



ARTICLE

Hybrid Deep Learning and Optimized Feature Selection for Oil Spill Detection in Satellite Images

Ghada Atteia^{1,*}, Mohammed Dabboor², Konstantinos Karantzos³ and Maali Alabdulhafith¹

¹Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh, 11671, Saudi Arabia

²Science and Technology Branch, Environment and Climate Change Canada, Dorval, QC H9P 1J3, Canada

³Laboratory of Remote Sensing, Department of Topography, School of Rural, Surveying and Geoinformatics Engineering, National Technical University of Athens, Athens, 10682, Greece

*Corresponding Author: Ghada Atteia. Email: geatteiaallah@pnu.edu.sa

Received: 13 January 2025; Accepted: 29 April 2025; Published: 09 June 2025

ABSTRACT: This study explores the integration of Synthetic Aperture Radar (SAR) imagery with deep learning and metaheuristic feature optimization techniques for enhanced oil spill detection. This study proposes a novel hybrid approach for oil spill detection. The introduced approach integrates deep transfer learning with the metaheuristic Binary Harris Hawk optimization (BHHO) and Principal Component Analysis (PCA) for improved feature extraction and selection from input SAR imagery. Feature transfer learning of the MobileNet convolutional neural network was employed to extract deep features from the SAR images. The BHHO and PCA algorithms were implemented to identify subsets of optimal features from the entire feature dataset extracted by MobileNet. A supplemented hybrid feature set was constructed from the PCA and BHHO-generated features. It was used as input for oil spill detection using the logistic regression supervised machine learning classification algorithm. Several feature set combinations were implemented to test the classification performance of the logistic regression classifier in comparison to that of the proposed hybrid feature set. Results indicate that the highest oil spill detection accuracy of 99.2% has been achieved using the logistic regression classification algorithm, with integrated feature input from subsets identified using the PCA and the BHHO feature selection techniques. The proposed method yielded a statistically significant improvement in the classification performance of the used machine learning model. The significance of our study lies in its unique integration of deep learning with optimized feature selection, unlike other published studies, to enhance oil spill detection accuracy.

KEYWORDS: Oil spill; machine learning; deep learning; classification; metaheuristic optimization

1 Introduction

Oil spill detection using Synthetic Aperture Radar (SAR) imagery and deep learning techniques has emerged as a promising solution for monitoring and mitigating environmental hazards. SAR imagery provides all-weather and day-and-night capabilities, making it an invaluable tool for detecting oil spills in marine environments. Oil spills are detected in SAR imagery by identifying areas of reduced backscatter, where the oil dampens surface capillary waves, creating smooth, dark patches on the water surface. These dark features contrast with the surrounding sea surface, allowing analysts or automated algorithms to flag potential oil contamination. Artificial intelligence (AI) tools have demonstrated significant effectiveness in analyzing satellite data for various environmental applications, including soil moisture retrieval,



significant wave height retrieval, crop classification, vehicle traffic monitoring, and land cover mapping. By employing advanced Machine Learning (ML) algorithms and data-driven approaches, these tools enable enhanced accuracy, scalability, and efficiency in extracting critical environmental insights, thereby supporting informed decision-making and sustainable resource management. Deep Learning (DL) algorithms, particularly Convolutional Neural Networks (CNNs), have shown remarkable effectiveness in extracting intricate patterns and features from SAR images [1].

Several studies investigated DL techniques for oil spill detection. A DL framework was developed for detecting and categorizing oil spills in [2]. In this framework, a full CNN for semantic segmentation was designed and trained for the binary classification of each pixel of a SAR image as either oil or no oil. Next, a second ANN was developed to assign the detected oil spills into one of 12 categories based on their shape, texture, and contrast [2]. In [3], utilizing a highly imbalanced dataset, a deep learning framework was introduced for the two-stage identification of oil spill occurrences. In the first stage, a novel 23-layer CNN is employed to classify image patches based on the percentage of oil spill pixels. The second stage involves semantic segmentation, implemented through a five-stage U-Net architecture. In [4], a deep learning framework was proposed, referred to as the faster region-based convolutional neural network. The framework is initiated by employing a deep CNN to extract feature maps from the input SAR imagery. A region proposal network then processes these extracted feature maps to generate predicted bounding boxes. Subsequently, the top-N ranked bounding boxes are integrated with pooled feature maps and passed through a fully connected layer, enabling precise identification of oil spill locations.

In [5], a DL oil spill detector utilizing the YOLOv4 object detection algorithm was developed and trained with a substantial SAR dataset to establish an early-stage surveillance system. The YOLOv4 object detector is a one-stage algorithm that combines object localization and classification into one step with only a single deep neural network. In this study, the developed model could distinguish between oil spills and look-alikes in proximity, but it failed when look-alikes covered oil spills. In [6], a multi-channel deep neural network segmentation model, integrating U-Net and ResNet architectures, was developed for oil spill detection. This pixel-level segmentation model was designed to accommodate multiple input channels, capture complex oil spill geometries at various scales and temporal dynamics, and address the challenges posed by look-alike features under low wind conditions. The model employed three input channels: SAR intensity, texture, and wind speed. In [7], a DL framework was proposed to detect oil spills automatically. To this end, SAR images were segmented into two classes: oil spill and background, using CNN and vision transformers. The proposed architecture was investigated by integrating Vision Transformer modules into different segments of the CNN backbone across various scenarios. Results of [7] showed that these model integrations can lead to improved detection of oil spills. In [8], a new self-evolving algorithm designed for automatic oil spill detection was introduced. The algorithm comprises interconnected modules responsible for oil spill detection, generating new training data, and improving deep learning models.

A modified segmentation network model was proposed in [9] for oil spill segmentation. The modified model was implemented in a conditional generative adversarial network. In [9], the modified model was utilized as a data generator within the network. In another study [10], a dual attention model was used to develop a DL oil spill detection system. This study enhances a conventional U-Net segmentation network by integrating the dual attention model. This integration aims to selectively identify the relevant and distinguishing global as well as local characteristics of oil spills within the SAR imagery. In [11], an online nonparametric Bayesian analysis method based on the Dirichlet process mixture of Gamma distributions was proposed for oil spill detection. This study involved a feature selection method based on the Resnet50 and Inception v3 pre-trained deep learning models. In another study [12], Topouzelis et al. developed two neural networks (NN) to detect and classify oil spills. The first neural network was designed to detect dark

formations in SAR imagery, while the second focused on distinguishing between actual oil spills and look-alikes. In a similar approach, Nieto-Hidalgo et al. [13] proposed a two-stage system utilizing three pairs of CNNs. Each pair of networks was trained to identify one specific class—ships, oil spills, or coast. The process involved two steps: the first CNN performed a coarse detection to identify regions of interest, followed by a second, more specialized CNN that refined the results to achieve precise pixel-level localization for each class. A comprehensive review of the most recent DL approaches for oil spill detection is provided in [14].

This study proposes a new hybrid DL and integrated BHHO-PCA approach for oil spill detection using SAR imagery. The proposed approach combines deep transfer learning, metaheuristic Binary Harris Hawk optimization, and Principal Component Analysis to optimize feature extraction and selection from input SAR images. Feature transfer learning was carried out with the MobileNet CNN to extract deep features from SAR images. The BHHO and PCA techniques were used to select subsets of optimum features from the complete feature dataset collected by MobileNet. An augmented hybrid feature set was created from the PCA and BHHO-produced features and utilized as input for oil spill detection via the Logistic Regression supervised machine learning classification. The classification performance of the LR classifier is tested using several metrics when trained by several combinations of input feature sets in several scenarios. The proposed BHHO-PCA approach improves upon existing methods, such as [2,3,5], by integrating an optimized feature selection process using the BHHO algorithm with PCA. This study significantly contributes to the field by integrating multiple feature extraction techniques to derive optimized features from SAR images. Existing DL-based oil spill detection models utilize standard paradigms that do not leverage advanced hybrid methods for feature selection. In hybrid feature selection approaches, metaheuristic optimization algorithms could be combined with traditional techniques to promote classification performance. Our study addresses the challenge of optimizing feature selection for oil spill detection in SAR imagery. The proposed approach addresses a critical gap in the existing literature, where such comprehensive methodologies for oil spill detection have been largely unexplored, offering a robust and innovative solution to enhance detection accuracy and reliability. To the best of our knowledge, this is the first work in the literature to propose a hybrid feature optimization approach where pre-trained CNN, PCA, and metaheuristic BHHO algorithms are combined for oil spill detection in SAR images. The proposed hybrid BHHO-PCA approach improves upon existing methods by effectively identifying the most discriminative features while reducing the dimensionality of the feature space. The main contributions of the study are:

- The proposal of a new hybrid DL-based approach for oil spill detection in SAR imagery.
- The development of a new feature engineering approach combines deep transfer learning, metaheuristic Binary Harris Hawk optimization, and Principal Component Analysis to optimize feature extraction and selection.
- A comparative study of the classification performance of the proposed oil spill detection system using several feature set combinations was conducted.
- An ablation study of the proposed oil spill approach will be conducted to highlight the contribution of its components.
- Statistical analysis of the introduced system to reveal the significance of the performance improvement due to the proposed feature engineering approach.
- Computational efficiency investigation of the proposed system.
- Benchmarking the proposed approach with other ML models and recent deep learning-based oil spill detection systems.

