Deep Learning-Based Approaches for Oil Spill Detection: A Bibliometric Review of Research Trends and Challenges

Rodrigo N. Vasconcelos 1,2,3,*; André T. Cunha Lima 2,4,5; Carlos A. D. Lentini 2,4,6,7; José Garcia V. Miranda 2; Luís F. F. de Mendonça 6,7,8; José M. Lopes 2,6; Mariana M. M. Santana 9; Elaine C. B. Cambuí 10; Deorgia T. M. Souza 1; Diego P. Costa 3,4; Soltan G. Duverger 3,11 and Washington S. Franca-Rocha 1

1 State University of Feira de Santana—UEFS, Feira de Santana 44036-900, Bahia, Brazil; dtmsouza@uefs.br (D.T.M.S.); wjsrocha@uefs.br (W.S.F.-R.)
2 Department of Earth and Environment Physics, Physics Institute, Campus Ondina, Federal University of Bahia—UFBA, Salvador 40170-280, Bahia, Brazil; acl.cunhalima@ufba.br (A.T.C.L.); cleninti@ufba.br (C.A.D.L.); viva@ufba.br (J.G.V.M.); josemi@ufba.br (J.M.L.)
3 GEODATIN—Data Intelligence and Geoinformation, Bahia Technological Park Rua Mundo, 121—Trobogy, Salvador 41301-110, Bahia, Brazil; diego.costa@ufba.br (D.P.C.); solta.galano@ufba.br (S.G.D.)
4 Interdisciplinary Center for Energy and Environment (CIEnAm), Federal University of Bahia UFBA, Salvador 40170-115, Bahia, Brazil
5 SENAI-Cimatec, Salvador 41650-010, Bahia, Brazil
6 Geosciences Institute (IGEO/UFBA), Federal University of Bahia—UFBA, Salvador 40170-115, Bahia, Brazil; luis.mendonca@ufba.br
7 Geosciences Institute (PPGEOF), Federal University of Bahia—UFBA, Salvador 40170-115, Bahia, Brazil
8 Department of Oceanography, Geoscience Institute, Campus Ondina, Federal University of Bahia—UFBA, Salvador 40170-280, Bahia, Brazil
9 Forest Engineering Institute (FEI/UEAP), State University of Amapá—UEAP, Av. Pres. Getúlio Vargas, 650 Centro, Macapá 68900-070, Amapá, Brazil; mariana.medeiros@ueap.edu.br
10 Institute of Biology, Federal University of Bahia—UFBA, Salvador 40170-115, Bahia, Brazil; elainebarbosa@ufba.com
11 Federal University of Bahia—UFBA, Salvador 40110-100, Bahia, Brazil
* Correspondence: rnnv5e5ppgm@gmail.com

Abstract: Oil spill detection and mapping using deep learning (OSDMDL) is crucial for assessing its impact on coastal and marine ecosystems. A novel approach was employed in this study to evaluate the scientific literature in this field through bibliometric analysis and literature review. The Scopus database was used to evaluate the relevant scientific literature in this field, followed by a bibliometric analysis to extract additional information, such as architecture type, country collaboration, and most cited papers. The findings highlight significant advancements in oil detection at sea, with a strong correlation between technological evolution in detection methods and improved remote sensing data acquisition. Multilayer perceptrons (MLP) emerged as the most prominent neural network architecture in 11 studies, followed by a convolutional neural network (CNN) in 5 studies. U-Net, DeepLabv3+, and fully convolutional network (FCN) were each used in three studies, demonstrating their relative significance too. The analysis provides insights into collaboration, interdisciplinarity, and research methodology and contributes to the development of more effective policies, strategies, and technologies for mitigating the environmental impact of oil spills in OSDML.

Keywords: SAR; remote sensing; oil spills
1. Introduction

Oil spills in the ocean have become one of the most significant environmental issues in modern times [1–6]. These spills cause severe damage to ecosystems, biodiversity, and the loss of essential ecosystem processes [7,8]. Furthermore, the impact of oil spills is not limited to the environment, as they can also adversely affect the economy and public health [9,10].

Despite the annual decline in oil spills, catastrophic releases still occur in oil production and transportation [11–13]. More recent examples include the incidents related to the Exxon Valdez oil spill in 1989, the Hebei Spirit oil spill in 2007, and the launch of the Deepwater Horizon in 2010.

Ocean oil spills are a serious environmental concern, given the recent increase in environmental disasters, [1,2,14]. Their consequences can be devastating, leading to the degradation of entire ecosystems, loss of biodiversity, and disturbance of critical ecological processes [15,16].

In this sense, remote sensing has become a popular science among researchers, who have increasingly used remote data for oil spill detection and mapping (OSDM) to monitor, survey, and manage its associated risks. Despite the advances, there is still a need to establish a consensus on the most effective methods for OSDM [3–6,17–19]. Therefore, a comprehensive overview of the existing literature in this field is essential to identify the most promising and effective techniques, for example [3,4].

