Enhancing Object Detection and Classification Using White Shark Optimization with Deep Learning on Remote Sensing Images

Reda Salama[1,*]

[1]Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, 21589, Saudi Arabia

Emails: rkhalifa@kau.edu.sa

Abstract

Remote sensing (RS) object detection is extensively applied in the fields of civilian and military. The important role of remote sensing is to identify objects like planes, ships, harbours airports, etc., and then it can attain position information and object classification. It is of considerable importance to use RS images for observing the densely organized and directional objects namely ships and cars parked in harbours and parking areas. The object detection (OD) process involves object localization and classification. Due to its wide coverage and longer shooting distance, Remote sensing images (RSIs) have hundreds of smaller objects and dense scenes. Deep learning (DL), in particular convolution neural network (CNN), has revolutionized OD in different fields. CNN is devised to automatically learn the hierarchical representation of data, which makes them fit for feature extraction. Hence, the study proposes a new white shark optimizer with DL-based object detection and classification on RSI (WSODL-ODCRSI) method. The purpose of the WSODL-ODCRSI model is to classify and detect the presence of the objects in the RSI. To accomplish this, the WSODL-ODCRSI model uses a modified single-shot multi-box detector (MSSD) for the OD process. The next stage of OD is the object classification process, which takes place with the use of the Elman Neural Network (ENN) algorithm. The WSO algorithm is exploited as a parameter-tuning model for improving the object classification results of the ENN approach. The stimulated study of the WSODL-ODCRSI algorithm has been established on the benchmark data set and the outcomes underlined the promising performance of the WSODL-ODCRSI model on the object process of classification.

Keywords: Remote sensing; Object detector; Deep learning; White shark optimizer; Computer vision

1. Introduction

During past years, computer vision (CV) technology has made tremendous efforts to overcome the challenges of object detection (OD). Different from image classification, OD inherited from the classification task needs to find the location of the object with a bounding box (BBox), which increases the requirement of the algorithm and makes the task more difficult [1]. Enormous amounts of remote sensing (RS) information were attained from imaging optical sensors on aerial platforms and artificial Earth satellites; both of these methods have the benefits of being realistic and accessible in real-world [2]. The data are categorized into visible, ultraviolet, infrared, hyperspectral, multispectral, or SAR images based on the imaging spectral range. This makes various contributions to the Earth Observation System, which facilitate people's activities and promote our understanding of the environment [3]. In recent times, with the tremendous growth of sensors and RS platforms, both the quality and quantity of RS data have increased a novel difficulties, viz., how to make efficient usage of present data and their application value [4]. As an image analysis application, OD using RSI has attracted much research interest.

DL, which is derived from artificial neural networks (ANN), extracts valuable data via a stacked deep network layer, thus imitating the learning mechanism of the human brain [5]. Due to its robust data analysis and mining abilities, DL has undergone rapid development, which is gaining more popularity among researcher workers [6].

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

Later, DL-based OD was proposed compared to DL-based image classification but developed more rapidly. The OD algorithm has emerged consecutively and has significantly expanded to the RS field [7]. Lately, data-driven deep learning (DL) approaches have endorsed considerable development in OD and segmentation tasks. The detection accuracy can be affected by the quality or quantity of the training dataset [8]. Some challenging and comprehensive data of the natural images, namely MS COCO and PASCAL VOC have driven the growth of OD. Still, OD remains challenging for optical satellite images. Initially, satellite photographs are captured from a bird's-eye view with a wider imaging range and have complete details in ground-based cameras with a horizontal view [9]. Due to the imaging settings, like lighting, occlusion, and views, objects in satellite images often have optical attributes and variable visual appearances [10]. At last, smaller objects often lack sufficient data compared to larger objects, which makes it increasingly difficult to categorize similar or background objects.

This manuscript offers the design of a new white shark optimizer with DL-based object classification and detection on the RSI (WSODL-ODCRSI) technique. The WSODL-ODCRSI method intends to classify and detect the existence of the objects in the RSI. To accomplish this, the WSODL-ODCRSI method employs a modified single-shot multi-box detector (MSSD) technique for the OD process. Besides, for the object classification, Elman Neural Network (ENN) model is used. The WSO technique is applied as a parameter-tuning approach to boost the object classification outcomes of the ENN technique. The stimulated study of the WSODL-ODCRSI technique is investigated on a benchmark dataset.

