduced a new semantic segmentation method for detecting oil spills. Although the enhanced U-Net model surpassed the original, the standard U-Net trained through adversarial learning achieved superior performance, reaching a Jaccard Index of 82% and an Accuracy Index of 98.3%.

Yekeen et al. (2020) introduced a novel deep learning model for automated marine oil spill detection. The model leverages instance segmentation through a Mask-Region-based Convolutional Neural Network (Mask R-CNN), using ResNet 101 on COCO with a Feature Pyramid Network for enhanced feature extraction. Tested over 30 epochs with a learning rate of 0.001, the model significantly outperformed traditional machine learning and semantic segmentation models in detecting oil spills and similar-looking substances (look-alikes) in SAR images. The model achieved an impressive overall accuracy of 96.6% for oil spill detection and 91.0% for look-alike segmentation, with ship detection achieving the highest accuracy at 98.3%. This suggests that deep learning instance segmentation can provide more reliable and accurate results in environmental monitoring applications like oil spill detection in marine settings. The research by Basit et al. (2021) employs deep learning methodologies to address the critical environmental issue of oil spills in marine and coastal ecosystems. The model was trained using a comprehensive dataset of Sentinel-1 SAR images, which have been segmented into distinct

categories including sea surface, oil spills, look-alikes, ships, and land areas. It achieved IoU of 95.69% for sea surface, 60.85% for oil spills, 54.90% for look-alikes, 70.27% for ships, and 96.79% for land. Collectively, these results contribute to a mean IoU of 75.70% across all categories, marking a nearly 10% improvement over previous state-of-the-art techniques. This significant enhancement in detection accuracy demonstrates the potential of using UNet and similar deep learning models for rapid and reliable oil spill response, helping to mitigate one of the major causes of water pollution.

Shaban et al. (2021) proposed a two-stage deep-learning model that relies on extracting specific handcrafted features from SAR images, such as object standard deviation and background standard deviation, to overcome the challenges posed by unbalanced data and enhance detection accuracy. The model uses a novel CNN structure for patch generation, emphasizing balanced data patches to reduce background bias and improve segmentation outcomes. The results have shown high accuracy, sensitivity, specificity, and a weighted Kappa of almost 99%, with the deep-learning model outperforming existing state-of-the-art methods in precision and Dice scores. However, the framework mainly focuses on oil spill detection and is not suitable for multi-class problems like detecting ships or other objects. Zhu et al. (2022) introduce the CBD-Net, a deep learning framework designed to improve the detection of oil spills in SAR images. This network tackles common challenges such as uneven intensity, high noise, and blurred boundaries in oil spill images, which are often exacerbated by natural phenomena like waves and currents. This method utilizes a spatial and channel squeeze excitation (scSE) block to increase the internal consistency of detected oil spill regions, making the system particularly adept at managing smaller, less conspicuous targets that other models might overlook. The results were promising, with CBD-Net achieving the highest mIoU of 83.42% and an F1 score of 87.87%, significantly outperforming comparative models. These metrics indicate robust and accurate extraction of oil spill regions from complex SAR images, which is crucial for effective marine environmental monitoring.

Li et al. (2023) introduced the Dual-Stream U-Net (DS-UNet), a novel architecture designed for oil spill detection in SAR images. The DS-UNet also includes an edge extraction branch that specifically targets the reduction of speckle noise, which is a common challenge in SAR image analysis. The performance of the DS-UNet was rigorously tested against two real-world datasets, Palsar and Sentinel, where it demonstrated superior performance over several state-

of-the-art semantic segmentation methods. Quantitative results showed that the DS-UNet outperforms other models in terms of Dice Similarity Coefficient (DSC), F1 score, and Hausdorff Distance (HD) across both datasets. These metrics indicate the DS-UNet's ability to more accurately segment oil spills from SAR images, achieving finer contour detection and more detailed information extraction compared to other models like AttnUNet, R2U-Net, and NestedUNet. In the study conducted by Soh et al. (2024), a sophisticated approach to detecting marine oil spills using SAR images is presented. This method leverages an optimized encoder-decoder network model, specifically a refined U-Net architecture, to effectively identify oil spills with reduced computational demands. The model employs advanced techniques such as depth wise separable convolutions, group normalization, and bilinear interpolation-based up sampling to enhance performance while maintaining a smaller model size. The effectiveness of this approach is validated using two public SAR datasets, and the inclusion of polarimetric data has further improved detection accuracy, achieving an F1-score of 91.65% and an Intersection over Union (IoU) of 84.59%.

Wang et al. (2024) conducted a comprehensive study on marine oil spill detection using an improved polarimetric feature derived from polarization SAR images. Emphasizing the utilization of SAR which operates under all weather conditions due to its fine spatial resolution. Their research introduced an enhanced polarimetric feature based on the Cloude-Pottier target decomposition. This feature was tested within three neural network models: U-Net, FCN-8s, and DeepLabv3+ResNet-18, with U-Net achieving the highest accuracy and dice scores. This development indicates a significant advancement in the detection of oil spills, particularly in challenging conditions like low wind, rain cells, and young ice that typically affect the accuracy of detection methods. Their findings underline the potential of using sophisticated imaging techniques to enhance the precision and speed of oil spill detection, which is critical given the ongoing risks associated with marine oil transportation and extraction activities. In their impactful study by Shanmukh et al. (2024) ,exploring the use of deep learning techniques to enhance oil spill detection in marine environments through SAR imagery. The research is crucial for the rapid detection and precise segmentation of oil spills, crucial for initiating swift response measures and mitigating environmental damage. The team utilized advanced deep learning algorithms such as PSPNet, DeepLabV3, and Fully Convolutional Networks (FCN) integrated with U-Net to perform semantic segmentation of SAR images. These models were meticulously trained on a dataset

specifically labeled to indicate regions affected by oil spills, ensuring the algorithms accurately identify and delineate these areas. Notably, U-Net displayed superior IoU scores and achieved an impressive accuracy rate of 95%, highlighting its effectiveness in segmenting detailed and complex image data. This study demonstrates the significant capabilities of deep convolutional neural networks in analyzing SAR imagery for efficient and reliable oil spill detection, offering substantial benefits for environmental protection and disaster response in marine contexts.

Despite significant advancements, the existing studies on oil spill detection using various deep learning models face several limitations and challenges. One primary challenge is the data scarcity and variability in SAR imagery. The limited availability of high-quality, annotated datasets restricts the training of robust models. To address this issue, this research aims to create a binary segmentation dataset specifically for oil spill detection. This newly developed dataset will not only aid in improving the current study but also serve as a valuable resource for future research efforts. The models used in these studies have achieved good accuracy in detecting oil spills, but their oil spill class detection is not sufficient for achieving highly accurate detection. This study aims to address this by creating a specific binary segmentation to focus more on detecting oil spills, thereby reaching higher accuracy.

Semantic segmentation is chosen over object detection for several reasons. One primary reason is that semantic segmentation allows for the precise calculation of the oil spill area, which is crucial for creating accurate masks in the binary-specific dataset. These masks confirm the correct location of oil spills, ensuring that the segmentation is both accurate and reliable. Accurate area measurement is vital for assessing the extent and potential impact of the spill, enabling more effective response strategies. Additionally, semantic segmentation provides pixel-level classification, offering a detailed and comprehensive understanding of the oil spill's spread and boundaries. This level of granularity is essential for environmental monitoring and remediation efforts. Moreover, semantic segmentation models are well-suited for handling the complex and often irregular shapes of oil spills, which can be challenging for traditional object detection methods that rely on bounding boxes. This approach also facilitates better integration with other remote sensing data and geospatial analyses, enhancing the overall robustness and applicability of the findings.

2.2. Oil Fingerprinting for Marine Oil Spill Identification

Since petroleum products are the world's primary energy source, their production and consumption are increasing at a rate that is having an increasing influence on the environment (Sharma & Shrestha, 2023). Despite significant advancements in reducing leakages through enhanced technological and regulatory measures and improved industry practices, the risk of major oil spills remains. Daily, hundreds to thousands of oil spills occur worldwide, involving everything from various types of crude oil to a wide array of refined products. These range from heavy, long-lasting fuels to lighter, ephemeral, yet highly toxic substances within marine environments. The outcomes, behaviors, and effects of oil spills in these settings depend on the chemical composition and physical properties of the oil, as well as the weathering processes involved. Marine oil spills continue to be a major concern because of their substantial financial implications and the lasting, profound harm they cause to marine ecosystems, local economies, and coastal communities.

There are ten primary techniques for addressing marine oil spills: oil booms, skimmers, sorbents, in-situ burning, dispersants, hot water treatments, high-pressure washing, manual labor bioremediation, chemical stabilization with elastomers, and natural recovery (Chezhian. et al., 2024). Each technique offers unique benefits and limitations and is suited for specific situations. For instance, oil booms are effective when the oil is localized or in a stable marine setting, where skimmers can then be deployed to extract the oil from the water. When oil booms are insufficient for containing a spill, dispersants are utilized to expedite the breakdown of oil. These agents increase the surface area of oil particles, promoting chemical interactions with water, hindering the spread of the oil slick, and facilitating microbial degradation of the oil. In managing oil spill incidents, considerations extend beyond merely cleaning up. Identifying the origin of the oil, distinguishing between different types of oils, and tracking the degradation and weathering of oils in various environments are crucial actions. Additionally, determining legal responsibilities is vital for effective oil spill remediation. To fulfill these roles, oil fingerprinting analysis serves as a robust tool. This technique, essential in the field, employs geochemical methods to analyze the composition of hydrocarbons released into the environment. It takes into account various

transformative processes that hydrocarbons undergo, such as biodegradation, gas flushing, water washing, and evaporation (Wang et al., 2006a).

Moreover, factors such as temperature, reservoir compartmentalization, and aquifer activity significantly impact the transformations occurring within. Consequently, hydrocarbons derived from a single source rock exhibit unique characteristics across different reservoirs (Mulabagal et al., 2013).The identification of hydrocarbons from various reservoirs involves analyzing changes in their composition or pinpointing distinct hydrocarbon "fingerprints." Oil fingerprinting methods are routinely employed to pinpoint the origins of an oil spill by comparing the compositional features of the spilled oil with those of potential sources. Determining the sources and properties of these oils is essential for assessing the consequences and environmental impacts of the spill, crafting appropriate responses, and allocating responsibilities and legal liabilities (Song et al., 2016).

2.2.1. Traditional methods

Recently, several sophisticated tools for detecting biomarkers have become accessible. These include comprehensive two-dimensional GC, GC/MS, isotopic resolution mass spectrometry (IRMS), electrospray ionization liquid chromatography-mass spectrometry (ESI-LC-MS), and ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICRMS) (Cho et al., 2012; Mansuy et al., 1997). GC/MS has been successfully used as a conventional technique for oil forensic identification, especially isomers and those with similar retention times. This technique allows for an effective analysis of the effects of physiochemical weathering on biomarkers (Yang et al., 2011) .In their pivotal study, Ismail et al. (2016) highlighted the significance of GC/FID and GC/MS in oil spill fingerprinting. The study demonstrated how chemometric fingerprinting could significantly enhance the efficiency and reduce the costs and time required for identifying oil spill sources, providing crucial independent validation in the assessment of oil spill pollution in Malaysia.

In the study conducted by Wang (2022) chemometric techniques are leveraged to identify oil spills. The research underscores the utility of methods like GC/MS alongside advanced statistical tools such as hierarchical cluster analysis (HCA), principal component analysis (PCA). This application of chemometric methods not only provides a robust framework for identifying oil

spill sources but also significantly enhances the efficiency and accuracy of oil spill forensics, proving essential for environmental impact assessments and mitigation strategies. Mirnaghi et al. (2019) utilized fluorescence spectroscopy coupled with excitation-emission matrix-parallel factor analysis-principal component analysis (EEM-PARAFAC-PCA) to analyze environmental samples from four notable oil spill sites in Canada. This methodological approach highlights the dynamic nature of environmental PAH changes post-spill and underscores the complexity of accurately classifying oil spill impacts over time. Diagnostic ratios serve as essential tools in oil fingerprinting, facilitating the identification, characterization, and tracking of oil weathering. One key benefit of using diagnostic ratios between spilled oil and potential source oil is their ability to minimize the effects of concentration variances, thereby autonomously normalizing the data (Song et al., 2018). Typically, oil source fingerprinting is performed using GC/MS techniques, where specific diagnostic biomarker ratios are calculated. These ratios are derived from a variety of published ratios and some newly developed ones, following similar principles. There are eight principal biomarker classes, each characterized by distinct diagnostic ratios that aid in distinguishing oil sources and conditions (Song et al., 2019).

2.2.2. Data-driven models

Machine learning (ML), a key part of artificial intelligence (AI), emerged as its own distinct field in the 1990s and has since seen significant growth (Song et al., 2018). While there are various definitions of ML, essentially, it is a branch of AI that enables software to enhance its predictive accuracy by identifying patterns within the data, even without specific programming for such tasks. In the realm of ML, data is categorized into two types: labeled and unlabeled. Labeled data is characterized by having both input and the corresponding output parameters in a format that can be processed by machines. There are three principal types of machine learning algorithms: supervised learning, unsupervised learning, and reinforcement learning, as identified by Oladipupo (2010).This discussion will primarily focus on supervised and unsupervised learning and their applications in analyzing environmental data.

Initially, machine learning proves highly efficient in handling labeled classification challenges Furthermore, it is tailored for analyzing large datasets and complex relationships among variables and samples, independent of confirming initial assumptions such as normality and linearity of data. Lastly, the selection of predictive algorithms in machine learning is based solely

on their empirical performance, rather than any prior knowledge of the data domain (Mieth et al., 2016a). The use of ML in addressing complex environmental issues is expanding. This includes applications such as predicting concentrations of fecal coliform in wastewater treatment processes (Khatri et al., 2020), classifying extensive seafloor habitats using acoustic and visual data (Leon et al., 2020), and forecasting water quality parameters in coastal regions (Alizadeh et al., 2018). However, few efforts have been reported in employing ML in oil fingerprinting to classify dispersed oil. In line with this advancement, a recent study introduced a binary classification framework employing machine learning to distinguish between weathered crude oil (WCO) and chemically dispersed oil (CDO), utilizing six machine learning algorithms alongside a dimensional reduction technique as PCA to solve the dispersed oil classification problem (Y. Chen et al., 2021). This framework underscores the potential of machine learning to enhance oil spill source identification through comprehensive diagnostic ratios derived from various types of biomarkers. This study employed data preprocessing and six ML algorithms, namely random forest classifier (RFC), Support Vector Classifier (SVC), K-nearest neighbor (KNN), Linear regression (LR), Ensemble vote classifier (EVC), and Decision tree (DT), for a comparative study. Similarly, a study by Hashemi-Nasab & Parastar (2020)developed a chemometric strategy using gas chromatographic and infrared spectroscopic fingerprints to classify crude oils more accurately, emphasizing the importance of sophisticated data analysis methods in petroleum forensics. The results of unsupervised classification were then used as a starting point for partial least squares-discriminant analysis (PLSDA) and counter propagation-artificial neural network (CP-ANN).

In the field of environmental forensics, especially in pinpointing the origins of oil spills, there's a notable gap in using ML to improve oil fingerprinting accuracy. Traditional approaches are essential but often don't fully utilize chemical data, especially when it's complex. Our study introduces the use of six machine learning algorithms, detailed in **Table 2.1**, alongside classic oil fingerprinting methods to enhance the analysis of oil spills, such as the one from the MV Manolis L shipwreck. Table 2.1 provides a comparative overview of each algorithm’s strengths and limitations, showing how they contribute to our innovative approach. For instance, while KNN and SVC offer robustness to nonlinear data patterns, they require careful consideration of dataset noise and computational demands. Similarly, the Random Forest and Ensemble Vote Classifiers demonstrate strong resistance to overfitting, making them suitable for complex environmental data. This balanced view of modern data analysis tools, combined with traditional techniques, aims

to significantly enhance the accuracy of oil spill source identification, and reflects our commitment to advancing environmental forensics to meet contemporary environmental challenges.

One of the primary goals of this study is to assess the accuracy of machine learning models in classifying oil samples based on their chemical components. The case of the MV Manolis L shipwreck serves as a practical scenario to test the effectiveness of our machine learning approach. By integrating advanced data analysis techniques with conventional oil fingerprinting methods, we seek to accurately identify the source of the oil recovered from the shipwreck. This approach allows us to evaluate the performance of various machine learning algorithms in distinguishing between different types of oil based on their unique chemical fingerprints. Ultimately, our research contributes to refining response and management strategies for oil spill incidents, emphasizing the significant role of machine learning techniques in environmental science for addressing complex challenges and enhancing environmental protection efforts.

Table 2.1 Overview of different ML algorithms

Algorithm	Pros	Cons	References
K-nearest neighbor (KNN)	Easy to understand and implement. Can handle both linear and nonlinear data.	Sensitive to noise in the data. Can be computationally expensive for large datasets.	(Boateng et al., 2020a; Taunk et al., 2019)
Support vector classifier (SVC) (SVC)	Can handle both linear and nonlinear data. Robust to noise in the data.	Can be computationally expensive for large datasets. Requires careful tuning of hyperparameters.	(Nalepa & Kawulok, 2018)
Random forest classifier (RFC)	Robust to overfitting. Can handle both linear and non-linear data.	Can be computationally expensive for large datasets. Requires careful tuning of hyperparameters.	(Azar et al., 2014)
Decision tree classifier (DTC)	Easy to understand and interpret. Can handle both linear and non-linear data.	Prone to overfitting. Sensitive to noise in the data.	(Swain & Hauska, 1977)
Logistic regression classifier (LRC)	Easy to understand and implement. Can be computationally	Not suitable for non-linear data. Prone to underfitting.	(H. Wang & Hao, 2012)
Ensemble vote classifier (EVC)	Robust to overfitting. Can handle both linear and non-linear data.	Can be computationally expensive for large datasets. Requires careful selection of classifiers to combine.	(Mohammed & Kora, 2023)

2.3. Summary

Remote sensing technologies, particularly SAR and satellite imagery, offer extensive spatial coverage and the ability to operate in various environmental conditions, making them invaluable for initial oil spill detection (Conceição et al., 2021; Girard-Ardhuin et al., 2005; Wang et al., 2021; Yang et al., 2021). However, the challenge lies in accurately segmenting and classifying these detected spills to distinguish between oil and other similar-looking substances,

like algae or water films. Semantic segmentation models, such as U-NET, DeepLabv3, and their variations, have shown promise in this domain by segmenting images at the pixel level. These models can differentiate between oil and non-oil pixels, but there are still significant gaps in their performance, particularly in multi-class and binary segmentation tasks. Multi-class segmentation involves categorizing pixels into multiple classes (e.g., oil, water, land), while binary segmentation classifies pixels into just two categories (e.g., oil vs. non-oil). Evaluating the efficiency of these models in varying conditions and across different types of oil spills remains a crucial challenge.

Oil fingerprinting, which involves analyzing the chemical composition of spilled oil to identify its source, is another vital component of the detection process. Traditional methods, such as GC/MS, provide detailed chemical profiles of oil samples, but these techniques are often time-consuming and require significant expertise (Bayona et al., 2015b; Hashemi-Nasab & Parastar, 2020; Tian et al., 2021). Recent advancements in machine learning offer new opportunities to enhance oil fingerprinting by automating the analysis process and improving accuracy.

This study aims to address the identified gaps in the literature by developing and validating an integrated approach that combines the strengths of remote sensing, advanced deep learning models for semantic segmentation, and robust oil fingerprinting techniques. Specifically, the study focuses on integrating oil spill detection and oil fingerprinting, creating a novel binary dataset to enhance oil spill detection, and improving multi-class segmentation using the latest model structures and advanced encoders. By leveraging deep learning, the research enhances the accuracy and efficiency of oil spill detection and classification. It will implement and optimize state-of-the-art deep learning models for both multi-class and binary segmentation tasks, training these models on extensive datasets to improve their generalization capabilities and robustness across various environmental conditions. Utilizing SAR and satellite imagery as primary data sources, the study provides comprehensive coverage, improving the initial detection accuracy of oil spills. Additionally, machine learning algorithms will be applied to automate and refine the oil fingerprinting process, using advanced chemometric methods to quickly and accurately analyze the chemical composition of oil samples, facilitating the identification of spill sources. Extensive testing and validation of the integrated framework using real-world datasets will ensure its practical applicability and reliability in various marine environments and spill scenarios. By addressing these key areas, this study aims to provide a more reliable and efficient tool for

environmental monitoring and protection, enabling rapid and accurate detection, classification, and source identification of oil spills, thus significantly enhancing response efforts and mitigating the environmental impact of such incidents.

Chapter 3 Methodology

3.1. Framework

This chapter provides a detailed description of the methodology used for oil spill detection and oil fingerprinting. It outlines the datasets, data preparation processes, and the models and techniques applied to achieve accurate detection and classification of oil spills. The comprehensive methodology presented ensures a systematic approach, integrating multiple advanced techniques to address the various challenges encountered in environmental monitoring and protection. The methodology ensures a robust framework by integrating advanced imaging techniques, thorough data processing steps, and sophisticated machine learning models (Girard-Ardhuin et al., 2005). These elements work together to address the complexities and challenges inherent in environmental monitoring, particularly for detecting and analyzing oil spills. As shown in **Fig 3.1**, Initially, remote sensing technologies, specifically Synthetic Aperture Radar (SAR) images, are employed for their ability to penetrate cloud cover and provide reliable data under various weather conditions (Basit et al., 2022). These images are processed using state-of -the-art semantic segmentation models, each selected for its unique strengths in handling the intricate details and noise present in SAR data. The models, including U-Net, LinkNet, UNet++, FPN, DeepLabv3+, and PSPNet, are trained and validated on meticulously prepared datasets to ensure they can accurately identify and segment oil spills.

The datasets themselves are a cornerstone of this research, ensuring that the models are trained with high-quality, relevant data. A well-established dataset created by Krestenitis et al. (2019a) serves as the primary multi-class dataset, containing annotated SAR images that capture oil-polluted sea areas. These images are tagged with geographical coordinates and timestamps to ensure each instance of an oil spill is accurately verified and mapped. Complementing this is a binary dataset specifically created for this study, derived from oil spill incidents recorded in the Gulf of Suez from 2017 to 2021, which provides an additional layer of robustness to the model evaluation (El-Magd et al., 2023). The use of these two datasets allows for a comprehensive evaluation of the models' performance in both binary and multi-class semantic segmentation tasks.

Both datasets undergo extensive preprocessing, including steps such as location verification, image cropping and resizing, radiometric calibration, speckle noise reduction, and luminosity adjustment, to ensure high-quality input data for the segmentation models. The model that performs better in multi-class segmentation will be chosen, and the effect of different encoders on these models will be analyzed to understand their impact on model performance.