In light of the recent growing occurrence and seriousness of oil spill incidents, there is an urgent requirement to enhance the dependability and precision of methods used in OSDM [3–6,17–19]. Techniques of detection, monitoring, and classifying oil spills in oceans and seas are challenging [3,4], so researchers have introduced machine learning algorithms to solve them.

Numerous machine learning techniques have been used for detecting oil spills. Decision trees [20], support vector machines (SVM) [21,22], random forests [23–25], and artificial neural networks (ANN) are the most common techniques described in the literature [26–30].

Among the machine learning techniques, deep learning has recently received more attention [14,30–33]. In general terms, deep learning is categorized as a subdivision of machine learning methods [34]. Unlike conventional machine learning algorithms that rely on predetermined features, deep learning algorithms gain knowledge directly from the data [31,33,34]. Several advances have recently been made in deep learning architectures for detecting oil spills [2,30,33]. In this sense, describing the current state and trends in the literature field of oil spill detection and mapping with deep learning is essential to consolidate practical analytical techniques and approaches to identify and monitor it and deepen the knowledge and maturation of this scientific field [6,14,30].

Bibliometrics is an influential tool researchers use to gain insights into global trends and developments within a specific topic [35,36]. This method involves applying mathematical and statistical tools to analyze the published literature across various academic disciplines identifying trends and patterns, and the impact of research efforts by individuals, research groups, institutions, countries, and journals [1,37–40]. Moreover, bibliometric reviews can help researchers to identify key contributors to a field of research and determine which research areas may require more attention or investment [36,41,42], which can help funding initiatives, collaboration strategies, and policy decisions [1,37,38].

This study helps to fill the knowledge gaps and identify the current state and trends in the scientific literature of oil spill detection and mapping with deep learning (OSDMDL) using bibliometric analysis. Several questions are addressed, including the most significant country contributions, publication trends, leading researchers, and influential journals. Overall, it provides a comprehensive overview of the scientific research on OSDMDL, identifying possible avenues for further research.

The manuscript is structured as follows. Section 2 provides a detailed description of the materials and methods employed in the study. It includes information about the
search strategies and conducted analysis. Section 3 presents the results related to the publishing trends of oil spill detection and mapping (OSDMDL), including co-occurrence networks, top-cited authors, countries, and journals. Section 4 discusses the findings and their implications. Finally, the manuscript concludes with the concluding remarks section, which summarizes the key points and highlights the significance of the study’s outcomes.

2. Materials and Methods

The proposed research framework is structured into two main steps, from database selection to graphical analysis, as described in Table 1, and is illustrated in the flowchart diagram (Figure 1).

<table>
<thead>
<tr>
<th>Questions</th>
<th>Analysis</th>
<th>Source data</th>
</tr>
</thead>
<tbody>
<tr>
<td>How do OSDMDL study publication trends behave?</td>
<td>General statistics/word Co-occurrence network/country collaboration spatial network</td>
<td>All papers</td>
</tr>
<tr>
<td>Which countries stand out in terms of OSDMDL production?</td>
<td>General statistics/word Co-occurrence network/country collaboration spatial network</td>
<td>All papers</td>
</tr>
<tr>
<td>What are the most influential papers in the OSDMDL field?</td>
<td>General description tables</td>
<td>Most cited papers</td>
</tr>
<tr>
<td>Which journals are most prominent in terms of the number of articles published in the OSDMDL field?</td>
<td>General statistics/general description and citation tables</td>
<td>All papers</td>
</tr>
<tr>
<td>What is the overall picture of collaboration between countries regarding the OSDMDL field?</td>
<td>General statistics/general description and citation tables/country collaboration spatial network</td>
<td>All papers</td>
</tr>
<tr>
<td>What is the central theme, focus, and approach most prominent in research in the OSDMDL field?</td>
<td>Word Co-occurrence network/general description</td>
<td>All</td>
</tr>
</tbody>
</table>
We employed a comprehensive methodological approach that incorporates the fundamental principles of traditional bibliometric analyses, utilizing both qualitative and quantitative descriptors. We integrated various analytical tools to provide a more comprehensive understanding of our target topic. Additionally, we sequentially scrutinized our findings by analyzing the co-occurrence network to comprehend better the patterns associated with implementing remote sensing technology and its advancements in OSDMDL.

2.1. Bibliographic Base

For our research, we have used the Scopus database, which Elsevier established in November 2004. It is a vast bibliographic resource covering various scientific literature from various fields [43]. The database includes citation analysis data from 1996, providing a comprehensive overview of the world’s research products. The Scopus database currently contains over 53 million published references from more than 24,000 scientific journals [43]. The web-based system of the Scopus database offers a range of tools that enable efficient and objective searching of the literature in specific fields using basic and
advanced search queries [43]. These features facilitate rapid and consistent information acquisition, offering a detailed and comprehensive view of scientific fields [43]. Given the characteristics of the Scopus database, we have selected it as the primary literature source for our study (refer to Figure 1). By utilizing this valuable resource, we anticipate obtaining a more thorough understanding of our research subject.