2. Literature Review

Devi et al. [11] presented an effectual feature extraction approach termed CNN-based features from accelerated segment test (FAST-CNN) and classification as SVM and DT. The presented FAST-CNN was usually creating a feature vector in which 2 situations have to be developed. Primarily, the dense layer extraction of CNN can be carried out for feature extraction. Secondarily, the last layer of FC was exchanged by FAST, and it can be utilized only to detect the stimulating point of images for creating a vector. In [12], Domain Adaptation Fast RCNN (DA Faster R-CNN) approach is projected to identify aircraft in RSIs. Adversarial training was executed for improving the domain shift. At last, the efficiency of the method was trained in the low brightness testing.

Gong et al. [13] examine a context-aware CNN (CA-CNN) method for OD. In the feature extraction, the authors present combining a Context-RoI mining layer as a CNN technique and removing context features by mapping Context-RoI mining in the fore-ground proposal to multi-level mapping features. The authors [14] introduced a multi-scale few-shot OD method for RSI that is performed Faster R-CNN structure. Primary, the authors create the entire backbone of the detector with an involution operator that improves the classifier capability of feature extraction. Afterwards, the presented detector acquires multi-scale features with the support of a path-aggregation component that utilizes semantic data of low-level features to localize and shortens the data broadcast way by the bottom-up flow.

Jiang et al. [15] establish an OD approach, which integrates deep features extraction in CNN for recognizing aircraft and ships in RSIs. The presented technique comprises 2 phases. In the training phase, images of objects with distinct kinds and equivalent labels can be utilized for fine-tuning a pre-trained CNN. These features can be integrated by concatenation and utilized for training SVM. In [16], proposed an effectual region-based VHR-RSI-OD structure termed a Double Multiscale FPN (DM-FPN) that employs inherent multiscale pyramidal features. The proposed method comprises of multi-scale RPN and multiscale OD network, these 2 components share convolution layers and is trained endwise.

In [17], developed a scale adaptive proposal network (SAPNet) for improving the performance of multiple OD in RSIs. The SAPNet contains multi-layer RPNs that are planned for creating multiscale object proposals and the fusion feature layer was executed for optimum multi-object detection. Cheng et al. [18] present an endwise cross-scale feature fusion (CSFF) structure that efficiently enhances OD accuracy. In detail, the authors utilize a feature pyramid network (FPN) for obtaining multiple level mapping features and next insert squeeze and excitation (SE) blocks as the topmost layer to model the connection among distinct feature channels. Afterwards, the authors utilize the CSFF element to acquire great and discriminative multi-level feature representation. Eventually, the authors execute the presented method in the structure of Faster R-CNN.

3. The Proposed Model

In this work, an automatic WSODL-ODCRSI method was determined for object detection and classification on RSI. The aim of the WSODL-ODCRSI approach is to classify and detect the existence of the objects in the RSI. To obtain this, the WSODL-ODCRSI model follows dual phases of operations such as MSSD-based object detector and WSO with ENN-based object classification. Fig. 1 represents the flow of the WSODL-ODCRSI methodology.