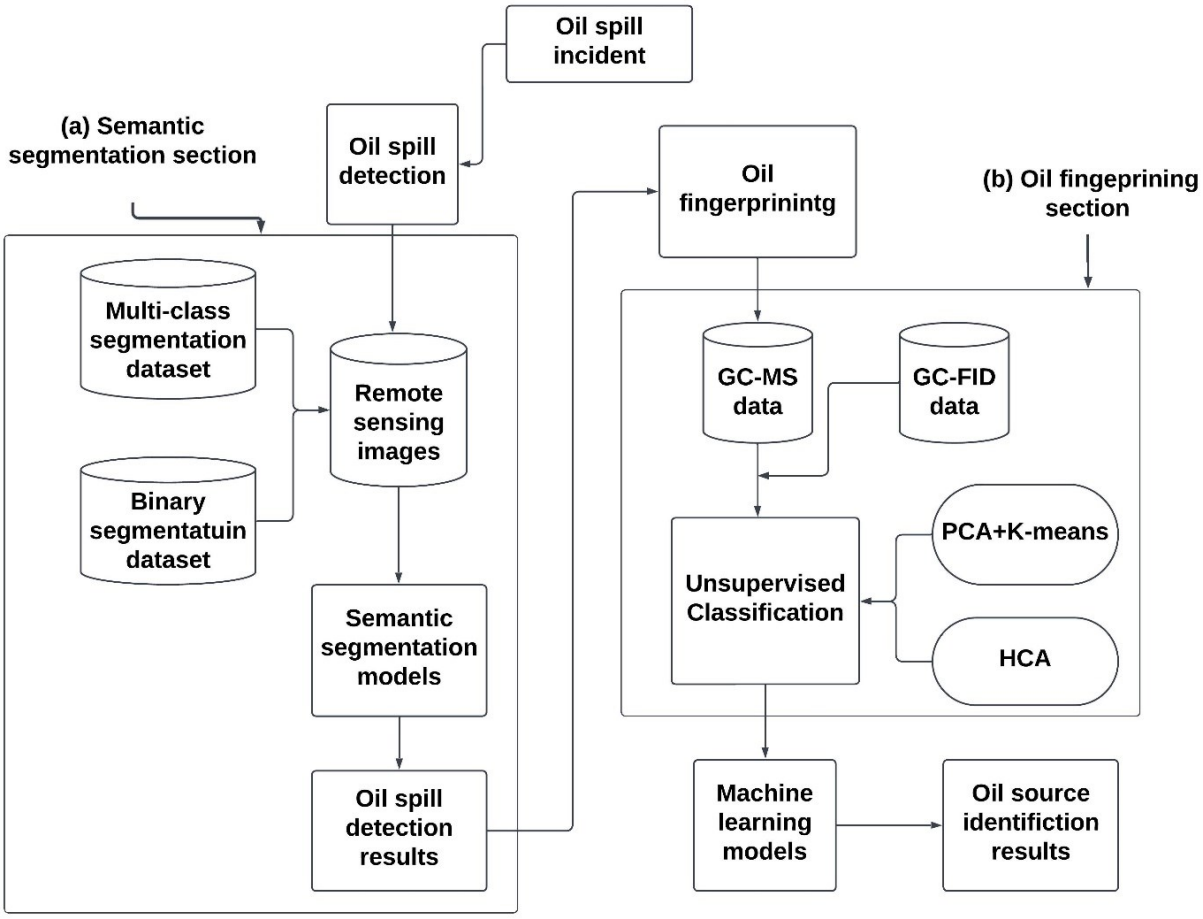


Fig. 3.1 Methodological Framework for Oil Spill Detection and Fingerprinting

Following the detection process, oil fingerprinting is conducted to analyze the chemical composition of the detected oil samples. This step is crucial for identifying the source of the oil spill, which in turn aids in effective response and remediation strategies. In this research, oil fingerprinting is based on a different case involving the sunken MV Manolis (Yang et al., 2020). The dataset for third case includes the chemical compositions of 17 analyzed oil samples collected from various portions of the MV Manolis over a four-year period (2013–2016), as well as three possible source samples and one weathered oil sample to assess the impact of weathering. Oil fingerprinting involves the collection and analysis of GC/MS and GC/FID (Gas Chromatography-Flame Ionization Detector) data. These analytical methods provide detailed insights into the chemical makeup of the oil samples, enabling the identification of specific hydrocarbons, biomarkers, and polycyclic aromatic hydrocarbons (PAHs). Unsupervised classification techniques such as Principal Component Analysis (PCA) and k-means clustering are applied to categorize the oil samples based on their chemical signatures (Murugan & Devi, 2019). These techniques help identify distinct patterns and group similar oil samples, which is crucial for accurate source identification. To further validate and refine these classifications, Hierarchical Cluster Analysis (HCA) is also utilized, ensuring a robust and reliable classification process.

To enhance the accuracy of oil spill source identification, various machine learning models are trained using the classified oil samples. Models such as k-Nearest Neighbors (KNN), Support Vector Classifier (SVC), Random Forest Classifier (RFC), Decision Tree Classifier (DTC), Logistic Regression Classifier (LRC), and Ensemble Vote Classifier (EVC) are employed. These models are optimized using GridSearch and cross-validation techniques to identify the best hyperparameters and prevent overfitting. The performance of these models is evaluated using cross-validation and the F-score metric, providing a comprehensive assessment of their accuracy, precision, recall, and overall balance. This rigorous evaluation ensures that the models are reliable and can generalize well to new, unseen data, making them suitable for real-world applications in oil spill detection and fingerprinting.

In summary, the overview presents the holistic approach taken in this research, starting from the detection of oil spills using advanced remote sensing techniques and semantic

segmentation models to the classification and analysis of oil samples through chemical fingerprinting and machine learning models. This integrated methodology is thoroughly explained in the subsequent sections of this chapter, detailing each step, process, and technique used to achieve accurate and efficient oil spill detection and analysis. By following this structured methodology, the research aims to provide a comprehensive framework for accurate oil spill detection and fingerprinting, contributing significantly to environmental protection and conservation efforts

3.2. Semantic Segmentation Models

Semantic segmentation is a fundamental task in computer vision, focusing on assigning a label to each pixel in an image, thereby enabling the precise localization and recognition of objects within a scene. It plays a crucial role in various applications, including environmental forensics, where it is used to detect and map oil spills from satellite images (Shelhamer et al., 2014). In the context of oil spill detection, semantic segmentation models are essential for accurately identifying and delineating oil-polluted areas in marine environments (Orfanidis et al., 2018). These models have demonstrated significant advancements in segmentation accuracy and efficiency, making them invaluable tools for environmental monitoring. By leveraging sophisticated techniques and architectures, these models help in precisely outlining the extents of oil spills, which is vital for effective response and remediation efforts.

The process involves training the models on datasets that include annotated satellite images, ensuring high precision in oil spill identification. The datasets undergo extensive preprocessing to improve the quality of the input data, involving steps such as location verification, image cropping and resizing, radiometric calibration, speckle noise reduction, and luminosity adjustment. This preprocessing ensures that the segmentation models receive high-quality data, enhancing their performance and accuracy. Through the application of these advanced segmentation models, environmental agencies and researchers can monitor and manage oil spills more effectively, ensuring that affected areas are promptly identified and addressed. In the following sections, we will delve into the specific architectures and methodologies employed by these models. Each model used in this study will be explained in detail, focusing on their unique structures and the specific aspects of their encoding and decoding processes that contribute to the

accurate and efficient detection of oil spills. The models discussed include U-Net, LinkNet, UNet++, Feature Pyramid Network (FPN), DeepLabv3+, and Pyramid Scene Parsing Network (PSPNet).

3.2.1. Unet

Originally designed for the segmentation of biomedical images, the U-Net architecture is versatile and extends beyond its initial application(Weng & Zhu, 2015). This architecture, part of the fully convolutional network (FCN) family, consists of two main segments: the contracting (encoder) path and the expansive (decoder) path. The design, known as U-Net due to its symmetric layout, involves the encoder capturing image content while the decoder focuses on precise localization through up sampling and reducing filter size, thereby creating a broader yet shallower feature representation (Shelhamer et al., 2014).

Based on the information illustrated in **Fig 3.2** , we can observe the structure of the U-Net architecture, which delineates the flow from the input image tile through the contracting path to the expansive path, culminating in the output segmentation map .The encoder employs alternating 3×3 convolutions and 2×2 max pooling with a stride of two, effectively down sampling the feature map while expanding the number of feature channels. In contrast, the decoder gradually restores spatial resolution by up sampling the feature map and applying a 2×2 convolution, reducing the number of channels. Each decoder step integrates up sampled maps with high-resolution features from the encoder, minimizing information loss. This process includes two subsequent 3×3 convolutions that reduce the channels further, concluding with a 1×1 convolution to assign a feature vector to each pixel, ultimately generating a segmentation mask for each class.

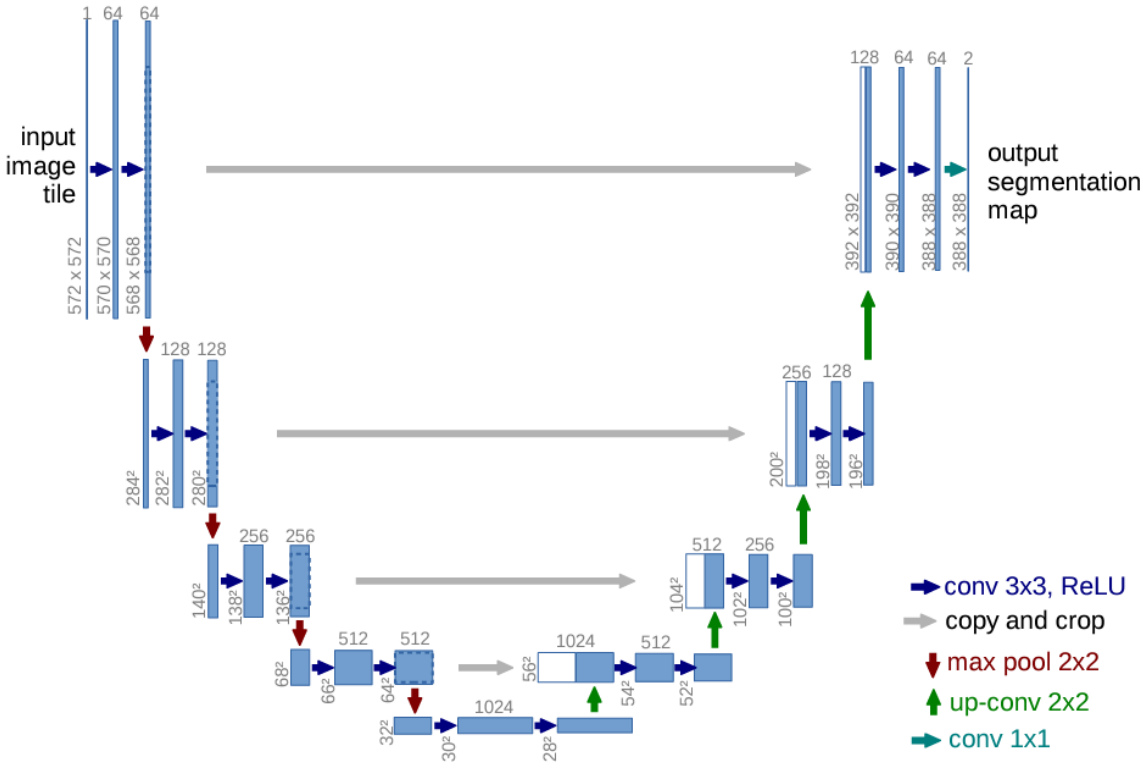


Fig. 3.2 U-net architecture (Weng & Zhu, 2015)

3.2.2. LinkNet

The primary aim of the LinkNet framework is to effectively restore spatial details that are typically reduced during the encoding phase by incorporating information from each corresponding encoder block into the decoding process(Chaurasia & Culurciello, 2017). This technique allows the decoder to use fewer parameters while reconstructing the image by leveraging the knowledge gained during the encoding stage.

The architecture utilizes ResNet-18(He et al., 2016) as its foundational structure, which begins the down sampling process through a 7×7 convolution filter with a stride of 2, immediately followed by a 3×3 max-pooling with the same stride. This encoder further contains four residual blocks that continue to decrease the size of the feature map. In parallel, the decoder includes four blocks each containing a 1×1 convolution, followed by a 3×3 convolution and an additional 1×1 convolution layer. These layers methodically decrease the number of filters, and the transposed convolutions in between enlarge the feature map's spatial dimensions using deconvolution filters

that are bilinearly initialized. The concluding segment of the decoder features a 3×3 transposed convolution, another 3×3 convolution layer, and a final 3×3 transposed convolution that reestablishes the input image's original spatial size (**Fig 3.3**).

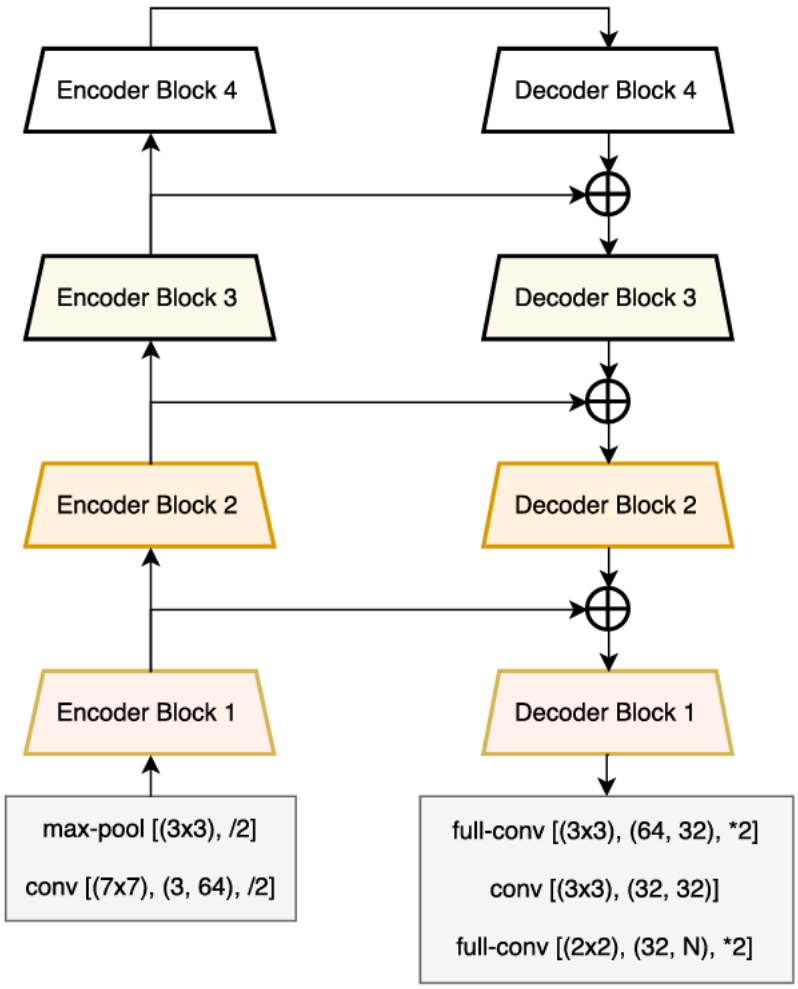


Fig. 3.3 LinkNet Structure (Chaurasia & Culurciello, 2017)

3.2.3. Unet++

UNet++ is an advanced segmentation model building upon the encoder-decoder framework like its predecessors (Zhou et al., 2018), U-Net and FCN, with the novel integration of nested and dense skip connections. These connections are crucial for combining detailed, low-level features from the encoder with the coarser, semantic-rich features from the decoder, thereby enhancing the

reconstruction of fine details in segmentation masks. This structure is particularly beneficial for medical imaging, where precision is paramount, as even slight errors can significantly impact clinical outcomes and diagnoses.

In UNet++, the architectural innovation lies in the re-designed skip pathways that densify the connections between the encoding and decoding processes. Instead of simple forwarding of features, as seen in U-Net, UNet++ strategically refines the encoder features through dense convolutional blocks before they merge with decoder features. This reduces the semantic gap, making the learning task more manageable and improving the model's ability to capture intricate details of medical images. Deep supervision in UNet++ further refines the training process by offering a dual-mode operation: an accurate mode averaging outputs from multiple segmentation branches, and a fast mode selecting from a single branch for quicker inference. The model demonstrates enhanced segmentation performance across various medical datasets, outperforming both U-Net and wide U-Net in intersection over union (IoU) metrics. Notably, while wide U-Net's improvements were partly due to an increased parameter count, UNet++'s gains are attributed to its architectural innovations. Overall, UNet++ marks a significant step forward in the precise and reliable segmentation of medical images, meeting the high accuracy demands of the field and potentially integrating into systems like Mask-RCNN to segment even occluded objects.

3.2.4. FPN

The Feature Pyramid Network (FPN) leverages a convolutional network's inherent hierarchical structure to enhance object detection frameworks, notably using Region Proposal Networks (RPN) and Fast R-CNN (Lin et al., 2017). FPN constructs a pyramid by integrating low-resolution, semantically strong layers with high-resolution, semantically weaker layers. This is achieved through a bottom-up pathway where traditional convolutions are performed, creating a series of feature maps at decreasing resolutions. These are primarily used for building the pyramid due to their detailed semantic information at various scales.

To enhance the resolution of the extracted features while maintaining their semantic strength, FPN introduces a top-down pathway with lateral connections. This pathway begins at the highest level of the pyramid and progressively increases the resolution of feature maps by a factor of two using nearest neighbor upsampling. These upsampled maps are then enhanced with lower-

level but higher-resolution maps from the bottom-up pathway through lateral connections. Each connection merges features of the same spatial size by element-wise addition after aligning the channels with 1×1 convolutions. This method ensures that each level of the output pyramid combines detailed spatial information with rich semantic information. The process continues until the finest resolution map is generated. This architecture allows the network to effectively detect objects at multiple scales and improve localization precision, crucial for tasks like bounding box proposal and object detection in various real-world scenarios.

3.2.5. DeepLabv3+

The recent iteration of the DeepLab series, DeepLabv3+(Chen et al., 2018), has advanced the original design by incorporating an adept yet straightforward decoder to enhance segmentation precision, particularly at object edges(Chen et al., 2017). It utilizes the framework established by DeepLabv3 and incorporates a tailored Aligned Xception model as its backbone, which has been adjusted for the task of semantic segmentation. In DeepLabv3+, the Xception model is expanded with additional layers to create a more profound network structure. Each max pooling step is replaced by depthwise separable convolutions that incorporate striding to maintain computational efficiency.

This form of convolution splits the process into two phases: depthwise convolution that operates on individual input channels, and pointwise convolution that uses a 1×1 kernel to integrate the channels' spatial data. When combined with atrous convolution, it becomes atrous separable convolution, which lessens the computational demand and yields feature maps with varying resolutions. Additionally, every 3×3 depth wise convolution in the Xception model is succeeded by batch normalization and ReLU activation. The encoder's final output then proceeds through an enhanced ASPP, which, aside from its four atrous convolutions, includes an extra branch for global context by applying average pooling to the feature map of the encoder backbone.

The decoder module in DeepLabv3+ integrates two sources of data: the encoded main branch's result, upscaled by a factor of four, and the low-level features from the encoder backbone, which are processed by an extra 1×1 convolution. This convolution harmonizes the contribution of the backbone's detailed features with the encoder's abstract semantic information. Following the

merging of both inputs, a series of two 3×3 convolutions refine the features, and the result is upscaled back to the original dimensions of the input image through bilinear upsampling **(Fig 3.4).**

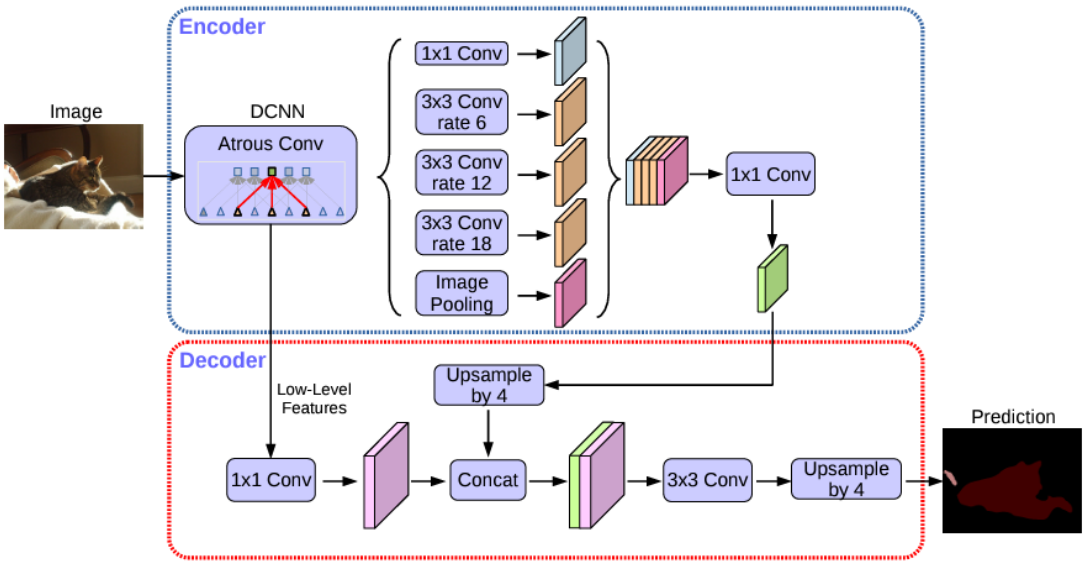


Fig. 3.4 DeepLabv3+ Structure (L. C. Chen et al., 2018)

3.2.6. PSPNet

The Pyramid Scene Parsing Network (PSPNet) was introduced to effectively capture the global context of scenes through a novel pyramid pooling module situated between the encoding and decoding stages (Zhao et al., 2016). Initially, an input image is processed by an encoder based on a ResNet architecture, which incorporates dilated convolutions to extract detailed feature maps. These maps are then directed to a pyramid pooling module consisting of four parallel branches, each employing pooling operations with progressively larger bin sizes—specifically, 1×1, 2×2, 3×3, and 6×6. As shown in Fig each branch outputs are subjected to a 1×1 convolutional layer to compress the feature maps to a fraction of their original dimensions, defined by the pyramid level. To restore these reduced maps to their original spatial dimensions, bilinear interpolation is utilized (**Fig 3.5**).

The enhanced feature maps are then merged with the initial feature maps from the encoder, creating a comprehensive feature representation. This combined map undergoes further interpolation and is processed through an additional convolutional layer to generate the final predictive segmentation map. To improve model training depth and effectiveness, an auxiliary loss

function is integrated earlier in the network, specifically after the fourth stage in the case of using a ResNet-101 architecture. This approach of deep supervision aids in training deeper network architectures more effectively. During the training phase, the loss from both the main and auxiliary functions is propagated backward through all previous layers, with a weighting system in place to ensure dominance of the main branch softmax loss in influencing the training outcomes.

