Bima Journal of Science and Technology, Vol. 7 (4) Dec, 2023 ISSN: 2536-6041

DOI: 10.56892/bima.v7i4.522

Oil Spill Detection Using Convolutional Neural Network

[1*]Musa Suleiman Hamza, [1]Suleiman Salihu Jauro, [1]Mustapha Ismail

[1]Department of Computer Science, Faculty of Science, Gombe State University, Nigeria.

Corresponding Author: musashamza@gmail.com

ABSTRACT

Oil spills from large tankers, vessels, and pipeline cracks lead to significant harm and damage to the marine ecosystem. The persistent threat of oil spills necessitates advanced detection methods to safeguard ecosystems and economies. In this study, we proposed a classification technique using a convolutional neural network (CNN) framework for the identification and classification of oil spill images. Traditional approaches grapple with complex patterns and varying conditions, prompting us to harness CNN's proficiency in image recognition. The model, feature extraction, and classification were rigorously evaluated using accuracy, precision, recall, and F1-score metrics. Employing a dataset containing labeled instances of oil and non-oil spills, the classification technique using CNN achieved an accuracy of 94.30%, precision of 83.01%, recall of 88.70%, and an F1-score of 85.70%. These results underscore the proposed potential for accurate oil spill detection. The comparative analysis revealed that while the proposed technique had a slightly lower accuracy than one existing model of 96% and 92%, it excelled in the precision of 79% and 76%, recall of 80% and 84%, and F1-score of both 80%. This highlights its potential as a valuable tool for oil spill detection, offering a more balanced approach to minimizing false positives and false negatives.

Keywords: Oil spill, Convolution neural network, neural network, ecosystem

INTRODUCTION

An oil spill is the discharge of oil onto the surface of a wide river. In the 1960s, marine pollutants became a serious ecological problem, principally as a result of increased oil exploration and production on the continental shelves, as well as the use of supertankers capable of transporting over 500,000 metric tons of oil. Because of rigorous maritime and environmental laws, large oil slicks from damaged or sunk supertankers are now uncommon. Oil slicks associated with well discharges and significant shipper duties cause several small and a few big ones each year. Annually, the world's oceans get more than a million metric tons of oil. Businesses and people greatly affect the entire natural condition when they unintentionally or carelessly bring crankcase oils and used gas solvents into the area. According to (Britannica T, 2023), these sources contribute approximately 3.5 to six million metric tons of oil to worldwide oil streams annually when coupled with ongoing seabed leaking.

Detecting and recognition of oil spills is a critical piece of possibility anticipating oil slicks. The exact discovery of oil slicks and expectation of their directions are useful to fisheries, natural life, obligation debate goals, and asset the board for checking and rationing the marine biological system. Due to the growing oil consumption by many sectors, there has been a significant expansion in ocean oil extraction and transportation during the last thirty years (Seydi et al., 2021). The oil spill, being one of the primary causes of ocean pollution, has detrimental effects on the coastal as well as deep-ocean habitats (Song et al., 2020). Oil spill detection is a topic that is widely studied; however, the majority of the study has focused on very large patches of offshore crude oil (De Kerf et al., 2020). Over expansive and distant ocean areas, satellites timely and at low cost gather data (Sun & Hu, 2019). Additionally, the availability of various remote sensing datasets, particularly open-access remote sensing images, has greatly aided in the frequent and real-time identification of oil slicks.

Through the analysis of photos of the oil spill, Convolutional Neural Networks (CNNs) are essential for the detection of oil spills. By learning to recognize the visual patterns connected to oil spills, these sophisticated algorithms function as extra eyes, allowing for quick and precise detection. As an observant defender, this technology quickly detects any environmental risks and enables timely action. It can be better in monitoring and protecting the environment by utilizing CNNs, which to quickly respond and efficiently ecosystems.

The rest of the paper is organized as follows: section I introduction, section II Literature review, and section three methodology. Results, discussion, and conclusion were drawn in sections IV and V respectively

LITERATURE REVIEW

Models based on machine learning are designed to deal with challenging classification problems via recursively and iteratively assessing probable solutions from samples used for training and data without having to be explicitly programmed to do the job. The following methods or approaches are offered:

The study by (Chehresa et al., 2016) recommended an algorithm for selecting an optimum set of features from Synthetic aperture radar (SAR). The proposed algorithm consists of five steps namely; Identifying dark spots, extracting features, normalizing them, choosing features, and finally classifying them. It was suggested Image advancement, dark area detection, and post-processing are the three methods used to extract dark spots from SAR photos, and dark spot detection suggests the likelihood of an oil spill image. Image enhancement contains some filters (Lee filter, local region filter, max-median filter, and gamma transform) to eliminate noises, prepare it to the dark spots are detected by employing thresholding value determined the dark spot by pixel with lower value, and post-processing is used to eliminate error (regions that mistakenly detect a dark spot or background). The proposed feature extraction used four classified groups to differentiate oil spills from lookalikes namely the geometrical aspect, physical condition, context, and texture feature of an oil spill. The proposed normalization method is interquartile range (IQR) to sidestep features with extensive rate ranges. The suggested feature selection used an evolutionary algorithm to produce candidate feature subsets popular with the execution loop. The proposed classification is Bayesian with Naive Bayes (NB) structure and the rate obtained by classification using an optimal set of features the accuracy is about 93.19%.