A search query was then formulated to include relevant keywords, phrases, and Boolean operators to obtain the most accurate and precise results. As an initial search strategy, we used the following search query "TITLE-ABS-KEY ("Oil Spill detection" OR "Oil Spill mapping") AND ("Neural network" OR "Deep neural network" OR "Convolutional neural network" OR "CNN" OR "deep belief network" OR "DBN" OR "recurrent neural network" OR "RNN" OR "conditional generative adversarial networks" OR "CGANs" OR "semantic segmentation model" OR "fully convolutional networks" OR "FCNs" OR "Feature Merge Networks" OR "FMNet" OR "U-Shaped Network" OR "U-Net" OR "deep convolutional neural networks" OR "DCNNs" OR "conditional adversarial network" OR "MCAN" OR "Conditional Generative Adversarial Network" OR "CGAN")").

The screening was performed manually, examining all titles and abstracts. When there was uncertainty about a paper’s relevance to our evaluation of the OSDMDP field, we thoroughly read the paper. Irrelevant papers, such as gray literature and review papers, were all excluded. We also excluded conference proceedings, book chapters, and books to ensure we obtained a more focused subset of OSDMDP-related papers published up to 2022. By doing so, we avoided any redundancy that may have arisen from the same content being published multiple times in different literary productions. The subsequent step was to analyze the most frequently cited papers to identify the most authoritative sources.[43]

2.2. Data Analysis

This study utilized the Bibliometrix package [40] and VOSviewer 1.6.17 [44–47] software to conduct quantitative and statistical publications analyses and generate a co-occurrence network of terms, respectively.

Bibliometrix is a statistical programming tool for analyzing scientometrics and bibliometrics data [48]. At the same time, VOSviewer is a specifically designed tool for bibliometric analysis used to create visualization maps of various aspects related to scientific publications, including scientific journals, researchers, research organizations, countries, keywords, and abstracts [44–47].

To create co-occurrence networks of words and terms, we incorporated all the information in the articles’ titles, abstracts, and keywords. We selected the “map based on text data” option in VOSviewer 1.6.17 [44–47] and used bibliographic database files to accomplish this. We employed a binary counting algorithm to count all occurrences of words and terms and then constructed a thesaurus file to avoid semantic errors resulting from a redundancy of meanings (see Table S3 in Supplementary Materials for more details).

To obtain the complete information for constructing the co-occurrence network, we used a threshold value of one associated with the minimum number of occurrences of terms and words. The complete set of terms and words was used to build the network [44–47].

To analyze author production, scientific production in different countries, collaborations between countries, trends over time, author production over time, most cited articles, and the number of publications based on the impact of the source, we utilized the Bibliometrix library [48].

Subsequently, we performed a second division of the literary data, selecting the most cited articles. This selection represented approximately 50% of the total number of selected articles. We extracted various attributes from each article, including data image, digital image processing, temporal data image usage, spatial data resolution, study site,
primary objective, study topic, spectral index, total citation, percentage of total citation, and total citation per year.

All analytical figures and analyses were conducted on R version 4.0.4 [49,50], using the Rstudio IDE, version 1.4.1106 [51], along with the ggplot2 version 3.3.5 [52] and Bibliometrix—version 3.1.4. libraries [48]. Table S3 contains a tabulated summary of the parameters employed in constructing the co-occurrence network, countries’ collaboration world map, and thesaurus file.

3. Results

3.1. Publishing Trends

After refining the database and carefully reviewing the literature, we identified 70 published documents using the OSDMDL methodology (see Table 2, Figure 2, and Table S1 for more details). Regarding publication trends, the production of OSDMDL-related articles has shown some inconsistency over the years. Specifically, the mean and standard deviation are ~2.6 ± 3.8 papers/year.

Table 2. General statistics are associated with production sources, papers by parents, funding institutions, authors, and affiliation. In the center of the graph is the co-occurrence network of terms and words. The different colors represent different clusters.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sources (Journals)</td>
<td>40</td>
<td>1</td>
<td>4</td>
<td>24</td>
<td>17</td>
</tr>
<tr>
<td>Documents</td>
<td>70</td>
<td>1</td>
<td>5</td>
<td>33</td>
<td>31</td>
</tr>
<tr>
<td>Annual growth rate %</td>
<td>11.76</td>
<td>0</td>
<td>8.01</td>
<td>16.65</td>
<td>60.36</td>
</tr>
<tr>
<td>Document contents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUTHORS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>225</td>
<td>1</td>
<td>13</td>
<td>107</td>
<td>120</td>
</tr>
<tr>
<td>Authors of single-authored docs</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Authors collaboration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-authored docs</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Co-authors per doc</td>
<td>4.07</td>
<td>1</td>
<td>3.8</td>
<td>3.73</td>
<td>4.58</td>
</tr>
<tr>
<td>International co-authorships %</td>
<td>15.71</td>
<td>0</td>
<td>60</td>
<td>12.12</td>
<td>12.9</td>
</tr>
<tr>
<td>Document types</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>article</td>
<td>70</td>
<td>1</td>
<td>5</td>
<td>33</td>
<td>31</td>
</tr>
</tbody>
</table>
Although the subject of machine learning has grown lately, it involves a robust and complex methodology, which restricts the number of researchers using it. Understanding all the deep learning process steps requires good scientific and programming skills to apply these methodologies. It is the reason for the low number of manuscripts published annually. Soon, with all the last decade’s technological advances, we see the number of publications in this area grow. Visually we can see 5 distinct production peaks in 2013, 2016, 2018, 2020, and 2022 (Figure 2). The highest productivity levels were observed in 2022, with 18 published papers, and in 2020, with 7 published papers, respectively.