This paper is structured as follows: [Section 2](#) provides an overview of the experimental sites and the dataset utilized in the study. [Section 3](#) outlines the study framework and the methodologies employed. [Section 4](#) discusses the results obtained, while [Section 5](#) concludes the study and highlights its key findings.

2 Experimental Sites and SAR Imagery

In this study, we examined six instances of oil spillage. The locations of these occurrences are depicted in Fig. 1. Six Sentinel-1 images were acquired over the locations of these instances (Fig. 1). Four incidents took place offshore of Ras Al-Khafji, Saudi Arabia, over which three Sentinel-1 SAR images were acquired on 17 July, 29 July, and 10 August 2017. Another spill happened offshore of northern Qatar, where a Sentinel-1 SAR image was acquired on 17 May 2018, while a separate event occurred in a sea region near Kuwait, where a Sentinel-1 SAR image was acquired on 17 June 2018. The final incident considered in our study occurred offshore of Jeddah, Saudi Arabia, where a Sentinel-1 SAR image was acquired on 13 October 2019 (see Fig. 1). All SAR images used in our study are in VV polarization. Furthermore, images were acquired at approximately similar radar incidence angles ranging from $\sim 31^\circ$ to $\sim 46^\circ$ (Table 1). The dataset utilized in this study was obtained from the Sentinel-1 satellite constellation and provided by the European Space Agency through the Copernicus Data Space Ecosystem [15]. Fig. 2 shows the Sentinel-1 VV imagery acquired over the oil spills (dark patches).

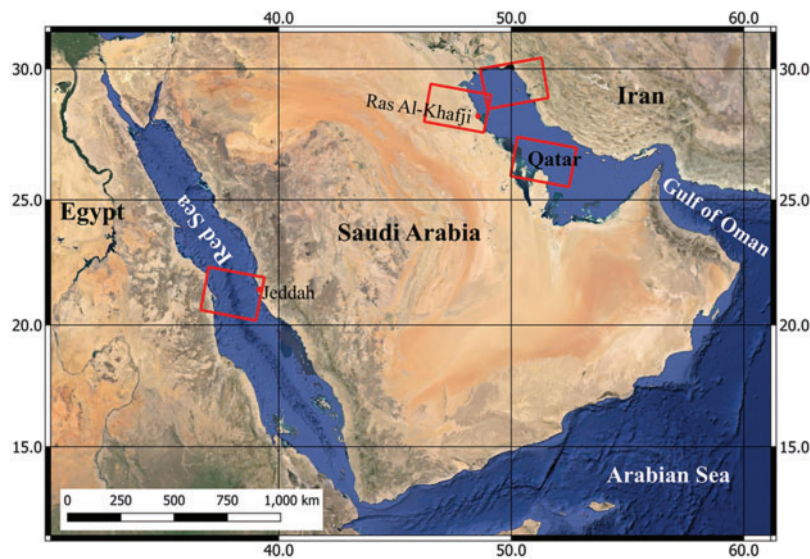


Figure 1: SAR imagery acquisitions (red frames) over the location of the oil spillages

Table 1: Sentinel-1 SAR imagery acquisition characteristics and the number of patches selected for oil spills and clean water

Product	Acquisition date	Orbit direction	Incidence angle near range	Incidence angle far range	Polarization	# Patches
Sentinel-1A IW GRDH	17/07/2017	Descending	30.77	46.18	VV	Oil:15 Water: 15
Sentinel-1A IW GRDH	29/07/2017	Descending	30.77	46.18	VV	Oil:3 Water:3
Sentinel-1A IW GRDH	10/08/2017	Descending	30.77	46.18	VV	Oil:12 Water:12

(Continued)

Table 1 (continued)

Product	Acquisition date	Orbit direction	Incidence angle near range	Incidence angle far range	Polarization	# Patches
Sentinel-1A IW GRDH	27/05/2018	Descending	30.75	46.18	VV	Oil:14 Water:14
Sentinel-1A IW GRDH	17/06/2018	Ascending	30.64	46.18	VV	Oil: 4 Water:4
Sentinel-1B IW GRDH	13/10/2019	Descending	30.78	45.97	VV	Oil:13 Water:13

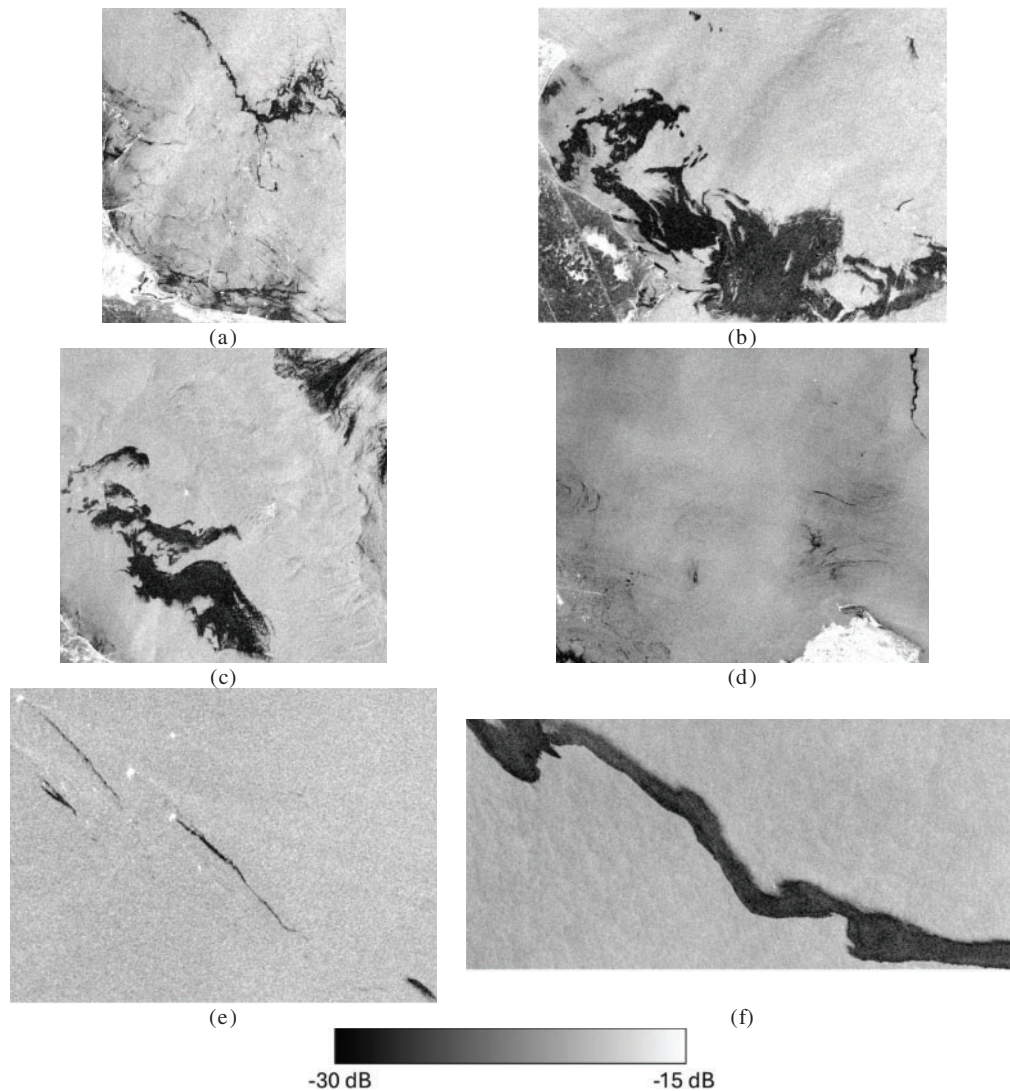


Figure 2: Sentinel-1 VV imagery acquired over the oil spills (dark patches) on (a) 17/07/2017, (b) 29/07/2017, (c) 10/08/2017, (d) 27/05/2018, (e) 17/06/2018, and (f) 13/10/2019

Sentinel-1 VV SAR imagery was selected for its free and open access, making it a reliable and widely used data source. VV polarization is particularly suited for ocean surface monitoring.

