IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 29, No. 6, December, 2024, pp. 2105-2113
Journal homepage: http://iieta.org/journals/isi

Assessing Offshore Oil Spills Using Remote Sensing and Geospatial Artificial Intelligence (GeoAI): A Systematic Literature Review

Alexander Inga Alva*, Ciro Rodríguez

Faculty of Systems and Computer Engineering, Universidad Nacional Mayor de San Marcos UNMSM, Lima 15081, Peru

Corresponding Author Email: alexander.inga@unmsm.edu.pe

https://doi.org/10.18280/isi.290602

Received: 10 May 2024
Revised: 25 July 2024
Accepted: 4 September 2024
Available online: 25 December 2024

Keywords:
artificial intelligence, deep learning, Geospatial Artificial Intelligence (GeoAI), machine learning, Remote Sensing (RS)

ABSTRACT

Artificial Intelligence (AI) cuts across all disciplines of human knowledge, which favors the development of new approaches to the care and preservation of the environment. Oil spills at sea, caused by the transportation and marketing of hydrocarbons, have been considered an unstoppable source of pollution in the last century. So, the present Systematic Review of the Literature (SRL) analyses the most recent publications in Remote Sensing (RS) and AI (Machine Learning, ML) and Deep Learning, DL) identifying a total of 3,355 published articles, and classified 50 articles based on new taxonomy proposed in the study. In this way, it identifies two main Offshore Oil Spill Datasets (OOSD): the HOSD and MKLab projects. Also, it classifies the literature that uses ML at 14%, and DL at 46%, which implies 40% for traditional RS at Offshore Oil Spills (OOS) Detection. This leads RS & AI to consolidate under the concept of Geospatial Artificial Intelligence (GeoAI), a new field of research that allows, in addition to detecting and predicting offshore oil spills under the best performance metrics, such as Precision, Machine Learning 96.5 %, and Deep Learning (99.19 %) using U-Net Model.

1. INTRODUCTION

Remote Sensing (RS) is a discipline that allows not only scanning objects at a distance but also natural phenomena on Earth's surface. For this purpose, the use of optical images under the visible spectrum RGB (3 layers), Multispectral (tens of layers), Hyperspectral (hundreds of layers), and Synthetic Aperture Radar Aperture (SAR) imagery are valuable. But traditional classification using Remote Sensing cannot handle the big data of satellite images because of poor robustness and low accuracy. So refined information also poses new challenges for smart interpretation of high resolution using semantic segmentation [1].

Although the concept of Artificial Intelligence (AI) was coined in 1955 [2] and the first LISP (LISt Processing) programming language was developed under this new approach in 1956 [3], it was not until the beginning of this new millennium that it has opened numerous gates to the academic world to find novel and fast solutions to problems in different areas of knowledge [4]. With its strong statistical foundation, specifically, ML allows the simulation of human learning, which can recognize and extract patterns from the dataset to solve problems of various kinds [5]. More recently, DL further granulates the learning of datasets through numerous dense layers and neural networks, such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Generative Models, Variational Autoencoder (VAE), and Generative Adversarial Network (GAN) [6]. Therefore, traditional RS enhanced with new artificial intelligence algorithms has driven the creation of a new discipline: Geospatial Artificial Intelligence (GeoAI) which addresses problems related to the perception of the geospatial environment [7]. Thus, it is evident that GeoAI can deal with one of humanity's biggest problems: marine pollution.

In this way, OOS are evidenced by an orange-brown emulsion of water in oil, also called foam, which pollutes coastal vegetation and beaches, generating a high environmental impact on the ecosystem, and toxicity for living organisms [8]. On a global scale, more than half of maritime transport corresponds to hydrocarbons, which without being an extremely harmful pollutant, its handling in billions of tons makes it a permanent threat.

Therefore, this article investigates various literature on RS with AI, specifically finding articles in specialized journals and OOS conferences, and GeoAI using ML and DL algorithms. In addition, it identifies (i) the use of the U-Net architecture for oil spills as a trend and (ii) valuable worldwide datasets. So, the study is divided as follows: (i) it will help new researchers continue analyzing and enhancing new GeoAI techniques and tools, and (ii) it will help to make good decisions in the face of environmental disasters. The remaining sections are organized in this way: Section 2 is about the Theoretical Background, Section 3 studies the Methodology, Section 4 outlines the study Results and Discussions, and Section 5 presents the Conclusions.

2. THEORETICAL BACKGROUND

This section presents a theoretical background about Oil Spills, and types of AI (ML-DL), Computer Vision (CV), GeoAI, Remote Sensing, Synthetic Aperture Radar, and Sentinel-1.

2.1 Oil spills

Oil spill refers to the accidental release or release of oil or its by-products into the environment, in oceans, seas, rivers, lakes, and on land. These spills, which can occur during transport by tanker trucks, pipelines, train cars, and/or ships, from the extraction of the deposits to the final consumer, can also occur in storage and cause devastating effects on ecosystems, causing damage to marine life, vegetation, and fauna, as well as to the human communities that depend on these natural resources [9].

2.2 Artificial intelligence

According to the study by Sidorov [10], AI is a science that focuses on the study and construction of models (or systems) of human intelligence that runs through computers. However, there are two perspectives on AI: one human-oriented and based on experience (empirical procedure), and the other is focused on rationality (scientific procedure). This research focuses on the following sublevels: ML-DL as shown in Figure 1. Also, CV is considered to explain additional process.

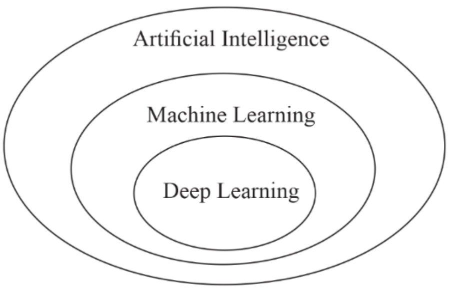

Figure 1. Artificial intelligence and its subsets

2.2.1 Machine learning

ML is the algorithmic construction of a statistical model based on a dataset to solve a practical problem. ML can be classified in Supervised Learning (SL), where the dataset is those labeled examples, and it is used to make a model that takes a feature vector as input and generates information that allows the label to be deduced, Semi-Supervised (SSL) where the dataset contains labeled and unlabeled examples, Unsupervised (UL) where the dataset is those unlabeled examples and Reinforcement Learning (RL) where the algorithm "lives" an environment and can perceive the state of that environment as a vector of features executing actions different actions that obtain different rewards to achieve an optimal action in each state [11].