The occurrence of oil spill events requires new research, and the consequent development of new detection and monitoring techniques. Therefore, manuscript production peaks usually follow major oil spill events in the last decade. These years’ publications represent 35.7% of the total articles published in the OSDMDL scientific field. In contrast, the production of articles between 1996 and 2012 only accounts for 14% of the published papers. However, we have noted a steady increase in cumulative publications in later years, particularly since 2013, as seen in Figure 2. These findings suggest that the OSDMDL scientific field has gained increasing popularity recently and remains a focal point for researchers. The use of machine learning and deep learning in complex data classification and decision-making in many academic areas has enabled the development of new algorithms capable of mathematically optimizing spatial data pattern detection and recognition systems [53].

The decade with the highest number of publications related to the OSDMDL methodology was the 2010s, with 33 papers published (Table 2). It corresponds to 47.1% of the published documents and has an average of 3.3 ± 1.8 standard deviation (SD). This process was strongly linked to the Deepwater Horizon oil spill case, owned and operated by Transocean, located in the Macondo basin (Mississippi Canyon) in the Mexican Gulf. It spilled more than 5 million oil barrels into the ocean. Naturally, due to the large scale of this spill, there was a significant increase in detection and monitoring studies at the time. Following the 2010s, 2020–2022 had 31 papers published, representing 44.3% of the total published documents with a mean of 10.3 ± 6.6 SD. We also found only 5 papers published during the 2000s, representing 7.1% of the publications with a mean of 0.5 ± 0.7
SD. Similarly, only 1 paper was published in the 1990s, representing 1.4% of the total published documents with a mean of 0.2 ± 0.5 SD. We believe that low technology has made applying such complex techniques through limited computer systems difficult.

Table 2 presents the trend statistics associated with production between 1996 and 2022. In terms of statistics related to the growth rate of the number of published papers, it is noted that, except for the period from 1996–1999 with a zero-growth rate, all other periods showed an increase, with 2020–2022 standing out with a growth rate of 60.36% (Table 2).

When analyzing the number of authors per decade, another pattern emerges, such as the increase in paper production, with a more significant number of authors in more recent decades. Specifically, the decades 2010–2019 and 2020–2022 had the highest numbers of authors, with 107 and 120, respectively.

Upon evaluating the co-authors per papers metric, it was possible to verify that the period from 2020–2022 had the highest values, with a score of 4.58. It was followed by the 2000–2009 decade, with a score of 3.8, then by the 2010–2019 period, with a score of 3.7, and finally, 1996–1999, with a score of 1 (Table 2). A higher value indicates a higher degree of collaboration between researchers during that period, with 2020–2022 having the highest degree of collaboration. It is worth noting that this metric only considers the paper’s quality or impact. However, it can be used as an indicator of collaboration and interdisciplinarity within a research field (Table 2).

Additionally, there was a noticeable increase in international collaboration, with the highest percentage of co-authorships (12.12%) found in 2000–2019 (Table 2). Interestingly, the 2020–2022 decade shows a similar percentage of co-authorships to the previous decade, with 12.9% (Table 2).

The co-occurrence network depicted in Figure 3 revealed various themes and research methods employed by authors in the OSDMDL literature. A total of 1423 items were identified from the network from 1996–2022. The top 5 most frequent unigrams were “Spill” (48 occurrences), “SAR” (35), “detection” (24), “Accuracy” (24), and “experiment” (14) (Figure 3). The term “accuracy” appears prominently as it involves training metrics and analysis of the results applied to neural networks. Accuracy represents the ability of the trained neural net to identify the targets precisely. Therefore, the authors need to know the effectiveness of the applied methodology.
The co-occurrence network was represented in 35 clusters during this period. The top 10 most frequent unigrams formed clusters that accounted for approximately 42.23% of the network. The 10 clusters with the highest number of items (Figure 3) were colored red (76), green (69), dark blue (64), khaki (62), purple (59), light blue (59), orange (56), brown (54), magenta (53), and pink (49).

Based on their representativeness, let us look at each of the ten main clusters dealing with various OSDMDL-related issues:

- Cluster 1: This cluster primarily focuses on neural networks, network approaches, experimental results, dark spot detection, and dark formation detection. Researchers in this cluster use various neural network models to detect and classify oil spills.
- Cluster 2: The second cluster revolves around oil spill detection, network methods, recall metrics, overall accuracy, and environmental factors. Researchers in this cluster develop and evaluate different network methods for detecting oil spills, considering environmental factors that affect oil spill detection accuracy.
- Cluster 3: This cluster is focused on synthetic aperture radar (SAR), accuracy, marine oil spill detection, convolutional neural networks (CNN), and marine oil pollution. Researchers in this cluster work on developing and improving SAR-based methods...
for detecting marine oil spills, using CNN models and other techniques to enhance accuracy.