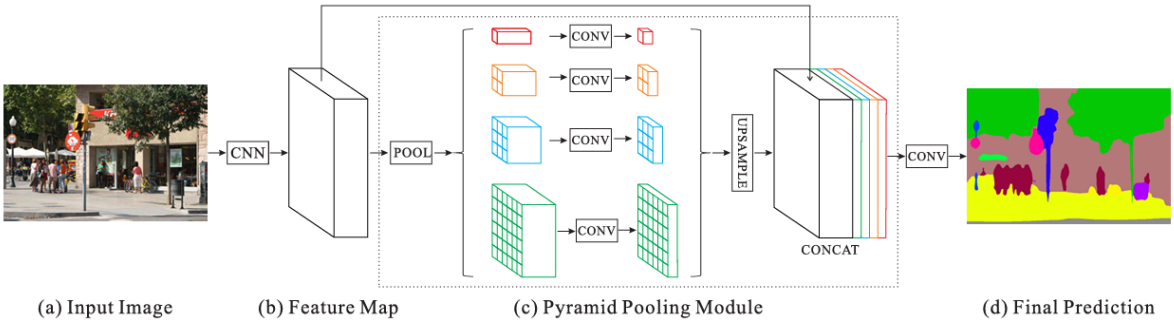


Fig. 3.5 Overview of PSPNet: Initially, (a) an input image is processed through a CNN to obtain the feature map from the final convolutional layer (b). This map is then processed by a pyramid parsing module that captures diverse regional details (c). This representation is then input into a convolutional layer to produce the per-pixel prediction (d) (Zhao et al., 2016)

The summarized **Table 3.1** highlights how various architectural techniques are tailored to enhance performance in different segmentation applications. Some models excel in straightforward tasks with simple yet effective designs, while others incorporate advanced features like residual blocks and dense skip connections to improve spatial detail restoration and complex feature merging. Multi-scale feature integration and pyramid pooling are employed to adeptly detect objects at varying scales. Techniques such as atrous and depthwise separable convolutions are integrated to capture multi-scale contextual information and refine object boundaries, making them suitable for high-precision tasks. Additionally, models that focus on capturing the global scene context excel in complex tasks requiring a broader understanding of the environment. The table provides a clear comparison of these capabilities and specialized features, offering a comprehensive overview of each model's performance and suitability for various segmentation tasks.

Table 3.1 Comparative analysis of semantic segmentation models

Model	Number of Layers	Key Components	Key Features
U-Net	23	Encoder-Decoder, Skip Connections	Simple and effective for biomedical segmentation (Weng & Zhu, 2015)
LinkNet	29	Encoder-Decoder, Residual Blocks	Efficient use of residual blocks (Chaurasia & Culurciello, 2017)
UNet++	Varies (approx. 38)	Nested U-Net architecture, Dense Skip Connections	Enhanced feature merging, deeper architecture (Zhou et al., 2018)
FPN	Varies	Pyramid pooling, Multi-scale Feature Integration	Multi-scale feature integration (Lin et al., 2017)
DeepLabv3+	Varies (50-101)	Atrous Convolutions, Encoder-Decoder, Depthwise Separable Convolutions	Effective multi-scale context capture (Chen et al., 2018)
PSPNet	Varies	Pyramid Pooling Module, Dilated Convolutions	Global scene context with pyramid pooling (Zhao et al., 2016)

3.2.7. Semantic segmentation performance evaluation

The performance of segmentation models is evaluated using the metric known as intersection-over-union (IoU). This metric calculates the overlap between the predicted area and the ground truth, providing an accurate measure of a model's ability to correctly identify and outline various classes such as sea surfaces and oil spills. It is computed by the formula:

$$\text{IoU} = \frac{\text{prediction} \cap \text{ground truth}}{\text{prediction} \cup \text{ground truth}} = \frac{\text{TP}}{\text{FP+TP+FN}} \quad (1)$$

The IoU is computed for each class in a dataset, and the average of these values results in the mean IoU (mIoU). MIoU serves as a crucial indicator of a model's overall precision across different environmental features.

Precision represents the proportion of predicted true positive samples and is obtained by dividing true positive pixels by the sum of true positive pixels and false positives:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

Recall represents the proportion of all true positive samples that are correctly predicted and is calculated as follows:

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

The F1 score is a metric that combines precision and recall, providing a single measure of a model's accuracy and is calculated as follows:

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

where TP, FP, and FN stand for the number of true positives, false positives, and false negatives, respectively. Using these metrics provides a comprehensive view of the oil spill detection method's performance, allowing for better improvement and optimization of the model. Precision, recall, and F1 score collectively offer insights into the model's capability to correctly identify oil spills while minimizing FP and FN.

3.3. Oil Fingerprinting Methods

After detecting an oil spill using remote sensing images and semantic segmentation models, environmental agencies can collect samples of the spilled oil to determine its source using oil fingerprinting techniques. This dual approach enhances the accuracy and reliability of oil spill source identification, combining advanced image analysis with detailed chemical characterization. Oil fingerprinting is a sophisticated technique used in environmental forensics to identify the source of oil spills (Wang & Wang, 2022).

One key method employed for identifying the source of spilled oil is oil fingerprinting. Oil fingerprinting relies on the recognition of specific groups of petroleum hydrocarbons, such as alkanes, biomarkers like terpenes and steranes, and PAHs (Wang et al., 2006b). Chemical composition of oils varies among different oil categories; therefore, their presence and distribution characteristics can serve as distinct signatures for identifying oil types (Wang et al., 2013). In recent years, oil fingerprinting has emerged as a valuable tool in source identification, as it allows for the detection of unique chemical signatures (fingerprints) within crude oil, enabling researchers to determine the source of an oil spill with a high degree of accuracy. Oil fingerprinting involves the use of analytical techniques to separate crude oil into its various chemical fractions, including aliphatic, aromatic, polar, and asphaltene compounds (Bayona et al., 2015b), which is a technique based on geochemical analysis of hydrocarbon fluids composition to provide valuable and unique

information for well and reservoir management(Pavlov & Vasiliev, 2017). Hydrocarbons in oil and gas deposits are affected by different processes, such as biodegradation, gas flushing, water washing, and evaporation. The degree of change depends on many factors such as temperature, reservoir compartmentalization, tectonics, aquifer activity, etc (Mohammadi et al., 2020).

Recent advancements in cluster analysis of crude oils have significantly contributed to the understanding of oil properties and behaviors based on their physicochemical characteristics. Sancho's work on cluster analysis using k-means based on these properties delineated groups of crude oils with similar characteristics, facilitating a more nuanced approach to oil categorization and refinery optimization (Sancho et al., 2022). Moreover, the use of unsupervised machine learning techniques such as principal component analysis and hierarchical clustering has also been explored to enhance the classification accuracy of crude oils. This is evident in the study by Jha et al., where such techniques were applied to analyze complex datasets from various crude oil sources, achieving a refined understanding of oil group dynamics which assists in better resource management and spill response strategies (Li et al., 2004). Zhan et al. (2019) and Liu et al. (2017) effectively used hierarchical clustering on geochemical biomarkers to distinguish crude oil samples by their geographic regions. Similarly, by integrating PCA, HCA, and feature selection techniques, Fernández-Varela et al. (2010) streamlined diagnostic ratios from GC/MS analyses, successfully reducing the number of essential features from 28 to 4. This effectively grouped and characterized the crude oils, highlighting the range of approaches and important outcomes in crude oil analysis and emphasizing the dynamic nature of environmental science and petrochemical analysis**.**

Recently, machine learning has emerged as a significant advancement over traditional statistics, providing robust tools for environmental analysis due to its numerous advantages (Saha et al., 2016). Machine learning excels in handling labeled classification challenges (Mieth et al., 2016b) and manages large datasets and complex variable relationships without relying on conventional assumptions such as normal distribution and linearity (Jordan & Mitchell, 2015). In line with this advancement, a recent study introduced a binary classification framework employing machine learning to distinguish between weathered crude oil (WCO) and chemically dispersed oil (CDO), utilizing six machine learning algorithms alongside a dimensional reduction technique (PCA) to solve the dispersed oil classification problem (Y. Chen et al., 2021). This framework

underscores the potential of machine learning to enhance oil spill source identification through comprehensive diagnostic ratios derived from various types of biomarkers. This study employed data preprocessing and six ML algorithms, namely RFC, SVC, KNN, LR, EVC, and DT, for a comparative study. Similarly, a study by Hashemi-Nasab & Parastar (2020)developed a chemometric strategy using gas chromatographic and infrared spectroscopic fingerprints to classify crude oils more accurately, emphasizing the importance of sophisticated data analysis methods in petroleum forensics. The results of unsupervised classification were then used as a starting point for partial least squares-discriminant analysis (PLSDA) and counter propagation-artificial neural network (CP-ANN).

In the field of environmental forensics, especially in pinpointing the origins of oil spills, there's a notable gap in using ML to improve oil fingerprinting accuracy. Traditional approaches are essential but often don't fully utilize chemical data, especially when it's complex. Our study introduces the use of six machine learning algorithms, detailed in Table 2.1, alongside classic oil fingerprinting methods to enhance the analysis of oil spills, such as the one from the MV Manolis L shipwreck. Table 2.1 provides a comparative overview of each algorithm’s strengths and limitations, showing how they contribute to our innovative approach. For instance, while KNN and SVC offer robustness to nonlinear data patterns, they require careful consideration of dataset noise and computational demands. Similarly, the Random Forest and Ensemble Vote Classifiers demonstrate strong resistance to overfitting, making them suitable for complex environmental data. This balanced view of modern data analysis tools, combined with traditional techniques, aims to significantly enhance the accuracy of oil spill source identification and reflects our commitment to advancing environmental forensics to meet contemporary environmental challenges.

The primary aim of this study is to assess the accuracy of machine learning models in classifying oil samples based on their chemical components. The case of the MV Manolis L shipwreck serves as a practical scenario to test the effectiveness of our machine learning approach. By integrating advanced data analysis techniques with conventional oil fingerprinting methods, we seek to accurately identify the source of the oil recovered from the shipwreck. This approach allows us to evaluate the performance of various machine learning algorithms in distinguishing between different types of oil based on their unique chemical fingerprints. Ultimately, our research contributes to refining response and management strategies for oil spill incidents, emphasizing the

significant role of machine learning techniques in environmental science for addressing complex challenges and enhancing environmental protection efforts.

3.3.1. Data entry and preprocessing

In this section, we detail the data preprocessing steps essential for clustering analysis, focusing on the application of PCA and k-means clustering, followed by HCA to enhance our chemometric approach.

Firstly, the initial data set derived from the MV Manolis L shipwreck contained missing values across four chemical compositions in the Excel datasheet. To manage this, variables exhibiting the highest frequency of missing data were removed to streamline the dataset for analysis. Subsequently, we performed data scaling on the remaining variables to normalize the data distribution. PCA was applied to simplify the complexity of the chemical data by focusing on the five most significant principal components for each chemical composition. These components were selected based on their ability to explain cumulatively less than 99% of the variance in the dataset. This criterion ensures that while most of the data's variability is captured, overfitting is avoided by not exceeding the 99% threshold. The PCA methodology, as outlined by Wetzel (2017), utilizes a formula (Eq. (1)) to extract these principal components, effectively reducing the dimensionality and retaining the most critical aspects of the data as:

$$PC_i = a_{1X_1} + a_{2X_2} + \ldots + a_{dX_d} \quad (5)$$

where, PC_i, principal component i; X_d, original feature d; a_d, numerical coefficient of X_d.

Following PCA, we applied the k-means clustering algorithm, a state-of-the-art, centroid-based method that partitions the data into k predefined clusters. This method aims to minimize the within-cluster sum-of-squares distances, also known as inertia (Lloyd, 1982). The algorithm's effectiveness depends on multiple runs to ensure the identification of the global solution, with the optimal number of clusters (k=3) determined through trial and error. To validate and refine the clustering results obtained from PCA and k-means, HCA was utilized. The dendrogram produced by HCA serves as a visual tool to verify the clustering results, providing a hierarchical structure that illustrates how closely samples are related to each other (Chanana et al., 2020). This method is especially useful in confirming the robustness of the initial clustering and in identifying any

nuanced patterns that might not be evident from k-means clustering alone. Upon verifying the clusters using HCA and PCA, the samples have been appropriately labeled. These labels now form a labeled dataset that is ready for the next phase, which involves supervised classification. In this subsequent part, we will employ six different machine learning algorithms to classify the samples based on the derived labels. This approach will enable us to further refine our analysis and enhance the precision of our oil spill source identification efforts.

3.3.2. Modelling development

In this study, the initial limitation of having only 21 original samples was encountered, leading to the utilization of a resampling approach for dataset expansion to enhance the robustness of the analysis. More specifically, resampling techniques, including Synthetic Minority Over-sampling Technique (SMOTE) (Sahlaoui et al., 2023) , were employed to significantly augment the dataset from its initial count of 21 samples to a robust total of 84 samples. The analytical approach was initiated by subjecting preprocessed GC/FID fingerprints to PCA. Hierarchical HCA with Ward's method was subsequently utilized as a supplementary validation technique for the PCA findings. The number of clusters derived from K-Means clustering was utilized to establish an initial y vector, representing class labels. This y vector played a pivotal role in the supervised classification process, wherein six machine learning models were employed, including KNN, SVC, RFC, DTC, LRC, and EVC. This holistic approach not only expanded the scope of the analysis but also effectively mitigated the limitations stemming from the initial shortage of samples.

In this research, a comprehensive comparative analysis utilized a set of six machine learning algorithms. To ensure robust evaluation, the dataset underwent preprocessing, followed by partitioning into training (80%) and test (20%) subsets. This approach aligns with established practices in the field (Chen et al., 2021; Medar et al., 2017). Below, concise descriptions of the six machine learning algorithms employed in this investigation are provided.

DTC basically constructs a decision tree to model the decision-making process, leading to a prediction or classification. Terminal nodes or leaves represent the outcome, based on traversing the tree through tests and decisions on the dataset's features (Swain & Hauska, 1977).

LRC estimates the probability that an example belongs to a particular class using a logistic function, making it ideal for binary classification tasks. The relationship between the features and the target is quantified using the logistic equation:

$$P_i = \frac{e^{a+bx}}{1+e^{a+bx}} \quad (6)$$

where, P_i, the probability of a label 1; e, the base of the natural logarithm; a and b are the model parameters (Robles-Velasco et al., 2021; H. Wang & Hao, 2012).

KNN is a non-parametric classifier that assigns a class based on the majority vote of the nearest neighbors. Distance between data points is key, commonly calculated using the Euclidean distance(Boateng et al., 2020a, 2020b):

$$d(x, x_i) = \sqrt{\sum_{k=1}^{n}(x_k - x_{ik})^2} \quad (7)$$

- Here, $d(x, x_i)$ represents the distance between the query point x and a data point x_i in the dataset.
- x_k and x_{ik} are the k-th feature of the query point and the data point, respectively.

SVC searches for the optimal hyperplane that maximizes the margin between different classes in the feature space, enhancing class separability (Nalepa & Kawulok, 2018).

RFC is an ensemble of decision trees, this classifier uses multiple trees to make a decision, taking the majority vote as the final prediction (Azar et al., 2014). The classification decision by the forest is an aggregate of decisions from individual trees:

$$f(x) = \sum_{m=1}^{M} C_m \prod(x, R_m) \quad (8)$$

- This is the equation for Random Forest, where the classification is an aggregation of decisions from multiple trees.

EVC combines the outputs of several models to make a final prediction, enhancing reliability and accuracy through the ensemble method (Mohammed & Kora, 2023). The final class is determined by the mode of the predictions:

$$y = \text{mode}[C1(x), C2(x), \ldots, Cm(x)] \quad (9)$$

This equation represents the mode (most common prediction) among different classifiers in EVC

3.3.3. Hyperparameter optimization and overfitting

Hyperparameter optimization plays a pivotal role in machine learning, especially in specialized fields like oil fingerprinting. Hyperparameters, which include settings such as the learning rate or the number of hidden layers in a neural network, are crucial in determining the learning process and the complexity of the model. These parameters significantly influence the performance of a machine learning model, as highlighted by Chen et al. (2021). Two main techniques in hyperparameter optimization are GridSearch and cross-validation (Badem et al., 2019). This ensures the identification of the optimal hyperparameters for a specific model. Cross-validation, on the other hand, is a technique for robustly assessing machine learning models by evaluating their performance across different data subsets, thus ensuring consistency in performance evaluation(Ranjan et al., 2019). To mitigate overfitting and maintain the model's predictive accuracy, techniques like cross-validation and regularization are employed. These strategies are essential in ensuring that the model generalizes well to new, unseen samples. Details the chosen parameters for each machine learning algorithm in GridSearch and cross-validation are shown in **Table 3.2**, based on previous studies of ML algorithms. This careful consideration in hyperparameter optimization is fundamental for tailoring machine learning models to specific applications like oil fingerprinting, ensuring their robustness and reliability in real-world scenarios.

Table 3.2 Hyperparameter tuning in different models

Algorithm Name	Parameters	Options
LRC	Solver Regularization (C)	Liblinear, 1, 10, 20, 30
RFC	n-estimators	50, 100, 150
DTC	Algorithm	CART (gini impurity), ID3 (entropy)
	Splitter	Best, Random
SVC	Kernel	Linear, RBF
	Regularization (C)	1, 10, 20
KNN	Gamma	Auto
	n-neighbors (K)	7, 10, 15

3.3.4. Performance evaluation

The evaluation of our model's performance is anchored in two key methodologies: cross-validation and the F-score metric, with an emphasis on shuffle split cross-validation implemented from the Scikit-learn library. Cross-validation is a rigorous technique where the data is divided into 'k' subsets (folds). The model is trained on 'k-1' folds and validated on the remaining fold, with this process repeating 'k' times. The results from each iteration are then averaged to derive a comprehensive performance measure, as detailed by Liu et al. (2019).

The F-score, also known as the F1-score, is a statistical measure used in the evaluation of machine learning models, particularly in the context of binary classification. It combines the precision and recall of the model into a single metric. Precision is the ratio of correctly predicted positive observations to the total predicted positives, while recall (sensitivity) measures the ratio of correctly predicted positive observations to all actual positives. The F-score is the harmonic mean of precision and recall, giving both metrics equal weight. The formula for the F-score, as previously detailed in Equation 4.

3.4. Summary

This chapter provides a detailed description of the methodology used for oil spill detection and oil fingerprinting, outlining the datasets, data preparation processes, and the models and techniques applied to achieve accurate detection and classification of oil spills. The comprehensive methodology presented ensures a systematic approach, integrating multiple advanced techniques

to address the various challenges encountered in environmental monitoring and protection. By integrating advanced imaging techniques, thorough data processing steps, and sophisticated machine learning models, the methodology ensures a robust framework capable of addressing the complexities and challenges inherent in environmental monitoring, particularly for detecting and analyzing oil spills (Girard-Ardhuin et al., 2005).

Initially, remote sensing technologies, specifically SAR images, are employed for their ability to penetrate cloud cover and provide reliable data under various weather conditions. These images are processed using state-of-the-art semantic segmentation models, each selected for its unique strengths in handling the intricate details and noise present in SAR data. The models, including U-Net, LinkNet, UNet++, FPN, DeepLabv3+, and PSPNet, are trained and validated on meticulously prepared datasets to ensure they can accurately identify and segment oil spills.

The datasets are a cornerstone of this research, ensuring that the models are trained on high-quality, relevant data. A well-established dataset created by Krestenitis et al. (2019a) serves as the primary multi-class dataset, containing annotated SAR images that capture oil-polluted sea areas. These images are tagged with geographical coordinates and timestamps to ensure each instance of an oil spill is accurately verified and mapped. Complementing this is a binary dataset specifically created for this study, derived from oil spill incidents recorded in the Gulf of Suez from 2017 to 2021, which provides an additional layer of robustness to the model evaluation (El-Magd et al., 2023). The use of these two datasets allows for a comprehensive evaluation of the models' performance in both binary and multi-class semantic segmentation tasks. Both datasets undergo extensive preprocessing, including steps such as location verification, image cropping and resizing, radiometric calibration, speckle noise reduction, and luminosity adjustment, to ensure high-quality input data for the segmentation models. The model that performs better in multi-class segmentation will be chosen, and the effect of different encoders on these models will be analyzed to understand their impact on model performance.

Following the detection process, oil fingerprinting is conducted to analyze the chemical composition of the detected oil samples. This step is crucial for identifying the source of the oil spill, which in turn aids in effective response and remediation strategies. In this research, oil fingerprinting is based on a different case involving the sunken MV Manolis (Yang et al., 2020). The dataset for this case includes the chemical compositions of 17 analyzed oil samples collected

from various portions of the MV Manolis over a four-year period (2013–2016), as well as three possible source samples and one weathered oil sample to assess the impact of weathering. Oil fingerprinting involves the collection and analysis of GC/MS and GC/FID data. These analytical methods provide detailed insights into the chemical makeup of the oil samples, enabling the identification of specific hydrocarbons, biomarkers, and PAHs. Unsupervised classification techniques such as PCA and k-means clustering are applied to categorize the oil samples based on their chemical signatures (Murugan & Devi, 2019). These techniques help identify distinct patterns and group similar oil samples, which is crucial for accurate source identification. To further validate and refine these classifications, HCA is also utilized, ensuring a robust and reliable classification process.

To sum up, the goal of this methodology is to define a comprehensive framework for handling oil spill incidents. This framework outlines how to detect oil spills using remote sensing and semantic segmentation trained models, followed by identifying the source of the oil through chemical fingerprinting. By demonstrating this integration, our research aims to provide a systematic approach that can be applied in real-world scenarios. This approach will help detect oil spills promptly and accurately, and identify their sources to halt further leakage, thereby preventing more damage to the ecosystem. This framework highlights the potential for combining advanced detection and identification methodologies to enhance the accuracy and reliability of oil spill response and environmental protection efforts.

Chapter 4 Case study & Results

4.1. Datasets

In this thesis, three set of datasets are utilized for the development and performance validation of semantic segmentation models aimed at detecting oil and identifying oil spill source. The datasets encompass diverse sources and methodologies, reflecting the comprehensive approach necessary to address the complexity of oil spill detection and oil fingerprinting. These datasets form the foundation for the subsequent analysis and model training, ensuring that the developed models are robust, accurate, and capable of performing in real-world scenarios.

The following sections will detail the datasets used for different aspects of oil spill research. First, the multi-class dataset developed by Krestenitis et al. (2019a)for detecting oil spills using satellite images will be explained. This dataset serves as a benchmark for future research in oil spill detection, providing a standardized framework for evaluating model performance. Next chapter will discuss the binary dataset created for training models based on binary segmentation for detecting oil, focusing on oil spill incidents in the Gulf of Suez. Last section will introduce the dataset used for oil fingerprinting for identifying the sources of marine oil spills and understanding their environmental impact. Each dataset's development, preparation, and significance will be explored in detail to highlight their contribution to advancing oil spill detection and analysis using machine learning techniques.

4.1.1. Dataset for oil Spill detection in multi-class segmentation (case 1)

For this thesis, a well-established dataset created by Krestenitis et al. (2019a) documented in their study on oil spill identification from satellite images using deep neural networks has been used. Recognizing the challenge of not having a standardized dataset for oil spill detection, Krestenitis et al. (2019a) aimed to provide a comprehensive dataset that could serve as a benchmark for future research in this area.

4.1.1.1. Dataset Description and Preparation

The dataset was meticulously assembled to include satellite Synthetic Aperture Radar (SAR) images capturing oil-polluted sea areas. These images were sourced from the European

Space Agency's Copernicus Open Access Hub, spanning from 28 September 2015 to 31 October 2017. Each image was tagged with geographical coordinates and timestamps provided by the European Maritime Safety Agency (EMSA) via the CleanSeaNet service, ensuring each instance of oil spill was verified and mapped accurately.