3 Methodology

In this study, the architecture of the proposed hybrid deep learning and BHHO-PCA approach for oil spill detection in SAR imagery comprises three stages, as illustrated in Fig. 3. The first stage consists of the preprocessing stage in which several preprocessing steps are applied to prepare the input data for the following stages. The second stage is the feature extraction and selection stage, and the third is the ML classification stage. Each stage is described in detail below.

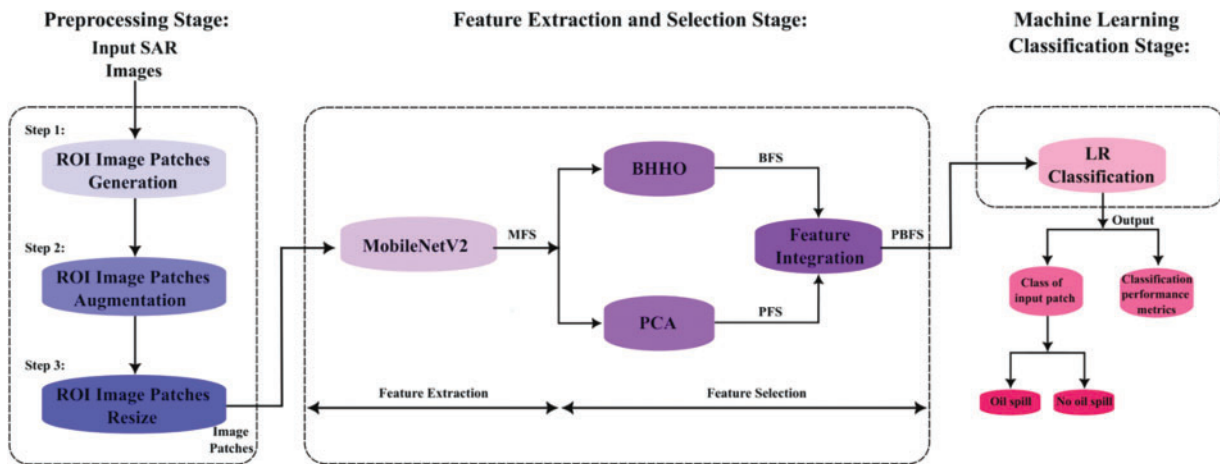


Figure 3: Architecture of the proposed hybrid deep learning model integrated with BHHO-PCA for oil spill detection in maritime SAR satellite imagery

3.1 Preprocessing Stage

This study uses a deep-learning CNN to extract image features for oil spill detection. Deep learning CNN requires enough input images to function effectively. In the first stage, the input SAR data undergoes preprocessing to augment the number of input images, as the number of available SAR images is limited. The following steps are applied to the input image set in this stage:

Step 1: Generation of Region of Interest (ROI) Image Patches

Step 2: ROI Image Patches Augmentation

Step 3: ROI Image Patch Resize

In the first step of the preprocessing stage, we generated multiple patches from each input SAR image by cropping several ROIs from it. This was feasible due to the large dimensions of the SAR images. Two sets of ROIs were generated. The first set was formed from image patches that contained only water pixels without oil spill pixels. However, all patches in the second set contain oil spill pixels and water pixels. The first set is labeled as the “No Oil Spill” (clean water) or negative class, while the second set is labeled as the “Oil Spill” or positive class. The process of cropping the ROI patches was performed manually based on polygon masks of oil spills available as ground truth data for each of the six input SAR images. A total of 61 ROI image patches were cropped for each class. Fig. 4 presents sample ROI patches for the positive and negative classes.

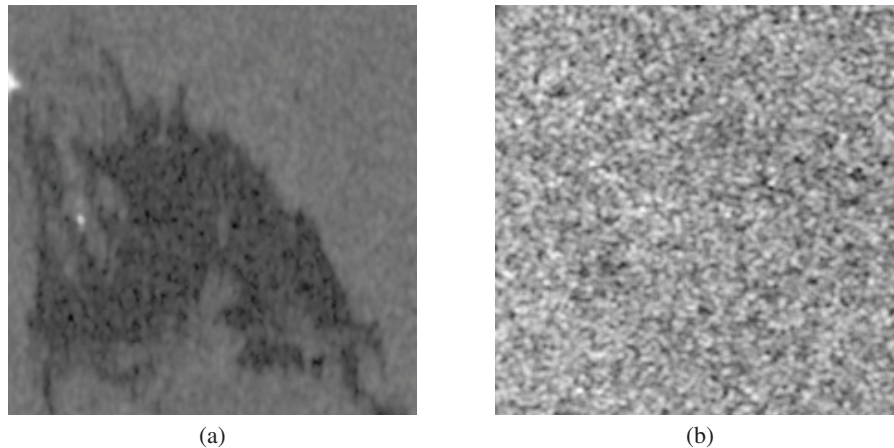


Figure 4: Sample ROI patches for the positive and negative classes: **(a)** Patch containing oil spill (positive class); **(b)** Patch containing clean water (negative class)

As the number of ROI patches is still small, we utilized two image transformation techniques in step 2 of this stage to augment the patch sets. In this step, vertical and horizontal image flips and image rotation with rotation angles starting from 20° to 340° with 20° increment steps have been utilized to supplement the number of ROI patches. A comprehensive survey on image data augmentation for DL can be found in [16]. Following the patch augmentation, the total number of positive and negative class patches became 2440, with 1220 patches for each class. In the third and last step of the preprocessing stage, the ROI patches are resized to have 224×244 pixels. This is the image size accepted by the input layer of the MobileNetV2 CNN used for feature extraction in the next stage.

3.2 Feature Extraction and Selection Stage

3.2.1 MobileNetV2

In the feature extraction phase of this stage, MobileNetV2 pre-trained CNN is utilized to extract deep image features. MobileNetV2 is a CNN architecture that utilizes inverted residual blocks consisting of a lightweight bottleneck layer followed by a linear layer. This structure helps capture richer features while reducing computational costs [17]. The linear bottleneck structure balances the computational cost and the representational power of the network. It allows for better information flow through the network while minimizing the number of parameters [17]. MobileNetV2 was chosen over other CNNs for its efficient balance of performance and computational efficiency. The architecture of MobileNetV2 features a streamlined design and memory efficiency. It utilizes depth-wise separable convolutions, significantly reducing model size and complexity while maintaining high accuracy. Its inverted residuals and linear bottleneck design enhance feature extraction, which makes it particularly suitable for high-dimensional data such as remote sensing images.

MobileNetV2's lightweight architecture is ideal for environments with limited computational resources, offering a competitive advantage over heavier CNN architectures like VGGNet or ResNet. Recent lightweight architectures, like EfficientNet, offer improved accuracy in end-to-end classification tasks. However, they typically introduce higher computational demands and longer inference times due to their increased number of model parameters. Given that the classification in the proposed pipeline is handled separately by the Logistic Regression classifier, the role of the backbone is purely to extract informative features. MobileNetV2 offers an optimal balance between efficiency and representational power in this context. It enables fast and

scalable feature extraction while preserving the effectiveness of subsequent processing stages. MobileNetV2 has been pre-trained on a subset of the ImageNet database, enabling the network to acquire rich feature representations across diverse images. MobileNetV2 is a 53-layer deep, lightweight CNN model.