- Cluster 4: This one deals with artificial neural networks, oil spill accidents, ocean waves, sunglint effects, and optical images. Researchers in this cluster explore the use of artificial neural networks for detecting oil spills, considering various factors that affect detection accuracy, such as ocean waves and sunglint effects.

- Cluster 5: The fifth cluster focuses on image data, classification, remote sensing analysis, time series images, and image segmentation. Researchers in this cluster develop and evaluate image-based methods for detecting and classifying oil spills using various data analysis techniques.

- Cluster 6: This cluster centers around pixel value, network development, polarimetric features, remote sensing, and Terrasar X images. Researchers in this cluster use polarimetric features of Terrasar X images and other remote sensing data to develop and optimize network-based approaches for detecting oil spills.

- Cluster 7: It deals with performance measures, sea surface, biogenic slick, polarimetric synthetic aperture radar images, and original SAR images. Researchers in this cluster develop and evaluate performance measures for detecting oil spills, taking into account various factors that affect detection accuracy, such as biogenic slicks and polarimetric SAR images.

- Cluster 8: This one focuses on data processing, improved fully convolutional network (FCN), satellite images, discrimination, and oil spill events. Researchers in this cluster use FCN-based methods for processing satellite images and detecting oil spills, considering factors that affect discrimination accuracy.

- Cluster 9: This cluster revolves around satellite data, wind conditions, film thickness, optimal classifier, and field observation. Researchers in this cluster develop and evaluate various methods for detecting and classifying oil spills using satellite data, considering factors that affect detection accuracy, such as wind conditions and oil film thickness.

- Cluster 10: The final cluster centers around support vector machines (SVM), stage, target type, network process, and optimum network. Researchers in this cluster explore SVM-based methods for detecting and classifying oil spills at different stages, dealing with different types of targets, and optimizing network processes for better detection accuracy.

Figure 4 shows how oil detection technologies have developed. Interestingly, the bigram frequency terms tend to occur after 2015, with a notable increase between 2018 and 2022. We can observe that, in previous years, there was a predominance of remote sensing data for spatial analysis techniques, especially SAR data. In the last 5 years, with the computational advance and the possibility of CNN analysis in parallel systems and GPUs (graphics processing units), detection studies in neural networks and multilayer textural analysis development have grown significantly.
However, it is worth mentioning that some bigram terms, such as “dark formation”, deserve special attention despite not having a high frequency in a particular year. It is because “dark formation” has been consistently present since 2008, indicating its significance in the broader context of the subject matter. As the identification of oil spills in SAR images is based on “dark formation”, despite the evolution of identification and monitoring techniques, we continue to use specular reflectance in SAR data as the main methodology to identify these targets.

3.2. Country Contribution

In the scientific field of OSDMDL, China emerged as the leading contributor among the top 10 countries, with a substantial share of 23.1% (33) of the total scientific production. Italy secured the second spot with 7% (10), followed by South Korea at 4.2% (6). The United Kingdom and the United States shared the fourth position, each with a 3.5% (5) contribution. Germany and Greece both recorded a 2.8% (4) contribution, while Canada, India, and Iran each contributed 2.1% (3) to the field (Figure 5).

**Figure 4.** Temporal trends of the most frequent bigrams abstract terms between 2008 and 2022. The blue circle size indicates the frequency, and blue lines indicate temporal trends of bigrams terms along time.

**Figure 5.** The co-authoring collaboration network by countries from documents published. The red lines indicate collaboration between authors from different countries, and the width indicates the frequency of collaborations.
It is important to note that scientific production is a crucial indicator of a country’s research output and contribution to the global knowledge base. The fact that China is leading the way in OSDMDL research suggests that the country is investing significantly in this field and has the necessary resources and expertise to produce high-quality research. Similarly, the presence of Italy, South Korea, the United Kingdom, the United States, Germany, Greece, Canada, India, and Iran in the top ten countries indicates their active involvement and interest in advancing research in OSDMDL.

Figure 5 depicts the flow and intensity of collaboration between countries based on author and institutional collaborations. The width of the edges in the figure indicates the strength of the links between countries. It provides valuable insights into the nature and extent of collaborative efforts in the OSDMDL research field, highlighting countries actively engaged in joint research and knowledge exchange.

Upon assessing the collaborative efforts among nations, it is noteworthy that the United Kingdom (UK) and China have displayed significant involvement, with a contribution of 19.23% (5) each, accounting for 38.46% of the total collaborations. On the other hand, Italy has exhibited a substantial level of engagement, representing 15.38% (4) of the collaborative activities. Canada, the United States of America (USA), and Greece have each displayed a comparable level of involvement, contributing to 11.54% (3). In contrast, Portugal, Spain, and Iran have shown a relatively lower level of collaborative activities, contributing to one-third of it each (1).

It is worth noting that the present study undertook an individualized evaluation of the collaborative efforts among nations, thus highlighting the degree of their involvement while also considering any collaboration among them.