In (Yang et al., 2017) suggested an algorithm using the bidimensional empirical mode decomposition's (BEMD) Hilbert spectrum to identify an oil spill from an SAR picture. When processing nonstationary as well as nonlinear signals, the BEMD involves splitting the signal into bidimensional IMFs (BIMF) and residual signals that may be utilized to identify regions of interest (ROI) from feature vectors. As a result, three IMFs and residuals made up the ROI image. Determine the Hilbert spectrum and the Hilbert marginal spectrum to convert the 2-D data into 1-D data. It will be simple to extract

Bima Journal of Science and Technology, Vol. 7 (4) Dec, 2023 ISSN: 2536-6041

DOI: 10.56892/bima.v7i4.522

the amplitude and frequency information, and a feature vector will be created to carry out texture analysis. Proposed Mahala-Nobis distances were used for classification with a rate of accuracy of more than 90%. (Singha et al., 2016) developed a methodology for processing chains using coherent dual-polarimetric terreSAR-X images for oil spill detection. The proposed methodology contained three steps namely dark spot segmentation, feature extraction, and output in the end. As a result, the methodology consists of routine processing and calibration. To recover dark items from the region of interest (ROI), segmentation was carried out using an adaptive thresholding technique. A set of traditional features and a set of polarimetric features were retrieved from each segmented dark spot. polarimetric feature extraction discovered between resolution preservation and speckle reduction. In addition, the support vector machine (SVM) extracted feature parameters used for training and calibration are presented in the methodology's classification accuracy, which indicates that the suggested approach accurately detects 90% of the oil spill and 80% of lookalikes from a dataset with an overall accuracy of 89%.

Study by (Mera et al., 2017) improved oil spill detection by introducing a general and methodical approach based on feature selection (FS) techniques to pick a condensed and pertinent set of features. Five feature selection (FS) methods were examined in the proposed techniques: Recursive Feature Elimination for Support Vector Machine (SVM-RFE), Consistency-based Filter, Information Gain, Relief, and Correlation-based feature selection (CFS). Additionally, a meaningful set of traits with good discriminatory control between oil spills and look-alikes was chosen using the suggested methods. As a result, the feature vector was obtained and decreased using feature selection (FS). SVM classifier was used to assess the techniques, and an overall accuracy of 87.1% was attained.

Table 1: surveys about various techniques associated with oil spillage detection using machine learning.

Technique	Author	Description	Limitations
Support vector machines (SVMs)	(Hassani et al., 2020).	Polarimetric Synthetic Aperture Radar (PolSAR) data with unique capabilities and useful characteristics is an immense source of data for oil spill detection on large scales. recommended dividing the discovered oil spill in the ocean into four types using PolSAR data: heavy oil, thin oil, oil/water mixture, and clear water. The robustness of the suggested SVM classifier is demonstrated by the acquired classification accuracies of 90.21% and 85.41% as well as Kappa coefficients of 0.8052 and 0.7905.	SVM does not perform very well on noisy dataset
	(Dabbiru et al., 2015).	Pixel-by-pixel implementation is achieved by merging the high-resolution SAR data with hyperspectral data and then assessing the combined data using the SVMs classification method. suggested that multi-sensor fusion had advantages, with the fused feature set's overall accuracy surpassing.	
Random forest	(Tong et al., 2019)	A random forest was proposed to improve oil spill detection accuracy using Radarsat-2 and UAVSAR polarimetric SAR datasets, furthermore accuracy of the proposed method reaches 92.99% and 82.25% respectively with two datasets. Therefore, improves the discrimination ability between look-alikes and oil slicks.	Random forest algorithm is ineffective for prediction on a large dataset.

To extract data across several high-level layers of abstraction, deep learning algorithms are a collection of distinct deep neural networks (DNNs) that automatically interpret complicated discriminative features from extraordinarily vast volumes of data in a hierarchy. Deep learning algorithms are remarkably capable and successful in many remote sensing and geoscience domains, partly because they are inspired by the structure and functions of the human brain (L. Zhang et al., 2016). Deep learning is completely data-driven, in contrast to traditional machine learning techniques, where feature representation qualities are only acquired from the data, and natural linkages between input and output data are automatically constructed (Deng & Yu, 2013). So, before the oil spill's classification phase, the feature extraction process that required specialist knowledge to create hand-crafted features was eliminated. Some deep learning models showed impressive performance in oil spill detection from SAR and optical images by automatically collecting discriminatory characteristics that were trained to distinguish between oil spills and lookalikes. Furthermore, the generalization ability of these models can address the case-specificity of traditional techniques.

The article (Yekeen & Balogun, 2020) created a deep learning system for radar imaging-based oil spill detection. For example, they segmented oil spills using the mask-region-based convolutional neural network (Mask R-CNN). Suggested Feature Pyramid Network (FPN) architecture combined with transfer learning on a pre-trained ResNet 101 with COCO data as a backbone and 30 iterations at a learning rate of 0.001 with the great accuracy obtained. (Jiao et al., 2019) proposed an algorithm for detecting oil spills using photos from unmanned aerial vehicles (UAVs). First, a deep convolutional neural network was used to identify the places with the highest potential for oil spills (DCNN). The Otsu thresholding algorithm was then used to enhance the detection outcome. The detailed polygon region from the detection box was then determined using the maximally stable extremal regions (MSER) algorithm. Built an oil completely convolution network, a deep learning architecture for oil spill detection utilizing radar data. The U-Net, encoder, and decoder served as the foundation for their proposed method. Used dense blocks to design a deep learning framework to distinguish between oil- and non-oil-producing locations (Bianchi et al., 2020).