The input image patches are delivered to the MobileNetV2 after being resized to its input layer size of 224×224 pixels. The CNN extracts image features at the global pooling layer after aggregating them across the network layers. The deep features of the input images are derived from the activations of this global pooling layer. The MobileNetV2 extracts 1280 features per input image. This feature set is denoted as MFS (MobileNetV2 Feature Set) throughout the study.

The feature selection phase of this stage aims to select significant features from all the extracted features in MFS. In this phase, we individually utilized the Principal Component Analysis and Binary Harris Hawk Optimization techniques to form two sets of selected features from diverse perspectives. These sets of selected features are then integrated to form a joint single feature set that captures the most informative image features, aiming to improve the classification performance of oil spills.

3.2.2 Principal Component Analysis

PCA is an unsupervised statistical method used to reduce the dimensionality of data while preserving most of its variability [18]. It achieves this by converting the original features into a new set of uncorrelated features known as principal components. These components are ranked according to the variance they capture in the data, enabling a more concise representation of the information [19]. We calculate the d -dimensional mean vector, μ , and the $d \times d$ covariance matrix, Σ , for the entire dataset in PCA. Then, we compute the eigenvectors and eigenvalues, sorting them in descending order based on eigenvalues. Next, we select the top k eigenvectors. We construct a $k \times k$ matrix, A , with columns containing these k eigenvectors. Finally, the data in the new space is obtained through the following transformation [19]:

$$X = A^T (x - \mu) \quad (1)$$

where the superscript T denotes the matrix transpose. The matrices x and X denote the data in the original and new spaces, respectively.

3.2.3 The Binary Harris Hawk Optimization

The Harris Hawk Optimization (HHO) is a metaheuristic optimization algorithm used to address feature selection tasks [20]. The core principles of HHO involve simulating the hunting behavior of Harris's Hawks, including the mechanisms of searching, capturing prey, and collaboration among individuals. The HHO algorithm is based on two phases: exploitation and exploration, which depend on the escaping energy of the prey, E

$$E = 2(2r - 1) \left(1 - \frac{t}{T}\right) \quad (2)$$

where t is the current iteration, T is the total number of iterations, and r is a random number in $[0, 1]$. The exploration phase occurs for $|E| \geq 1$, allowing the hawks to search globally in different regions for potential solutions to the optimization problem. The exploitation phase occurs for $|E| < 1$, suggesting a local search around the neighborhood for the best solutions. The HHO algorithm was designed for continuous optimization problems. However, a binary HHO, BHHO, version has been developed for feature selection, representing solutions in a binary form. Binary Harris Hawks Optimization enhances feature selection by providing a balanced and adaptive approach compared to traditional methods. By mimicking

the hunting strategies of Harris hawks, BHHO effectively navigates the feature space while maintaining an optimal balance between exploration and exploitation. This dynamic approach helps avoid the pitfalls of local optima and high computational costs, which are often associated with filter-based and wrapper-based traditional feature selection techniques. Moreover, BHHO is an adaptive mechanism that allows the algorithm to adjust search intensity based on the performance of the current feature subset. This adaptability contributes to achieving higher classification accuracy with fewer selected features and improves the model's effectiveness. Additionally, BHHO employs a stochastic selection process that enables the discovery of non-linear feature interactions that traditional deterministic methods may not capture. Furthermore, BHHO is a population-based approach that evaluates multiple potential solutions simultaneously. This feature enhances computational efficiency, making it particularly suitable for high-dimensional datasets (such as remote sensing data) where traditional methods might struggle. Additionally, BHHO utilized an objective function to guide the feature optimization process. This approach ensures that the selected features directly improve the model's performance and allows for a customized and precise selection of features that enhance the model's predictive performance and generalizability. The BHHO algorithm was selected in this study over other metaheuristic optimization approaches for its adaptive exploration-exploitation mechanism and its simplicity in implementation. As BHHO allows the dynamic switch between exploration and exploitation based on the prey's escape energy, it enables more effective search behavior without the need for complex control parameters as required by other optimization techniques, such as Genetic Algorithms, which require crossover and mutation operations, or Particle Swarm Optimization which depends on velocity and position updates influenced by global and local best solutions. The binary version of HHO directly handles discrete search spaces, and therefore, it is particularly well-suited for feature selection tasks. Thus, BHHO offers a computationally efficient and robust alternative to traditional metaheuristic algorithms.

The BHHO uses an S-shaped or V-shaped transfer function to convert continuous variables into binary ones. In our study, an S-shaped transfer function is used, which has the following form:

$$T(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

A detailed description of the algorithm can be found in [20].

A fitness function is needed in wrapper feature selection to assess every solution for feature optimization. Reducing the number of features and improving prediction accuracy are the main objectives of feature optimization. In this study, the fitness function used for the BHHO is described as

$$FT = \alpha * \frac{F_{Pr}}{T_1} + (1 - \alpha) \frac{|S|}{|F|} \quad (4)$$

where F_{Pr} is the number of false predictions of a K-Nearest Neighbor (KNN) classifier, T_1 is the total number of instances, and the ratio $\frac{F_{Pr}}{T_1}$ is the error rate. $|S|$ is the feature subset length, $|F|$ is the total number of features, and α is a parameter used to manage the effect of classification. For the BHHO algorithm, the error rate in the fitness function is evaluated using the Euclidean distance for five neighbors. The parameter α is set to 0.99 as the classification performance is of great importance in this application. The maximum number of iterations has been set to 100, and the number of hawks has been specified as 10. The dimension of the search space in this problem equals the number of features in the MFS, and hence, the search space contains 1280 features or $|F| = 1280$. The output set of features selected by the BHHO is denoted as the BHFS throughout the paper. After the PFS and BHFS have been generated, a comprehensive feature set is formed by integrating these two feature sets to promote the performance of the ML classifier of oil spills. This feature set is called PBFS.

3.3 ML Classification Stage

For the oil spill binary classification problem, the features from the PBFS have been used to train and evaluate a Logistic regression classifier, and the classification performance has been recorded.

Logistic Regression Machine Learning Classification

The logistic regression ML classification is a supervised approach for binary classification tasks. It predicts the probability of a binary outcome based on one or more predictor features. It is commonly applied in binary classification tasks where the dependent variable to be predicted is categorical with two possible outcomes. Thus, logistic regression classifies a dependent variable into one of two classes. Consider an input observation, denoted as x , presented as a feature vector $[x_1, x_2, \dots, x_n]$. The classifier yields an output, y , which can either be 1 (indicating the observation belongs to the class) or 0 (suggesting the observation does not belong to the class). Our objective is to determine the probability, $P(y = 1|x)$, that this observation belongs to the class. So, a training dataset is utilized to derive a vector of weights $w = [w_1, w_2, \dots, w_n]$ and a bias term b . To decide on a test observation, a single number z expressing the weighted sum of the evidence for the class is estimated as follows:

$$z = \left(\sum_i^n w_i x_i \right) + b \quad (5)$$

A sigmoid function, also known as a logistic function, is used to estimate the probability of the test observation x belonging to a class as follows:

$$P(y = 1|x) = \sigma(x) = 1/(1 + \exp(-z)) \quad (6)$$

Thus, for an observation x , we assert that it belongs to class 1 if $P(y = 1|x) > 0.5$; otherwise, it does not.

The model used in this study is an optimized model resulting from optimizing an efficient linear model that the grid search algorithm has selected. The model's hyperparameters have been tuned using grid search with 6-fold cross-validation. The optimizable hyperparameters are the regularization type, strength, and learner type. The ranges of the optimizable hyperparameters are:

- Regularization Type: Lasso-Ridge.
- Regularization Strength: $[4.0 \times 10^{-9} - 40]$
- Learner Type: SVM-Logistic regression.

The logistic regression model with Lasso regularization and 0.0014729 regularization strength was the best model selected by the search grid and, therefore, has been used in the current study. Logistic Regression machine learning classifiers are preferred over deep learning classifiers due to their simplicity, computational efficiency, and robustness on small datasets such as the one used in this study. As a relatively simple ML classifier, it is particularly suitable to focus on the impact of the proposed hybrid feature optimization approach on the classification performance.

The classification performance is evaluated using the classification accuracy (AC) and recall (R) derived from the classifier's confusion matrix in a six-fold cross-validation scheme. Mathematical expressions for R and AC are provided in Eqs. (7) and (8), respectively:

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

$$AC = \frac{TP + TN}{TN + TP + FN + FP} \quad (8)$$

In this context, TP denotes true positives, FN denotes false negatives, TN denotes true negatives, and FP denotes false positives.