3.3. Most Influential Publication

Upon analyzing the literature in the OSDMDL field, it was observed that a select subset of 37 articles stood out due to their high citation count, representing 52.8% of the articles in the selected field (see Table S2 for more details). The remarkable finding was that this subset had accumulated 93% of the total citations, accounting for 1551 out of 1667 citations in the OSDMDL field.

This result shows the impact of these 37 articles in the OSDMDL field, which researchers have widely recognized and cited. The higher citation count for these articles could be attributed to their contribution to advancing the understanding of OSDMDL and their utility in guiding future research.

Upon analyzing the most highly cited articles in the field of interest, it was observed that the top 10 articles accounted for 56.5% of the total citations, amounting to 943 citations, as depicted in Figure 6. This observation suggests that a small fraction of the literature in the field has garnered disproportionate attention and influence.
Figure 6. The figure shows the ten most impactful papers based on total citations. The blue circles on the right side indicate the respective citation numbers.

When evaluating the top 5 neural network architectures most commonly used in the 37 selected studies focused on the OSDMDL field, it was observed that multilayer perceptrons (MLP) stood out, being present in 11 studies. It was followed by the convolutional neural network (CNN) architecture used in five studies, while U-Net, DeepLabv3+, and fully convolutional network (FCN) were all used in three studies. It is worth noting that among the 37 studies evaluated, only 12 (32.4%) utilized more than 1 neural network architecture to generate their results. Among these studies, the multilayer perceptrons (MLP) and deep convolutional neural network (DCNN) architectures were the most commonly used.

The high frequency of using MLPs can be attributed to their widely used and well-established architecture for neural networks, known for their ability to perform well in various machine learning tasks, including image and signal processing, classification, and regression. On the other hand, the use of DCNNs can be explained by their superior performance in image processing tasks, as they can learn increasingly complex and abstract features from input images through multiple convolutional layers. It is noteworthy that while MLPs and DCNNs are the most used neural network architectures in the evaluated studies, other architectures, such as CNNs, U-Net, DeepLabv3+, and FCN, also have their strengths and are frequently used in specific applications.

When evaluating the top five sensor systems primarily used, it was observed that the combination of Envisat ASAR and RADARSAT-2 was used in seven articles, indicating a significant preference over other systems. Next, ERS-SAR 2 was used in six articles, demonstrating similar popularity. Sentinel-1, on the other hand, was used in five articles, a common choice among researchers. Finally, TerraSAR-X was used in three articles, indicating that its use is still less every day than the other options. Free SAR data were few and did not have continuous acquisition.

Machine learning studies require a significant number of samples to train the codes. Therefore, most publications directly reflected the RADARSAT and Envisat ASAR most extensive time series data. Moreover, since 2014, we have had a significant database of Sentinel SAR data, which now has more than 15 confirmed oil spill cases in its global
records. The TerraSAR-X and COSMO-SkyMed data also presented official records of spills, but their paid platform makes most of the studies applied to this theme unfeasible. Upon analyzing the use of features in the studies, it is clear that 40.5% of the papers utilized it. It is a surprising finding, as features are essential to image analysis and machine learning algorithms. Among the studies that used features, some categories were the most frequently used.

One of the most utilized categories was texture features. Texture features are a way to describe the spatial arrangement of pixels within an image, and they are often used to help differentiate between objects or regions of interest. In addition, geometrical and statistical features were also used, providing valuable information about the shape and distribution of objects within the image. For many years, this was the principal methodology applied to identify oil targets in the ocean using SAR images. However, it became less effective over the years with the development of deep learning technologies. Spectral bands were another prominent feature utilized.

Upon analyzing the patterns related to image preprocessing, it could be observed that a vast majority, around 65%, employed one or more preprocessing techniques. It indicates the importance of image preprocessing in remote sensing studies to obtain accurate results. The most commonly used techniques were related to image normalization, which reduces the presence of speckles, and the data rescaling to improve the image quality.

3.4. Influential Journals

Figure 7 illustrates the distribution of articles among various journals that have published OSDMDL-related research. The analysis of the results indicates that 40 different journals have contributed to the knowledge in this field. Among these journals, the 10 most notable sources in terms of the number of articles published are “Remote Sensing,” with 12 articles, representing approximately 17.1% of the total, followed by “IEEE Transactions on Geoscience and Remote Sensing”, with 7 articles accounting for 10% of the total. “IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing” comes in third with 5 articles, approximately 7.1% of the total. “Marine Pollution Bulletin” had 4 papers, about 5.7% of the total.

![Figure 7. The figure shows the ten most impactful papers based on total citations. The blue circles on the right side indicate the respective citation numbers.](image-url)
In comparison, “The Journal of Coastal Research” and “Sensors” had 3 articles each, representing roughly 4.3% of the total for each one. “Applied Sciences” and “International Journal of Remote Sensing” had 2 articles each, accounting individually for about 2.86%. Finally, “Acta Oceanologica Sinica” and “Archives of Environmental Contamination and Toxicology” had 1 article each, representing approximately 1.43% individually. These findings provide insights into the most influential journals in the OSDMDL research domain and can help researchers identify critical sources for future research in this area.