A convolution neural network-based technique was proposed for oil spill detection that uses satellite photos and the superpixel technique of simple linear iterative clustering (SLIC). First, they retrieved some Polari metric features. SLIC superpixels were computed using three channels (HH, HV, and VV). The convolution neural network approach was then regarded to have the extracted features and superpixels as inputs. Lastly, the semantic segmentation technique using a convolution neural network produced the oil spill result, The greatest MIoU of 90.5% was the result. Multiscale learning techniques, which process input images for feature extraction using several convolution kernels with various receptive fields, are used because the size and extent of oil spill black spots vary (J. Zhang et al., n.d.).

Another method for detecting oil spills that uses a one-dimensional (1-D) Convolution Neural Network (CNN) and spectral index-based feature selection to combine hyperspectral data was introduced. They evaluated the suggested method's effectiveness to that of the RF and support vector machine (SVM) classifiers. the outcome of 1-D-based oil spill detection Compared to the other two specified machine

Bima Journal of Science and Technology, Vol. 7 (4) Dec, 2023 ISSN: 2536-6041

DOI: 10.56892/bima.v7i4.522

learning techniques; CNN had a greater accuracy (Liu et al., 2019). introduced adversarial learning of an f-divergence function to produce the segmentation mask of a processed SAR image. The authors first segmented the input image using a deep convolutional neural network (DCNN), and then they utilized a second DCNN to lessen the difference between the segmentation result and the ground truth. However the strategy can not fully utilize the pixel-wise categorization that semantic segmentation approaches may offer; it can only segment one class (oil spills) (Yu et al., 2018).

A convolutional auto-encoder network was suggested that can semantically separate scan lines from images captured by the Side-Looking Airborne Radar (SLAR) system that displays marine classes and oil spills. However the system can only use one parallel autoencoder per class, and the robustness of deep convolutional neural network (DCNN) segmentation models cannot be fully exploited due to the lack of SLAR data (A. J. Gallego et al., 2019).

MATERIALS AND METHODS

Proposed Framework/Method

The proposed framework for oil spill detection involves using Convolutional Neural Networks (CNNs). CNNs are a type of deep learning algorithm known for their remarkable performance in image recognition tasks, making them suitable for identifying patterns and features in imagery of oil spills.

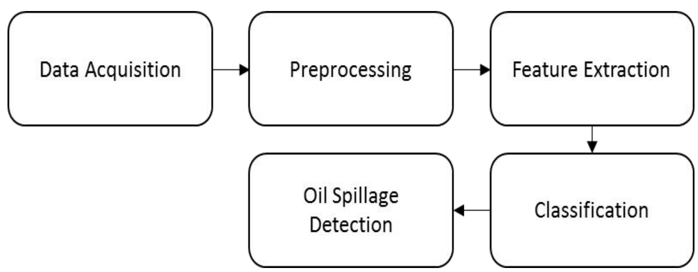


Figure 1: Block diagram of the proposed oil spill detection using neural network

1. Data acquisition

Gather images or photographs of the target area that include instances of both oil spills and non-oil spills. The dataset from Kaggle contains images and labels for oil spill detection.

2. Preprocessing

I. Resize and standardize the images to a common resolution to ensure consistent input for the CNN. Resize images to a common size:

$$X_{resized} = resize(X, height, width) \qquad Eq - i$$

II. Normalize pixel values to a common scale (e.g., [0, 1]). To ensure values are within the range

$$X_{normalizaed} = \frac{X_{resized}}{255} \qquad Eq - ii$$

III. Augment the dataset to boost variety and enhance model generalization by using methods like flipping, rotation, and brightness modifications.

$$X_{aug} = augmentation\_function(X_{norm}) \qquad Eq - iii$$

3. Feature extraction

Utilize a pre-trained CNN architecture (e.g., VGG16, ResNet) as a feature extractor. Remove the classification head and use the intermediate layers to capture high-level features. These features will represent textures, patterns, and shapes in the images. Remove the final classification layer and denote the feature extractor as F. Extract features from input images using:

$$Z = F(X_{aug}) \qquad Eq-iv$$

4. Classification

In this phase, we will build upon the features extracted from the pre-trained CNN architecture and perform binary classification to distinguish between images with or without oil spills.

i. Fully Connected Layers for Classification:

We all add fully connected (dense) layers that take the extracted features and learn to make predictions based on them. These layers will be responsible for capturing complex relationships and patterns present in the data.

ii. Flattening the Extracted Features:

The features extracted from the pre-trained CNN are typically in the form of a 3D tensor. To prepare them for the dense layers, we flatten them into a 1D vector.

iii. Linear Transformation for Each Layer:

For each fully connected layer, we perform a linear transformation by multiplying the input data with weight matrices ($W^{[l]}$) and adding bias terms ($b^{[l]}$) to calculate intermediate values ($Z^{[l]}$).

$$Z^{[l]} = W^{[l]}.A^{[l-1]} + b^{[l]} \qquad Eq-v$$

iv. Applying ReLU Activation Function:

We apply the Rectified Linear Unit (ReLU) activation function element-wise to present non-linearity.

$$A^{[l]} = ReLU(Z^{[l]}) \qquad Eq-vi$$

v. Output Layer for Binary Classification:

The final fully connected layer will have one neuron, representing binary classification (0 non-oil spill, 1 for oil spill). We compute the linear transformation and apply the sigmoid activation to obtain the predicted probability.

$$Z^{[L]} = W^{[L]}.A^{[L-1]} + b^{[L]} \qquad Eq-vii$$

$$A^{[L]} = sigmoid\left(Z^{[L]}\right) \qquad Eq-viii$$

vi. Sigmoid Activation for Probability:

The output of the last neuron is compressed by the sigmoid activation function into a number between 0 and 1, which indicates the likelihood that the input image contains an oil spill.