This study also assesses the proposed classification model using the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve. The ROC curve represents the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) of the classification model across various threshold values. For binary classification problems, as in the problem at hand, AUC measures the capability of the classifier to separate between classes. A flowchart of the pipeline proposed in this study is presented in Fig. 5.

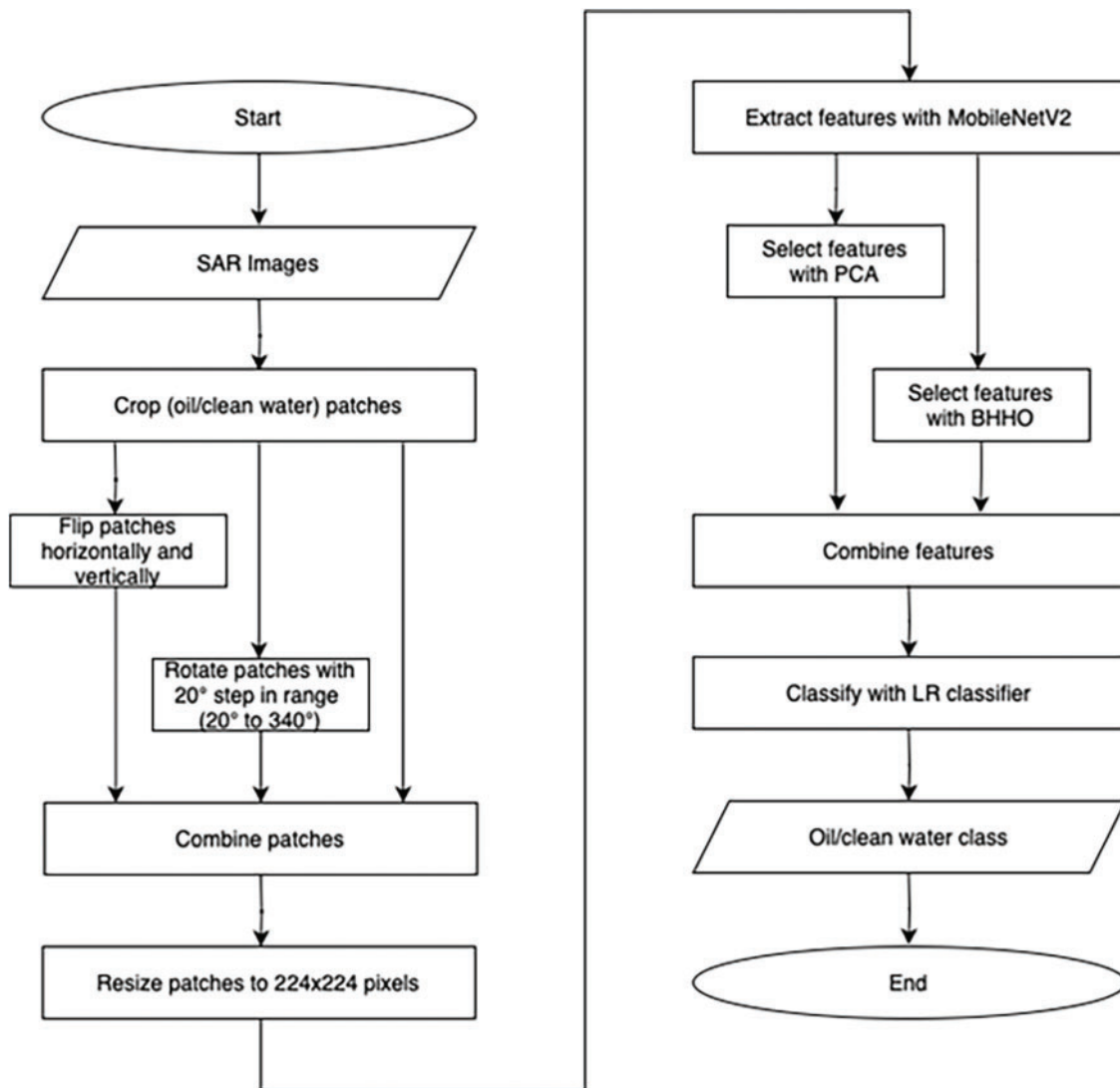


Figure 5: Flowchart of the proposed hybrid deep-learning oil spill detection approach

4 Results and Discussion

4.1 Classification Performance

In this study, five experiments were conducted to evaluate the classification performance of the proposed oil spill detection approach. To generate the PCA Feature Set (PFS), the PCA algorithm was adjusted to select principal components from the input feature set at a Cumulative Percent Variance of 95%. The analysis revealed that the MFS yielded 40 Principal Components (PCs), which account for 95% of the variability in the input data. Fig. 6 shows the ranked features based on the variance percentages of the PCs forming the PFS, which reflects the importance of the PC features.

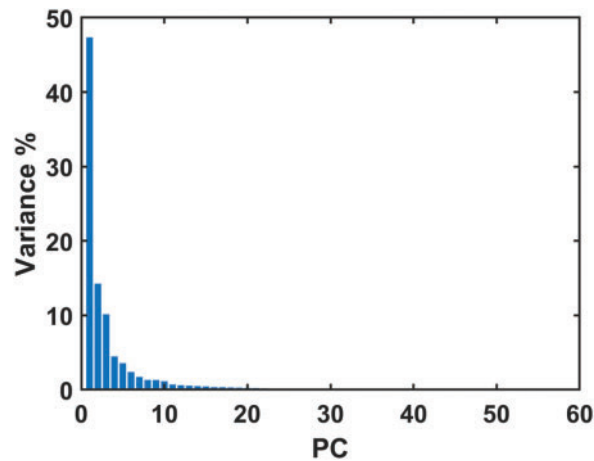


Figure 6: Percentage of variance explained by each principal component in the PFS

Separately, the MFS is directed to the BHHO algorithm to select the most significant features according to the Harris Hawk optimization criterion. Fig. 7 shows a plot of fitness values over the run iterations of the BHHO feature optimization process and the generation of the BHHO-driven feature set, denoted as BHFS. The minimum fitness value recorded is 0.01844 at the 13th iteration and remains steady until the 100th iteration.

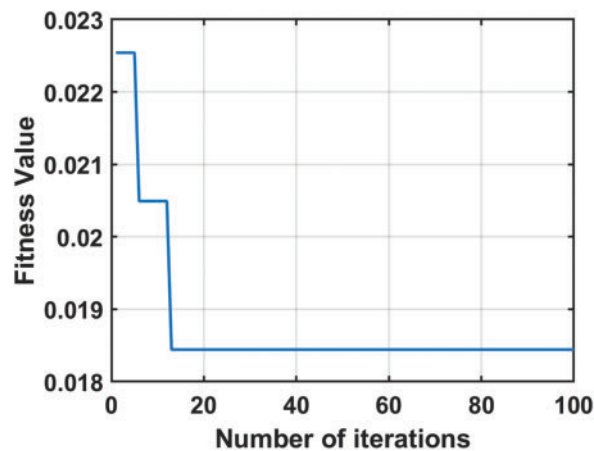


Figure 7: Performance of the fitness function throughout the iterations of the BHHO algorithm during feature optimization

The first experiment used the MFS set to train the LR classifier without any feature selection technique. The second experiment involved training the LR classifier with PCA-extracted features called PFS. The LR classifier was trained using the BHHO-derived feature set for the third experiment. In the fourth experiment, the LR classifier was trained by a hybrid feature set, PBFS, which integrates both PCA and BHHO-driven features. In the fifth experiment, the BHHO features were processed through the PCA algorithm for additional refinement, resulting in the feature set known as CPBFS (Cascaded PCA and BHHO feature set), which was then used for LR classifier training.

Table 2 illustrates the feature engineering method employed, the name of each feature set, the number of features for every experiment, and the resulting classification accuracy. Accuracy is used to assess classification performance due to the balance between the two classes. In this study, the classification performance of the MFS serves as a benchmark for accuracy. When trained with the complete MFS set, the LR classifier achieved a reference accuracy of 97.6%. An improved accuracy of 98.5% was recorded by the LR classifier when trained by the PFS, which contains 40 PCA-derived features. The BHHO-derived feature set, BFS, recorded a better accuracy of 98.8% in the third experiment. It was found from experiment five that cascading the BHHO with the PCA yielded a slightly reduced classification accuracy of 98.7% when the classifier was trained with the CPBFS. As indicated in Table 2, the LR classifier achieved the highest classification accuracy of 99.2% when trained with the BHHO-PCA hybrid feature set, PBFS, which comprises 564 features. The entries corresponding to the highest performance are highlighted in bold font in Table 2.