2.2.2 Deep learning (DL)

DL is a particular subfield of ML that emphasizes data-driven learning of reality through using of successive or "deep" layers, which can number in the tens, hundreds, or even thousands. They train on said data, unlike other learning models in which one or two layers are enough and which are called: superficial layers. DL uses models called Neural Networks (NN) composed of stacked layers. This model was partly inspired by Neurobiology and partly by the understanding processes of the human brain (and its visual cortex), although ultimately, DL is a framework based more than anything on mathematics and learning representations from data [12].

2.2.3 Computer vision (CV)

CV seeks to describe the world that surrounds the human being, in one or more images, reconstructing it in a 3D model that preserves its properties, such as Shape, Illumination, and Color Range. Its main applications are Optical Character Recognition (OCR), Human Recognition, Medical Image (MI), Self-driving Vehicles, and 3D Construction (Photogrammetry) [13].

2.2.4 U-Net

U-Net is one of the most used NN for image classification that was initially created for MI segmentation. This architecture consists of two branches: the first is Coding or contraction, in which all image is captured through a set of convolution layers and "max pooling" layers that create a map of characteristics of an image and reduce its size (fewer network parameters). The last branch is the Decoding or symmetric expansion, which allows precise localization using transposed convolution. In DL, is necessary to use large datasets to train models, which implies: facing the difficulty of gathering those volumes, processing time, hardware resources and high budget to solve an image classification problem as is shown in Figure 2. This is where U-NET takes advantage of conventional models by being effective even with a limited dataset, obtaining greater precision and preserving characteristics by keeping the output size equal to the input [14].

2.3 Geospatial Artificial Intelligence (GeoAI)

Due to the big data of very-high-resolution geospatial data and efficient computer architectures, DL techniques enhanced the development of the geospatial system by providing it with an almost human level of perception, which is known today as GeoAI.

This new concept was coined during the First International ACM SIGSPATIAL Workshop [15]. GeoAI is a new discipline in which the innovations of Artificial Intelligence (ML/DL), data mining, and high computational performance converge to generate knowledge, based on Geography Information Systems (GIS) and exposure modeling in spatial, temporal, and spatiotemporal environments, since it includes spatial non-stationarity and scalability in its analysis of the enormous amounts of spatial-data (Spatial Big Data) that the various geographical environments possess [16].

2.4 Remote Sensing

RS is a discipline that uses sensors and cameras installed on space platforms that record specific scenes of our planet (or others) to distance. To do this, the RS uses the Electromagnetic Spectrum: radio, microwave, radar (SAR images), thermal, infrared, visible (Multispectral and Hyperspectral images), ultraviolet, X-rays, and Gamma rays. If the electromagnetic interaction comes from the Sun energy is considered the passive type and if it comes from an artificial beam is the active type) [17].

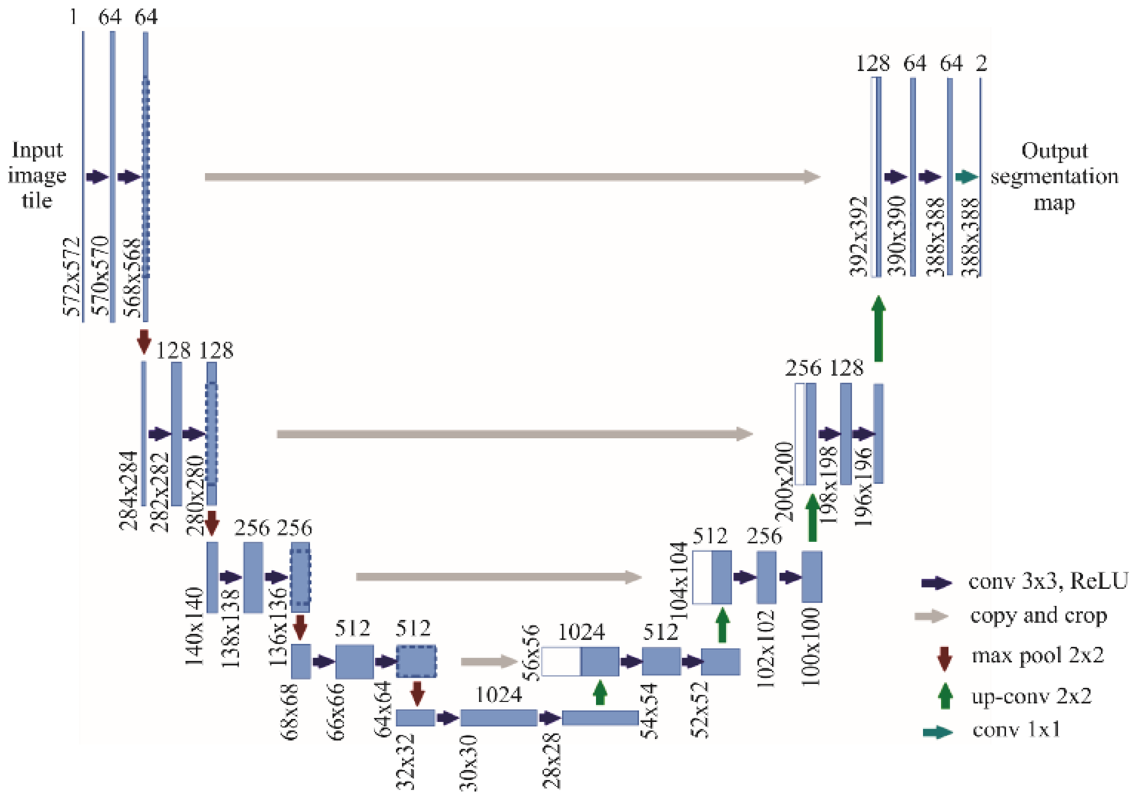

Figure 2. U-Net Model

2.5 Synthetic aperture radar

SAR is a microwave imaging sensor consisting of a radar located on an aerial platform. This radar achieves high resolution (in the longitudinal/azimuthal direction) by mixed echoes from varied pulses while the platform moves, simulating a synthetic antenna larger than the real one.