a. Loss function: Binary Cross-Entropy Loss

The difference between true labels (y(i)) and predicted probabilities (A[L]) for each training example is quantified by the binary cross-entropy loss:

DOI: 10.56892/bima.v7i4.522

$$J = -\frac{1}{m}\sum_{i=1}^{m}\left(y^{(i)}\log\left(a^{[L](i)}\right) + \left(1 - y^{(i)}\right)\log\left(1 - a^{[L](i)}\right)\right) \qquad Eq - ix$$

a. **Regularization: L2 Regularization**

L2 regularization discourages large weights ($W^{[l]}$) by adding a penalty term to the loss function:

$$J_{regularized} = J + \frac{\lambda}{2m}\sum_{l=1}^{L} \| W^{[l]} \|^2 \qquad Eq - x$$

b. **Optimization: Gradient Descent or Adam**

To minimize the loss function, Gradient Descent (or alternative optimization techniques like Adam) modifies weights and biases:

$$W^{[l]} = W^{[l]} - \propto . dW^{[l]} \qquad Eq - xi$$

$$b^{[l]} = b^{[l]} - \propto . db^{[l]} \qquad Eq - xii$$

Where $\propto$ is the learning rate. The trained model can then classify new satellite images as either clean water or containing an oil spill, based on the learned features and adjusted weights and biases.

CNN architecture

The architecture of a Convolutional Neural Network (CNN) entails a series of layers designed to process and extract features from input images.

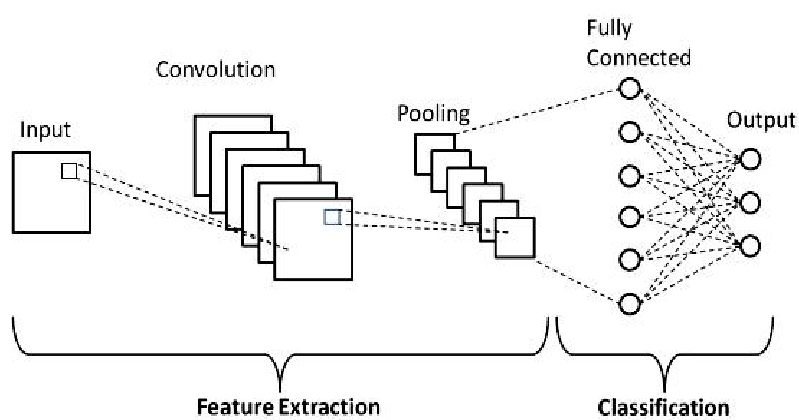


Figure 2: CNN architecture

1. **Input Layer:**

The input for this layer is the unprocessed picture data. Usually, images are represented as matrices of pixel values, with a different matrix for each channel (for example, red, green, and blue).

2. **Convolutional Layers:**

Convolutional layers are the heart of CNNs. They entail multiple filters (also named kernels) that slide over the input image. Each filter performs element-wise multiplication and accumulation, highlighting specific features like edges, textures, or patterns. Convolutional layers are responsible for feature extraction.

By creating a unique kernel K(x, y) (mask), which may be applied locally throughout the image to collect edges, a high responsiveness would be obtained if this particular feature materialized. Applying a filter to an image I(x, y) could be useful to extract some pertinent features or information, such as edges.

We apply the kernel K of size mxn to the M N image I using the convolution technique, otherwise indicated by an asterisk:

$$S(x,y) = (I * K)(x,y) = \sum_{x=1}^{m}\sum_{y=1}^{n} I(x - m, y - n).k(m,n) \qquad Eq - xiii$$

This is carried out for every pixel in the (x, y) picture. The goal is to apply the kernel K on image I in the form of a sliding window to apply it to the entire image. Each pixel's convolution operation is applied locally as the kernel moves over the image. This implies that the kernel is applied to each pixel in the image, updating each pixel as a result.

3. Activation Function:

Following convolution, non-linearity is sometimes introduced element-wise by applying an activation function such as ReLU (Rectified Linear Unit), which enables the network to learn intricate correlations in the data.

4. Pooling Layers:

The spatial dimensions of the feature maps are decreased while crucial information is kept via pooling layers (like MaxPooling or AveragePooling). This aids in reducing computation and managing overfitting.

5. Fully Connected Layers:

Based on the features that have been extracted, these layers are utilized to create classifications or predictions. They allow complicated decision-making by connecting all of the neurons in the current layer to all of the neurons in the previous layer.

6. Flattening Layer:

The feature maps usually flatten into a 1D vector before reaching the completely connected layers. This gets the data ready for layers that are fully connected.

7. Dropout Layer:

A regularization technique called dropout involves ignoring certain neurons at random during training. It enhances generalization and helps avoid overfitting.

8. Output Layer:

The network's predictions or classifications are generated by the last layer. Depending on the task, this layer has a different number of neurons. One neuron may have a sigmoid activation function, for example, in binary classification. Several neurons with softmax activity would be present in a multiclass classification.

CNN architectures can vary greatly in terms of their depth, width, and specific configurations of layers. More recent architectures often incorporate design elements to improve training stability, gradient flow, and computational efficiency. Preprocessing, data augmentation, and hyperparameter tuning are also essential for achieving good performance in oil spillage. Because CNNs can automatically extract pertinent features from images, they have shown impressive performance in a variety of computer vision tasks, including the detection of oil spills.