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

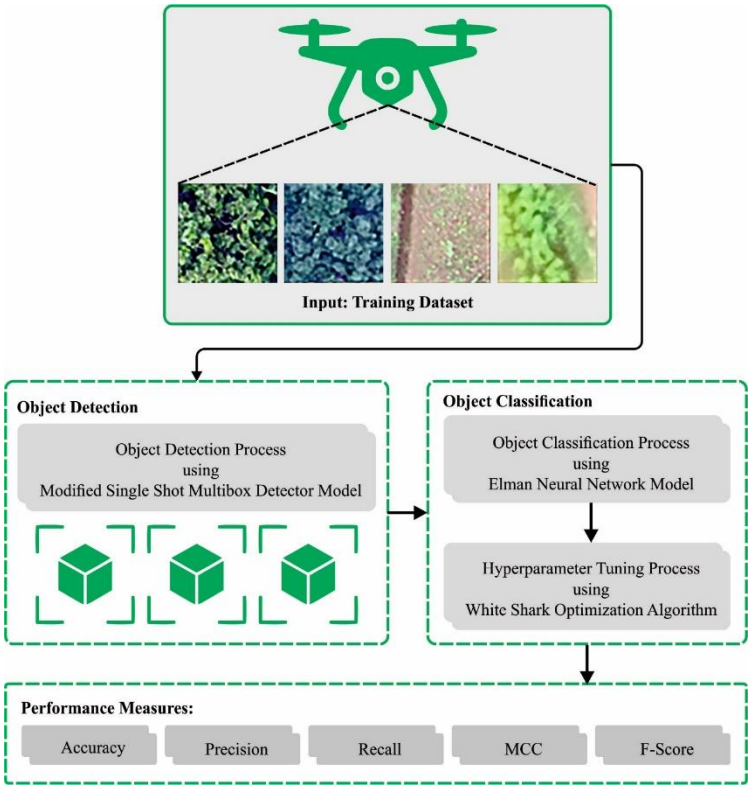

Figure 1. Workflow of WSODL-ODCRSI methodology

A. Object Detection Module

Initially, the MSSD model properly recognizes the occurrence of the objects in the RSI. The regression concept of YOLO was adopted by the SSD model. YOLO model exploits the high-level feature maps for the prediction [19]. SSD model adopts multiscale feature maps for border regression and classification. The detection effects of the smaller object are better than YOLO.

First, the SSD model fine-tune the unique input images to the permanent dimension. SSD model includes the adapted VGG-16 and five-level convolution layers which extract the feature maps. Also, it obtains the multiscale feature maps with the pixels of 64×64,32×32,16×16,8×8,4×4, 2×2 and 1×1 size. SSD model set numbers of prior boundary boxes and various aspect ratios for the feature graph to produce a sequence of boundary boxes. According to the labelled box, the prior box performs IOU matching. The computation is given as follows:

$$F_{IOU} = \frac{S_A \cap S_B}{S_A \cup S_B}, \tag{1}$$

In Eq. (1), the area of ground truth and prior boxes can be correspondingly denoted as S_B and S_A.

In MSSD, the amalgamation of the adapted inception model and deep feature improvement model improves the SSD technique for OD in a complicated environment. Based on the Inception, an enhanced inception model is developed and fused to the dilated convolution with dissimilar dilated amounts that increase the effective receptive area and accelerate the computation speed. Therefore, object feature data was extracted. But the lower-level feature layers with strong object data are not recycled effectively, the portion of object data is lost within the multilayer transmission method, and the attained feature data at all the layers are unbalanced. To overcome these shortcomings, a deep feature enhancement (DFE) model was introduced for boosting the detection effects.

Traditional SSD exploits 1x1 and 3x3 convolutional kernels for single convolution in the cascaded convolutional layer of VGG-16. Its ability to feature extraction becomes weak. An improved inception module was introduced for enhancing the feature extraction. This model has encompassed dilated convolution based on InceptionV4. A cascade convolutional layer comprised of three convolutional sizes with dissimilar kernels and a max-pooling layer combined simultaneously. It increases the network width to enhance the model performance. Firstly, 1x1 convolution decreases the dimension beforehand the convolution process and after the pooling process, the computation cost can be decreased; next, to extract complex features and enhances the computation speed. 5×5 convolutional layer is replaced by two 3×3 convolution kernels, which bring a similar receptive field. Next, to

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

resolve object data loss and the internal data structure loss problems caused by the convolution kernel to attain dissimilar perceptive fields, the dilated rates $(1, 3, 5, 7)$ are presented for replacing the typical convolution in similar convolution. The receptive field can be extended with a similar amount of parameters. Simultaneously, the data of the receptive area is combined to upsurge the network ability of feature representation. At last, to increase the accuracy of the model, a splicing branch has been included for retaining the original feature data of the object. Furthermore, $ReLu$ activation functions and Batch Normalization (BN) has been applied, afterwards the convolution layer to enhance the robustness and accelerate network training speed.

B. Object Classification Module

ENN is a recursive network introduced by the interior layer of self-referencing. The ENN model is exploited for the object classification process. [20]. Hidden layer (HL), context, input, and output layers are the four major components of ENN. Before the HL, the contextual layer is intended to memorize or store the output value. Through the context layer, the outcomes of HL are transported to the input layer. Further, these models process dynamic messages, which realize the dynamic operations. Fig. 2 represents the structure of ENN.