Now, concerning the enhancements that this sensor has over optical and infrared, it can be noted that when operating at microwave frequencies in bands L: 1 to 2 GHz, C: 3.75 to 7.5 GHz, and: 7.5 to 12 GHz, these can penetrate clouds, smoke or fog and register images of the ground/sea.

Furthermore, being an active sensor (independent lighting that transmits its electromagnetic pulses toward the ground or sea being photographed), it works not only during the day but also at night and in various climates. This allows constant monitoring of the state of the sea, land, weather, and of course OOS [18].

2.6 Sentinel-1

Sentinel-1 is a constellation of 2 polar (180° offset) orbiting satellites that orbit our planet every 6 days and which take radar images to study marine glacial displacement in the Arctic, monitor the marine environment that includes oil spills and ship detection for maritime safety, monitoring of the planet to detect risks of movement, mapping for the management of marine and terrestrial ecosystems and routes for humanitarian support in disaster situations. Sentinel-1 is administered by the European Space Agency (ESA); and it was launched in 2014, and Sentinel-1B in 2016 (which ended in 2022).

Both satellites travelled to space in a Soyuz rocket from the French Guiana. The Sentinel-1C, will be launched in the second half of 2024 on Vega-C Rocket [19].

3. METHODOLOGY

This research aims to prevent oil spills at sea and to achieve this goal, Remote Sensing will be used to identify the factors that allow the occurrence of oil spills at sea and GeoAI will be used, at the level of ML & DL, for the detection or mitigation of spills. Table 1 shows the research questions identified to achieve these objectives.

Table 1. Questions and objectives of the research paper

	Questions	Objectives
Q1	How are RS techniques utilized to detect and monitor OOS?	Identify RS techniques utilized to detect and monitor OOS?
Q2	What role does GeoAI play in enhancing the accuracy and efficiency of OOS Detection.	Highlight GeoAI roleplay in enhancing the accuracy and efficiency of OOS Detection?
Q3	What are the latest advancements in the use of RS and GeoAI for OOS management; and what is the future outlook for research in this field?	Identify the latest advancements in the use of RS and GeoAI for OOS management; and future outlook for research in this field.
Q4	What is the World's main OOS Datasets?	Highlight the main OOS Datasets in the World.

In principle, to have an overview of the research, the PICOC (Population, Intervention, Comparison, Outcomes, Context) method has been used (see Table 2) for the systematic review of the literature. It should be noted that the PARSIFAL platform was used.

In addition, keywords, synonyms, interventions, and inclusion and exclusion criteria were considered in 3 databases: Scopus, IEEE, and Scielo.

Table 2. PICOC method

Population	Papers About Oil Spills, Oil Spill Datasets
Intervention	GeoAI, Sustainable Development Goals
Comparison	Remote Sensing
Outcome	Prevent environmental impacts from oil spills
Context	Environment

Table 3. Keys and queries

Keys	Queries
Search String	(spill AND oil AND machine AND learning) AND PUBYEAR > 2018 AND PUBYEAR < 2024 AND (LIMIT-TO (EXACTKEYWORD, "Oil Spill") OR LIMIT-TO (EXACTKEYWORD, "Oil Spills")
Inclusion Criteria	GeoIA
Exclusion Criteria	Papers more than 5 years old

The main queries are Oil Spills and Machine Learning, inclusion criteria refer to the new technology named GeoAI, and its exclusion criteria are considered over the last 5 years as shown in Table 3. The sources analyzed were reviewed until November 31, 2023.

On the other hand, it was found that, although several countries have indeed updated information, such as Greece, Turkey, Iran, Italy, Sweden, and China.

It is the last one that has greater relevance for this research as shown in Figure 3.

Subsequently, the PRISMA method was used, which allowed a systematic review and meta-analysis of the research to be carried out [20].

The search retrieved 3406 documents after applying the filters: "Oil Spill" and "Machine Learning". Then 51 duplicates were detected, decreasing to 3355. Next, the filter of only considering articles, not books, was applied, yielding 2445.

Subsequently, it was filtered by study areas achieving 2042, then it was filtered by specific technologies finding 1064, then it opted for SAR images, reducing to 130 of which 50 additional articles that support this research were considered as shown in Figure 4.

Below is a summary of the main publications, datasets, algorithms, and metrics of ML & DL models and SAR in recent years (Table 4).