Table 2: Classification performance of the logistic regression classifiers in the experiments conducted in this study, with the highest accuracy values highlighted in bold

Experiments	Feature engineering method	Feature set	Number of features	Accuracy %
Experiment 1	Feature transfer (MobileNetV2)	MFS	1280	97.6
Experiment 2	PCA	PFS	40	98.5
Experiment 3	BHHO	BHFS	524	98.8
Experiment 4	Integration of PCA and BHHO	PBFS	564	99.2
Experiment 5	Cascaded PCA and BHHO	CPBFS	36	98.7

Table 3 presents the classification performance for each class achieved by the LR classifier across the five experiments. Therefore, the recall for the ‘Oil spill’ and ‘clean water’ classes is recorded individually. The recall value discloses the classifier’s capability of correctly predicting the existence of the positive class. In Table 3, the recall values were computed using a one-versus-all classification technique. It is observed that the highest recall for the ‘No oil spill’ class was achieved in Experiment 4, where the LR classifier was trained by the integrated PCA-BHHO set (PBFS). The utmost recall for the ‘Oil spill’ class was also observed when the LR classifier was trained individually with either the PBFS or CPBFS feature sets. The lowest recall for the ‘Oil spill’ class was recorded when the classifier was trained with the complete feature set (MFS), while the lowest recall for the ‘No oil spill’ class was noted when using the CPBFS. Table 3 shows that the LR classifier achieved the highest recall values for both classes when trained with the proposed hybrid PBFS.

To discuss the classification error of the best model in Experiment 2 (i.e., the proposed model), we refer to Fig. 8, which presents the model’s confusion matrix regarding the false negative rate (FNR) and TPR. The FNR represents the proportion of actual positive cases (oil spills) incorrectly predicted as negative (clean water). For the Oil Spill class, the FNR is 1.2%, which means that 1.2% of actual oil spill images were misclassified as clean water images. This indicates a high sensitivity to detecting oil spills, as the TPR

(or recall) is 98.8%. For the clean water class, the FNR is 0.4%, demonstrating the model's robustness in identifying non-oil (clean water) images, with a TPR of 99.6%. The proposed model demonstrates high performance with a low FNR for both classes. This indicates the model is highly effective in detecting oil spills, with a TPR of 98.8% for oil spill images and 99.6% for clean water images. The slightly higher FNR for oil images suggests a marginal challenge in distinguishing oil spills. Still, the model maintains a balanced and reliable performance, which is crucial for environmental monitoring.

Table 3: Recall obtained for each class by the Logistic Regression classifiers across the five experiments, with the highest recall values highlighted in bold

Experiment	Feature set	# Features	Recall per class (%)	
			Oil spill class	Clean water class
Experiment 1	MFS	1280	96.1	99
Experiment 2	PFS	40	98.3	98.7
Experiment 3	BHFS	524	98.1	99.4
Experiment 4	PBFS	564	98.8	99.6
Experiment 5	CPBFS	36	98.8	98.5

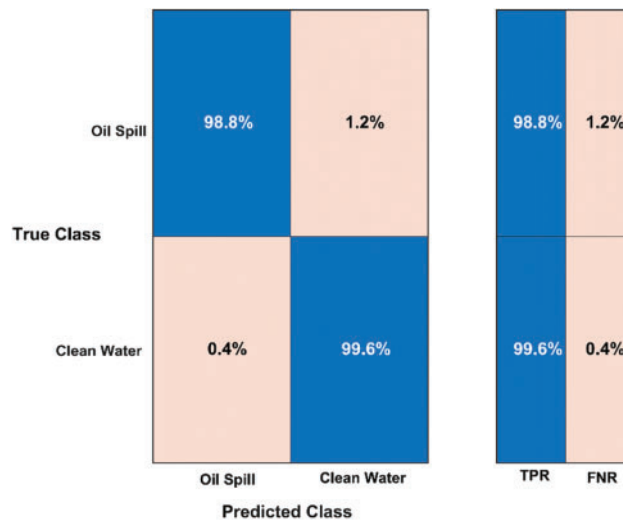


Figure 8: Confusion matrix of the proposed hybrid deep learning model in Experiment 4

Given the binary classification problem, AUC has been employed further to assess the classification performance of the LR classifier. The LR model trained by the proposed feature set (PBFS) demonstrated a higher AUC value of 1 than that of the base model trained with the MFS, which recorded 0.99 for AUC. This result confirms the enhanced capability of the proposed oil spill detection model in differentiating between positive and negative classes, i.e., oil spills vs clean water in the input images.

Our model's performance is validated for radar incidence angles ranging from 31° to 46°. Additionally, as steeper incidence angles are generally more favorable for oil spill detection [21], the model's performance is expected to improve at lower incidence angles.

4.2 Statistical Analysis

In this subsection, we perform a statistical analysis to investigate whether the improvement in classification accuracy of the LR model trained by the proposed feature sets is statistically significant. The study evaluated the LR models trained by the five feature sets in experiments 1 through 5 of Table 2. As in all experiments, we used the LR classifier, and statistical analysis evaluates the significance of the feature sets or the feature engineering approach and its impact on the classification accuracy of the LR model.

To determine the appropriate significance test, we first checked the normality of the distribution of the models' accuracy scores using the Shapiro-Wilk test with a 0.05 significance level. The test results showed that all five models had p -value > 0.05 , meaning the scores followed a normal distribution. Therefore, the One-Way ANOVA test was used to compare the significance of the feature sets. The One-Way ANOVA test returned a highly significant p -value < 0.00000002 . This means that at least one feature set showed a statistically significant difference in model classification accuracy compared to the others. As the ANOVA test only determines whether differences between models exist and does not specify which feature sets have differences, a post-hoc analysis using Tukey's Honest Significant Difference (HSD) test has been performed to identify specific feature sets with significant differences. Tukey's HSD test provided pairwise comparisons between the models to determine which differences in model accuracy were statistically significant. Table 4 presents Tukey HSD Test results for the LR model of Experiment 4 trained and evaluated using the proposed integrated PCA and BHHO feature set (PBFS). It is noticed that PBFS significantly outperformed all other feature sets ($p < 0.05$) in all comparisons. This analysis reveals that the LR model trained with the proposed integrated PCA and BHHO feature set (PBFS) is statistically superior, and its adoption is recommended for improved classification accuracy.

Table 4: Tukey HSD Test results for the LR model of experiment 4 trained and evaluated using the proposed integrated PCA and BHHO feature set (PBFS)

Feature set	Feature set	p -value	Significant difference
MFS	PBFS	2.921718×10^{-9}	TRUE
PFS	PBFS	0.002757643	TRUE
BHFS	PBFS	0.001065096	TRUE
PBFS	CPBFS	0.000659553	TRUE

4.3 Feature Importance

To understand how features in PBFS contribute to the model predictions, Fig. 8 presents the Shapley feature importance bar chart of the LR classifier trained by the proposed PBFS feature set in experiment 4. The Shapley feature importance bar chart presents the average absolute magnitude of Shapley values per feature. It ranks the features by importance and helps in model interpretability. The higher the Shapley value, the more influential the feature is to the ML model. The Shapley chart of Fig. 9 presents the importance bars of the first 10 features separately. In contrast, the sum of the Shapley values of the remaining features is accumulated and presented as the 'Fsum' bar. It is clear from the chart that the model relies heavily on multiple features to make predictions, as evidenced by the high sum of Shapley values. This suggests that the proposed PBFS feature set, selected by the proposed feature engineering approach, contains influential features that contributed significantly to the ML model. The top-ranked features, such as Feature 1 and Feature 2, have the highest mean absolute Shapley values, indicating that they consistently play a significant role in pushing predictions toward either class.