The non-linear formula of the ENN is given below:

$$y(k) = g\left(w^3 x(k)\right) \tag{2}$$

$$x(k) = f\left(w^1 x_c(k) + w^2 u(k-1)\right) \tag{3}$$

$$x_c(k) = x(k-1) \tag{4}$$

From the expression, y signifies the n-dimensional output vector; k denotes the training amount of ENN; x represents the HL neuron output vector; f shows the transfer function of the HL neurons, x_c symbolizes the feedback state vector; g refers to the transfer function of the output neurons; u implies the input vector; and w^1, w^2 and w^3 denotes the connection weight matrices.

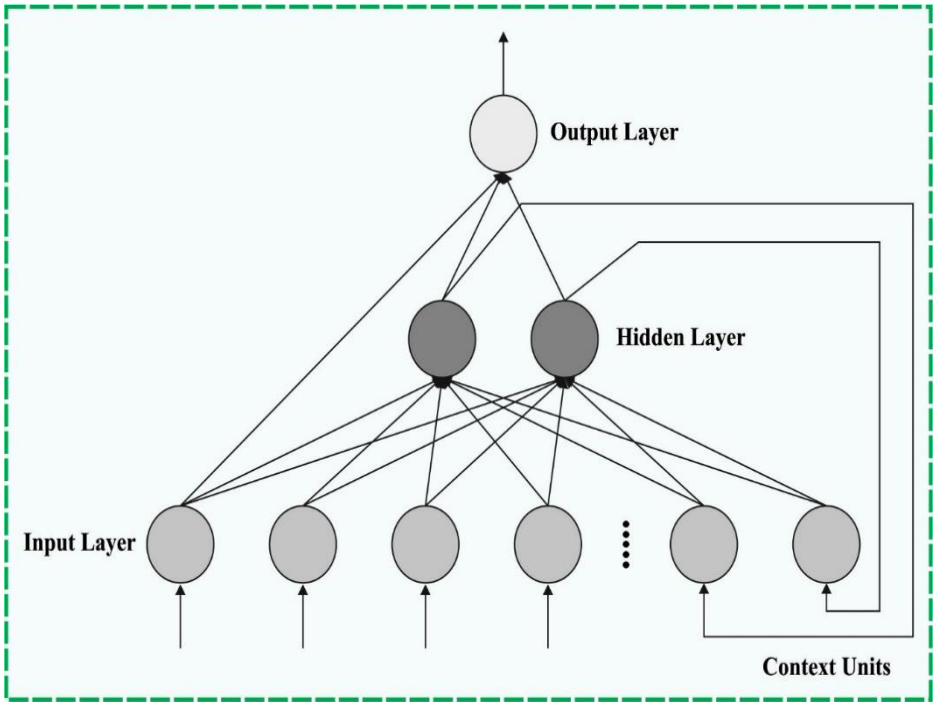

Figure 2. Architecture of ENN

The WSO algorithm has been used to enhance the accuracy of the ENN-based classification model. WSO is inspired by the foraging behaviour of the white shark [21]. By moving the waves, great white sharks in the ocean capture the target and other characteristics for catching the prey kept deep in the ocean. The velocity of sharks to capture the prey, search for the optimum food source, and other shark movements, which are nearby to the optimum food source are the three strategies of white sharks while catching the prey. The initial population of white sharks is given below:

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

$$W_q^p = lb_q + r \times \left(up_q - lb_q\right) \quad (5)$$

In Eq. (5), W_q^p signifies the initial parameter of the p_{th} white shark at q_{th} dimension. r represents a random integer within [0,1]. The lower and upper boundaries in the q_{th} dimension are correspondingly represented as lb_q and up_q.