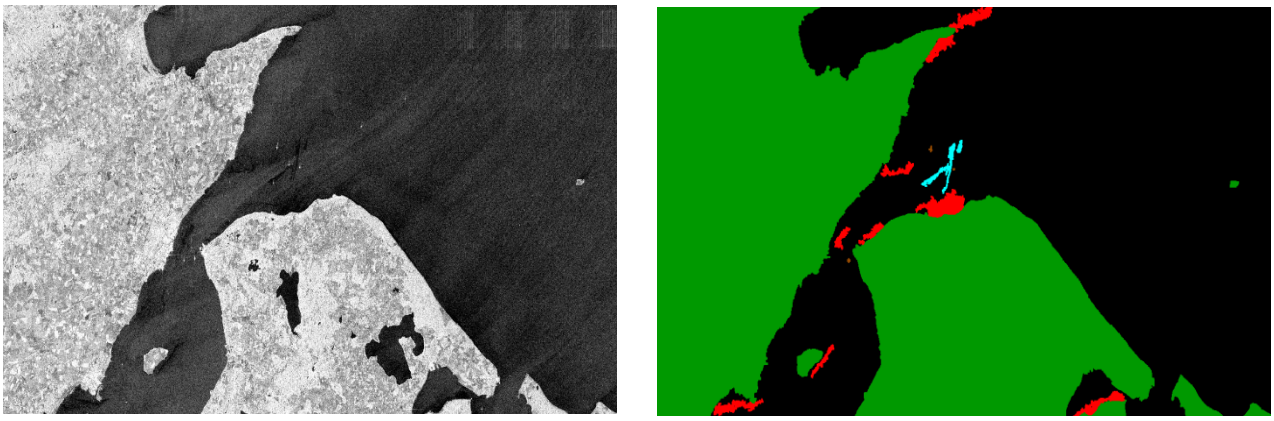
4.1.1.2. Technical specifications and processing

The SAR system used operated at C-band with a coverage range of about 250 km and a pixel spacing of 10 × 10 m. This setup was chosen for its ability to cover large maritime areas while capturing details necessary for accurate oil spill detection. Data processing included:

1. **Location Verification**: Matching SAR images with EMSA records to pinpoint oil spill locations.
2. **Image Cropping and Resizing**: Focusing on regions of interest and adjusting image resolution to 1250 × 650 pixels.
3. **Radiometric Calibration**: Ensuring consistent image quality and comparability.
4. **Speckle Noise Reduction**: Applying a 7 × 7 median filter to clean up the images.
5. **Luminosity Adjustment**: Transforming dB values to real luminosity for better visual interpretation

4.1.1.3. Dataset annotation and utilization

The dataset includes 1112 images classified into five categories: oil spills, look-alikes, ships, land, and sea surface, with the sea surface always serving as the background. To aid semantic segmentation, each class was assigned a distinct color in the RGB spectrum, and for model training, 1D label masks were provided. This dataset was split into training (90%) and testing (10%) subsets to facilitate model evaluation. The sample of SAR images and its RGB mask is shown in **Fig 4.1**. As shown if Fig 4.1 (b), Cyan color corresponds to oil spills, red to look-alikes, brown to ships, green to land and black is for sea surface.



(a) Synthetic aperture radar (SAR) image. (b) Ground truth mask

Fig. 4.1 Sample of SAR images after preprocessing (a) SAR image and (b) Corresponding annotated image. blue corresponds to 'oil spills,' red to 'look-alikes,' brown to 'ships,' green to 'land,' and black to the 'sea surface' class in the multi-class dataset

4.1.1.4. Contribution to research

This dataset not only standardizes the evaluation of oil spill detection models but also enriches the research community's resources, allowing for more consistent and comparative studies in the detection of marine oil spills using machine learning. The dataset is of very high quality and is publicly accessible, making it an excellent benchmark for future studies. Researchers can use this dataset to compare their results and improve their models. The inclusion of multiple classes, especially oil and look-alikes, presents a significant challenge for models to accurately detect oil spills. This complexity requires models to distinguish between oil spills and other objects that may have similar appearances in SAR images. By providing a robust dataset with these challenging classes, this research fosters the development of more sophisticated and accurate oil spill detection algorithms, ultimately advancing the field of environmental protection and monitoring.

4.1.2. Dataset for oil spill detection in binary segmentation (case 2)

In the development of semantic segmentation models for oil spill detection—utilizing U-Net and U-Net++ architectures and other models with various encoders—a critical component was

the construction of a robust binary dataset. The primary goal of using this binary dataset was to evaluate the performance of the models on binary segmentation tasks, specifically to differentiate between oil and non-oil regions. By focusing on binary classification, it was aimed to ascertain how the accuracy of the models would improve compared to multi-class segmentation tasks. This binary approach simplifies the model's task, potentially leading to higher accuracy and more reliable detection of oil spills (El-Magd et al., 2023).This dataset was derived using SAR imagery, specifically focusing on oil spill incidents recorded in the Gulf of Suez from 2017 to 2021. A case study was conducted to identify the precise locations and times of oil spills. This process allowed for the gathering of relevant SAR images that accurately captured these events. The detailed mapping of oil spill incidents provided a solid foundation for developing the binary dataset, ensuring that the models trained on this data could effectively learn and predict the presence of oil spills in maritime environments.

4.1.2.1. Case study

The Gulf of Suez is recognized as an area with a high potential for repeated oil spill events, primarily due to the extensive petroleum exploration and production activities. Furthermore, the Gulf is a vital maritime route for shipping and oil transportation. Situated between the African continent and the Sinai Peninsula, it connects the Red Sea to the Mediterranean Sea, forming a crucial link for international trade (Omar et al., 2021) . Its strategic positioning facilitates the movement of oil tankers, enabling efficient oil transport from the Middle East to global markets. The Gulf has a northwest-southeast orientation and spans a length of 320 km**.** Its width varies between 30 and 80 km, with water depth ranging from 40 to 60 m**.** The Gulf is known for its consistently warm sea surface temperatures, maintaining a range of 20 to 30°C year-round**.** Additionally, it experiences strong prevailing winds, typically fluctuating between 5 and 12 meters per second**.**

The study area encompasses the entire Gulf area and part of the Red Sea, extending from Suez at latitude 30°N southward to Quseer at 26°N and from longitude 32°20′E eastward to Sharm El-Sheikh at 34°30′E. This specific region was chosen as the study's focus due to extensive oil exploration, production, and shipping activities in cities like Suez, Ain Sokhna, Ras Ghareb, Hurghada, Safaga, and Quseer. **Fig 4.2** shows the study area, highlighting major coastal cities and the locations of oil spills from 2017 to 2021 (El-Magd et al., 2023). The red areas on the map

indicate regions affected by oil spills, providing a visual representation of the extent and distribution of oil pollution in the Gulf of Suez and adjacent Red Sea areas. The inset map offers additional context by showing the location of the study area within Egypt (**Fig 4.2**).

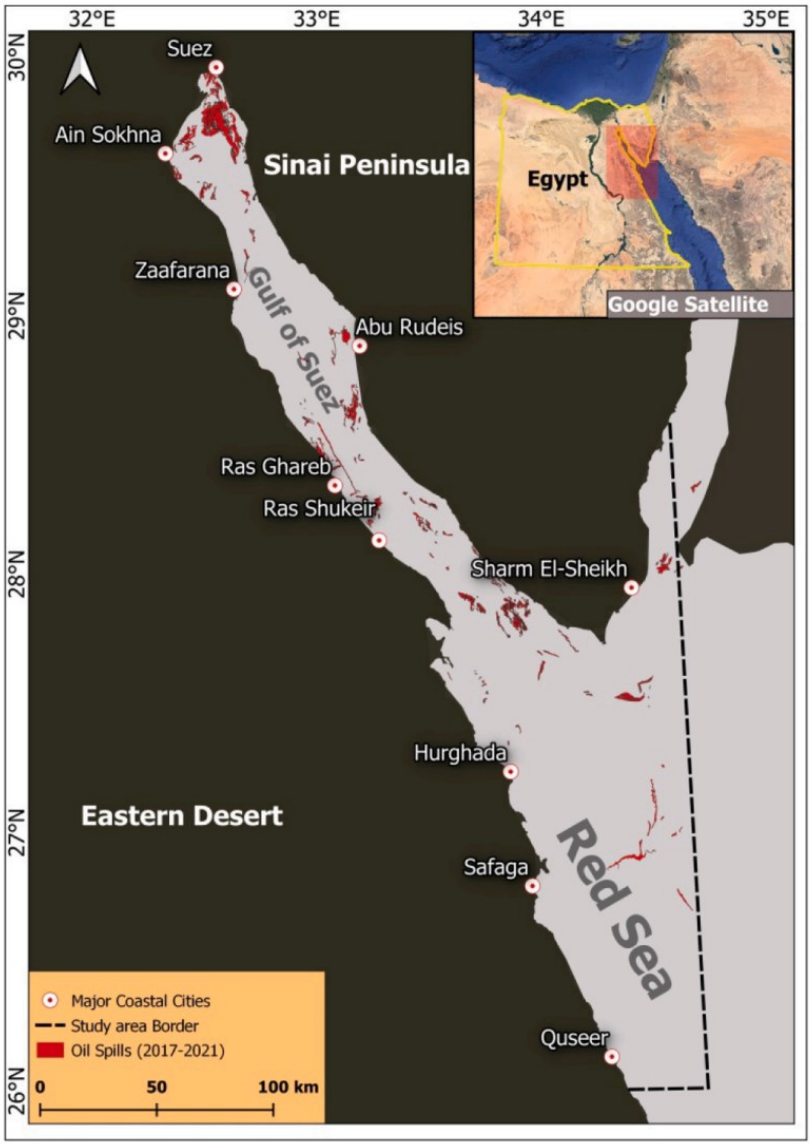


Fig. 4.2 Spatial distribution of oil spill incidents in the Gulf of Suez over five years (El-Magd et al., 2023).

4.1.2.2. Data preparation steps

The creation of the dataset involved several steps, executed using the ESA's SNAP (Sentinel Application Platform) software, to ensure the SAR data was appropriately processed for use in machine learning models:

1. **Downloading SAR imagery**: The initial step involved downloading Level 1 Ground Range Detected (GRD) products, which contain detailed radar backscatter information necessary for detecting oil spills.

2. **Image subsetting**: Utilizing GIS mapping data and the specific spill locations outlined by Zakzouk et al., subsets of the larger SAR images were created to focus on affected areas, enhancing the efficiency of the analysis.

3. **Orbit file application**: Accurate satellite positioning was ensured by applying precise orbit files to the SAR data, which corrects for any satellite orbital deviations at the time of image capture.

4. **Radiometric correction**: This step normalized the radar backscatter values across all images, ensuring consistency in the data input to the segmentation models.

5. **Speckle filtering**: A Lee-Sigma filter with a 7x7 window was applied to reduce speckle noise, a common problem in radar imagery, which can obscure or distort the appearance of oil spills.

6. **Decibel scaling**: Radar image pixel values were converted from linear to decibel scale to enhance the dynamic range and contrast of features within the image.

7. **Mask creation**: After converting the subsisted images to decibel scaling, the SAR oil spill detection tool within the SNAP software was used to generate binary masks indicating spill presence. This involved using GIS maps of oil spills based on the source (**Fig 4.3 (b))** Finally, the SNAP SAR oil spill detection tool was used to generate binary masks indicating spill presence, employing a trial-and-error method to adjust the threshold between 2 to 5 to match the oil spill record GIS maps available (**Fig 4.3 (c)).**

8. **Exporting data**: The final processed images and their corresponding masks were exported—images in JPEG format and masks in PNG format—for integration into the machine learning dataset.

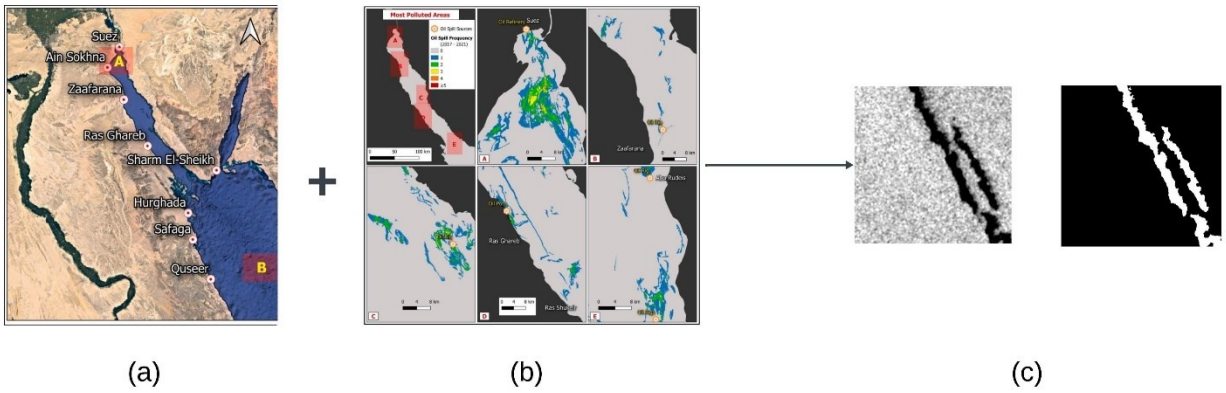


Fig. 4.3 (a) Case study area of the Suez Canal; (b) Oil spill mapping created by El-Magd et al. (2023); (c) Sample of dataset created for binary semantic segmentation.

This comprehensive dataset comprises a total of 202.14 million non-oil pixels and 5.99 million oil pixels. For training, 165.09 million non-oil pixels and 5.31 million oil pixels were used, while the test set included 37.06 million non-oil pixels and 0.69 million oil pixels. After processing the images, a total of 794 images, each 512 by 512 pixels, were created. Of these, 80% were used for training and 20% for testing.

4.1.3. Dataset for oil fingerprinting (case 3)

In cases of marine oil spills, accurately identifying the source of spilled oil and comprehending its behavior in the marine environment is of utmost importance. These factors play a crucial role in evaluating the environmental impact of such incidents and formulating an effective response strategy. In 1985, the MV Manolis L wrecked on Blow Hard Rock near Change Islands in Notre Dame Bay, Newfoundland and Labrador, sinking with significant amounts of heavy and diesel fuel oils. Initial oil recovery was hampered by harsh winter conditions, leading to no visible pollution reports for 28 years. In 2013, a severe storm exposed cracks in the ship's hull, leaking oil into the ocean. The Canadian Coast Guard responded by sealing the cracks and installing a cofferdam, which allowed for the oil's eventual recovery by 2018 (Yang et al., 2020).The dataset used in this article was the chemical compositions of total 17 samples of analyzed oil collected from different portions of the sunken MVManolis over a 4-year period (2013–2016). The chemical composition of the analyzed oil included total petroleum hydrocarbons (TPH), n-alkanes fromn-C9 to n-C40, petroleum-related biomarkers including terpenes, hopene’s, and steranes, as well as

non-alkylated polycyclic aromatic hydrocarbons (PAHs) and their alkylated congeners (APAHs) which was reported by Yang et al. (2020). Information about the samples is shown in **Table 4.1** which describe of total 21 sample including 17 MVManolis samples and 3 possible sources and one possible source sample that is weathered oil to check the effect of weathering on the samples.

Table 4.1 Sample Information of MVManolis (Yang et al., 2020)

Samples code in 3D PCA plot	ESTS code
a	2016-09-28-3510
b	2016-09-28-3511
c	2016-09-28-3515
d	2016-09-28-3518
e	2016-09-28-3520
f	2016-09-28-3522
g	2016-09-28-3524
h	2016-09-28-3525
i	2016-09-28-3529
j	2016-09-28-3530
k	2016-09-28-3532
l	2016-09-28-3534
m	2016-09-28-3535
n	2013/05/13-2244
o	2014/08/19-2542.4
p	2014/08/19-2542.3.1_10.12%
q	2015/06/12-2932
r	2015/10/29-3072
s	2015/04-17-2714 Fresh
t	2015/04/17-2714 W3
u	2015/09/23-3048

4.2. Oil Spill Detection Results

This section presents an overview of the results from evaluating the performance of various segmentation models used in this study. The evaluation focuses on assessing the accuracy and effectiveness of these models in detecting and segmenting oil spills using two key datasets: the multi-class dataset and the binary dataset. The performance metrics and qualitative analyses provide insights into how well each model performs, highlighting their strengths and weaknesses.

4.2.1. Accuracy evaluation

The performance comparison of different semantic segmentation models in the multi-class dataset reveals distinct strengths and weaknesses tied to their architectural designs shown in **Table 4.2**. DeeplabV3+ demonstrates the highest mIoU of 68.29%, excelling particularly in complex categories like oil spills and look-alikes with IoUs of 64.63% and 44.59%, respectively. This highlights DeeplabV3+'s ability to handle intricate textures and distinguish between similar patterns, which is crucial for environmental monitoring tasks. Additionally, DeeplabV3+ achieves an F1 score of 73.78%, a recall of 76.69%, and a precision of 81.34%, reflecting its overall robustness.

Unet++ shines in the sea surface category with an IoU of 96.37%, benefiting from its sophisticated architecture that enhances information flow and gradient preservation, making it highly effective in consistently textured and expansive areas. Unet++ also posts a mean IoU of 66.28%, an F1 score of 71.54%, a recall of 75.85%, and a precision of 81.40%, showing balanced performance across various metrics. The Feature Pyramid Network (FPN), with its multi-scale feature-building capability, stands out in the segmentation of ships and land, registering IoUs of 40.99% and 88.44%, respectively. FPN’s design enables effective detection of objects across different scales, making it ideal for varied environmental topographies. However, its performance dips in the segmentation of oil spills and look-alikes, recording lower scores of 61.38% and 41.16%, respectively, due to its less effective capture of subtle textural nuances required in these complex categories. FPN achieves a mean IoU of 65.64%, an F1 score of 72.25%, a recall of 77.56%, and a precision of 79.25%.

On the lower end, LinkNet, despite its design for speed and efficiency with less computationally demanding components, posts a mean IoU of 64.99%. Its architecture, while fast, does not provide the necessary depth or broad receptive fields needed for the detailed feature capture required in high-precision tasks, resulting in reduced efficacy in distinguishing between closely similar categories like oil spills and look-alikes. LinkNet's performance metrics include an F1 score of 71.01%, a recall of 77.56%, and a precision of 74.80%. Unet shows strong performance in sea surface segmentation with an IoU of 95.66%, but it struggles more with look-alike detection, achieving an IoU of 43.68%. The overall mean IoU for Unet is 64.21%, with an F1 score of 70.36%, a recall of 76.62%, and a precision of 78.59%. PSPNet performs moderately with a mean IoU of 60.70%, showing balanced capabilities across various classes but not leading in any specific area. It achieves an F1 score of 67.32%, a recall of 71.64%, and a precision of 79.20%.

Table 4.2 Results of models on test set for multi-class dataset

Model Name	**Sea Surface IoU**	**Oil Spill IoU**	**Look-alike IoU**	**Ship IoU**	**Land IoU**	**Mean IoU**	**F1 Score**	**Recall**	**Precision**
Unet	0.9566	0.6253	0.4368	0.3071	0.8849	0.6421	0.7036	0.7662	0.7859
LinkNet	0.9633	0.6044	0.4421	0.3897	0.8500	0.6499	0.7101	0.7756	0.7480
Unet++	0.9637	0.6320	0.4375	0.3961	0.8848	0.6628	0.7154	0.7585	0.8140
FPN	0.9624	0.6138	0.4116	0.4099	0.8844	0.6564	0.7225	0.7756	0.7925
Deeplab V3+	0.9625	0.6463	0.4459	0.4329	0.9270	0.6829	0.7378	0.7669	0.8134
PSPNet	0.9532	0.5220	0.3979	0.3047	0.8573	0.6070	0.6732	0.7164	0.7920

Fig 4.4 depicts the training performance of various models across different classes for the multi-class dataset, providing insights into their learning behaviors and effectiveness over time. All models show an improvement in their IoU scores as training progresses, indicating effective

learning. However, classes such as oil spill and look-alike exhibit more variability and slower convergence, highlighting the difficulty in distinguishing these closely related categories. In contrast, the sea surface and land classes consistently achieve high IoU values early in the training process and maintain these high levels throughout the epochs, suggesting these classes are well-represented in the dataset. This is consistent with the data showing 797.7 million sea surface pixels and 45.7 million land pixels in the dataset. The ship class also shows variability, with models like FPN and DeeplabV3+ performing better due to their ability to capture multi-scale features effectively**.** However, even these models do not achieve the high IoU values of sea surface and land, indicating fewer instances or more variability within the ship class, which has only 0.3 million pixels**.**

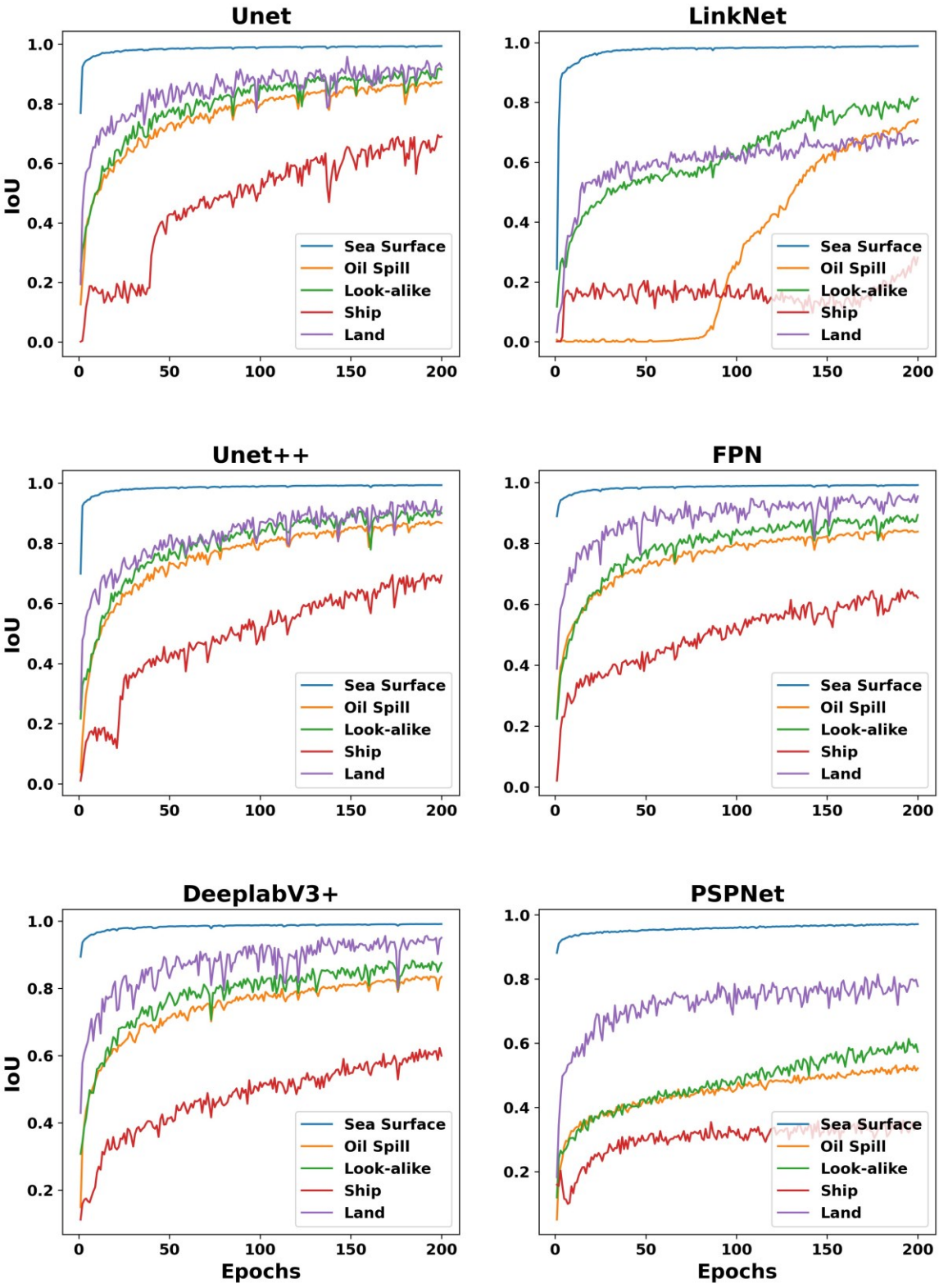


Fig. 4.4 Comparison of different architectures in terms of IoU measured in Case 1 for each class

Based on the evaluation of different semantic segmentation models on the binary dataset, which involves distinguishing between oil spills and the background, distinct strengths and weaknesses tied to their architectural designs are evident. The performance comparison is summarized in **Table 4.3**.