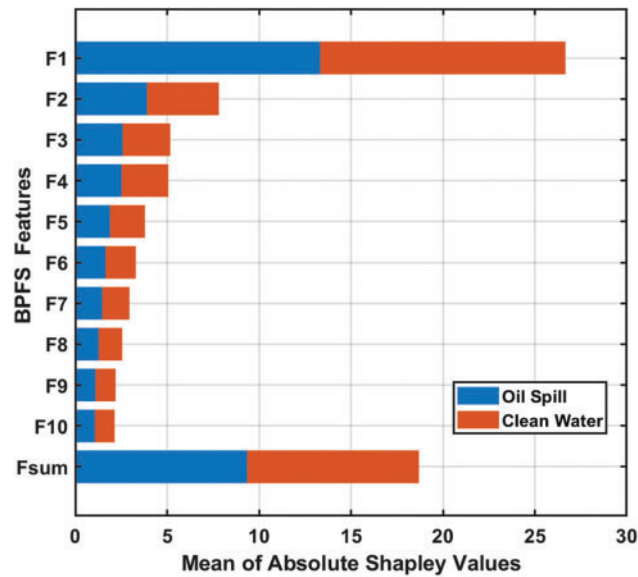


Figure 9: Shapley feature importance bar chart for LR classifier trained by the proposed BPFS feature set. ‘Fsum’ presents the summation of the SHAP values for the remaining 554 features in the PBFS

4.4 Ablation Study

In this subsection, we assess the relative contribution of each component of the proposed hybrid model (MobileNetV2, PCA, BHHO) through an ablation study. In this study, each component is systematically removed while the rest of the model architecture is unchanged to isolate the impact of each component on the overall classification performance. To evaluate the contribution of each component, the following model configurations are tested:

- Model A (Baseline CNN performance): Keeping only MobileNetV2 without PCA and BHHO.
- Model B: Removing PCA and keeping MobileNetV2 + BHHO.
- Model C: Removing BHHO and keeping MobileNetV2 + PCA.
- Model D: The proposed complete hybrid model with MobileNetV2 + PCA + BHHO.

For each model variant, we measured the accuracy of the overall classification performance, as depicted in Table 5. MobileNetV2 has been kept to form the baseline model as it is the feature extraction tool in the study. The improvement percentages in the classification accuracy for the ablated models when using Model A as the baseline model are also provided in Table 5.

Table 5: Classification performance comparison between variant models of the proposed hybrid deep learning in the Ablation study. Entries with the highest values are highlighted in bold

Model configuration	Accuracy %	Contribution analysis	Accuracy improvement %
MobileNetV2 Only—Baseline CNN performance without PCA and BHHO (A)	97.6	Evaluate the impact of the integrated feature selection approach (PCA + BHHO)	–
Without PCA (B)	98.8	Shows the impact of PCA	1.23%

(Continued)

Table 5 (continued)

Model configuration	Accuracy %	Contribution analysis	Accuracy improvement %
Without BHHO (C)	98.5	Shows the impact of BHHO	0.92
Complete Hybrid Model (D)	99.2	All components are included	1.64%

The ablation study demonstrated that the integration of PCA improved the accuracy of the baseline model by 0.92%, while the BHHO feature selection method contributed a 1.23% improvement in the classification accuracy. The integrated feature selection approach (using both PCA and BHHO) is considered the most significant component of the proposed system as it yielded the highest improvement of 1.64% in accuracy over the baseline model. These findings validate the architectural choices and emphasize the effect of combining deep learning with the integrated PCA and BHHO feature selection approach.

4.5 Computational Efficiency of the Proposed Pipeline

The proposed pipeline has been implemented using MATLAB R2024 (b). All experiments have been executed on a laptop with a 13th Gen Intel(R) Core (TM) i7-13620H-2.40 GHz processor and 32 GB RAM. Using the described environment, the processing time for each pipeline phase is depicted in Table 6. The total processing time for the proposed pipeline (experiment 4) is 293.31 s (approximately 5 min). The pipeline includes feature extraction using deep learning CNN, feature optimization using BHHO and PCA, and ML model hyperparameter optimization and training. The phase of hyperparameter optimization and the 6-fold cross-validation training scheme of the LR classifier consumed the largest processing time, followed by the BHHO feature optimization phase. While traditional machine learning models generally require less computational time if trained with the same dataset, the processing time of our approach is influenced by deep learning-based feature extraction, feature selection, and optimization steps. Given the complexity of the approach and available resources, the reported processing time is reasonable. Additionally, computational efficiency can be further improved in real-world deployment through hardware acceleration (GPUs/TPUs), parallel processing, and model quantization. In future work, we will explore efficiency optimizations for real-time applications.

Table 6: Processing time of the phases of the proposed hybrid learning and integrated PCA-BHHO feature selection pipeline

Pipeline phase	Processing time (s)
Feature extraction with MobileNetV2	22.13
PCA feature selection	0.53
BHHO feature optimization	73.63
PCA-BHHO feature integration	0.003348
LR model hyperparameter optimization, training with 6-fold cross-validation	197.02
Entire pipeline	293.31

4.6 Classification Performance Benchmarking

Benchmarking with other ML and NN models has been conducted to understand the proposed model's performance clearly. [Table 7](#) depicts the classification accuracy of several models trained and evaluated using the BPFS. The comparison shows that the LR model achieves the highest accuracy over the other ML models. It is noticeable that the SVM and the wide NN exhibit slightly lower performance than the LR model, but outperform all other models.

Table 7: Comparison between the classification performance of the proposed LR-based model and other ML models trained and evaluated using the PBFS for oil spill detection

Model	AC %
Fine tree	96
Medium tree	95.9
Coarse tree	93.6
Fine KNN	98.1
Medium KNN	97.3
Coarse KNN	89
Ensemble of boosted trees	98
Ensemble of subspace KNN	98.9
Ensemble of bagged trees	97.5
Linear SVM	99.1
LR	99.2
Wide neural network	99.1
Tryliered neural network	98.9

Next, the classification results of our proposed oil spill detection approach are compared to a few studies in the literature that utilized deep learning approaches to detect oil spills in Sentinel-1 SAR imagery. [Table 8](#) lists the studies, their oil spill detection methods, and the corresponding classification accuracy. The comparison reveals that our proposed deep learning-based PCA-BHHO approach achieves the highest classification performance and outperforms the state-of-the-art approaches applied to Sentinel-1 SAR images for oil spill detection.

Table 8: Comparison of the classification performance between the proposed approach and other deep learning-based methods for oil spill detection in Sentinel-1 SAR images

Study	Deep learning-based oil spill detection method in Sentinel-1 SAR images	Classification accuracy
[11]	Pretrained Resnet-50 CNN integrated with Infinite Gamma Mixture models for feature selection.	94.5%
[12]	Custom Artificial Neural Network.	87%
[3]	Custom 23-layer CNN model employing U-Net for oil spill segmentation.	92%

(Continued)

Table 8 (continued)

Study	Deep learning-based oil spill detection method in Sentinel-1 SAR images	Classification accuracy
[13]	A two-stage CNN using a one-vs-rest strategy is followed by a refined pixel-wise classification of the areas detected in the first stage.	98.1%
Our proposed approach	Hybrid Deep Learning and Integrated BHHO-PCA Approach.	99.2%

Our proposed DL model was developed using C-band SAR imagery and is expected to generalize to other C-band datasets from satellites with comparable system characteristics, such as the RADARSAT Constellation Mission. The similarity in frequency and polarization across these sensors supports this generalizability. However, applying the model to SAR datasets from other frequency bands (e.g., L-band or X-band) would require further evaluation. Future work could explore transformer techniques to extend the model's applicability beyond C-band sensors.

It is worth mentioning that in real-world operational scenarios, challenges such as false positives and false negatives are critical considerations. While the DL model proposed in this study shows promising results, its performance may vary in real-world conditions, and further validation is necessary to assess its robustness for operational deployment. Additionally, ethical concerns regarding the potential misuse of satellite-based monitoring must be acknowledged. Although satellite data provides a powerful tool for environmental monitoring, it is essential to ensure that results are accurately interpreted and used responsibly. Proper validation and contextual understanding are essential to prevent the misuse of the technology in regulatory or legal settings. The operational implementation of the proposed DL model could be integrated into marine oil spill response frameworks similar to those outlined in [22,23]. Regarding computational efficiency for operational implementation, our model could leverage parallel processing to handle large datasets effectively. Additionally, hardware acceleration, such as Graphics Processing Units, could optimize performance, making the model suitable for real-world implementation.