The velocity of a white shark to discover the prey by using the movement of the sea waves is given as follows:

$$vl_{s+1}^p = \mu \left[vl_s^p + F_1\left(W_{gbest_s} - W_s^p\right) \times C_1 + F_2 \left(W_{best}^{vl_s^p} - W_s^p\right) \times C_2 \right] \quad (6)$$

Where $s = 1,2, \ldots, m$ refers to the index of white sharks with a population dimensions of m. The global optimal location accomplished by any p_{th} shark in s_{th} step is represented as W_{gbest_s}. The new velocity of p_{th} shark can be signified as vl_{s+1}^p in $(s+1)_{th}$ step. The first location of p_{th} shark in s_{th} step is symbolized as W_s^p. The optimum location of the $P\tau h$ sharks and the index vector on attaining the optimum location is signified as $W_{best}^{vl_s^p}$ and vc^i. The initial speed of p_{th} sharks in the s_{th} step is signified as vl_s^p. μ is used to explore the convergence factor of the white shark. F_1 and F_2 denote the force of the white shark for controlling the effects of W_{gbest_s} and $W_{best}^{vl_s^p}$ on W_s^p. C_1 and C_2 are the generation of homogeneously distributed arbitrary integers amongst (0 ,1). The white shark's index vector can be described by:

$$vc = [t \times rand(1,t)] + 1 \quad (7)$$

In Eq. (7), $rand(1,t)$ is a vector of random integers attained with a standard distribution within [0,1]. The following formula is used to control the effect;

$$F_1 = F_{\max} + (F_{\max} - F_{\min}) \times e^{-(4u/U)^2} \quad (8)$$

$$F_2 = F_{\min} + (F_{\max} - F_{\min}) \times e^{-(4u/U)^2} \quad (9)$$

Where u and U denote the initial and maximal amount of iteration, $F_{\min}$ and $F_{\max}$ denote the white shark's present and subordinate velocities. The convergence factor is given as follows;

$$\mu = \frac{2}{\left|2 - \tau - \sqrt{\tau^2 - 4\tau}\right|} \quad (10)$$

In Eq. (10), τ represents the acceleration coefficients. The location updating strategy of the white shark can be given in Eq. (11):

$$W_{s+1}^p = \begin{cases} W_s^p . \neg \oplus W_o + up.c + lo.d; rand < MV \\ W_s^p + vl_s^p / fr; rand \geq MV \end{cases} \quad (11)$$

The newest location of the p_{th} shark at $(s+1)$ iterations, $\neg$ characterizes the negation operators, and c and d signify the binary vector. lo and ub denotes the lower and upper limitations of the search space:

$$c = sgn\left(W_s^p - up\right) > 0 \quad (12)$$

$$d = sgn\left(W_s^p - 1\right) > 0 \quad (13)$$

$$W_o = \oplus (c,d) \quad (14)$$

The frequency of the shark movement is given as follows:

$$fr = fr_{\min} + \frac{fr_{\max} - fr_{\min}}{fr_{\max} - fr_{\min}} \quad (15)$$

In Eq. (15), $fr_{\max}$ and $fr_{\min}$ signify the maximal and minimal frequency rates:

$$MV = \frac{1}{(c_0 + e^{(s/2-S)/c_1})} \quad (16)$$

In Eq. (16), MV shows the weight of terms in the document:

$$W_{s+1}^{'p} = W_{gbest_s} + r_1 \vec{D} is_W sgn(r_2 - 0.5) r_3 < Str_{sns} \quad (17)$$

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

In Eq. (17), $W_{s+1}^{'p}$ denotes the position updating the food source of p_{th} the white shark. The $sgn(r_2 - 0.5)$ produces 1 or -1 to adapt the searching direction. The power of white shark succeeding additional sharks closer to the food source Str_{sns} and shark distance $\vec{D}is_w$ is expressed by Eqs. (18) and (19):

$$\vec{D}is_W = |rand \times \left(W_{gbest_s} - W_s^p\right| \tag{18}$$

$$Str_{sns} = \left|1 - e^{(c_2 \times s/S)}\right| \tag{19}$$

The initial optimum solution remains constant, and the location of other sharks can be updated based on both constant optimum solutions. The fish school behavior of the white shark is given in Eq. (20):

$$W_{S+1}^P = \frac{W_s^p + W_{s+1}^{'p}}{2 \times rand} \tag{20}$$

The weight factor j_{we} is signified as follows:

$$p_{we} = \frac{1}{m-1} * \left(\frac{\sum_{Y=,Y\neq j}^{\rho} q_{fit}}{\sum_{Y=1}^{\rho} q_{fit}}\right) \tag{21}$$

In Eq. (21), q_{fit} is represented as fitness of all the terms in the text document.

$$p_{we} = \frac{1}{m-1} * \frac{[^1fit + ^2fit + \cdots + ^{q+1}fit + \cdots + ^{m-1}fit + ^m fit]}{1fit + ^2fit + \ldots^{q-1} + ^q fit + + ^{q+1} + \cdots + ^{m-1}fit + ^m fit} \tag{22}$$

In the WSO, the concatenation of HM (hybrid mutation) is used for fast convergence. The WSO output is symbolized as $(sel) = \{sel^1, sel^2, \ldots . . sel^m\}$, which is a new sub-group. Simultaneously, m signifies a new number of similar features.