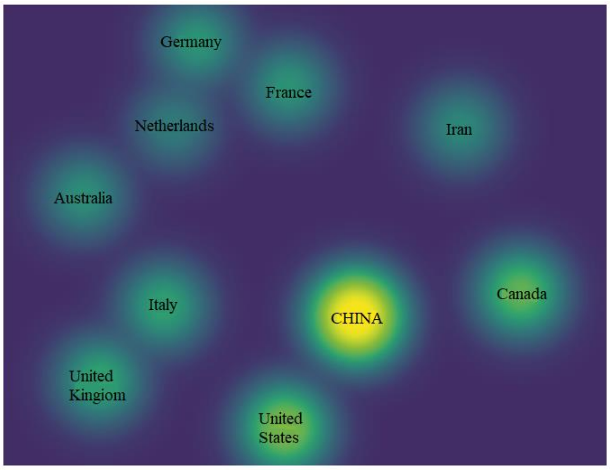

Figure 3. Network display of articles

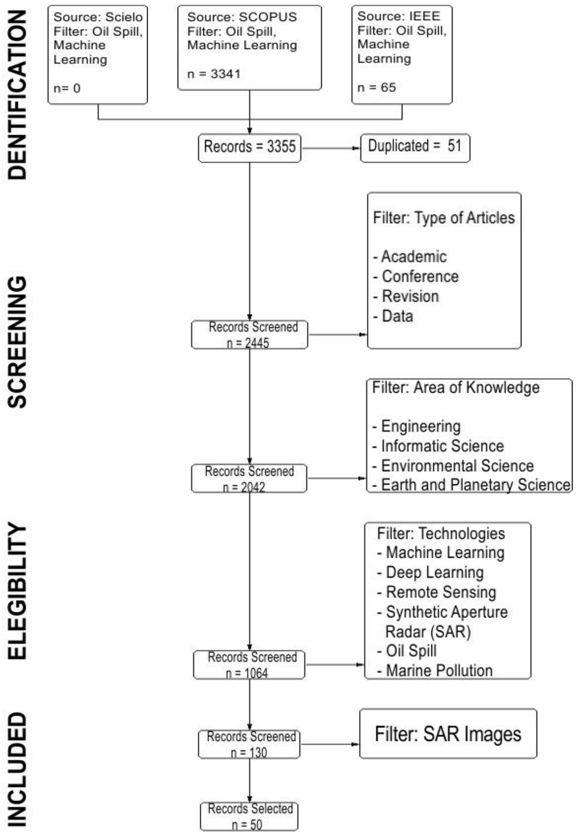

Figure 4. Prisma model

Table 4. Comparative summary of the main publications in SAR, ML and DL between 2018-2023

Cite	Datasets	Algorithm	Metric
[21]	SAR	ML-DL	Mixed
[22]	Multiples	ML-DL	Mixed
[23]	SAR	Extremo ML	---
[24]	LIF-LiDAR	ML + NN	100.0 %
[25]	SAR- HOSD	ML SVM+ERW	96.50%
[26]	SAR - AIS	ML + DT	42.55 %
[27]	SAR - NOWPAP	CNN + Adversarial F-Divergence	96.2%
[28]	SAR 23 SLAR images	DL CNN Two Stages	98.34%
[29]	SAR *Dataset 1 **Dataset 2 ***Dataset 3	CNN-SVM	*99.19% **99.89% ***97.56%

[30]	SAR UAVSAR	DL CNN + Sensitivity to dielectric constant	99.0%
[31]	SAR - 783 internet oil images	DL CNN + GoogleNet + VGG16	97.5%
[32]	SAR C band - Radarsat-2	DL CV CNN	97.69%
[33]	10 Oil products	DL RPNet	83.67%
[34]	SAR	CNN 1-D	F1 score = 97.90%
[35]	SAR Sentinel-1	DL CNN + UNET	mIoU = 75.7%.
[36]	SAR Sentinel-1	DL CNN U-Net + DenseNet201	96.53%
[37]	SAR Sentinel-1	UNET+ DenseNet	80.31%
[38]	Images from oil spill video	CNN + SSWT-UNet++	98.5 %
[39]	SAR Sentinel 1	DL CNN Deeplabv3 + con Resnet101	MioU = 66.5%
[40]	SAR Sentinel 1	DL CNN Deeplabv3 + Resnet101	98.92%
[41]	SAR Sentinel 1	DCNN	92.3%
[42]	SAR - 839 Images Radarsat-2 y Sentinel-1	DL+ CNN U-Net++	98.3%
[43]	SAR AVIRIS - HOSD	DL CNN + SSTNet	95.96%
[44]	SAR *Simulated **Penglai ***DWH	DL	*100.00% **93.86% ***85.84%
[45]	SAR AVIRIS y Landsat 8	DNN	92.0 %
[46]	SAR *PU **SV ***IP ****UH2018	CNN + 3-D-CAE	*63.9% **85.5% ***51.7% ****29.4%
[47]	SAR	DL CNN + YOLO	95.3%
[48]	SAR - 2487 satellite images	YOLOX-S + FFEDNet	86.76%
[49]	SAR ERS, Envisat	ML	90.5%

4. RESULTS AND DISCUSSION

Recent publications have made it possible to categorize Artificial Intelligence into several levels, one of them being ML and its respective sublevel DL which are diverse and specialized algorithms that allow, for example, segmenting areas, detection of edges, delimit thresholds, and monitoring, and even more, predict more and more accurately oil spills. This section seeks to answer the research questions defined earlier in this article.

4.1 Q1: How are RS techniques utilized to detect and monitor offshore oil spills?

Traditional RS uses Satellite Imagery (SI), images or radar sensors, text, and atmosphere-ocean models from Satellites or Unmanned Aerial Vehicles (UAVs) or Drones to detect spills in oceans, seas, rivers, lakes, lagoons, and other bodies of water [50, 21]. Traditional RS studies use pixel-based methods for detecting OOS [51] as evidenced below: classification of OOS and lookalikes using qualitative analysis methods and contextual features [52], monitoring using SAR SI and Interferometric Wide Swath (IW) mode [53], monitoring using dynamic window scanning of SI (and human expertise), and extraction of information [54].

4.2 Q2: What role does GeoAI play in enhancing the accuracy and efficiency of offshore oil spill detection?

Concerning GeoAI, it is the integration of AI (ML/DL) methods and techniques, geodata, and spatial inference methods, such as Kriging, that allow the generation of new predictive geospatial models [55].

One of the advantages of GeoAI is that it fosters greater access to a large and varied amount of geospatial data (Geospatial Big Data), and it is linked to Information and Communication Technologies (ICT), such as, for example, through Spatial Data Infrastructures (SDI), it would allow the interaction between private geodata and government data, reducing duplication and promoting better geospatial analysis by users and researchers [22].

The development of GeoAI improves Remote Sensing (RS) [56-58] allowing such ML/DL models to solve RS problems by identifying hidden patterns, performing segmentation and feature detection in satellite images, both natural (vegetation) and artificial (roads) and detecting oil spills, committed by man or nature.