5 Conclusions

This study proposed a new hybrid deep learning and integrated BHHO-PCA feature selection approach for oil spill detection using SAR imagery. MobileNet CNN was applied to extract deep features. The PCA and BHHO techniques were utilized to provide a hybrid feature optimization approach to promote the classification performance of a logistic regression classifier. The performance of the LR classifier trained using various feature sets has been examined. The proposed hybrid feature set recorded the highest classification accuracy over all other feature combinations. The feature importance analysis revealed that the feature set selected by the proposed feature engineering approach contains influential features that contributed significantly to the ML model. Moreover, the findings of the ablation study validate the architectural choices and emphasize the significant effect of combining deep learning with the integrated PCA and BHHO feature selection approach. The proposed model showed statistical significance in improving classification accuracy over all other models tested in this study. Additionally, the proposed approach outperforms the state-of-the-art approaches applied to Sentinel-1 SAR images for oil spill detection. The reported processing time of the proposed pipeline is reasonable given the complexity of the approach, and future work will explore computational efficiency optimizations for real-time applications. Future work will consider using datasets from

different satellite sensors (such as multi-spectral imagery), diverse environmental conditions, and varying oil spill sizes to refine the proposed DL model further. Transformer networks will be explored to extend the proposed approach's applicability beyond C-band sensors. Furthermore, the operational implementation of the proposed DL model in real-world operational scenarios will be explored in future work.

Acknowledgement: The authors would like to extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number RI-44-0456.

Funding Statement: The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number RI-44-0456.

Author Contributions: The authors confirm their contributions to the paper as follows: Ghada Atteia and Mohammed Daboor conceived the study. Konstantinos Karantzas and Mohammed Daboor provided the SAR imagery. Ghada Atteia implemented the DL framework. Ghada Atteia, Mohammed Daboor, Konstantinos Karantzas and Maali Alabdulhafith analyzed the results. All authors contributed to the writing of the paper. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The dataset used in this study was acquired by the Sentinel-1 satellites and distributed by the European Space Agency (ESA) via the Copernicus Data Space Ecosystem at <https://dataspace.copernicus.eu> (accessed on 1 January 2025).

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

Abbreviations

AI	Artificial Intelligence
ROC	Receiver Operating Characteristics
SAR	Synthetic Aperture Radar
LR	Logistic Regression
AUC	Area Under Curve
PCA	Principal Component Analysis
CNN	Convolutional Neural Network
BHHO	Binary Harris Hawk Optimization

References

1. Li J, Yu Z, Yu L, Cheng P, Chen J, Chi C. A comprehensive survey on SAR ATR in deep-learning era. *Remote Sens.* 2023;15(5):1454. doi:10.3390/RS15051454.
2. Bianchi FM, Espeseth MM, Borch N. Large-scale detection and categorization of oil spills from SAR images with deep learning. *Remote Sens.* 2020;12(14):2260. doi:10.3390/RS12142260.
3. Shaban M, Salim R, Khalifeh HA, Khelifi A, Shalaby A, El-Mashad S, et al. A deep-learning framework for the detection of oil SPILLS from SAR data. *Sensors.* 2021;21(7):2351. doi:10.3390/S21072351.
4. Huang X, Zhang B, Perrie W, Lu Y, Wang C. A novel deep learning method for marine oil spill detection from satellite synthetic aperture radar imagery. *Mar Pollut Bull.* 2022;179(1):113666. doi:10.1016/J.MARPOLBUL.2022.113666.
5. Yang YJ, Singha S, Mayerle R. A deep learning based oil spill detector using Sentinel-1 SAR imagery. *Int J Remote Sens.* 2022;43(11):4287–314. doi:10.1080/01431161.2022.2109445.
6. Hasimoto-Beltran R, Canul-Ku M, Díaz Méndez GM, Ocampo-Torres FJ, Esquivel-Trava B. Ocean oil spill detection from SAR images based on multi-channel deep learning semantic segmentation. *Mar Pollut Bull.* 2023;188(1):114651. doi:10.1016/J.MARPOLBUL.2023.114651.

7. Dehghani-Dehcheshmeh S, Akhoondzadeh M, Homayouni S. Oil spills detection from SAR earth observations based on a hybrid CNN transformer network. *Mar Pollut Bull.* 2023;190(5):114834. doi:10.1016/J.MARPOLBUL.2023.114834.
8. Li C, Kim DJ, Park S, Kim J, Song J. A self-evolving deep learning algorithm for automatic oil spill detection in Sentinel-1 SAR images. *RSEnv.* 2023;299:113872. doi:10.1016/J.RSE.2023.113872.
9. Ahmed S, ElGharbawi T, Salah M, El-Mewafi M. Deep neural network for oil spill detection using Sentinel-1 data: application to Egyptian coastal regions. *Geomat Nat Hazards Risk.* 2023;14(1):76–94. doi:10.1080/19475705.2022.2155998.
10. Mahmoud AS, Mohamed SA, El-Khoriby RA, AbdelSalam HM, El-Khodary IA. Oil spill identification based on dual attention UNet model using synthetic aperture radar images. *J Indian Soc Remote Sens.* 2023;51:121–33. doi:10.1007/S12524-022-01624-6/TABLES/9.
11. Almulihi A, Alharithi F, Bourouis S, Alroobaea R, Pawar Y, Bouguila N. Oil spill detection in SAR images using online extended variational learning of dirichlet process mixtures of Gamma distributions. *Remote Sens.* 2021;13(15):2991. doi:10.3390/RS13152991.
12. Topouzelis K, Karathanassi V, Pavlakis P, Rokos D. Detection and discrimination between oil spills and look-alike phenomena through neural networks. *ISPRS J Photogramm Remote Sens.* 2007;62(4):264–70. doi:10.1016/J.ISPRSJPRS.2007.05.003.
13. Nieto-Hidalgo M, Gallego AJ, Gil P, Pertusa A, Nieto-Hidalgo M, Gallego A-J, et al. Two-stage convolutional neural network for ship and spill detection using SLAR images. *IEEE Trans Geosci Remote Sens.* 2018;56(9):5217–30. doi:10.1109/TGRS.2018.2812619.
14. Vasconcelos RN, Lima ATC, Lentini CAD, Miranda JG V, de Mendonça LFF, Lopes JM, et al. Deep learning-based approaches for oil spill detection: a bibliometric review of research trends and challenges. *J Mar Sci Eng.* 2023;11:1406. doi:10.3390/JMSE11071406/S1.
15. Copernicus Data Space Ecosystem. Europe's eyes on earth n.d. [cited 2025 Jan 7]. Available from: <https://dataspace.copernicus.eu/>.
16. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data.* 2019;6:1–48. doi:10.1186/S40537-019-0197-0/FIGURES/33.
17. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; 2018 Jun 18–22; Salt Lake City, UT, USA. p. 4510–20. doi:10.1109/CVPR.2018.00474.
18. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans R Soc A Math Phys Eng Sci.* 2016;374(2065):20150202. doi:10.1098/RSTA.2015.0202.
19. Duda RO, Hart PE, Stork DG. Chapter 9—algorithm-independent machine learning. In: *Pattern classification*. Hoboken, NJ, USA: John Wiley & Sons Inc.; 2001. 680 p.
20. Too J, Abdullah AR, Saad NM. A new quadratic binary Harris Hawk optimization for feature selection. *Electronics.* 2019;8(10):1130. doi:10.3390/ELECTRONICS8101130.
21. Dabboor M, Singha S, Montpetit B, Deschamps B, Flett D. Pre-launch assessment of Radarsat constellation mission medium resolution modes for sea oil slicks and look-alike discrimination. *Can J Remote Sens.* 2019;45:530–49. doi:10.1080/07038992.2019.1659722.
22. Tewari S, Sirvaiya A. Oil spill remediation and its regulation. *Int J Res Sci Eng.* 2015;6:1–7.
23. Purohit BK, Tewari S, Prasad KSNV, Talari VK, Pandey N, Choudhury P, et al. Marine oil spill clean-up: a review on technologies with recent trends and challenges. *Reg Stud Mar Sci.* 2024;80(4):103876. doi:10.1016/J.RSMA.2024.103876.