The impact of the top 10 journals in the context of OSDMDL, based on the number of citations, is illustrated in Figure 7. The two most influential journals in this field are IEEE Transactions on Geoscience and Remote Sensing, with significantly more citations than the rest. The total number of citations for the remaining journals in the top 10 is relatively similar.

3.5. Authors Contributing

When analyzing Figure 8, we can observe the top 10 authors who have published the most in the context of OSDMDL. Our analysis of the selected articles revealed the presence of 225 different authors, with an average of 4.07 authors per article, but only 3 articles had a single author. The top 5 most prominent authors were Li Y., who ranked first with 7 papers, representing 10% of the analyzed articles. Zhang J. followed in second place with 5 papers and approximately 7.1%. Del Frate F., Jung H-S., and Gong P. tied in third place with 4 papers, which accounts for approximately 6% for each author (Figure 8).

![Figure 8](image_url)

**Figure 8.** Temporal trends of essential authors between 1990 and 2021. The blue circle size indicates the number of published papers, and the red lines indicate temporal trends of papers published for each author over time.

When examining the temporal trends of authors in the context of the OSDMDL, it became evident that some authors had a significant presence in the number of published articles while also showing a consistent production level over time. The three authors who stood out the most in this regard were Del Frate F., Li J., and Li Y. As shown in Figure 8, these authors have maintained a relatively steady level of productivity throughout the
years, which may indicate their continued interest and contribution to this research topic. Their consistent production also suggests a high level of expertise and knowledge in OSDMDL.

4. Discussion

We conducted a comprehensive literature review and identified 70 documents that have been published using the OSDMDL scientific target topic. Although our results emphasized the importance and relevance of OSDMDL-related papers, they revealed that its production has been inconsistent over the years, with five distinct peaks. The highest peak was observed between 2020 and 2022. As we explained, due to the computational advance and the possibility of CNN analysis in parallel systems and GPUs (graphics processing units), the development of detection studies in neural networks and multilayer textural analysis grew significantly. Overall, there has been an increase in published papers across all periods, except for 1996–1999, which showed a decline in production. The detailed statistics in Table 2 highlight the changes in production over the years, with recent years showing a significant increase in OSDMDL-related publications. These findings suggest that this topic is rapidly expanding, and more research is needed to address this field’s challenges and opportunities [1,2,4,14,33].

In this study, the authors analyzed the OSDMDL methodology trends over 26 years. The first works deal with radar remote sensing techniques for detecting oil spills on the ocean surface. These studies showed that known parameters, such as polarization, backscatter, and damping ratio, helped apply deep learning techniques effectively today. The results show that the number of papers published in this field has consistently increased over time, with the highest growth rate observed in the last three years. It suggests a growing interest among researchers in the field of OSDMDL.

Another interesting finding is the increase in the number of authors per decade, with the highest numbers found in 2010–2019 and 2020–2022. It indicates that the field is becoming more collaborative and interdisciplinary, with researchers from different backgrounds working together [1,2,4,14,33]. This increasing collaboration reflects the highest score observed in the last three years [1,2,4,14,33]. This trend of international collaboration has a positive development in the OSDMDL-related research topic as it allows for exchanging a broader range of perspectives and expertise. Overall, the outcomes of this study offer a significant understanding of the patterns of collaboration and interdisciplinary approaches within the field of OSDMDL.

The diversity of OSDMDL-conducted research is highlighted by the various themes and methods identified here. Policymakers and researchers can use this information to understand the field better and identify areas for further research. Moreover, the co-occurrence network created in this study provides a valuable tool for researchers to identify collaborations, trends, and gaps, helping them plan and prioritize future research efforts.

Words such as “oil spill,” “oil,” and “SAR” are some of the most common recurrent bigram terms in the literature after 2015, indicating promising opportunities for future research and innovation in detecting and mitigating oil spills. However, specific bigram terms such as “dark formation” have consistently appeared since 2008, highlighting their significance in the broader context of OSDMDL. The study suggests that researchers should consider these recurring terms while researching OSDMDL.

Our results show a detailed analysis of the scientific production and collaborative efforts in the field of OSDMDL. One notable finding is that China has emerged as the leading contributor, accounting for 23.1% of the total scientific production, followed by Italy and South Korea. The UK and the USA shared the fourth position with a contribution of 3.5% each. It highlights the dominance of certain countries in this field and emphasizes the need for more diverse contributions from other countries. Furthermore, the collaborative efforts among nations were also discussed, with the UK and China displaying significant involvement while Portugal, Spain, and Iran showed relatively
lower levels of collaborative activities. It indicates that more collaborative efforts are needed to facilitate the exchange of knowledge and expertise among researchers from different countries.

Another significant finding is that highly cited OSDMDL articles represent only a tiny fraction but have accumulated most of the citations. The top 10 most cited articles accounted for 56.5% of total citations, indicating that a small fraction of the literature in the field has gained disproportionate attention and influence. It highlights the importance of producing high-quality research that can significantly advance the understanding of OSDMDL and guide future research.