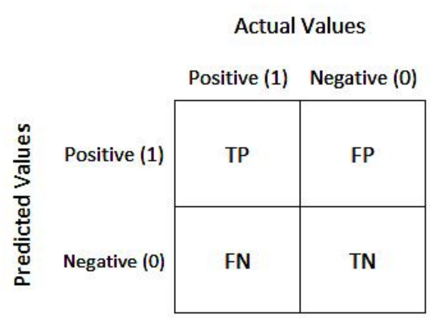
Performance Evaluation and Evaluation Metrics:

Performance evaluation is essential to assess the effectiveness of the proposed oil spill detection method. Common evaluation metrics for binary classification tasks like oil spill detection are presented in Table 2 below.

RESULTS

The section presents the results obtained after the experiments using Google Collab. Table 3 describes the proposed classification technique using CNN model layers and the output shape. The results are presented in tabular and graphical forms which are analyzed using standard performance evaluation metrics as specified during the design. Table 4 below shows the result of oil spill detection on training time, step loss, binary accuracy, validation loss, validation accuracy, and based on the number of epochs. The results obtained were also evaluated using the performance evaluation metrics.

Table 2: show the different Performance Evaluation and Evaluation Metrics

S/No	Performance evaluation	Description	Formula	
1	Accuracy	The proportion of correctly classified instances.	$= \frac{(TP + TN)}{(TP + TN + FP + FN)}$	$Eq - xiv$
2	Precision	The proportion of true oil spill instances among the predicted oil spill instances.	$= \frac{TP}{(TP + FP)}$	$Eq - xv$
3	Recall	The proportion of true oil spill instances correctly identified by the model.	$= \frac{TP}{(TP + FN)}$	$Eq - xvi$
4	F1-score	The harmonic mean of precision and recall provides a balance between the two.	$= 2\frac{(Precision * Recall)}{(Precision + Recall)}$	$Eq - xvii$
5	Confusion Matrix	A matrix representing the number of TP, TN, FP, and FN.		

where TP, FP, TN, and FN represent true positive, false positive, true negative, and false negative respectively.

Bima Journal of Science and Technology, Vol. 7 (4) Dec, 2023 ISSN: 2536-6041

DOI: 10.56892/bima.v7i4.522

Table 3: Convolutional Neural Network (CNN) model

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 254, 254, 32)	896
max_pooling2d (MaxPooling2D)	(None, 127, 127, 32)	0
conv2d_1 (Conv2D)	(None, 125, 125, 64)	18496
max_pooling2d_1 (MaxPooling 2D)	(None, 62, 62, 64)	0
conv2d_2 (Conv2D)	(None, 60, 60, 64)	36928
flatten (Flatten)	(None, 230400)	0
dense (Dense)	(None, 64)	14745664
dense_1 (Dense)	(None, 1)	65
Total params: 14,802,049		
Trainable params: 14,802,049		
Non-trainable params: 0		

Table 4: Results of the Convolutional Neural Network model

No.	Training Time	Step Loss	Binary Accuracy	Validation Loss	Validation Accuracy
1	142s 7s/step	0.8025	0.6980	0.6303	0.6011
2	141s 8s/step	0.5252	0.7828	0.4406	0.7724
3	132s 7s/step	0.1295	0.8918	0.3936	0.8537
4	132s 7s/step	0.3600	0.7097	1.1555	0.8237
5	132s 7s/step	0.7221	0.7258	0.4446	0.8264
.	.	.	.	.	.
.	.	.	.	.	.
6	131s 7s/step	0.9306	0.7559	0.4944	0.8512
7	132s 7s/step	0.1063	0.7989	0.4693	0.8512
8	132s 7s/step	0.0783	0.9900	0.5546	0.8347
9	132s 7s/step	0.7754	0.9730	0.2959	0.7355
10	131s 7s/step	0.7106	0.9794	0.9633	0.9430

The test dataset was used to evaluate the trained model classifier's effectiveness in identifying oil spills. Moreover, the assessment measures were calculated to gauge the model's performance. Based on the evaluation metrics, the following outcomes were attained: The overall accuracy of the model's forecasts is measured by accuracy. It is computed by dividing the total number of instances in the dataset by the number of successfully predicted instances. With an accuracy of 0.9433 in this instance, CNN successfully classified 94.33% of the dataset's instances. The capability of the model to accurately identify positive cases among the instances it predicted as positive is measured by a metric called precision. It is computed by dividing the total number of true positives (positive instances accurately predicted) by the sum of false positives (positive instances anticipated but really negative). With a precision score of 0.8301, it means that 83.01% of the cases that CNN anticipated to be positive were indeed positive. Recall gauges the model's accuracy in classifying positive cases among all of the real positive instances in the dataset. It is sometimes stated as sensitivity or true positive rate. It is computed by dividing the total number of false negatives (positive cases that are mistakenly reported as negative) by the number of true positives. With a recall score of 0.8866, CNN was able to recognize roughly 88.66% of the dataset's real positive instances. The F1-score is calculated as the precision and recall harmonic means. By including recall as well as precision, it provides a balanced evaluation of the model's effectiveness. The following formula is used to calculate it: 2 * (precision * recall) /

(precision + recall). With an F1-score of 0.8571 in this instance, the CNN performed fairly well in terms of precision and recall.

These metrics indicate that the accuracy of the CNN suggested model, which was 94.33%, was comparatively high. The slightly decreased precision and recall, however, suggests that there might be some misclassifications or challenges in precisely predicting positive events. Because the F1 score takes both recall and precision into consideration, it provides a fair evaluation of the model's performance.