ML build and train models that identify and classify (label) such satellite images using several approaches such as the ability to perform regression processes [59] so they are considered the latest technology within the development of the Industry 4.0 (4IR) [60, 61], the texturing of the characteristics or characterization of Tamura that resembles the sense of human vision [23], the personalized vision on a cloud-based computing platform, such as the Microsoft Azure Machine-Learning Service [49] the measurement of the thicknesses of the spectra of diverse kind of oil through Laser-Induced Fluorescence Lidar (LIF-LiDAR) [24], the development of an unsupervised Isolation Forest (iForest) method for Hyperspectral Segmentation Images (HSI) [25] and oil spill prediction using algorithms such as Linear Regression (LR), Support Vector Machine (SVM) and Decision Tree based on the dataset generated by Automatic Identification Systems (AIS), i.e. position, course and speed of the ship, container characteristics, classes of oil, skimming rate, etc. [26].

4.3 Q3: What are the latest advancements in use of RS and GeoAI for oil spill management, and what is the future outlook for research in this field?

Conventional Neural Networks (CNNs) are presented as a more sophisticated Deep Learning (DL) model to perform semantic segmentation, having among its main exponents the techniques of: Adversarial F-Divergence Learning, which allows detecting spills without manual initialization (automatically) [27], the Side-Looking Airborne Radar (SLAR) to generate hyperspectral images on which the first CNN performs an approximate detection of objects (or spills) and then, a second specialized one, precisely locates the pixel correspondence by class [28]; the Radial Basis Function Kernel (RBF-SVM) method is used for the classification of three sets of polarimetric RADARSAT-2 SARs based on PolSAR data converted into a 9-channel data block entered into 5 layers from which 2 high-level features are extracted and merged by principal component analysis (PCA) for validation [29]; the learning of non-linear characteristics, textural and statistical shapes and patterns, which achieves better accuracy in classification based on the parameters of sensitivity to the dielectric constant and the damping properties of waves [30]; the VGG16 enhanced with Transfer Learning, developed on the GoogleNet platform and trained on a proprietary dataset obtained from spill images acquired from the internet [31]; the Complex Value (CVCNN) that outperforms the Real Value (RVCNN) both in the detection of spills in the ocean, the classification of crude oil and biogenic [32]; the RPnet (Random Patch Network) with fusion of multiple features in hyperspectral images, which through a drone, evaluates the monitoring of spills and through a sea sample of five petroleum products, under controlled conditions in containment facilities, optically characterizing the different types of contamination by oil spills [33] and the 1-D model with rapid inference capability that directly processes the data of the absorption spectrum of Terahertz (THz) radiation waves of the sample, demonstrating that this CNN improves accuracy, has high consistency and great robustness in the detection of contamination [34].

However, the most frequently used CNN in classification and segmentation is U-Net, which was originally proposed for the analysis of biomedical images and later modified to be used to find spills [35]. In the following years, based on this U-Net, considerable improvements were made to the model: the level of the IoU (Intersection on Junction) of the oil spill detection systems was increased through the processing of a dataset in 3 stages: image resizing, data augmentation and application of a hot encoder (in its mask) and use of the DenseNet201 model as part of the encoder and U-Net as a decoder [36]; challenging the characterization of objects (ships, pipelines) and events (oil slicks and/or spills) from unbalanced datasets by using a U-Net model with DenseNet as an encoder [37], and concerning the scarcity and poor quality of underwater images to recognize oil spills, new network architecture suitable superpixels with texture (SSWT)-UNet++ based on texture superpixels is proposed., processed them as input, adjusted the resolution of each layer and changed its output to intermediate layers to improve its classification compared to the previous methods mentioned [38].

4.4 Q4: What is the world's main OOSD?

Regarding datasets, it should be noted that one of the most widely used to classify possible oil leaks at sea, ships, and soils are the Sentinel-1 SAR images [62], which have been processed with different methods, such as DL semantic segmentation with Deeplabv3 + network with Resnet101 [39] and, improving on the previous segmentation, the use of polarimetric SAR (PolSAR) [40]; the inventory of temporal redundancy of natural oil spills in singular zones in sea, discriminating between natural seepage and man-made spillage [41].

Further, other widely used datasets are Radarsat-2 [42] and HOSD, Hyperspectral Remote Sensing Oil Spill benchmark Database) based on the Gulf of Mexico (GoM) spill [43] and, in a private way, the spill that occurred on the Penglai 19-3 oil platform in China [44].

In the line of Optical Sensors, AVIRIS is widely used for spill mapping, after segmentation of hyperspectral images, in which reflectance spectra can be measured at different resolutions with ML models [45] or with Neural Networks (NN) for Unsupervised models using a new Spectral Loss (SL, Spectral Loss function) [46].

4.6 Results and Discussion

Traditional RS has been enhanced with the development of AI (ML/DL). Nowadays real-time tracking of spills from ships and SAR satellites is possible, and even more so, the delimitation of spills in images processed with CCN using (YOLO) You Only Look Once [47] as in the case of noise reduction by speckle [48].

On the other hand, statistical models such as logistic regression, regularized regression (elastic network), and SVM are used to predict the occurrence of failures in the machines, while Cox regression is used to know how good equipment is through time.

However, in the hydrocarbon industry, predictive maintenance should be the first choice to avoid disasters. To this end, in the phase of oil transport and distribution, variables such as weight, volume, and temperature must be collected, analysed, associated, and contrasted [63].

In addition, it should be considered that there is much work in far areas or carried out at sea by vessels whose safety will depend on their preparation for failures and acquisition of specialized equipment [64].

This new area of research, GeoAI, based on the systematic review, is sustained in traditional disciplines such as RS-GIS and in turn directed toward the future considering the models of Machine Learning to make predictions with precisions of 42.55 % [38], 90.5 % [49] and 96.5 % [25].

On the other hand, to determine oil spills with Deep Learning, precisions reached 99.0 % [30], (98.92 %) [40], and (99.19 %) [29].

5. CONCLUSIONS

Detection of offshore oil spills at sea has traditionally been carried out using Remote Sensing by analysis of oceanic atmospheric models from RGB, multispectral, and hyperspectral optical satellites, and SAR images whose sensors are transported on satellites, airplanes, and UAVs.