Unet++ demonstrates robust performance with a mean Intersection over Union (mIoU) of 87.47%, particularly excelling in detecting oil spills with an IoU of 75.98%. This underscores Unet++'s sophisticated architecture, enhancing information flow and gradient preservation, making it highly effective for this application. Additionally, Unet++ achieves an F1 score of 89.30%, recall of 92.03%, and precision of 90.36%. LinkNet shows the highest overall performance with an mIoU of 89.88%, excelling in both background (IoU of 99.41%) and oil spill detection (IoU of 80.36%). Its efficient architecture allows for high-precision tasks, as evidenced by an F1 score of 91.10%, recall of 91.54%, and precision of 92.18%. PSPNet and Unet both exhibit strong performance with mIoUs of 85.17% and 85.12%, respectively. PSPNet's strength lies in its balanced performance, achieving an F1 score of 87.74%, recall of 88.72%, and precision of 90.05%. Similarly, Unet achieves an F1 score of 87.41%, recall of 96.12%, and precision of 84.67%.

DeeplabV3+ and FPN show slightly lower performance in oil spill detection, with IoUs of 68.71% and 65.24%, respectively. DeeplabV3+'s performance indicates its proficiency in dealing with intricate textures and fine details, crucial for precise oil spill detection, resulting in an mIoU of 83.93%, an F1 score of 87.00%, recall of 86.11%, and precision of 91.04%. FPN, while effective for varied environmental topographies, struggles with detailed textural nuances required for oil spill detection, resulting in a lower mIoU of 81.91%, an F1 score of 84.71%, recall of 91.11%, and precision of 85.46%. The models' mIoU scores are significantly better in the binary dataset compared to the multi-class dataset. This is due to the binary classification involving only two classes: oil spill and background.

Table 4.3 Results of models on test set for binary dataset

Model Name	Background IoU	Oil Spill IoU	Mean IoU	F1 Score	Recall	Precision	Accuracy
Unet	0.9845	0.7180	0.8512	0.8741	0.9612	0.8467	0.9848
LinkNet	0.9941	0.8036	0.8988	0.9110	0.9154	0.9218	0.9942
Unet++	0.9896	0.7598	0.8747	0.8930	0.9203	0.9036	0.9897
FPN	0.9857	0.6524	0.8191	0.8471	0.9111	0.8546	0.9859
DeeplabV3+	0.9915	0.6871	0.8393	0.8700	0.8611	0.9104	0.9917
PSPNet	0.9919	0.7115	0.8517	0.8774	0.8872	0.9005	0.9921

The plots depict the training performance of various models on the binary dataset, which involves binary segmentation of oil spills and the background. All models show improvement in their Intersection over Union (IoU) scores as training progresses, indicating effective learning. The background class consistently achieves high IoU values early in the training process and maintains these high levels throughout the epochs (Fig 4.5). This suggests that the background class is well-represented in the dataset and is easier for the models to learn and segment accurately. In contrast, the oil spill class exhibits more variability and slower convergence across all models, highlighting the challenge in distinguishing oil spills from the background.

LinkNet demonstrates the highest IoU for the oil spill class among the models, achieving an IoU of 80.4%. This indicates LinkNet's proficiency in capturing detailed features necessary for oil spill detection, despite its lighter architecture. Similarly, DeeplabV3+ shows steady and high performance across the oil spill class, aligning with its test set results where it performs exceptionally well in detecting oil spills. This consistency indicates that DeeplabV3+'s architecture effectively learns complex features without overfitting.

.

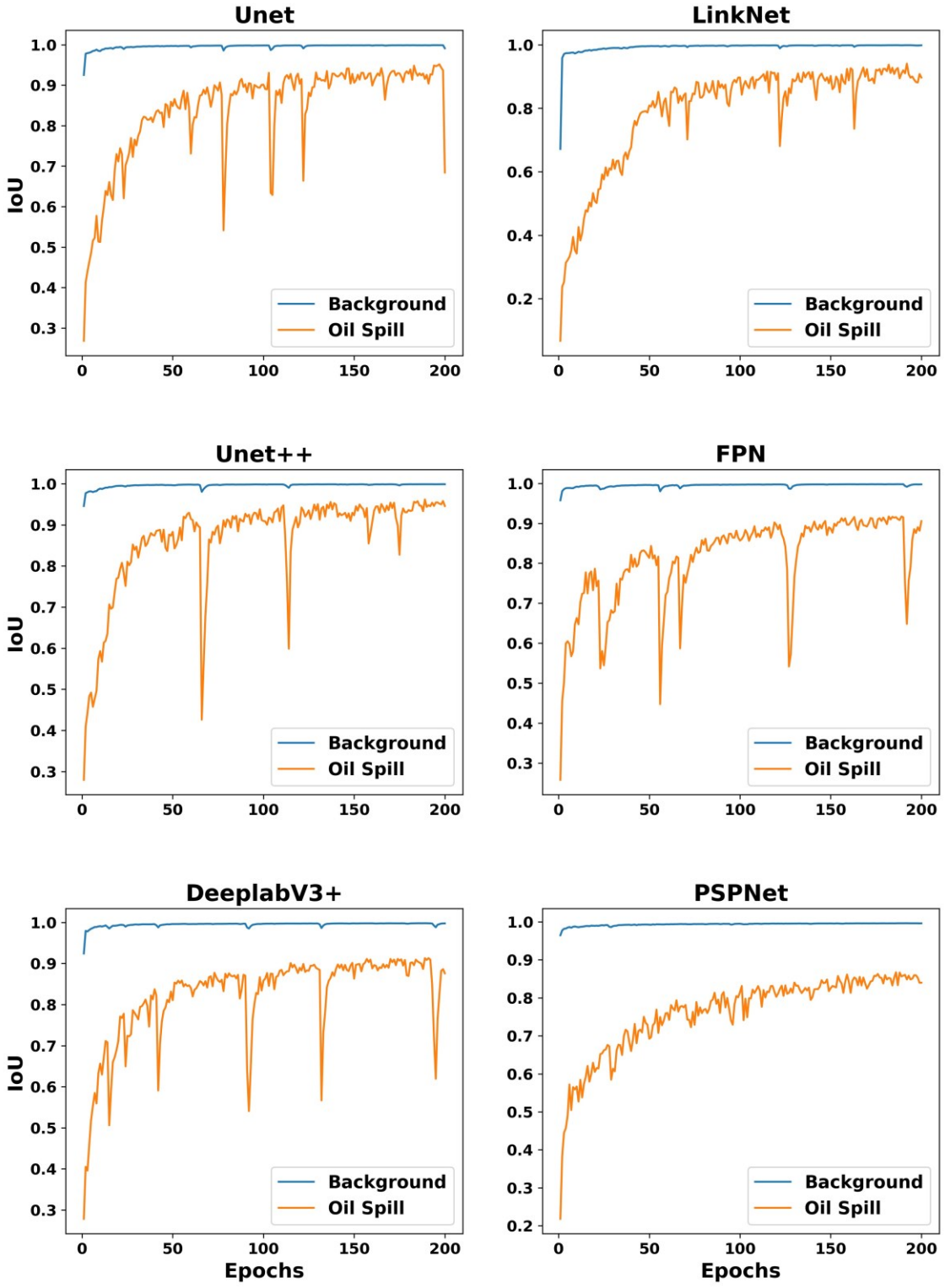


Fig. 4.5 Comparison of different architectures in terms of IoU measured in Case 2 for each class, over a range of training.

The mean Intersection over Union (mIoU) plots illustrate the learning behaviors and effectiveness of various models across two datasets: a multi-class segmentation problem (case 1) and a binary segmentation task (case 2) in **Fig 4.6**.

In the multi-class dataset (case 1), all models show a steady increase in mIoU over the epochs. DeeplabV3+ and Unet++ achieve higher mIoU values, indicating their superior handling of multiple classes. FPN also performs well, while LinkNet and PSPNet struggle more with the complex multi-class task. The variability in mIoU is more pronounced in the multi-class dataset, highlighting the challenges of differentiating between more complex categories.

In contrast, the binary dataset (case 2) shows higher mIoU values across all models due to the simpler classification task. DeeplabV3+, Unet++, and PSPNet achieve mIoU values close to or above 90%, showcasing strong performance and quick convergence. LinkNet, though improved, still lags behind the top performers, indicating its relative inefficiency in capturing detailed features even in a binary setting.

To summarize, the performance of different architectures on the multi-class dataset reveals their various strengths and weaknesses in multi-class segmentation tasks, with DeeplabV3+ and Unet++ standing out as the top performers. The binary dataset, on the other hand, shows higher and more consistent mIoU values across models due to the simpler task, with models like DeeplabV3+, Unet++, and PSPNet leading in accuracy. These insights underscore the importance of model selection based on the specific segmentation task, with more complex models excelling in multi-class scenarios and maintaining high performance in binary tasks.

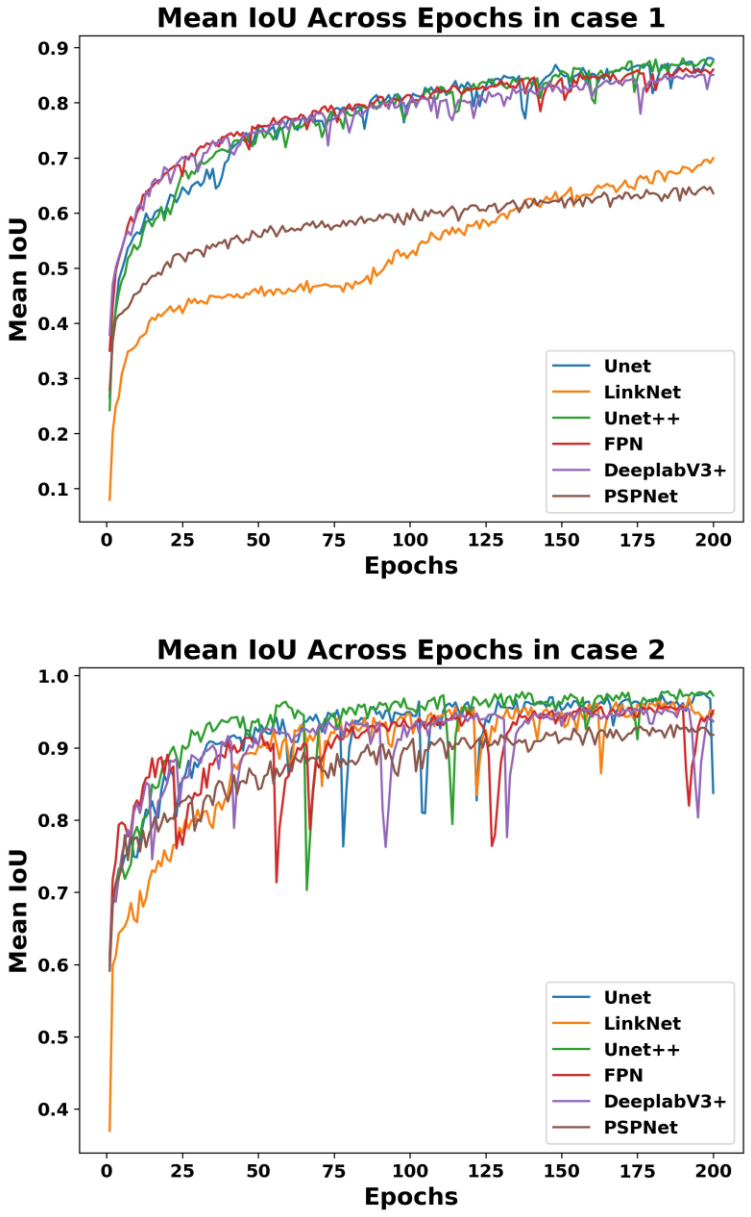


Fig. 4.6 Mean IOU during training in multi-class dataset (case 1) and Binary dataset (case 2)

4.2.2. Qualitative results

In the provided Figs, visual results from the evaluated segmentation models for multi-class dataset (**Fig 4.7**) and Binary dataset (**Fig 4.8**) are compared. The Unet model shows reasonable segmentation performance but often includes fragmented and noisy predictions, struggling to delineate the spill boundaries accurately. LinkNet's predictions are relatively cleaner with fewer

false positives, but it still faces challenges in capturing fine details. Unet++ demonstrates improved performance with better edge detection and fewer false positives, although it occasionally misses smaller spill areas.

The Feature Pyramid Network (FPN) performs well in capturing multi-scale features, resulting in consistent segmentation of larger spill areas but struggling with finer details and smaller regions. DeeplabV3+ stands out with robust performance, accurately segmenting both large and small spill areas with well-defined boundaries, effectively handling complex textures. PSPNet's performance is similar to DeeplabV3+, providing precise segmentation with clear boundaries, capturing varying scales and complexities in the imagery.

Comparing the models' qualitative results, DeeplabV3+ and PSPNet exhibit the best performance, providing the most precise and comprehensive segmentation. Unet and LinkNet, while effective in some instances, show limitations in handling finer details and reducing false positives. The visual results align with the IoU values, indicating that models like DeeplabV3+ and PSPNet excel in identifying oil spills with higher accuracy and fewer errors, making them suitable for detailed environmental monitoring tasks.

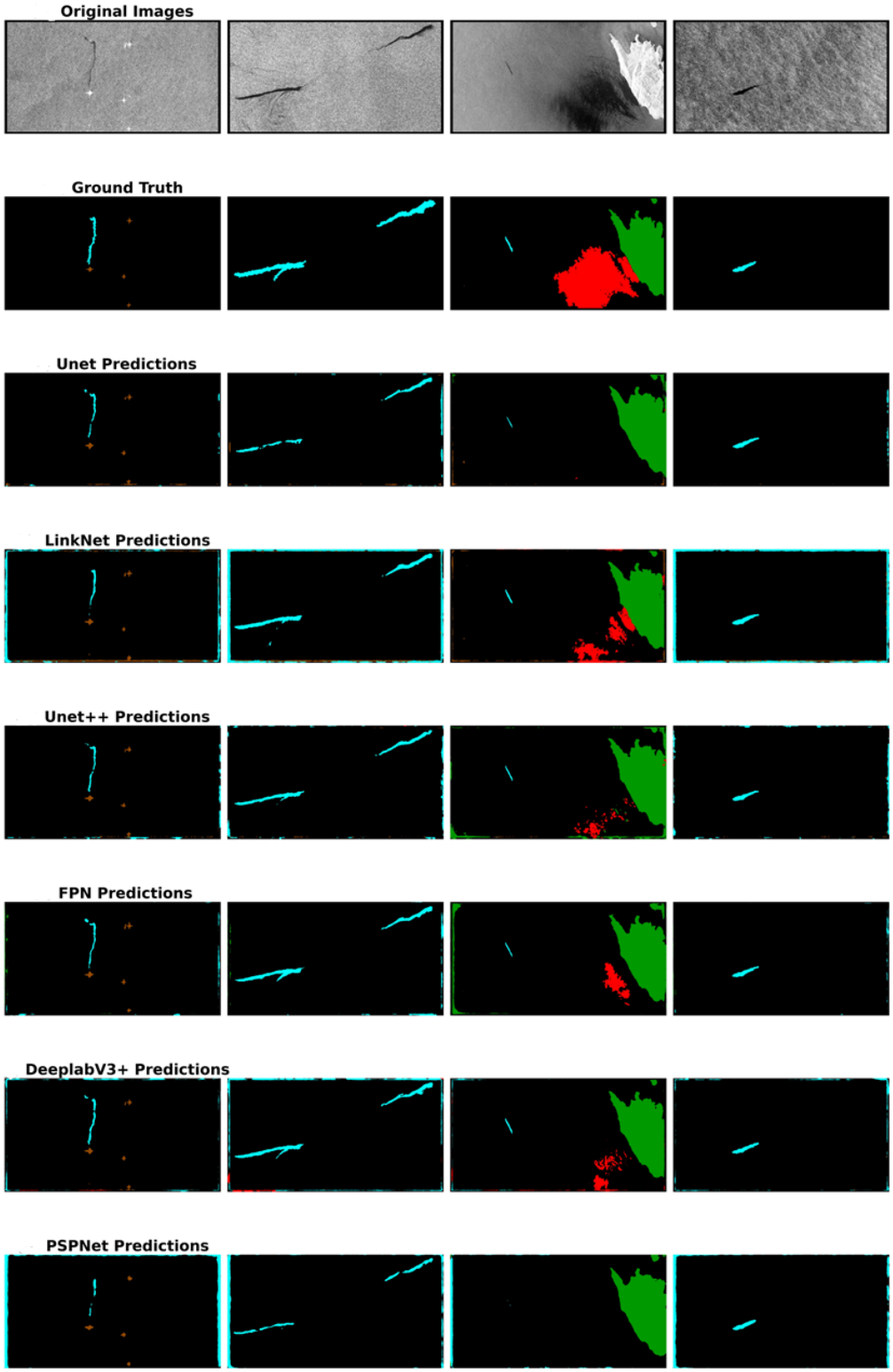


Fig. 4.7 Qualitative results of the examined segmentation models on the presented oil spill multi-class dataset. blue corresponds to 'oil spills,' red to 'look-alikes,' brown to 'ships,' green to 'land,' and black to the 'sea surface' class in the multi-class dataset

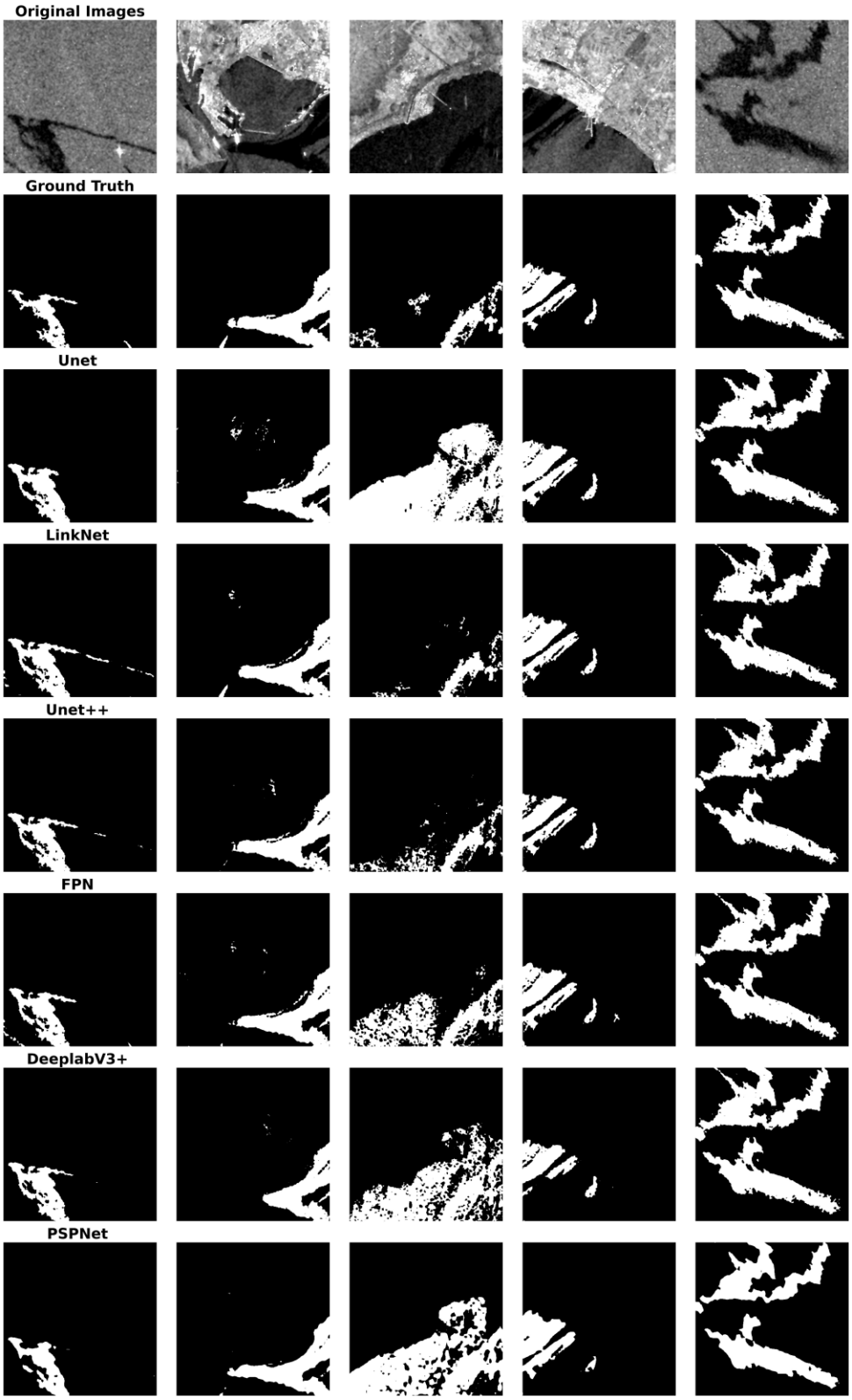


Fig. 4.8 Qualitative results of the examined segmentation models on the presented oil spill Binary dataset, where white corresponds to oil spill and black to non-oil class.

4.2.3. Comparison of Deeplabv3 models

DeepLabv3+ was chosen for its advanced capabilities in semantic segmentation, particularly its ability to handle complex scenarios with high accuracy. Its use of atrous convolution allows for effective control over the field of view, capturing multi-scale contextual information without significantly increasing computational load. This makes DeepLabv3+ particularly suitable for tasks requiring detailed and precise segmentation, such as distinguishing between closely related classes or detecting fine details in high-resolution images.

The choice of encoder significantly impacts the performance of DeepLabv3+, as different encoder architectures have varying strengths in feature extraction. The performance data shows that EfficientNet-B5 achieved the highest mIoU of 68.29%, with an F1 score of 0.7378, recall of 0.7669, and precision of 0.8134, indicating its superior ability to capture detailed features across different classes. This encoder also has 29.49 million parameters, providing a balanced performance with a reasonable model size.