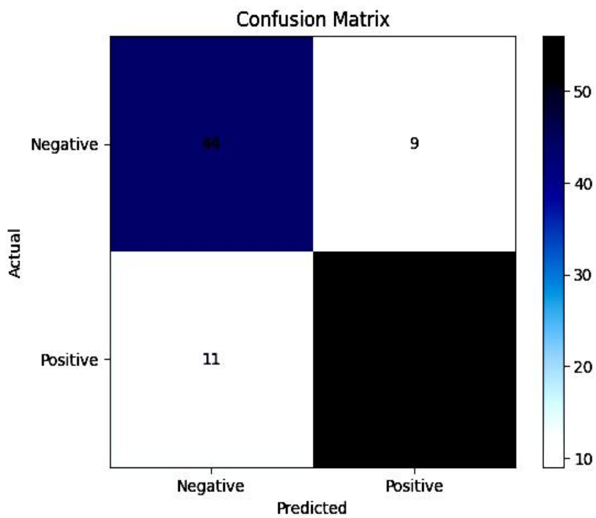


Figure 2: Confusion matrix

A Receiver Operating Characteristic (ROC) curve is a graphical depiction of a classification model's performance at various categorization criteria.

As can be seen in the figure below, the suggested model performs exceptionally well overall and has an excellent Area Under the Curve (AUC) value of 0.96. The model's ROC curve is closer to the top-left quadrant of the graph, showing a high sensitivity and low specificity, or false positive rate, across a range of classification thresholds, with an AUC of 0.96. This shows that the model can accurately distinguish between positive and negative cases, producing a low number of false positives and a high percentage of genuine positives.

With an AUC of 0.96, the model has good predictive ability and consistent performance across a range of categorization thresholds. It implies that the model is capable of making accurate predictions because it has picked up significant patterns and features from the data. AUC values this high are frequently regarded as exceptional outcomes and are suggestive of a strong and trustworthy classification model.

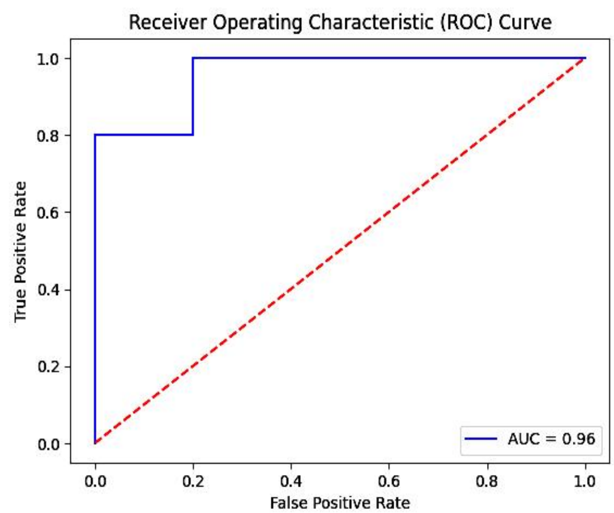


Figure 3: Receiver Operating Characteristic Curve

DISCUSSION

The result of the proposed and the existing oil spill techniques were evaluated using accuracy, precision, recall, and f1 score using tabulation and graphical representation. The table below shows a representation of the proposed oil spill detection with the existing model (Huby Alaa Akram, Raaid Alubady, 2022) and (Shaban et al., 2021).

DOI: 10.56892/bima.v7i4.522

Table 5: Comparison of the proposed models with the existing models for oil spill detection.

Author	Techniques	Accuracy	Precision	Recall	F1 score
Proposed	Classification technique using CNN	94.33%	83.01%	88.66%	85.71%
(Huby Alaa Akram, Raaid Alubady, 2022)	U-net semantic segmentation technique	96.00%	79.00%	80.00%	80.00%
(Shaban et al., 2021)	Two-stage deep-learning framework	92.00%	76.00%	84.00%	80.00%

1. Accuracy:

The proposed model achieved an accuracy of 94.33%. This falls slightly below the accuracy of the highest-performing existing model at 96.00% but still exceeds the accuracy of the other existing model, which stands at 92.00%. While accuracy is important in oil spill detection, it is crucial to consider other evaluation metrics for a comprehensive assessment. The figure below illustrates the comparison of the accuracy of the proposed model to the existing ones.

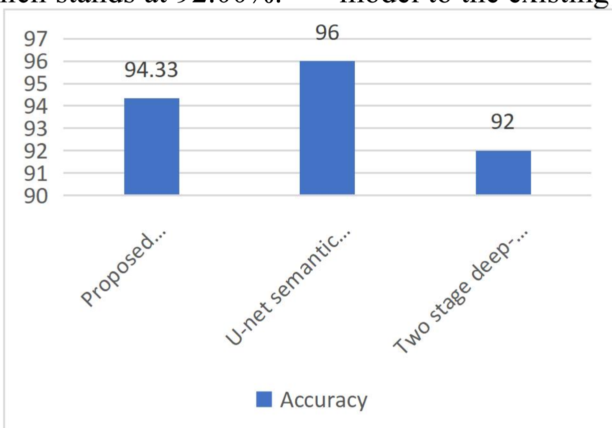


Figure 4: Comparison of accuracy of the proposed model

2. Precision

The suggested model's precision, at 83.01%, beats the two current models of oil spills, which have precision ratings of 79.00% and 76.00%. An increased precision indicates that the suggested model is more adept at detecting real positive cases while reducing the number of false positive predictions related to oil spills. The proposed model's precision is compared to the current models in Figure 5 below.