The predominant role of GeoAI is that it allows to use of methods, techniques, and performance evaluation metrics such as accuracy, precision, sensitivity, IoU, F1-Score, Confusion Matrix, and ROC-AUC of Artificial Intelligence (ML-DL), to be used within the growing offshore oil spills Big Data to improve identification of hidden patterns, segmentation and feature detection in satellite images.

In the same way that OOS were initially studied using RGB optical images, to be later evaluated with multi- and hyperspectral images and currently analysed with SAR images, GeoAI has been evolving from ML models (e.g. Support Vector Machine, iForest) to do segmentation and feature detection, towards DL Models, through CNN using different architectures such as DenseNet, SSTNet, SSWTNet, and U-Net (U-Net++) being last one best performance. Future research will focus on improving CCN runtime by evolving toward Quantum Convolutional Neural Networks (QCNN).

Although it is an unfortunate reality that oil spills at sea continue to grow, this data is the property of transportation companies and oil companies, which are private. However, in the present investigation, two valuables public OOSD have been identified: HOSD project composed of 18 hyperspectral satellite images (2.19 Gb .mat files) and MKLab project composed of 1112 SAR images (336 Mb .jpg, and .png files) from Sentinel-1 Mission (Copernicus-ESA) and Clean SeaNet Service. Both datasets are available to the scientific community upon request to the respective research institutes.

ACKNOWLEDGMENT

This work is supported by the Faculty of Systems and Computer Engineering Universidad Nacional Mayor de San Marcos UNMSM, Lima-Peru.

REFERENCES

[1] Jiang, N., Li, J.Y. (2020). An improved semantic segmentation method for Remote Sensing images based on neural network. Traitement du Signal, 37(2): 271-278. https://doi.org/10.18280/ts.370213

[2] McCarthy, J., Minsky, M., Rochester, N., Shannon, C.E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence. http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html.

[3] McCarthy, J. (1960). Recursive functions of symbolic expressions and their computation by machine, Part I. Association for Computing Machinery, 3(4): 184-195. https://doi.org/10.1145/367177.367199

[4] Ponce, J., Torres, A., Quezada, F., Silva, A., Martínez, E., Casali, A., Scheihing, A., Tupac, Y., Torres, M., Ornelas, F., Hernández, J., Zavala, C., Vakhnia, N., Pedreño, O. (2014). Inteligencia Artificial. Iniciativa Latinoamericana de Libros de Texto Abiertos (LATIn), 53-53. https://doi.org/10.13140/2.1.3720.0960

[5] Wolf, A. (2022). Machine Learning Simplified: A Gentle Introduction to Supervised Learning. Machine Learning, 6-6. https://themlsbook.com/.

[6] Campesato, O. (2020). Artificial Intelligence, Machine Learning, and Deep Learning. Mercury Learning and Information, 98-99. https://www.perlego.com/es/book/3026669/artificial-intelligence-machine-learning-and-deep-learning-pdf.

[7] Gao, S., Hu, Y., Li, W. (2023). Handbook of Geospatial Artificial Intelligence. CRC Press. https://doi.org/10.1201/9781003308423

[8] Jolliff, J.K., Smith, T.A., Ladner, S., Arnone, R.A. (2014). Simulating surface oil transport during the Deepwater Horizon oil spill: Experiments with the BioCast system. Ocean Modelling, 75: 84-99. https://doi.org/10.1016/j.ocemod.2014.01.004

[9] Fingas, V. (2012). The Basics of Oil Spill Cleanup. CRC Press. https://doi.org/10.1201/b13686

[10] Sidorov, G. (2011). Métodos modernos de Inteligencia Artificial. Centro de Investigación en Computación. https://www.cic.ipn.mx/~sidorov/libroIA.pdf.

[11] Burvov, A. (2019). The Hundred-Page Machine Learning Book, 3-4. https://themlbook.com/.

[12] Chollet, F. (2021). Deep Learning with Python, Second Edition. Manning Publications. https://www.manning.com/books/deep-learning-with-python-second-edition

[13] Szeliski, R. (2022). Computer Vision Algorithms and Applications, 2nd ed. Springer. https://szeliski.org/Book/.

[14] Ronneberger, O., Fischer, P., Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. Computer Vision and Pattern Recognition, 234-241. https://doi.org/10.48550/arXiv.1505.04597

[15] Mao, H., Hu, Y., Kar, B., Gao, S., McKenzie, G. (2018). GeoAI 2017 workshop report: the 1st ACM SIGSPATIAL International Workshop on GeoAI: @AI and Deep Learning for Geographic Knowledge Discovery: Redondo Beach, CA, USA - November 7, 2016. Sigspatial Special, 9(3): 25. https://doi.org/10.1145/3178392.3178408

[16] Vopham, T., Hart, J., Laden, F., Chiang, Y. (2018). Emerging trends in geospatial artificial intelligence (geoAI): Potential applications for environmental epidemiology. Environmental Health, 17(40): 3. https://doi.org/10.1186/s12940-018-0386-x

[17] Campbell, J.B., Wynne, R.H., Thomas, V.A. (2023). Introduction to Remote Sensing Sixth Edition. The Guilford Press. https://www.guilford.com/books/Introduction-to-Remote-Sensing/Campbell-Wynne-Thomas/9781462549405.

[18] Di Martino, G., Iodice, A. (2021). Maritime Surveillance with Synthetic Aperture Radar. The Institution of Engineering and Technology, 4. https://doi.org/10.1049/SBRA521E

[19] European Space Agency. (2023). Sentinel-1 Radar visión for Copernicus. https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1.