ResNet-based encoders also demonstrated strong performance. For instance, ResNet101 achieved a mean mIoU of 62.82%, with an F1 score of 0.7046, recall of 0.7380, and precision of 0.8057. It has 45.67 million parameters, showcasing its robustness in feature extraction but with a higher computational load compared to EfficientNet-B5. ResNet50, with 26.68 million parameters, achieved a mean mIoU of 60.85%, F1 score of 0.6842, recall of 0.7423, and precision of 0.7721, balancing performance and model complexity. EfficientNet models generally outperformed ResNet models, likely due to their optimized balance between depth, width, and resolution, enhancing their capacity to capture diverse features at multiple scales. EfficientNet-B4, for instance, achieved a mean mIoU of 65.46%, F1 score of 0.7201, recall of 0.7632, and precision of 0.8100 with 18.62 million parameters. Similarly, EfficientNet-B2, with only 8.64 million parameters, achieved a mean mIoU of 65.36%, F1 score of 0.7217, recall of 0.7561, and precision of 0.8155, demonstrating its efficiency.

The inclusion of metrics such as F1 score, recall, and precision provides a comprehensive evaluation, highlighting EfficientNet-B5's balanced performance across these critical measures, making it a top choice for detailed segmentation tasks. This variability in performance underscores

the importance of selecting the appropriate encoder to maximize the segmentation accuracy of DeepLabv3+ for specific tasks. (**Table 4.4**).

Table 4.4 Effect of encoders on accuracy of DeepLabv3+

Encoder Name	Mean mIoU	F1 Score	Recall	Precision	Parameters (M)
resnet50	0.6085	0.6842	0.7423	0.7721	26.68
resnet18	0.6231	0.6944	0.7351	0.7832	12.33
resnet101	0.6282	0.7046	0.7380	0.8057	45.67
resnext50_32x4d	0.6179	0.6935	0.7534	0.7731	26.15
resnet152	0.6077	0.6794	0.7191	0.7784	61.31
efficientnet-b2	0.6536	0.7217	0.7561	0.8155	8.64
efficientnet-b3	0.6346	0.7044	0.7512	0.7939	11.68
efficientnet-b4	0.6546	0.7201	0.7632	0.8100	18.62
Efficientnet-b5	0.6829	0.7378	0.7669	0.8134	29.49
efficientnet-b6	0.6467	0.7147	0.7663	0.8063	41.97
efficientnet-b7	0.6511	0.7196	0.7721	0.8019	65.11
timm-efficientnet-b5	0.6334	0.7017	0.7664	0.7787	29.49

4.2.4. Comparison of Unet and Unet++

The number of parameters is a crucial metric in evaluating the complexity and computational requirements of deep learning models for semantic segmentation. **Table 4.5** demonstrates how these factors impact performance and efficiency for different UNet and UNet++ models with various encoders.

For instance, UNet++ with EfficientNet-B5 achieves the highest mIoU of 65.48% with 89.45 million parameters, indicating an excellent balance between performance and model complexity. In contrast, UNet++ with DenseNet161, while achieving a respectable mIoU of 58.67%, has 191.09 million parameters, reflecting the increased computational load. This suggests that more sophisticated encoders can improve segmentation accuracy but may also significantly increase the number of parameters.

UNet and UNet++ are popular models for semantic segmentation due to their ability to capture fine-grained details by utilizing skip connections that merge low-level and high-level

features. The table highlights how different encoders impact these models' performance. For instance, EfficientNet-B5 used with UNet++ not only achieves the highest mIoU but also shows high accuracy (96.08%) and superior precision (78.76%), recall (78.59%), and F1 score (71.39%). This superior performance is due to EfficientNet's optimized balance of network depth, width, and resolution, which enhances feature extraction and generalization. In contrast, UNet++ with ResNet50 achieves a lower mIoU of 60.17%, indicating its relatively lesser ability to capture fine details compared to EfficientNet.

DenseNet-based encoders also perform well, with UNet and UNet++ showing varying degrees of effectiveness. For instance, UNet with DenseNet161 achieves an mIoU of 62.36% and a decent accuracy of 94.82% with 72.67 million parameters, highlighting its robust performance. However, UNet++ with DenseNet161 has a higher parameter count (191.09 million) while achieving a lower mIoU of 58.67%. DPN92 with UNet++ shows an mIoU of 62.36% with 151.62 million parameters, suggesting its feature extraction capabilities are less effective for this task. These observations underscore the importance of selecting the appropriate encoder to optimize both accuracy and model efficiency, with EfficientNet-B5 emerging as the best overall choice for detailed segmentation tasks.

Table 4.5 Effect of encoders on accuracy of Unet and Unet++

Model	**Encoder**	**Mean IoU**	**F1 Score**	**Recall**	**Precision**	**Accuracy**	**Parameters (M)**
Unet	densenet121	0.6125	0.6754	0.7449	0.7648	0.9423	62.76
Unet	densenet169	0.6126	0.6738	0.7410	0.7671	0.9529	62.87
Unet	densenet201	0.6170	0.6802	0.7423	0.7786	0.9425	69.52
Unet	densenet161	0.6236	0.6845	0.7363	0.7693	0.9482	72.67
Unet++	resnet50	0.6017	0.6605	0.7222	0.7715	0.9388	107.2
Unet++	densenet161	0.5867	0.6498	0.7292	0.7226	0.9369	191.09
Unet++	efficientnet-b5	0.6548	0.7139	0.7859	0.7876	0.9608	89.45
Unet++	dpn92	0.6236	0.6846	0.7477	0.7645	0.9518	151.62

4.3. Oil Fingerprinting Results

4.3.1. Principal components analysis (PCA)

A variety of parameters, including total petroleum hydrocarbons (TPH), n-alkanes, biomarkers (terpanes and steranes), and PAHs, were analyzed and present. The results showed that the oil had not experienced significant weathering. This is because the patterns of biodegradation and photo-oxidation, which are two major weathering processes, were not evident from the analysis of the characteristics of PAHs and n-alkanes.

PCA was applied for feature selection and dimensionality reduction. As mentioned in section 3.3.1, all datasets from four classes of chemical compositions were standardized before conducting PCA, and 99% variance was chosen to ensure the retention of most information from the original datasets. Scree plots were drawn for PCs against the percentage of explained variance foe each chemical composition is shown **in Fig. 4.9**. The scree plots offer a comprehensive breakdown of the PCs and their associated contributions to the variance in the dataset. Notably, n-alkanes take precedence with the highest PC score, denoted as while TPH and PAHs closely follow, making substantial contributions to dataset variance. In contrast, biomarkers exhibit the least significant PC presence, with only 2 PCs capturing less than 99% of the total variance. Notably, the first PC across all components represents over 51% of the variance, signifying its substantial influence on the dataset. Specifically, the first PC for PAHs and biomarkers impressively accounts for approximately 83% of the variance, underscoring their pivotal roles. To enhance the interpretability of the data and facilitate subsequent model predictions, a strategic approach was employed to streamline and simplify the information. Instead of transforming all features into a lower dimension, we selectively chose features with the highest loading values to represent the PCs. This careful selection process aims to highlight the most influential and informative aspects of the data, making it more amenable for analysis and model predictions. By focusing on the selective representation of these principal components, we effectively retain the critical elements while reducing the overall dimensionality of the dataset. This approach ensures that the streamlined data remains highly informative for further analyses and modeling in oil fingerprinting.

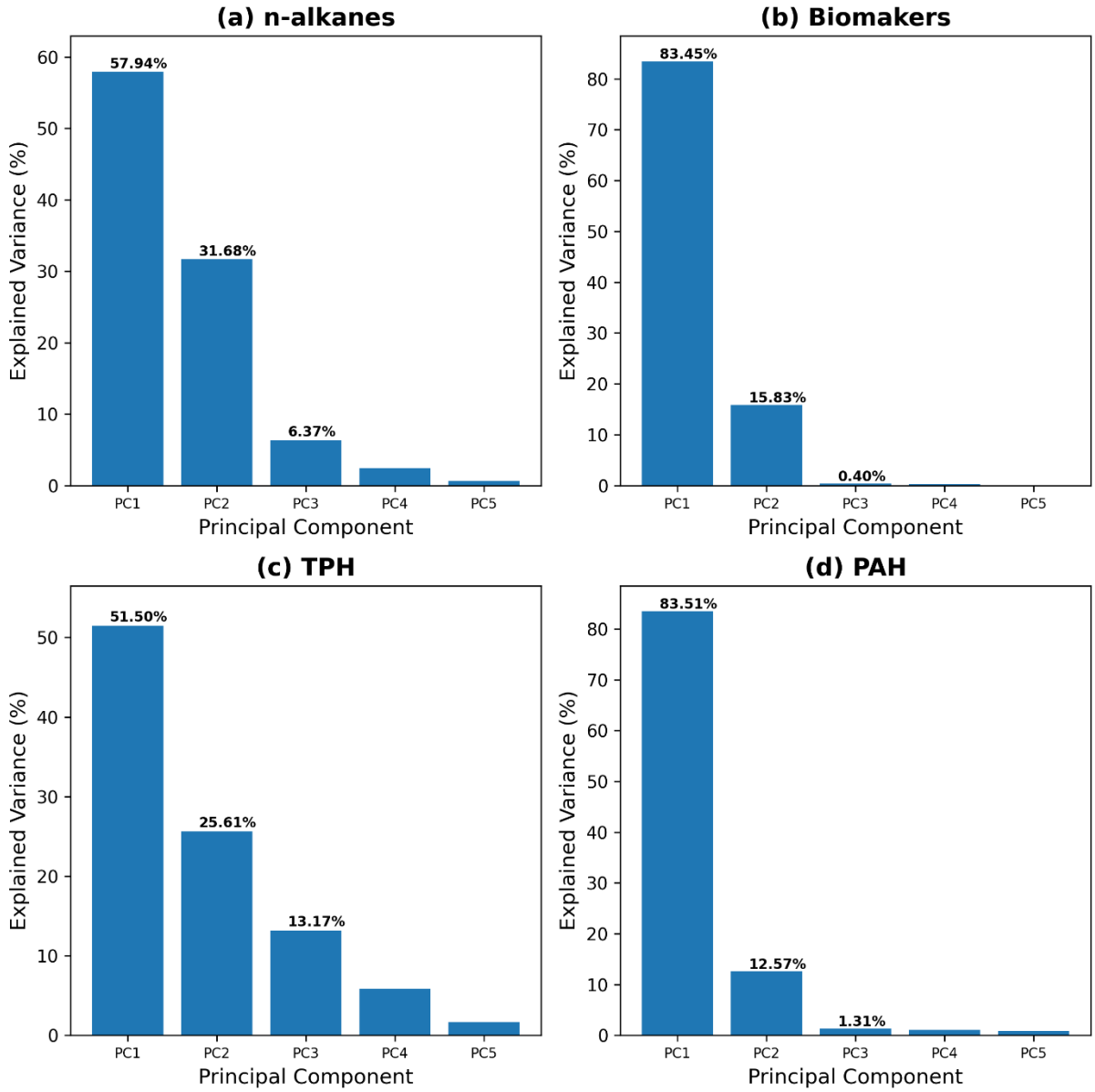


Fig. 4.9 Scree plots of principal components under 99% variance for each chemical composition: (a): n-alkanes, (b): Biomarkers, (c): TPH, (d): PAH

Based on the Principal Component Analysis (PCA) results (Fig. 4.9), it is evident that the top three PCs effectively encapsulate the primary variances across the analyzed chemical categories. PC1 is predominantly characterized by n-C16 from the n-alkanes group, PHC F1 (<n-C10) representing Total Petroleum Hydrocarbons (TPH), 17α(H), 21β(H)-hopane as a biomarker,

and Fluoranthene for PAH (**Table 4.6**). These components exhibit substantial loadings, signifying their critical role in capturing the variance within their respective domains. PC2 includes variables such as n-C24, another significant n-alkane, PHC F2 (<n-C10<n16), C23 tricyclic terpane from biomarkers, and Biphenyl from PAHs, highlighting their influential correlation with this principal component. PC3 comprises n-C38, TAH/PHC (%), which are essential in TPH analysis, 17α(H),21β(H)-30-norhopane (H29) as a biomarker, and Anthracene from PAHs, indicating their substantial contributions to this component. Collectively, these variables underscore the dynamics of environmental impact and the complexity of chemical interactions within the PCA framework.

Table 4.6 Key Chemical Indicators Across top three Principal Components for each chemical composition

Principal Component	n-alkanes	TPH	Biomarkers	PAH
PC1	n-C16	PHC F1 (<n-C10)	17□(H),21□(H)-hopane	fluoranthene
PC2	n-C24	PHC F2 (<n-C10<n16)	C23 tricyclic terpane	biphenyl
PC3	n-C38	TAH/PHC (%)	17□(H),21□(H)-30-norhopane (H29)	anthracene

4.3.2. Visualizing data

The visuals in **Fig 4.10** reinforce the robust analytical approach using PCA and Hierarchical Cluster Analysis (HCA) to discern the complex chemical composition of various oil samples. **Fig 4.10 (a)** displays a PCA scatter plot for n-alkanes, where samples #3515 and #3522 emerge as distinct from the others, suggesting their classification as diesel oils with notably higher TPH values and a significant unresolved complex mixture (UCM) hump characteristic of such oils. These patterns, supported by Yang et al. (2020), indicate that these samples may originate from a different source compared to the others in the dataset.

Expanding upon this, **Fig 4.10 (b)** provides a PCA scatter plot for TPH, where the same two samples are distinctly grouped, further affirming their unique chemical signature, possibly due to their higher TPH content. This level of differentiation is crucial in environmental forensics for tracing the origins of oil spills and determining appropriate remediation strategies. The clustering revealed in the dendrograms for n-alkanes (**Fig 4.10 (c)**) and TPH (**Fig 4.7(d)**) is consistent with the PCA results, where the hierarchical structure shows clear separation between different oil types. The dendrogram illustrates that while samples #3515 and #3522 form a branch indicating their similarity, other samples like #3534 and #3535 appear to fall into a different cluster, possibly indicative of weathered oils, as seen in their specific dendrogram branches. For visualizing the samples, these two metrics were chosen because they effectively highlight the significant variations and clustering within the dataset.

Together, these data visualizations form a coherent narrative, elucidating the intricate relationships within the chemical profiles of oil samples. By integrating PCA and HCA, the methodology transcends mere pattern recognition, offering a comprehensive framework that enhances the precision in identifying oil types and sources, particularly valuable in the context of the MV Manolis L shipwreck oil analysis. This combined approach not only aids in the environmental assessment of oil spills but also serves as a strategic tool in environmental protection and conservation efforts.

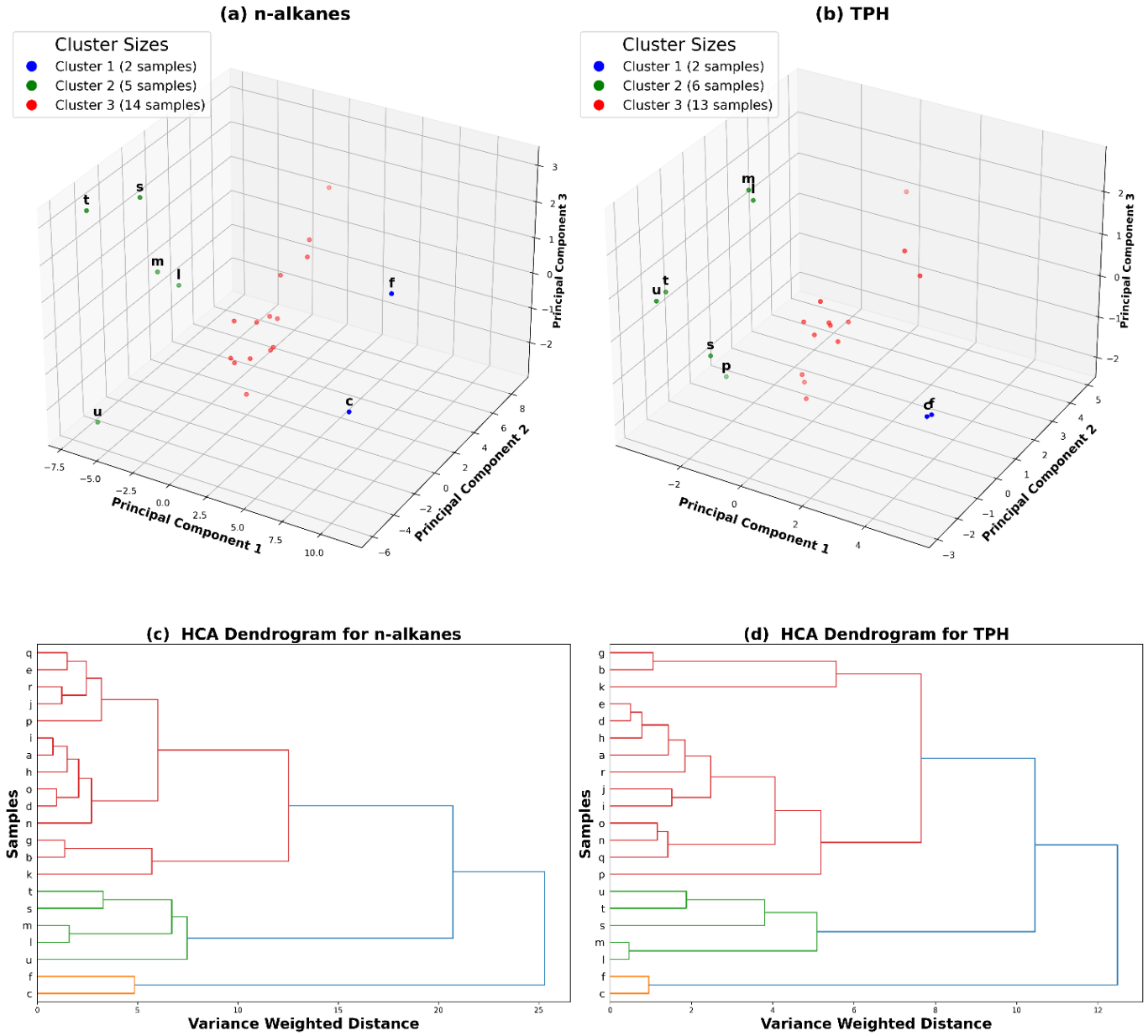


Fig. 4.10 PCA score plot (a) 3D PCA score plot of n-alkanes, (b) 3D PCA score plot of TPH, (c) HCA dendrogram from n-alkanes and (d) HCA dendrogram from TPH.

4.3.3. Models development

In this study, we employed a comprehensive approach to model training and comparative analysis by utilizing six distinct ML algorithms and exploring three diverse feature spaces. The goal was to develop a robust and highly predictive model. To ensure the reliability of our models, we employed a cross-validation test size of 20% within the GridSearchCV framework. This

approach effectively allocated 20% of the dataset to the test datasets, enabling us to evaluate the model's performance on unseen data. For the convenience of prediction and evaluation, we transformed the labels into numerical values, specifically 0, 1, and 2. This numerical representation of labels streamlines the modeling process and aids in the accurate assessment of our algorithms' performance. The outcomes of our hyperparameter optimization efforts are meticulously documented and presented in Table 6. This table encapsulates the key parameters that were fine-tuned to achieve optimal model performance, providing valuable insights into the settings that yielded the best results for each algorithm and feature space. This comprehensive analysis serves as a foundation for our model development, allowing us to make informed decisions about the most effective approach for our specific problem.

The SVC model achieved the highest accuracy (1.000) for predicting n-alkane composition in the X dataset with all PCs. The SVC model also achieved the highest accuracy (0.988) for predicting PAH composition in the X dataset with all PCs. For TPH composition, the SVC model achieved the highest accuracy (0.929) in the X dataset with all PCs. Finally, for biomarker composition, the KNN model achieved the highest accuracy (0.965) in the X dataset with all PCs. Overall, the SVC model performed the best for predicting the composition of all four chemical groups (n-alkanes, PAHs, TPH, and biomarkers) in the X dataset. This suggests that the SVC model is a robust and reliable tool for predicting the chemical composition of oil samples (**Table 4.7**).

Table 4.7 Outputs of different ML models in each Chemical Composition

Chemical composition	Dataset	Model	Variable	Accuracy
n-alkanes	X (all PCs)	DTC	34	0.976
		KNN	34	0.988
		SVC	34	1
		RFC	34	0.98
		LRC	34	1
		EVC	34	0.988
PAH	X (all PCs)	DTC	15	0.965
		KNN	15	0.893
		SVC	15	0.988
		RFC	15	0.918
		LRC	15	0.964
		EVC	15	0.917
TPH	X (all PCs)	DTC	10	0.953
		KNN	10	0.965
		SVC	10	0.929
		RFC	10	0.976
		LRC	10	0.893
		EVC	10	0.941
Biomakers	X (all PCs)	DTC	6	0.893
		KNN	6	0.965
		SVC	6	0.869
		RFC	6	0.917
		LRC	6	0.869

4.3.4. F-score

The resulting F-scores are summarized in **Table 4.8**, representing the outcomes of the top-performing models under the specific conditions outlined in this study. Among the various

chemical compositions investigated in this study, it's worth noting that certain models stood out as the most proficient. The Random Forest model exhibited the highest F-score, indicating its exceptional predictive power and precision in identifying and classifying the target elements. Following closely in performance were the SVM and KNN models, which also demonstrated strong predictive capabilities in this context. While these models excelled in their performance, it's essential to acknowledge that the Decision Tree and Logistic Regression models, while still offering valuable insights, exhibited comparatively lower F-scores. This suggests that their performance might be more suitable for specific scenarios or may benefit from further fine-tuning to reach their full potential. The selection of the most appropriate model should be contingent on the specific requirements and objectives of the given application or study.

Table 4.8 F-score for the top-performing models

Chemical composition	Model	F-score
n-alkanes	SVC	1.0
PAH	SVC	0.987
TPH	RFC	0.975
Biomarkers	KNN	0.963

4.4. Summary of Results

The evaluation of various segmentation models focused on assessing their accuracy and effectiveness in detecting and segmenting oil spills using the multi-class and binary datasets. DeeplabV3+ demonstrated steady and high performance across most classes, particularly excelling in mIoU and challenging categories. This model's architecture effectively learns complex features without overfitting. FPN showed high IoU scores for ships and land, thanks to its multi-scale feature extraction capability, but struggled with oil spills and look-alikes, likely due to the need for more detailed textural differentiation. LinkNet, while improving over epochs, showed the lowest overall performance, reflecting its design for efficiency rather than capturing intricate details.

In the binary segmentation task, models typically exhibited higher mIoU scores due to the reduced complexity, less class confusion, and simpler data distribution. Unet++ performed well, benefiting from its sophisticated architecture that enhances information flow and gradient preservation, making it highly effective in this binary segmentation task. FPN performed well for the background class but showed lower IoU scores for the oil spill class, indicating a need for more detailed textural differentiation. PSPNet, while improving over epochs, showed variability in its performance, reflected in its test set mean IoU scores. DeepLabv3+ was chosen for its advanced capabilities in semantic segmentation, particularly its ability to handle complex scenarios with high accuracy. Its use of atrous convolution allows for effective control over the field of view, capturing multi-scale contextual information without significantly increasing computational load. The performance data showed that EfficientNet-B5 achieved the highest mean IoU of 68.29%, indicating its superior ability to capture detailed features across different classes. This encoder also exhibited high F1 scores, recall, and precision, making it a top choice for detailed segmentation tasks. In contrast, UNet and UNet++ are popular models for semantic segmentation due to their ability to capture fine-grained details by utilizing skip connections that merge low-level and high-level features. UNet++ with EfficientNet-B5 not only achieved the highest mIoU but also showed high accuracy and superior precision, recall, and F1 score. This superior performance is due to EfficientNet's optimized balance of network depth, width, and resolution, which enhances feature extraction and generalization. UNet++ with DenseNet161, while achieving a respectable mIoU, reflects the increased computational load, underscoring the importance of selecting the appropriate encoder to optimize both accuracy and model efficiency.