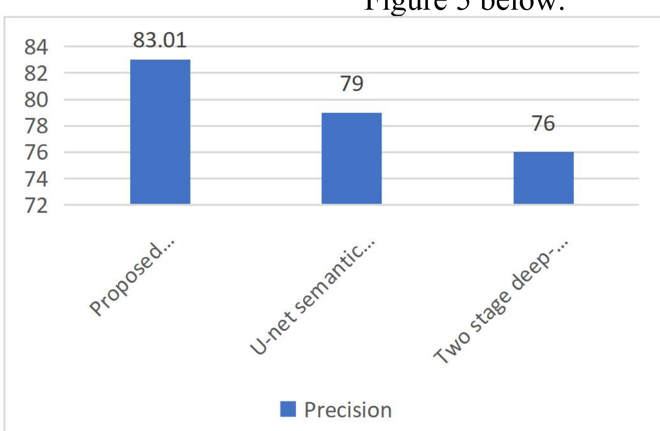


Figure 5: Comparison of precision of the proposed model

DOI: 10.56892/bima.v7i4.522

3. Recall

The suggested oil spill detection model has a recall of 88.66%, which is greater than the recall of the current models, which are 80.00% and 84.00%. This suggests that, as seen in the recall comparison in Figure 6 below, the suggested model is more successful at accurately detecting a higher percentage of real oil spill-positive cases.

4. F1 score

The F1-score of the proposed oil spill model obtained was 85.71%, showing improvement over the F1-scores of the existing models, both of which are at 80.00%. An impartial assessment of the model's efficacy in identifying oil spills is given by the F1-score, which takes precision and recall into account. Figure 7 showcases the comparison of the f1 score of the proposed models.

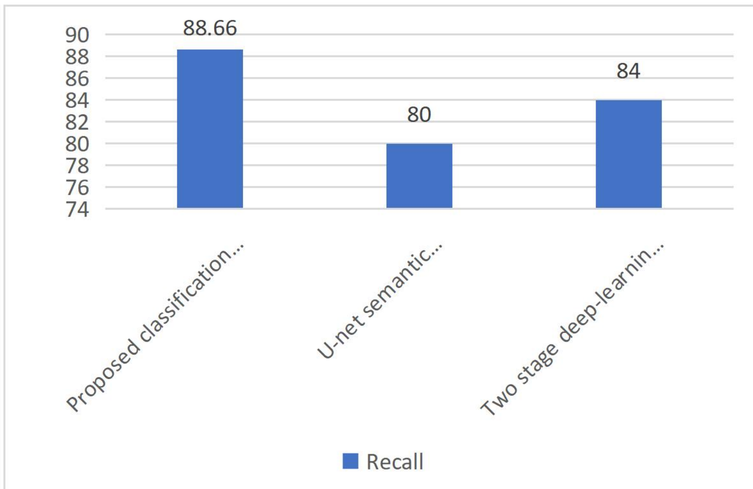


Figure 6: Comparison of recall of the proposed model

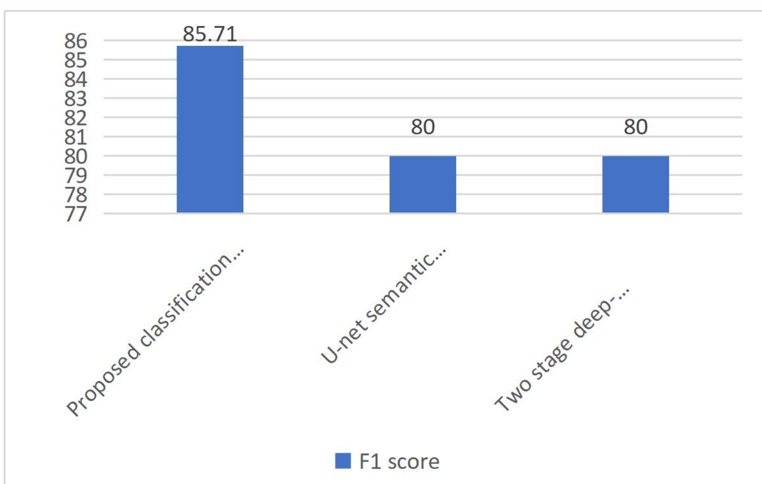


Figure 7: Comparison of the f1-score of the proposed model

Based on the thorough analysis of accuracy, precision, recall, and F1-score, the classification techniques using the CNN model for oil spill detection showcase strong performance. While its accuracy is slightly below that of the highest existing model, the proposed model's higher precision, recall, and F1 score collectively demonstrate its proficiency in identifying oil spills. These results suggested that the model can be a valuable tool in detecting oil spills effectively. However, practical implementation considerations, dataset size, computational requirements, and other domain-specific factors would also be taken into reason when making a final decision.

CONCLUSION

In conclusion, the Convolutional Neural Network (CNN) model achieved promising results in the classification task, with an accuracy of 94.33%. This indicates that the approach was able to correctly classify a significant portion of the cases in the dataset. However, it is worth perceiving that the precision and recall scores were slightly higher, standing at 0.8301 and 0.8866, respectively. This suggests that there might be some limited or negligible misclassifications or challenges in accurately identifying positive instances. Further investigation and fine-tuning of the model could be beneficial to improve its performance in this regard. The result was 0.8571 for the F1-score, which accounts for both recall and precision. By taking into account both the precision with which positive occurrences may be identified and the general accuracy of the forecasts, this score offers a fair assessment of the model's performance. It's crucial to keep in mind that the interpretation of these findings needs to take into account the particular classification task as well as the dataset's class distribution. The results obtained with this CNN model show promise but may require further refinement and evaluation to ensure reliable and robust predictions.

REFERENCE

Bianchi, F. M., Espeseth, M. M., & Borch, N. (2020). Large-scale detection and categorization of oil spills from SAR images with deep learning. *Remote Sensing*, *12*(14).