[20] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ Publishing Group, 372. https://doi.org/10.1136/bmj.n71

[21] Huby, A.A., Sagban, R., Alubady, R. (2022). Oil spill detection based on machine learning and deep learning: A Review. In IICETA 2022 - 5th International Conference on Engineering Technology and Its Applications, Al-Najaf, Iraq, pp. 85-90. https://doi.org/10.1109/IICETA54559.2022.9888651

[22] Nugroho, Y.S., Supangkat, S.H. (2021). Spatial data infrastructure integrated with geospatial artificial intelligence: A systematic literature review. In 2021

International Conference on ICT for Smart Society (ICISS), Bandung, Indonesia, pp. 1-5. https://doi.org/10.1109/ICISS53185.2021.9533208
[23] Lyu, X. (2018). Oil spill detection based on features and extreme learning machine method in SAR images. In 2018 3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Huhhot, China, pp. 559-563. https://doi.org/10.1109/ICMCCE.2018.00123
[24] Yin, S., Sun, F., Liu, W., Bi, Z., Liu, Q., Tian, Z. (2023). Remote identification of oil films on water via laser-induced fluorescence LiDAR. IEEE Sensors Journal, 23(12): 13671-13679. https://doi.org/10.1109/JSEN.2023.3271370
[25] Duan, P., Kang, X., Ghamisi, P., Li, S. (2023). Hyperspectral Remote Sensing benchmark database for oil spill detection with an isolation forest-guided unsupervised detector. IEEE Transactions on Geoscience and Remote Sensing, 61: 1-11. https://doi.org/10.1109/TGRS.2023.3268944
[26] Sandhiya, S., Ramya, R., Sukitha, M., Vikhashni, S. (2023). Detection of oil spill events at sea using machine learning. In 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, pp. 53-58. https://doi.org/10.1109/ICIRCA57980.2023.10220760
[27] Yu, X., Zhang, H., Luo, C., Qi, H., Ren, P. (2018). Oil spill segmentation via adversarial f-divergence learning. IEEE Transactions on Geoscience and Remote Sensing, 56(9): 4973-4988. https://doi.org/10.1109/TGRS.2018.2803038
[28] Nieto-Hidalgo, M., Gallego, A.J., Gil, P., Pertusa, A. (2018). Two-stage convolutional neural network for ship and spill detection using SLAR images. IEEE Transactions on Geoscience and Remote Sensing, 56(9): 5217-5230. https://doi.org/10.1109/TGRS.2018.2812619
[29] Song, D., Zhen, Z., Wang, B., Li, X., Gao, L., Wang, N., Xie, T., Zhang, T. (2020). A novel marine oil spillage identification scheme based on convolution neural network feature extraction from fully polarimetric SAR imagery. IEEE Access, 8: 59801-59820. https://doi.org/10.1109/ACCESS.2020.2979219
[30] De Laurentiis, L., Jones, C.E., Holt, B., Schiavon, G., Del Frate, F. (2021). Deep learning for mineral and biogenic oil slick classification with airborne synthetic aperture radar data. IEEE Transactions on Geoscience and Remote Sensing, 59(10): 8455-8469. https://doi.org/10.1109/TGRS.2020.3034722
[31] Feinauer, D.M., Tayem, N., Latif, G., Alghazo, J., Alenazy, A.M., Alzubaidi, L. (2022). Oil spill identification using deep convolutional neural networks. In 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), Al-Khobar, Saudi Arabia, pp. 240-245. https://doi.org/10.1109/CICN56167.2022.10008373
[32] Li, Y., Yang, J., Yuan, Z., Zhang, Y. (2022). Marine oil spills detection and classification from Polsar images based on complex-valued convolutional neural network. IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, pp. 7085-7088. https://doi.org/10.1109/IGARSS46834.2022.9883991
[33] Yang, J., Hu, Y., Ma, Y., Li, Z., Zhang, J. (2022). Research on oil spill pollution type identification using RPNET deep learning model and airborne hyperspectral image. IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, pp. 807-810. https://doi.org/10.1109/IGARSS46834.2022.9883530
[34] Zhan, H., Meng, Z., Ren, Z., Miao, X., Bao, R., Zhao, K. (2022). Terahertz spectroscopy combined with deep learning for predicting the depth and duration of underground sand pollution by crude oil. IEEE Transactions on Instrumentation and Measurement, 71: 1-8. https://doi.org/10.1109/TIM.2021.3105237
[35] Basit, A., Siddique, M.A., Sarfraz, M.S. (2021). Deep learning based oil spill classification using unet convolutional neural network. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, pp. 3491-3494. https://doi.org/10.1109/IGARSS47720.2021.9553646
[36] Huby, A.A., Alubady, R., Sagban, R. (2022). Oil spill segmentation from SAR images using deep neural networks. In 2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, pp. 282-287. https://doi.org/10.1109/ISMSIT56059.2022.9932853
[37] Ghonaime, T.W., Al-Sadek, A.F. (2023). OilSpillNet: Deep learning model for synthetic aperture radar imagery oil spill detection. In 2023 Intelligent Methods, Systems, and Applications (IMSA), Giza, Egypt, pp. 412-417. https://doi.org/10.1109/IMSA58542.2023.10217532
[38] Song, H., Gao, L., Song, J. (2023). SSWT-UUNet++ Submarine oil pipeline spill detection approach. IEEE Geoscience and Remote Sensing Letters, 20: 1-5. https://doi.org/10.1109/LGRS.2023.3268348
[39] Luo, D., Gu, C., Chen, P., Yang, J., Yuan, Y., Zheng, G., Ren, L. (2021). Oil spill segmentation of SAR image based on improved deep convolution neural network. 2021 Photonics & Electromagnetics Research Symposium (PIERS), Hangzhou, China, pp. 2232-2237. https://doi.org/10.1109/PIERS53385.2021.9694767
[40] Ma, X., Xu, J., Wu, P., Kong, P. (2022). Oil spill detection based on deep convolutional neural networks using polarimetric scattering information from Sentinel-1 SAR images. IEEE Transactions on Geoscience and Remote Sensing, 60: 1-13. https://doi.org/10.1109/TGRS.2021.3126175
[41] Scardigli, A., Risser, L., Haddouche, C., Jatiault, R. (2023). Integrating unordered time frames in neural networks: Application to the detection of natural oil slicks in satellite images. IEEE Transactions on Geoscience and Remote Sensing, 61: 1-14. https://doi.org/10.1109/TGRS.2023.3241681
[42] Ronci, F., Avolio, C., Di Donna, M., Zavagli, M., Piccialli, V., Costantini, M. (2020). Oil spill detection from SAR images by deep learning. In IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, pp. 2225-2228. https://doi.org/10.1109/IGARSS39084.2020.9323590
[43] Kang, X., Deng, B., Duan, P., Wei, X., Li, S. (2023). Self-supervised spectral-spatial transformer network for hyperspectral oil spill mapping. IEEE Transactions on Geoscience and Remote Sensing, 61: 1-10. https://doi.org/10.1109/TGRS.2023.3260987
[44] Kang, X., Wang, Z., Duan, P., Wei, X. (2022). The potential of hyperspectral image classification for oil