Our results suggest that researchers must carefully select appropriate neural network architectures, sensor systems, features, and preprocessing techniques to produce reliable and accurate results. For example, the multilayer perceptrons (MLP) and convolutional neural network (CNN) architectures were the most generally used, and the combination of Envisat ASAR and RADARSAT-2 was the most frequently used sensor system.

In this study, we investigated the distribution of articles related to OSDMDL across different journals and identified the most impactful sources for future research. Our results revealed that 40 journals have contributed to this field, with Remote Sensing, IEEE Transactions on Geoscience and Remote Sensing, and IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing being the top 3 journals in terms of the number of published articles. Interestingly, these three journals also accounted for the most citations, with IEEE Transactions on Geoscience and Remote Sensing and Remote Sensing having a significantly higher number of citations than the other journals. The dominance of these few top journals in terms of published articles and citations highlights the influence of these journals on the field. Therefore, researchers should publish their work in these high-impact journals to increase the visibility and impact of their research.

Through the analysis of the selected articles, the top 10 authors who had published the most in the context of OSDMDL revealed that the authors’ collaboration was paramount. We found 225 authors involved, with an average of 4.07 authors per article. Researchers working in this domain are more likely to collaborate with other researchers, which can lead to more comprehensive and diverse research outcomes. Notably, the most influential authors identified in this study are from different countries, including China, Italy, and South Korea. This finding suggests that the OSDMDL research domain is an international research area that attracts scholars from diverse geographic regions.

We also discovered that only three articles had a single author, indicating that collaborative work is a common practice in this research domain. Li Y. emerged as the most prominent author with seven papers, followed by Zhang J. with five papers. Del Frate F., Jung H-S., and Gong P. tied for third place with four papers each. These authors have significantly contributed to the developing knowledge in the OSDMDL field.

When analyzing the authors’ temporal trend, three authors stood out with a consistent level of productivity over time: Del Frate F., Li J., and Li Y. They may possess a high level of expertise and interest in the field, based on the consistency in their contribution level to the field, as indicated by Figure 6. Identifying authors with a consistent level such as this is significant as it indicates their long-term commitment to the field and their sustained efforts in advancing the understanding of OSDMDL. Their research can guide future studies and shape the direction of OSDMDL-related topics.

Pinpointing authors such as these three can help identify potential collaborators and mentors for early career scientists and potential research groups or institutions that are actively involved in OSDMDL research.

5. Conclusions

This article is the first to use bibliometric review methods to assess the evolution of the OSDMDL literature. Through this study, we provide insights into scientific production related to countries, journals, and methods of analysis and evaluation that focus on remote sensing in OSDMDL.
This article examines the literature on OSDMDL over the past 26 years (1996–2022). The study uses a qualitative and quantitative word association network approach to provide an overview of the research trends in this field. The authors conducted a bibliometric analysis with systematic review elements and identified significant and exciting aspects of OSDMDL research. The findings indicate a significant increase in published articles in this field. However, there still needs to be more opportunities to expand conceptual and theoretical studies and methodological aspects of OSDMDL.

Several areas should be prioritized in the future. One crucial area is the investigation of novel neural network architectures, as recent developments in this field have yet to be fully explored. Researchers can discover new and more effective ways to detect and monitor oil spills by delving into these new techniques.

Another promising area for theoretical and methodological advances is the exploration of preprocessing and feature space analysis techniques. These techniques play a crucial role in the accuracy of results and the overall effectiveness of the models used in OSDMDL. By improving these techniques, researchers can enhance the accuracy of their models and thus provide better tools for detecting and monitoring oil spills.

Advances in the OSDMDL field allow for improving the effectiveness of oil spill detection and monitoring tools. By leveraging the latest technologies and techniques, researchers can develop more effective and efficient methods for detecting and responding to oil spills, ultimately minimizing their impact on the marine environment.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/jmse11071406/s1, Table S1: The table describes the details associated with selected papers in OSDMDL. The columns contain the information SR—(First Author, year, and paper), LA—(Language), TI—(Paper title), PY—(Publication year), and TC—(Total citation); Table S2: The table describes the details associated with 37 most influential publication selected papers in OSDMDL, Table S3: The table describes the summary of Parameter Values in Co-occurrence Network, Countries’ Collaboration World Map, and Thesaurus File. References [54–117] are cited in the Supplementary Materials.


**Funding:** This work was funded by the Brazilian Navy, the National Council for Scientific and Technological Development (CNPQ), and the Ministry of Science, Technology, and Innovation (MCTI) in the CNPQ/MCTI 06/2020 call—Research and Development for Coping with Oil Spills on the Brazilian Coast—Ciências do Mar Program, grant #440852/2020-0. During this work, RNV was supported by the CNPQ research fellowship (Process #81330/2021-4), JGVM by the CNPQ research fellowship (Process #308758/2021-8) and W.S.F.-R. Technology in Interdisciplinary and Transdisciplinary Studies in Ecology and Evolution (INCT IN-TREE).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We appreciate comments and suggestions from the anonymous reviewers that helped improve the quality and presentation of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.
References


50. R Core Team R. A Language and Environment for Statistical Computing; R Core Team R: Vienna, Austria, 2020.


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