Britannica T. (2023). *Oil spill*. Encyclopedia Britannica. https://www.britannica.com/science/oil-spill.

Chehresa, S., Amirkhani, A., Rezairad, G. A., & Mosavi, M. R. (2016). Optimum Features Selection for Oil Spill Detection in SAR Image. *Journal of the Indian Society of Remote Sensing*, *44*(5), 775–787.

Chen, G., Guo, H., & An, J. (2017). Research on SAR oil spill image classification based on DBN in small sample space. *2017 4th International Conference on Systems and Informatics, ICSAI 2017*, *2018-Janua*, 479–483.

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(4), 834–848.

Dabbiru, L., Samiappan, S., Nobrega, R. A. A., Aanstoos, J. A., Younan, N. H., & Moorhead, R. J. (2015). Fusion of synthetic aperture radar and hyperspectral imagery to detect impacts of oil spill in Gulf of Mexico. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2015-Novem*, 1901–1904.

De Kerf, T., Gladines, J., Sels, S., & Vanlanduit, S. (2020). Oil spill detection using machine learning and infrared images. *Remote Sensing*, *12*(24), 1–13.

Deng, L., & Yu, D. (2013). Deep learning: Methods and applications. *Foundations and Trends in Signal Processing*, *7*(3–4), 197–387.

Gallego, A.-J., Gil, P., Pertusa, A., & Fisher,

R. B. (2018). Segmentation of Oil Spills on Side-Looking Airborne Radar Imagery with Autoencoders. *Sensors*.

Gallego, A. J., Gil, P., Pertusa, A., & Fisher, R. B. (2019). Semantic segmentation of SLAR imagery with convolutional LSTM selectional autoencoders. *Remote Sensing*, *11*(12).

Hassani, B., Sahebi, M. R., & Asiyabi, R. M. (2020). Oil Spill Four-Class Classification Using UAVSAR Polarimetric Data. *Ocean Science Journal*, *55*(3), 433–443.

Huby Alaa Akram, Raaid Alubady, R. S. (2022). Oil Spill Segmentation from SAR Images Using Deep Neural Networks. *IEEE*, 282–287.

Jiao, Z., Jia, G., & Cai, Y. (2019). A new approach to oil spill detection that combines deep learning with unmanned aerial vehicles. *Computers and Industrial Engineering*, *135*, 1300–1311.

Liu, B., Li, Y., Li, G., & Liu, A. (2019). A spectral feature based convolutional neural network for classification of sea surface oil spill. *ISPRS International Journal of Geo-Information*, *8*(4).

Mera, D., Fernández-Delgado, M., Cotos, J. M., Viqueira, J. R. R., & Barro, S. (2017). Comparison of a massive and diverse collection of ensembles and other classifiers for oil spill detection in SAR satellite images. *Neural Computing and Applications*, *28*, 1101–1117.

Oprea Sergiu, Gil Pablo, Mira Martínez Damián, A. S. B. (2017). Candidate Oil Spill Detection in SLAR Data: A Recurrent Neural Network-based Approach. *SciTePress*, 372–377.

Seydi, S. T., Hasanlou, M., Amani, M., & Huang, W. (2021). Oil Spill Detection Based on Multiscale Multidimensional Residual CNN for Optical Remote Sensing Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 10941–10952. Shaban, M., Salim, R., Khalifeh, H. A., Khelifi, A., Shalaby, A., El-Mashad, S., Mahmoud, A., Ghazal, M., & El-Baz, A. (2021). A deep-learning framework for the detection of oil spills from SAR data. *Sensors*, *21*(7).

Singha, S., Ressel, R., Velotto, D., & Lehner, S. (2016). A Combination of Traditional and Polarimetric Features for Oil Spill Detection Using TerraSAR-X. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(11), 4979–4990.

Song, D., Zhen, Z., Wang, B., Li, X., Gao, L., Wang, N., Xie, T., & Zhang, T. (2020). A Novel Marine Oil Spillage Identification Scheme Based on Convolution Neural Network Feature Extraction from Fully Polarimetric SAR Imagery. *IEEE Access*, *8*, 59801–59820.

Sun, S., & Hu, C. (2019). The Challenges of Interpreting Oil-Water Spatial and Spectral Contrasts for the Estimation of Oil Thickness: Examples From Satellite and Airborne Measurements of the Deepwater Horizon Oil Spill. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(5), 2643–2658.

Tong, S., Liu, X., Chen, Q., Zhang, Z., & Xie, G. (2019). Multi-feature based ocean oil spill detection for polarimetric SAR data using random forest and the self-similarity parameter. *Remote Sensing*, *11*(4). https://doi.org/10.3390/RS11040451

Yang, Y., Li, Y., & Zhu, X. (2017). A novel oil spill detection method from synthetic aperture radar imageries via a bidimensional empirical mode decomposition. *Acta Oceanologica Sinica*, *36*(7), 86–94.

Yekeen, S. T., & Balogun, A. L. (2020). Automated marine oil spill detection using deep learning instance segmentation model. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *43*(B3), 1271–1276. https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-1271-2020

Yu, X., Zhang, H., Luo, C., Qi, H., & Ren, P. (2018). Oil Spill Segmentation via Adversarial f-Divergence Learning. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(9), 4973–4988.

Zhang, J., Feng, H., Luo, Q., Li, Y., Wei, J., & Li, J. (n.d.). *remote sensing Oil Spill Detection in Quad-Polarimetric SAR Images Using an Advanced Convolutional Neural Network Based on SuperPixel Model*.

Zhang, L., Zhang, L., & Du, B. (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, *4*(2), 22–40.