spill mapping. IEEE Transactions on Geoscience and Remote Sensing, 60: 1-15. https://doi.org/10.1109/TGRS.2022.3205966
[45] Xie, M., Li, Y., Dong, S., Zhang, B., Gou, T. (2022). Fine-grained oil types identification based on reflectance spectrum: implication for the requirements on the spectral resolution of hyperspectral remote sensors. IEEE Geoscience and Remote Sensing Letters, 19: 1-15. https://doi.org/10.1109/LGRS.2022.3176494
[46] Perez-Garcia, A., Paoletti, M.E., Haut, J.M., Lopez, J.F. (2023). Novel spectral loss function for unsupervised hyperspectral image segmentation. IEEE Geoscience and Remote Sensing Letters, 20: 1-15. https://doi.org/10.1109/LGRS.2023.3288809
[47] Adhikary, S., Tiwari, S.P., Banerjee, S. (2022). Realtime oil spill detection by image processing of synthetic aperture radar data. In 2022 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, pp. 1-5. https://doi.org/10.1109/CONECCT55679.2022.9865768
[48] Xu, F.Y., An, X.Z., Liu, W.Q. (2022). Oil spill detection in SAR images based on improved YOLOX-S. In 2022 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI), Shijiazhuang, China, pp. 261-265. https://doi.org/10.1109/ICCEAI55464.2022.00061
[49] Almeer, S., Albalooshi, F., Alhajeri, A. (2021). Oil spill detection system in the Arabian gulf region: An azure machine-learning approach. In 2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Zallaq, Bahrain, pp. 418-422. https://doi.org/10.1109/3ICT53449.2021.9581841
[50] Lavallin, A., Downs, J.A. (2021). Machine learning in geography–Past, present, and future. Geography Compass, 15(5): e12563. https://doi.org/10.1111/gec3.12563
[51] Jiang, W., Zhang, L. (2018). Geospatial data to images: A deep-learning framework for traffic forecasting. Tsinghua Science and Technology, 24(1): 52-64. https://doi.org/10.26599/TST.2018.9010033
[52] Holstein, A., Kappas, M., Propastin, P., Renchin, T. (2018). Oil spill detection in the Kazakhstan sector of the Caspian Sea with the help of ENVISAT ASAR data. Environmental Earth Sciences, 77: 1-11. https://doi.org/10.1007/s12665-018-7347-0
[53] Eronat, A.H., Bengil, F., Neser, G. (2019). Shipping and ship recycling related oil pollution detection in Candarh Bay (Turkey) using satellite monitoring. Ocean Engineering, 187: 106157. https://doi.org/10.1016/j.oceaneng.2019.106157
[54] Samra, R.M.A., Eissa, R., El-Gammal, M. (2018). Applying the environmental sensitivity index for the assessment of the prospective oil spills along the Nile Delta Coast, Egypt. Geocarto International, 35(6): 589-601. https://doi.org/10.1080/10106049.2018.1533592
[55] Babaan, J., Hsu, F.T., Wong, P.Y., Chen, P.C., Guo, Y.L., Lung, S.C., Chen, Y.C., Wu, C.D. (2023). A Geo-AI-based ensemble mixed spatial prediction model with fine spatial-temporal resolution for estimating daytime/nighttime/daily average ozone concentrations variations in Taiwan. Hazardous Materials, 446: 130749. https://doi.org/10.1016/j.jhazmat.2023.130749
[56] Li, D., Shao, Z., Zhang, R. (2020). Advances of geo-spatial intelligence at LIESMARS. Geo-Spatial Information Science, 23(1): 40-51. https://doi.org/10.1080/10095020.2020.1718001
[57] Usery, E.L., Arundel, S.T., Shavers, E., Stanislawski, L., Thiem, P., Varanka, D. (2022). GeoAI in the US Geological Survey for topographic mapping. Transactions in GIS, 26(1): 25-40. https://doi.org/10.1111/tgis.12830
[58] Richter, K.F., Scheider, S. (2023). Current topics and challenges in geoAI. KI-Künstliche Intelligenz, 37(1): 11-16. https://doi.org/10.1007/s13218-022-00796-0
[59] Maxwell, A., Warner, T., Fang, F. (2018) Implementation of machine-learning classification in Remote Sensing: an applied review. International Journal of Remote Sensing, 39(9): 2784-2817. https://doi.org/10.1080/01431161.2018.1433343
[60] Ślusarczyk, B. (2018). Industry 4.0-Are we ready? Polish Journal of Management Studies, 17(1): 232-248. https://doi.org/10.17512/pjms.2018.17.1.19
[61] Sarker, I.H. (2021). Machine learning: Algorithms, real-world applications and research directions. In SN Computer Science, 2(3): 160. https://doi.org/10.1007/s42979-021-00592-x
[62] Krestenitis, M., Orfanidis, G., Ioannidis, K., Avgerinakis, K., Vrochidis, S., Kompatsiaris, I. (2019). Oil spill identification from satellite images using deep neural networks. Remote Sensing, 11: 1762. https://doi.org/10.3390/rs11151762
[63] Desai, J.N., Pandian, S., Vij, R.K. (2021). Big data analytics in upstream oil and gas industries for sustainable exploration and development: A review. Environmental Technology and Innovation, 21: 101186. https://doi.org/10.1016/j.eti.2020.101186
[64] Cadei, L., Montini, M., Landi, F., Porcelli, F., Michetti, V., Origgi, M., Tonegutti, M., Duranton, S. (2018). Big data advanced anlytics to forecast operational upsets in upstream production system. In Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE. https://doi.org/10.2118/193190-MS