The oil fingerprinting results highlight the analysis of various chemical parameters to understand the composition of oil samples and the absence of significant weathering. PCA was applied for feature selection and dimensionality reduction, identifying that n-alkanes, TPH, and PAHs significantly contributed to the dataset variance, while biomarkers contributed the least. The first PC captured over 51% of the variance, underscoring its substantial influence, particularly with PAHs and biomarkers accounting for approximately 83% of the variance. This strategic approach to PCA not only retained critical elements of the data but also streamlined it for further analysis and modeling. The key chemical indicators across the top three principal components included n-C16, PHC F1, 17α(H),21β(H)-hopane, and Fluoranthene, highlighting their critical roles in capturing dataset variance and contributing to a robust framework for subsequent analyses. In the

model development phase, a comprehensive approach was employed using six distinct ML algorithms and exploring three diverse feature spaces. The Support Vector Classifier (SVC) model achieved the highest accuracy for predicting n-alkane composition (1.000), PAH composition (0.988), and TPH composition (0.929). For biomarker composition, the K-Nearest Neighbors (KNN) model achieved the highest accuracy (0.965). This comprehensive analysis highlighted the robustness and reliability of the SVC model in predicting the chemical composition of oil samples across multiple chemical groups, demonstrating its effectiveness in handling complex datasets and ensuring accurate predictions.

Chapter 5 Discussion

5.1. Semantic Segmentation

5.1.1. Network analysis

Starting with our multi-class segmentation results, DeeplabV3+ consistently shows high performance across most classes, leading in mean Intersection over Union (mIoU) and excelling particularly well in challenging categories with an mIoU of 68.29%. This consistency indicates that our model's architecture effectively learns complex features without overfitting. This aligns with the findings of Alpers et al. (2017), who highlighted the robustness of advanced segmentation models in handling complex scenarios. In our experiments, FPN shows high IoU scores for ships (40.99%) and land (88.44%) during training, thanks to its multi-scale feature extraction capability. However, it struggles with oil spills (61.38%) and look-alikes (41.16%), likely due to the need for more detailed textural differentiation. Similar observations were made by Basit et al. (2022) in their comparison of segmentation models for SAR images. LinkNet, while improving over epochs, shows the lowest overall performance with a mean IoU of 64.99%, reflecting its design for efficiency rather than capturing intricate details in our dataset. The imbalance in our dataset can lead to models being biased towards more frequently represented classes, reducing their ability to accurately segment underrepresented classes. PSPNet's architecture, while powerful for certain applications, might not be as effective in capturing the fine-grained details necessary for differentiating between closely related classes like oil spills (52.20%) and look-alikes (39.79%) (Saha et al., 2016). The challenges in segmenting oil spills and look-alikes in our research underscore the need for architectures that can capture fine-grained details and handle complex textures effectively. Addressing this imbalance through data augmentation or re-sampling techniques could help improve the models' performance on the less represented classes, leading to more accurate and robust semantic segmentation in our specific context.

Following our multi-class segmentation, binary segmentation was performed. This simpler classification task generally leads to higher mIoU scores due to reduced complexity and less class

confusion. In our research, Unet++ performed well, achieving a mean IoU of 87.47%. Its sophisticated architecture enhances information flow and gradient preservation, making it highly effective in this binary segmentation task. FPN also performed well for the background class with an IoU of 98.57% but showed lower IoU scores for the oil spill class at 65.24%, indicating a need for more detailed textural differentiation (Basit et al., 2021). PSPNet, while improving over epochs, exhibited variability in its performance, reflected in its test set mean IoU scores of 85.17%. Notably, the IoU for the non-oil spill class achieved the highest number compared to all the latest research on oil spill detection using DCNNs, demonstrating the effectiveness of these models in accurately identifying and segmenting non-oil regions (Conceição et al., 2021; Shaban et al., 2021; Shanmukh et al., 2024). The training plots for our binary dataset align well with the test set results, providing a comprehensive view of how each model learns and performs. The challenges in segmenting oil spills in our study underscore the need for architectures that can capture fine-grained details and handle complex textures effectively. These promising results highlight the robustness and accuracy of the models used in our research, significantly advancing the state of oil spill detection.

Given the prominent results of DeeplabV3+ in our multi-class segmentation tasks, we focused specifically on this model to achieve even better results by experimenting with different encoders. DeeplabV3+ was chosen for its advanced capabilities in semantic segmentation, particularly its ability to handle complex scenarios with high accuracy within our study. Its use of atrous convolution allowed us to effectively control the field of view, capturing multi-scale contextual information without significantly increasing computational load. Our performance data revealed that the EfficientNet-B5 encoder achieved the highest mean IoU of 68.29%, indicating its superior ability to capture detailed features across different classes in our dataset (Chen et al., 2017). This encoder also exhibited high F1 scores, recall, and precision, making it a top choice for our detailed segmentation tasks. Other EfficientNet variants, such as EfficientNet-B2, B3, and B4, also performed well, demonstrating the efficiency of this architecture within the context of our research. In comparison, ResNet-based encoders showed respectable performance but did not match the results of EfficientNet-B5 in our tests, highlighting the importance of choosing the right encoder to optimize model performance specifically for our study.

In our research, UNet and UNet++ demonstrated their widely recognized ability to capture fine-grained details in semantic segmentation by using skip connections that merge low-level and high-level features. Among the encoders we tested, UNet++ with EfficientNet-B5 stood out, achieving the highest mean IoU of 0.6548 and exhibiting high accuracy, precision, recall, and F1 score. This superior performance is attributed to EfficientNet-B5's optimized balance of network depth, width, and resolution, which enhances feature extraction and generalization within our dataset (Weng & Zhu, 2015). In contrast, UNet++ with DenseNet161, while still respectable with a mean IoU of 0.5867, demonstrated the impact of increased computational load, having a high parameter count of 191.09 million. These observations underscore the importance of selecting the appropriate encoder to optimize both accuracy and model efficiency in our specific context, with EfficientNet-B5 emerging as the best overall choice for detailed segmentation tasks. Other configurations, such as UNet with various DenseNet encoders, showed consistent but lower performance in our experiments, further highlighting EfficientNet-B5's effectiveness for our research objectives.

5.1.2. Overcoming challenges in oil spill detection

The integration of advanced remote sensing techniques, such as SAR images, with state-of-the-art semantic segmentation models has significantly enhanced the accuracy and reliability of oil spill detection. Models like U-Net, LinkNet, UNet++, FPN, DeepLabv3+, and PSPNet effectively manage the complex and noisy nature of SAR data. For instance, DeepLabv3+ achieved a mean IoU of 68.29% with EfficientNet-B5, showcasing its superior ability to capture detailed features across different classes. These models excel in capturing intricate details and minimizing noise, crucial for environmental monitoring (Alizadeh et al., 2018; Jafarzadeh et al., 2021).

High-quality datasets, such as those created by Krestenitis et al. (2019a) and validated with incidents in the Gulf of Suez, provide a robust foundation for model training and evaluation. These meticulously annotated datasets ensure accurate verification and mapping of each oil spill instance, enabling the models to generalize well to real-world scenarios and enhancing detection reliability. However, this integration comes with challenges. The advanced techniques and models require significant computational resources, which can be a barrier for smaller institutions. Preprocessing steps like radiometric calibration and noise reduction are time-consuming and require expertise. Environmental variations, such as weather conditions and similar-looking natural phenomena, can

introduce noise and artifacts, necessitating continuous model refinement and validation (El-Magd et al., 2023; Dabboor et al., 2018).

Despite these advancements, the study faced limitations, including the time-consuming nature of SAR image preprocessing and the computational resources required for extensive model training. Additionally, creating a binary dataset using GIS maps, while achieving approximately 94% accuracy, posed challenges in further improving accuracy. Uncertainties related to environmental factors, such as wind creating look-alike images and the weathering of oil spills, also impacted model performance. Continuous refinement and validation are essential to maintain model robustness and reliability in diverse real-world scenarios (Huang et al., 2022).

5.2. Oil Fingerprinting

5.2.1. Model analysis

The results of the PCA showed that in all four chemical compositions (n-alkanes, TPH, PAHs, and biomarkers), the top principal component (PC) covered at least 51% of the variance in the dataset. This high level of variance explained by the top PCs underscores their substantial influence on the dataset and their importance in the analysis. For n-alkanes, n-C16 emerged as the top PC, reflecting its dominance in the composition of crude oil and its stability in various environmental conditions, making it a reliable marker for oil fingerprinting (Wang et al., 1999). In the case of polycyclic aromatic hydrocarbons (PAHs), Fluoranthene was the top PC. Fluoranthene is a significant PAH due to its prevalence and persistence in the environment. Its stable chemical structure and widespread presence in different oil samples make it a critical component for distinguishing between various oil sources. Its selection as the top PC highlights its importance in capturing the key variance in PAH profiles across the samples (Mirnaghi et al., 2019). For biomarkers, 17α(H),21β(H)-hopane was identified as the top PC. Hopanes are a class of triterpenoids that are highly resistant to weathering and biodegradation. The specific structure of 17α(H),21β(H)-hopane makes it a robust indicator of oil origin and history. Its stability and resistance to environmental changes allow it to serve as a reliable marker for identifying and differentiating oil samples (Wang et al., 2006).

To ensure reliability, a cross-validation test size of 20% was used within the GridSearchCV framework, which effectively allocated part of the dataset for testing and enabled performance evaluation on unseen data. Labels were transformed into numerical values (0, 1, and 2) to streamline the modeling process and facilitate accurate assessment. The outcomes of the hyperparameter optimization efforts were meticulously documented, providing insights into the settings that yielded optimal performance. The Support Vector Classifier (SVC) model achieved the highest accuracy for predicting n-alkane (1.000), PAH (0.988), and TPH (0.929) compositions. For biomarker composition, the KNearest Neighbors (KNN) model achieved the highest accuracy (0.965). Overall, the SVC model emerged as the best performer across most chemical groups. This superior performance can be attributed to SVC's ability to handle high-dimensional spaces and its robustness in separating classes with clear margins, making it particularly effective for the supervised classification of complex chemical datasets. The combination of Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) in clustering, followed by the application of SVC for classification, provided a robust and reliable framework for predicting the chemical composition of oil samples

5.2.2. Strengths and limitations of data-driven oil fingerprinting approaches

The oil fingerprinting method employed in this research offers significant advantages, particularly its novel use of data-driven models to classify samples without predefined labels. This approach enhances the ability to identify unique oil types from complex datasets, providing a more flexible and accurate means of analysis. Employing six different ML algorithms ensured a comprehensive evaluation of the data, leading to high prediction accuracy and robust model performance. For instance, the SVC model achieved remarkable accuracy across multiple chemical groups, showcasing its robustness and reliability in handling complex chemical datasets. This method facilitated effective clustering and classification of oil samples, contributing significantly to the advancement of oil spill detection and analysis. High-quality datasets and the combination of PCA and HCA laid a strong foundation for the models, ensuring they were well-calibrated and capable of generalizing to real-world scenarios. However, there are notable challenges. While clustering in all chemical compositions was similar, slight differences indicate the need for further analysis of unsupervised classification based on oil chemical composition to better understand the relationships among these parameters. The limited sample necessitating the increase of sample

numbers using synthetic datasets to avoid overfitting and improve generalizability. Despite these challenges, the continuous refinement and validation of models with diverse datasets will enhance their broad applicability and reliability.

Chapter 6 Conclusion

6.1. Summary

This research aims to develop a comprehensive framework for oil spill detection and oil fingerprinting, leveraging advanced remote sensing techniques, semantic segmentation models, and machine learning methods. The methodology integrates various cutting-edge techniques to address the challenges in oil spill detection and fingerprinting. The process begins with the use of Synthetic Aperture Radar (SAR) imaging for its reliability under various weather conditions. The study utilized two datasets: a multi-class dataset by Krestenitis et al. (2019a) and a binary dataset from the Gulf of Suez oil spill incidents (2017-2021). These datasets underwent extensive preprocessing to ensure high-quality input for training state-of-the-art semantic segmentation models, including U-Net, LinkNet, UNet++, FPN, DeepLabv3+, and PSPNet.

Key results from the evaluation of semantic segmentation models demonstrated that DeepLabv3+ achieved the highest mean Intersection over Union (mIoU) of 68.29% in multi-class segmentation, while Unet++ and LinkNet performed exceptionally well in binary segmentation tasks, with mIoU scores of 87.47% and 89.88%, respectively. For oil fingerprinting, the chemical compositions of oil samples from the MV Manolis L shipwreck were analyzed using GC/MS and GC/FID techniques. Principal Component Analysis (PCA) and clustering methods were applied to categorize oil samples based on their chemical signatures, with the Support Vector Classifier (SVC) model achieving the highest accuracy across multiple chemical groups, highlighting its robustness and reliability in predicting oil composition.

In conclusion, this research presents a robust framework for oil spill detection and fingerprinting, combining remote sensing, advanced segmentation models, and machine learning techniques. The developed methodologies significantly enhance the precision and reliability of oil spill detection and source identification, crucial for effective environmental monitoring and protection. By integrating diverse data sources and analytical techniques, this study offers a comprehensive approach to addressing the challenges of environmental forensics, contributing to more accurate and timely responses to oil spill incidents and advancing the field of environmental monitoring.

6.2. Contributions

This research significantly advances the field of environmental forensics and oil spill detection by introducing an integrated approach that combines advanced imaging techniques, comprehensive data processing, and sophisticated machine learning models. The novel integration of oil spill detection and oil fingerprinting frameworks sets a new standard for comprehensive environmental monitoring. A notable contribution of this research is the creation of a novel binary segmentation dataset specifically developed for this study. This dataset enhances the accuracy of oil spill detection models and can serve as a benchmark for future research, enabling the evaluation of other semantic segmentation models.

This research has notably enhanced model performance, increasing the mean Intersection over Union (mIoU) from Krestenitis' study at 65.06% to 68.29%, a 4.96% improvement. Furthermore, the binary segmentation model achieved an IoU of 80.36% for the oil spill class, the highest mIoU recorded in previous studies on oil spill detection. These contributions are crucial for advancing environmental monitoring and protection, providing a solid foundation for future research and development.

The oil fingerprinting conducted in this research successfully demonstrated the rapid identification of oil spill sources based on their chemical composition using machine learning algorithms. For instance, the Support Vector Classifier (SVC) model achieved the highest accuracy in predicting the composition of n-alkanes (1.000), PAHs (0.988), and TPH (0.929). Additionally, the K-Nearest Neighbors (KNN) model achieved an accuracy of 0.965 for biomarkers. These results underscore the effectiveness of machine learning models in handling complex chemical datasets, ensuring precise and rapid identification of oil spill sources based on their chemical profiles.

6.3. Recommended Future Studies

Future studies could focus on developing real-time oil spill detection systems by leveraging cloud computing for data processing. Integrating cloud technologies would facilitate rapid processing and analysis of SAR images and other relevant data, enabling timely responses to oil spill incidents. This approach would significantly enhance the ability to monitor vast areas

continuously and efficiently, improving overall environmental protection efforts. Real-time detection and processing would also allow for immediate mitigation strategies, potentially reducing the environmental impact of oil spills significantly. Additionally, further research should aim to improve the accuracy of detecting oil spills and distinguishing them from look-alike substances, such as algal blooms or seaweed. This can be achieved by exploring more advanced and complex machine learning models, including deep learning architectures that can capture finer details and handle the nuances of different oil spill scenarios. Incorporating additional data sources could significantly enhance detection capabilities, making it possible to differentiate more accurately between oil spills and other similar substances.

Lastly, continuous refinement and optimization of machine learning models should be a primary focus, particularly in addressing the challenges of imbalanced datasets and improving the generalizability of models across different geographical locations and environmental conditions. Techniques such as transfer learning and domain adaptation could be explored to enhance model performance in diverse scenarios. By addressing these areas, future research can further advance the field of oil spill detection and environmental forensics, contributing to more effective monitoring, response, and remediation strategies.

References

Alizadeh, M. J., Kavianpour, M. R., Danesh, M., Adolf, J., Shamshirband, S., & Chau, K. W. (2018). Effect of river flow on the quality of estuarine and coastal waters using machine learning models. *Engineering Applications of Computational Fluid Mechanics, 12*(1), 810–823. https://doi.org/10.1080/19942060.2018.1528480

Alpers, W., Holt, B., & Zeng, K. (2017). Oil spill detection by imaging radars: Challenges and pitfalls. *Remote Sensing of Environment, 201*, 133–147. https://doi.org/10.1016/J.RSE.2017.09.002

Azar, A. T., Elshazly, H. I., Hassanien, A. E., & Elkorany, A. M. (2014). A random forest classifier for lymph diseases. *Computer Methods and Programs in Biomedicine, 113*(2), 465–473. https://doi.org/10.1016/J.CMPB.2013.11.004

Badem, H., Turkusagi, D., Caliskan, A., & Cil, Z. A. (2019). Parkinson Hastalığı Teşhisi için Yapay Arı Kolonisi Temelli Öznitelik Seçimi. *TIPTEKNO 2019 - Tip Teknolojileri Kongresi*. http://openaccess.iste.edu.tr/xmlui/handle/20.500.12508/1224

Basit, A., Siddique, M. A., & Sarfraz, M. S. (2021). Deep Learning Based Oil Spill Classification Using UNET Convolutional Neural Network. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2021-July, 3491–3494. https://doi.org/10.1109/IGARSS47720.2021.9553646

Basit, A., Siddique, M. A., Bhatti, M. K., & Sarfraz, M. S. (2022). Comparison of CNNs and Vision Transformers-Based Hybrid Models Using Gradient Profile Loss for Classification of Oil Spills in SAR Images. *Remote Sensing, 14*(9), 2085. https://doi.org/10.3390/RS14092085

Bayindir, C., David Frost, J., & Barnes, C. F. (2018). Assessment and Enhancement of SAR Noncoherent Change Detection of Sea-Surface Oil Spills. *IEEE Journal of Oceanic Engineering, 43*(1), 211–220. https://doi.org/10.1109/JOE.2017.2714818

Bayona, J. M., Domínguez, C., & Albaigés, J. (2015a). Analytical developments for oil spill fingerprinting. *Trends in Environmental Analytical Chemistry, 5*, 26–34. https://doi.org/10.1016/J.TEAC.2015.01.004

Bayona, J. M., Domínguez, C., & Albaigés, J. (2015b). Analytical developments for oil spill fingerprinting. *Trends in Environmental Analytical Chemistry, 5*, 26–34. https://doi.org/10.1016/J.TEAC.2015.01.004

Bayona, J. M., Domínguez, C., & Albaigés, J. (2015c). Analytical developments for oil spill fingerprinting. *Trends in Environmental Analytical Chemistry, 5*, 26–34. https://doi.org/10.1016/J.TEAC.2015.01.004

Boateng, E. Y., Otoo, J., & Abaye, D. A. (2020a). Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review. *Journal of Data Analysis and Information Processing, 8*(4), 341–357. https://doi.org/10.4236/JDAIP.2020.84020

Boateng, E. Y., Otoo, J., & Abaye, D. A. (2020b). Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review. *Journal of Data Analysis and Information Processing, 8*(4), 341–357. https://doi.org/10.4236/JDAIP.2020.84020

Chanana, S., Thomas, C. S., Zhang, F., Rajski, S. R., & Bugni, T. S. (2020). HCAPCA: Automated hierarchical clustering and principal component analysis of large metabolomic datasets in R. *Metabolites, 10*(7), 1–15. https://doi.org/10.3390/METABO10070297

Chaurasia, A., & Culurciello, E. (2017). LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation. *2017 IEEE Visual Communications and Image Processing, VCIP 2017*, 2018-January, 1–4. https://doi.org/10.1109/VCIP.2017.8305148

Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11211 LNCS*, 833–851. https://doi.org/10.1007/978-3-030-01234-2_49

Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. arXiv 2017, arXiv:1706.05587. https://arxiv.org/abs/1706.05587v3

Chen, Y., Chen, B., Song, X., Kang, Q., Ye, X., & Zhang, B. (2021). A data-driven binary-classification framework for oil fingerprinting analysis. *Environmental Research, 201*, 111454. https://doi.org/10.1016/J.ENVRES.2021.111454

Chezhian A, Mukesh M, Sureshkumar P (2024) A Review on Marine Oil Pollution and Cleanup Strategies. J Exp Zool India 27(1):77– 86. https://doi.org/10.51470/jez.2024.27.1.77

Cho, Y., Na, J. G., Nho, N. S., Kim, S., & Kim, S. (2012). Application of saturates, aromatics, resins, and asphaltenes crude oil fractionation for detailed chemical characterization of heavy crude oils by fourier transform ion cyclotron resonance mass spectrometry equipped with atmospheric pressure photoionization. *Energy and Fuels, 26*(5), 2558–2565. https://doi.org/10.1021/EF201312M/ASSET/IMAGES/LARGE/EF-2011-01312M_0008.JPEG

Conceição, M. R. A., de Mendonça, L. F. F., Lentini, C. A. D., da Cunha Lima, A. T., Lopes, J. M., de Vasconcelos, R. N., Gouveia, M. B., & Porsani, M. J. (2021). SAR Oil Spill Detection System through Random Forest Classifiers. *Remote Sensing, 13*(11), 2044. https://doi.org/10.3390/rs13112044

Cristianini, N., & Shawe-Taylor, J. (2000). An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. https://doi.org/10.1017/CBO9780511801389

Dabboor, M., Singha, S., Montpetit, B., Deschamps, B., & Flett, D. (2018). Assessment of simulated compact polarimetry of the RCM medium resolution SAR modes for oil spill detection. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2018-July, 2416–2419. https://doi.org/10.1109/IGARSS.2018.8517756

El-Magd, I. A., Zakzouk, M., Ali, E. M., Abdulaziz, A. M., Rehman, A., & Saba, T. (2023). Mapping oil pollution in the Gulf of Suez in 2017–2021 using Synthetic Aperture Radar. *The Egyptian Journal of Remote Sensing and Space Sciences, 26*(3), 826–838. https://doi.org/10.1016/J.EJRS.2023.08.005

Farias, C. O., Hamacher, C., Wagener, A. de L. R., & Scofield, A. de L. (2008). Origin and degradation of hydrocarbons in mangrove sediments (Rio de Janeiro, Brazil) contaminated by an oil spill. *Organic Geochemistry, 39*(3), 289–307. https://doi.org/10.1016/J.ORGGEOCHEM.2007.12.008

Fazeres-Ferradosa, T., Rosa-Santos, P., Taveira-Pinto, F., Vanem, E., Carvalho, H., & Correia, J. (2019). Editorial: Advanced research on offshore structures and foundation design: part 1. *Proceedings of the Institution of Civil Engineers - Maritime Engineering, 172*(4), 118–123. https://doi.org/10.1680/jmaen.2019.172.4.118

Fernández-Varela, R., Andrade, J. M., Muniategui, S., & Prada, D. (2010). Selecting a reduced suite of diagnostic ratios calculated between petroleum biomarkers and polycyclic aromatic hydrocarbons to characterize a set of crude oils. *Journal of Chromatography A, 1217*(52), 8279–8289. https://doi.org/10.1016/J.CHROMA.2010.10.043

Fustes, D., Cantorna, D., Dafonte, C., Arcay, B., Iglesias, A., & Manteiga, M. (2014). A cloud-integrated web platform for marine monitoring using GIS and remote sensing. Application to oil spill detection through SAR images. *Future Generation Computer Systems, 34*, 155–160. https://doi.org/10.1016/J.FUTURE.2013.09.020

Gaines, R. B., Frysinger, G. S., Hendrick-Smith, M. S., & Stuart, J. D. (1999). Oil Spill Source Identification by Comprehensive Two-Dimensional Gas Chromatography. *Environmental Science and Technology, 33*(12), 2106–2112. https://doi.org/10.1021/ES9810484

Girard-Ardhuin, F., Mercier, G., Collard, F., & Garello, R. (2005). Operational oil-slick characterization by SAR imagery and synergistic data. *IEEE Journal of Oceanic Engineering, 30*(3), 487–495. https://doi.org/10.1109/JOE.2005.857526

Hashemi-Nasab, F. S., & Parastar, H. (2020). Pattern recognition analysis of gas chromatographic and infrared spectroscopic fingerprints of crude oil for source

identification. *Microchemical Journal, 153*, 104326. https://doi.org/10.1016/J.MICROC.2019.104326

Hasimoto-Beltran, R., Canul-Ku, M., Díaz Méndez, G. M., Ocampo-Torres, F. J., & Esquivel-Trava, B. (2023a). Ocean oil spill detection from SAR images based on multi-channel deep learning semantic segmentation. *Marine Pollution Bulletin, 188*, 114651. https://doi.org/10.1016/J.MARPOLBUL.2023.114651

Hasimoto-Beltran, R., Canul-Ku, M., Díaz Méndez, G. M., Ocampo-Torres, F. J., & Esquivel-Trava, B. (2023b). Ocean oil spill detection from SAR images based on multi-channel deep learning semantic segmentation. *Marine Pollution Bulletin, 188*, 114651. https://doi.org/10.1016/J.MARPOLBUL.2023.114651

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-December*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Hoffman, A. J., & Devereaux Jennings, P. (2011). The BP Oil Spill as a Cultural Anomaly? Institutional Context, Conflict, and Change. *Journal of Management Inquiry, 20*(2), 100–112. https://doi.org/10.1177/1056492610394940

Huang, X., Zhang, B., Perrie, W., Lu, Y., & Wang, C. (2022). A novel deep learning method for marine oil spill detection from satellite synthetic aperture radar imagery. *Marine Pollution Bulletin, 179*, 113666. https://doi.org/10.1016/J.MARPOLBUL.2022.113666

International Tanker Owners Pollution Federation (ITOPF). (2018). *Oil tanker spill statistics*. International Tanker Owners Pollution Federation Limited, London, United Kingdom. Retrieved from https://www.itopf.org/knowledge-resources/documents-guides/document/tanker-spill-statistics-2018/

Ismail, A., Toriman, M. E., Juahir, H., Kassim, A. M., Zain, S. M., Ahmad, W. K. W., Wong, K. F., Retnam, A., Zali, M. A., Mokhtar, M., & Yusri, M. A. (2016). Chemometric techniques in oil classification from oil spill fingerprinting. *Marine Pollution Bulletin, 111*(1–2), 339–346. https://doi.org/10.1016/J.MARPOLBUL.2016.06.089

Jafarzadeh, H., Mahdianpari, M., Homayouni, S., Mohammadimanesh, F., & Dabboor, M. (2021). Oil spill detection from Synthetic Aperture Radar Earth observations: a meta-analysis and comprehensive review. *GIScience & Remote Sensing, 58*(7), 1022–1051. https://doi.org/10.1080/15481603.2021.1952542

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260. https://doi.org/10.1126/SCIENCE.AAA8415/ASSET/AB2EF18A-576D-464D-B1B6-1301159EE29A/ASSETS/GRAPHIC/349_255_F5.JPEG

Khatri, N., Khatri, K. K., & Sharma, A. (2020). Artificial neural network modelling of faecal coliform removal in an intermittent cycle extended aeration system-sequential batch reactor based wastewater treatment plant. *Journal of Water Process Engineering, 37*, 101477. https://doi.org/10.1016/J.JWPE.2020.101477

Krestenitis, M., Orfanidis, G., Ioannidis, K., Avgerinakis, K., Vrochidis, S., & Kompatsiaris, I. (2019a). Oil Spill Identification from Satellite Images Using Deep Neural Networks. *Remote Sensing, 11*(15), 1762. https://doi.org/10.3390/RS11151762

Krestenitis, M., Orfanidis, G., Ioannidis, K., Avgerinakis, K., Vrochidis, S., & Kompatsiaris, I. (2019b). Oil Spill Identification from Satellite Images Using Deep Neural Networks. *Remote Sensing, 11*(15), 1762. https://doi.org/10.3390/RS11151762

Leon, A. Z., Huvenne, V. A. I., Benoist, N. M. A., Ferguson, M., Bett, B. J., & Wynn, R. B. (2020). Assessing the Repeatability of Automated Seafloor Classification Algorithms, with Application in Marine Protected Area Monitoring. *Remote Sensing, 12*(10), 1572. https://doi.org/10.3390/RS12101572

Li, C., Wang, M., Yang, X., & Chu, D. (2023). DS-UNet: Dual-Stream U-Net for Oil Spill Detection of SAR Image. *IEEE Geoscience and Remote Sensing Letters, 20*. https://doi.org/10.1109/LGRS.2023.3330957

Li, J., Fuller, S., Cattle, J., Way, C. P., & Hibbert, D. B. (2004). Matching fluorescence spectra of oil spills with spectra from suspect sources. *Analytica Chimica Acta, 514*(1), 51–56. https://doi.org/10.1016/J.ACA.2004.03.053

Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-January, 936–944. https://doi.org/10.1109/CVPR.2017.106

Liu, Q., Song, Y., Jiang, L., Cao, T., Chen, Z., Xiao, D., Han, G., Ji, W., Gao, F., Wang, P., & Zhang, X. (2017). Geochemistry and correlation of oils and source rocks in Banqiao Sag, Huanghua Depression, northern China. *International Journal of Coal Geology, 176–177*, 49–68. https://doi.org/10.1016/J.COAL.2017.04.005

Lloyd, S. P. (1982). Least Squares Quantization in PCM. *IEEE Transactions on Information Theory, 28*(2), 129–137. https://doi.org/10.1109/TIT.1982.1056489

Ma, F., Zhang, F., Xiang, D., Yin, Q., & Zhou, Y. (2022). Fast Task-Specific Region Merging for SAR Image Segmentation. *IEEE Transactions on Geoscience and Remote Sensing, 60*. https://doi.org/10.1109/TGRS.2022.3141125

Mahmoudi Ghara, F., Shokouhi, S. B., & Akbarizadeh, G. (2022). 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, 15*, 8834–8844. https://doi.org/10.1109/JSTARS.2022.3213768

Mansuy, L., Philp, R. P., & Allen, J. (1997). Source identification of oil spills based on the isotopic composition of individual components in weathered oil samples. *Environmental Science and Technology, 31*(12), 3417–3425. https://doi.org/10.1021/ES970068N/ASSET/IMAGES/LARGE/ES970068NF00007.JPEG

Mdakane, L. W., & Kleynhans, W. (2022). Feature Selection and Classification of Oil Spill from Vessels Using Sentinel-1 Wide-Swath Synthetic Aperture Radar Data. *IEEE Geoscience and Remote Sensing Letters, 19*. https://doi.org/10.1109/LGRS.2020.3025641

Medar, R., Rajpurohit, V. S., & Rashmi, B. (2017). Impact of Training and Testing Data Splits on Accuracy of Time Series Forecasting in Machine Learning. *2017 International Conference on Computing, Communication, Control and Automation, ICCUBEA 2017*. https://doi.org/10.1109/ICCUBEA.2017.8463779

Mieth, B., Kloft, M., Rodríguez, J. A., Sonnenburg, S., Vobruba, R., Morcillo-Suárez, C., Farré, X., Marigorta, U. M., Fehr, E., Dickhaus, T., Blanchard, G., Schunk, D., Navarro, A., & Müller, K. R. (2016a). Combining Multiple Hypothesis Testing with Machine Learning Increases the Statistical Power of Genome-wide Association Studies. *Scientific Reports, 6*(1), 1–14. https://doi.org/10.1038/srep36671

Mieth, B., Kloft, M., Rodríguez, J. A., Sonnenburg, S., Vobruba, R., Morcillo-Suárez, C., Farré, X., Marigorta, U. M., Fehr, E., Dickhaus, T., Blanchard, G., Schunk, D., Navarro, A., & Müller, K. R. (2016b). Combining multiple hypothesis testing with machine learning increases the statistical power of genome-wide association studies. *Scientific Reports, 6*. https://doi.org/10.1038/SREP36671

Mirnaghi, F. S., Pinchin, N. P., Yang, Z., Hollebone, B. P., Lambert, P., & Brown, C. E. (2019). Monitoring of polycyclic aromatic hydrocarbon contamination at four oil spill sites using fluorescence spectroscopy coupled with parallel factor-principal component analysis. *Environmental Science: Processes & Impacts, 21*(3), 413–426. https://doi.org/10.1039/C8EM00493E

Mohammadi, L., Rahdar, A., Bazrafshan, E., Dahmardeh, H., Susan, M. A. B. H., & Kyzas, G. Z. (2020). Petroleum Hydrocarbon Removal from Wastewaters: A Review. *Processes, 8*(4), 447. https://doi.org/10.3390/PR8040447

Mohammed, A., & Kora, R. (2023). A comprehensive review on ensemble deep learning: Opportunities and challenges. *Journal of King Saud University - Computer and Information Sciences, 35*(2), 757–774. https://doi.org/10.1016/J.JKSUCI.2023.01.014

Mulabagal, V., Yin, F., John, G. F., Hayworth, J. S., & Clement, T. P. (2013). Chemical fingerprinting of petroleum biomarkers in Deepwater Horizon oil spill samples collected

from Alabama shoreline. *Marine Pollution Bulletin, 70*(1–2), 147–154. https://doi.org/10.1016/J.MARPOLBUL.2013.02.026

Murugan, N. S., & Devi, G. U. (2019). Feature extraction using LR-PCA hybridization on twitter data and classification accuracy using machine learning algorithms. *Cluster Computing, 22*(6), 13965–13974. https://doi.org/10.1007/S10586-018-2158-3/FIGURES/5

Nalepa, J., & Kawulok, M. (2018). Selecting training sets for support vector machines: a review. *Artificial Intelligence Review, 52*(2), 857–900. https://doi.org/10.1007/S10462-017-9611-1

Taiwo, O. A. (2010). Types of Machine Learning Algorithms, New Advances in Machine Learning, Yagang Zhang (Ed.), ISBN: 978-953-307-034-6, InTech, University of Portsmouth United Kingdom. Pp 3 – 31. Available at InTech open website: http://www.intechopen.com/books/new-advances-inmachine-learning/types-of-machine-learning-algorithms. https://doi.org/10.5772/9385

Omar, M. Y., Shehada, M. F., Mehanna, A. K., Elbatran, A. H., & Elmesiry, M. M. (2021). A case study of the Suez Gulf: Modelling of the oil spill behavior in the marine environment. *Egyptian Journal of Aquatic Research, 47*(4), 345–356. https://doi.org/10.1016/J.EJAR.2021.10.005

Orfanidis, G., Ioannidis, K., Avgerinakis, K., Vrochidis, S., & Kompatsiaris, I. (2018). A Deep Neural Network for Oil Spill Semantic Segmentation in Sar Images. *Proceedings - International Conference on Image Processing, ICIP*, 3773–3777. https://doi.org/10.1109/ICIP.2018.8451113

Pavlov, D., & Vasiliev, A. (2017). Oil Fingerprinting Technology for Well and Reservoir Management. *Society of Petroleum Engineers - SPE Russian Petroleum Technology Conference 2017*. https://doi.org/10.2118/187781-MS

Raeisi, A., Akbarizadeh, G., & Mahmoudi, A. (2018). Combined Method of an Efficient Cuckoo Search Algorithm and Nonnegative Matrix Factorization of Different Zernike Moment Features for Discrimination between Oil Spills and Lookalikes in SAR Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11*(11), 4193–4205. https://doi.org/10.1109/JSTARS.2018.2841503

Ranjan, G. S. K., Kumar Verma, A., & Radhika, S. (2019). K-Nearest Neighbors and Grid Search CV Based Real Time Fault Monitoring System for Industries. *2019 IEEE 5th International Conference for Convergence in Technology, I2CT 2019*. https://doi.org/10.1109/I2CT45611.2019.9033691

Robles-Velasco, A., Muñuzuri, J., Onieva, L., & Rodríguez-Palero, M. (2021). Trends and applications of machine learning in water supply networks management. *Journal of Industrial Engineering and Management, 14*(1), 45–54. https://doi.org/10.3926/JIEM.3280

Ronci, F., Avolio, C., Di Donna, M., Zavagli, M., Piccialli, V., & Costantini, M. (2020). An adversarial learning approach for oil spill detection from SAR images. *IEEE National Radar Conference - Proceedings, 2020-September*. https://doi.org/10.1109/RADARCONF2043947.2020.9266475

Saha, D., Deo, M. C., Joseph, S., & Bhargava, K. (2016). A combined numerical and neural technique for short term prediction of ocean currents in the Indian Ocean. *Environmental Systems Research, 5*(1), 1–14. https://doi.org/10.1186/S40068-016-0057-2

Sahlaoui, H., Alaoui, E. A. A., Agoujil, S., & Nayyar, A. (2023). An empirical assessment of smote variants techniques and interpretation methods in improving the accuracy and the interpretability of student performance models. *Education and Information Technologies*. https://doi.org/10.1007/S10639-023-12007-W

Sancho, A., Ribeiro, J. C., Reis, M. S., & Martins, F. G. (2022). Cluster analysis of crude oils with k-means based on their physicochemical properties. *Computers & Chemical Engineering, 157*, 107633. https://doi.org/10.1016/J.COMPCHEMENG.2021.107633

Shaban, M., Salim, R., Khalifeh, H. A., Khelifi, A., Shalaby, A., El-Mashad, S., Mahmoud, A., Ghazal, M., & El-Baz, A. (2021). A Deep-Learning Framework for the Detection of Oil Spills from SAR Data. *Sensors, 21*(7), 2351. https://doi.org/10.3390/S21072351

Shanmukh, M. P., Priya, S. B., & Madeswaran, T. (2024). Improving Oil Spill Detection in Marine Environments Through Deep Learning Approaches. *2024 4th International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies, ICAECT 2024*. https://doi.org/10.1109/ICAECT60202.2024.10468888

Sharma, B., & Shrestha, A. (2023). Petroleum dependence in developing countries with an emphasis on Nepal and potential keys. *Energy Strategy Reviews, 45*, 101053. https://doi.org/10.1016/J.ESR.2023.101053

Shelhamer, E., Long, J., & Darrell, T. (2014). Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(4), 640–651. https://doi.org/10.1109/TPAMI.2016.2572683

Soh, K., Zhao, L., Peng, M., Lu, J., Sun, W., & Tongngern, S. (2024). SAR marine oil spill detection based on an encoder-decoder network. *International Journal of Remote Sensing, 45*(2), 587–608. https://doi.org/10.1080/01431161.2023.2299274

Song, X., Lye, L. M., Chen, B., & Zhang, B. (2019). Differentiation of weathered chemically dispersed oil from weathered crude oil. *Environmental Monitoring and Assessment, 191*(5), 1–10. https://doi.org/10.1007/S10661-019-7392-5/FIGURES/8

Song, X., Zhang, B., Chen, B., & Cai, Q. (2016). Use of Sesquiterpanes, Steranes, and Terpanes for Forensic Fingerprinting of Chemically Dispersed Oil. *Water, Air, and Soil Pollution, 227*(8), 1–15. https://doi.org/10.1007/S11270-016-2981-1/FIGURES/7

Song, X., Zhang, B., Chen, B., Lye, L., & Li, X. (2018). Aliphatic and aromatic biomarkers for fingerprinting of weathered chemically dispersed oil. *Environmental Science and Pollution Research International, 25*(16), 15702–15714. https://doi.org/10.1007/S11356-018-1730-Y

Sudha, V., & Saro Vijendran, A. (2024). Oil Spill Detection and Recognition Utilizing Faster R-CNN with Enhanced Mobilenetv2 Architecture. *International Journal of Intelligent Systems and Applications in Engineering IJISAE, 17*(s), 304–321. https://ijisae.org/index.php/IJISAE/article/view/4876

Swain, P. H., & Hauska, H. (1977). Decision Tree Classifier: Design and Potential. *IEEE Transactions on Geoscience Electronics, 15*(3), 142–147. https://doi.org/10.1109/TGE.1977.6498972

Taunk, K., De, S., Verma, S., & Swetapadma, A. (2019). A brief review of nearest neighbor algorithm for learning and classification. *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, 1255–1260. https://doi.org/10.1109/ICCS45141.2019.9065747

Temitope Yekeen, S., Balogun, A. L., & Wan Yusof, K. B. (2020). A novel deep learning instance segmentation model for automated marine oil spill detection. *ISPRS Journal of Photogrammetry and Remote Sensing, 167*, 190–200. https://doi.org/10.1016/J.ISPRSJPRS.2020.07.011

Tian, C., Yao, X., Lu, J., Shen, L., & Wu, A. (2021). GC–MS fingerprints of essential oils from agarwood grown in wild and artificial environments. *Trees - Structure and Function, 35*(6), 2105–2117. https://doi.org/10.1007/S00468-021-02177-W/TABLES/5

Tong, S., Liu, X., Chen, Q., Zhang, Z., & Xie, G. (2019). Multi-Feature Based Ocean Oil Spill Detection for Polarimetric SAR Data Using Random Forest and the Self-Similarity Parameter. *Remote Sensing, 11*(4), 451. https://doi.org/10.3390/RS11040451

Wang, C., Chen, B., Zhang, B., He, S., & Zhao, M. (2013). Fingerprint and weathering characteristics of crude oils after Dalian oil spill, China. *Marine Pollution Bulletin, 71*(1–2), 64–68. https://doi.org/10.1016/J.MARPOLBUL.2013.03.034

Wang, D., Song, S., Yang, J., Xu, M., Song, D., Guo, J., Wan, J., & Liu, S. (2024). Marine oil spill detection using improved polarimetric feature based on polarization SAR image. *International Journal of Remote Sensing, 45*(3), 911–929. https://doi.org/10.1080/01431161.2024.2305181

Wang, H., & Hao, F. (2012). An efficient linear regression classifier. *2012 IEEE International Conference on Signal Processing, Computing and Control, ISPCC 2012*. https://doi.org/10.1109/ISPCC.2012.6224355

Wang, M., & Wang, C. (2022). Chemometric techniques in oil spill identification: A case study in Dalian 7.16 oil spill accident of China. *Marine Environmental Research, 182*, 105799. https://doi.org/10.1016/J.MARENVRES.2022.105799

Wang, X., Liu, J., Zhang, S., Deng, Q., Wang, Z., Li, Y., & Fan, J. (2021). Detection of Oil Spill Using SAR Imagery Based on AlexNet Model. *Computational Intelligence and Neuroscience, 2021*(1), 4812979. https://doi.org/10.1155/2021/4812979

Wang, Z., Fingas, M., & Page, D. S. (1999). Oil spill identification. *Journal of Chromatography A, 843*(1–2), 369–411. https://doi.org/10.1016/S0021-9673(99)00120-X

Wang, Z., Stout, S. A., & Fingas, M. (2006a). Forensic Fingerprinting of Biomarkers for Oil Spill Characterization and Source Identification. *Environmental Forensics, 7*(2), 105–146. https://doi.org/10.1080/15275920600667104

Wang, Z., Stout, S. A., & Fingas, M. (2006b). Forensic Fingerprinting of Biomarkers for Oil Spill Characterization and Source Identification. *Environmental Forensics, 7*(2), 105–146. https://doi.org/10.1080/15275920600667104

Weng, W., & Zhu, X. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *IEEE Access, 9*, 16591–16603. https://doi.org/10.1109/ACCESS.2021.3053408

Wetzel, S. J. (2017). Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders. *Physical Review E, 96*(2), 022140. https://doi.org/10.1103/PHYSREVE.96.022140/FIGURES/10/MEDIUM

Yang, Y. J., Singha, S., & Mayerle, R. (2021). Fully Automated SAR Based Oil Spill Detection Using YOLOV4. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2021-July, 5303–5306. https://doi.org/10.1109/IGARSS47720.2021.9553030

Yang, Z., Mirnaghi, F., Shah, K., Lambert, P., Hollebone, B., Yang, C., Brown, C. E., Thomas, G., & Grant, R. (2020). Source identification and evolution of oils recovered from the MV Manolis L shipwreck. *Fuel, 271*, 117684. https://doi.org/10.1016/J.FUEL.2020.117684

Yang, Z., Yang, C., Wang, Z., Hollebone, B., Landriault, M., & Brown, C. E. (2011). Oil fingerprinting analysis using commercial solid phase extraction (SPE) cartridge and gas chromatography-mass spectrometry (GC-MS). *Analytical Methods, 3*(3), 628–635. https://doi.org/10.1039/C0AY00715C

Zhan, Z. W., Lin, X. H., Zou, Y. R., Li, Z., Wang, D., Liu, C., & Peng, P. (2019). Chemometric differentiation of crude oil families in the southern Dongying Depression, Bohai Bay Basin,

China. *Organic Geochemistry, 127*, 37–49. https://doi.org/10.1016/J.ORGGEOCHEM.2018.11.004

Zhang, J., Ai, B., Shang, H., & Li, B. (2023). Oil spill detection in SAR images based on improved mask R-CNN model. *International Conference on Remote Sensing, Mapping, and Geographic Systems (RSMG 2023), 12815*, 74–79. https://doi.org/10.1117/12.3010303

Zhang, S., Xing, J., Wang, X., & Fan, J. (2022). Improved YOLOX-S Marine Oil Spill Detection Based on SAR Images. *2022 12th International Conference on Information Science and Technology, ICIST 2022*, 184–187. https://doi.org/10.1109/ICIST55546.2022.9926772

Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2016). Pyramid Scene Parsing Network. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-January, 6230–6239. https://doi.org/10.1109/CVPR.2017.660

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). UNet++: A nested U-Net architecture for medical image segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11045 LNCS*, 3–11. https://doi.org/10.1007/978-3-030-00889-5_1/FIGURES/3

Zhu, Q., Zhang, Y., Li, Z., Yan, X., Guan, Q., Zhong, Y., Zhang, L., & Li, D. (2022). Oil Spill Contextual and Boundary-Supervised Detection Network Based on Marine SAR Images. *IEEE Transactions on Geoscience and Remote Sensing, 60*. https://doi.org/10.1109/TGRS.2021.3115492