The fitness selection is a prime feature in the WSO system. The encoder solution can be employed to evaluate the integrity of the candidate solution. Now, the precision value represents essential state utilized to improve a FF.

$$Fitness \ = \ \max(P) \tag{23}$$

$$P = \frac{TP}{TP + FP} \tag{24}$$

Here, TP and FP indicate the values of true and false positive.

4. Results and Discussion

In this section, the object detection and classifier outcomes of the WSODL-ODCRSI algorithm are examined on the datasets [22], containing 6450 instances and six different classes as described in Table 1.

Table 1: Details of database

Class	No. of Images
Maize	2075
Banana	1661
Forest	1270
Other	750
Legume	363
Structure	331
Total Images	6450

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

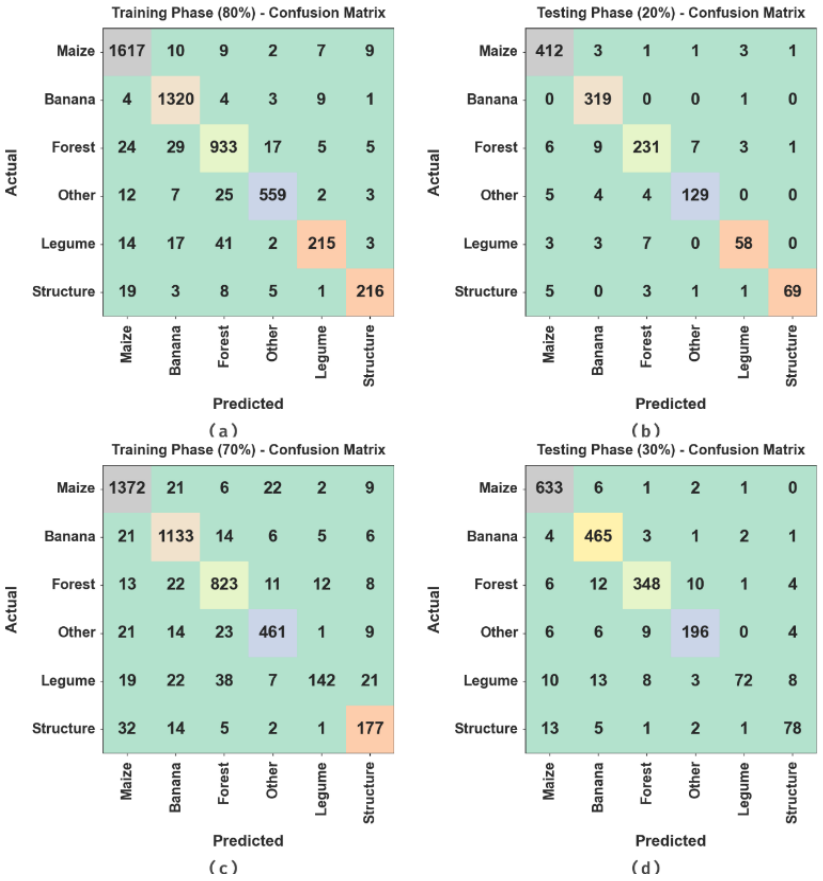

Figure 3. Confusion matrices of (a-b) 80:20 of TRST/TSST and (c-d) 70:30 of TRST/TSST

In Fig. 3, the object classification outcomes of the WSODL-ODCRSI approach are depicted in the confusion matrix form. The figure demonstrates that the WSODL-ODCRSI algorithm properly recognizes the objects.

In Table 2 and Fig. 4, the complete classification results of the WSODL-ODCRSI system are established very clearly. The outcomes recognized the effectual recognition performance of the WSODL-ODCRSI technique. On 80% of TRST, the WSODL-ODCRSI method obtains average $accu_y$ of 98.06%, $prec_n$ of 93.09%, $reca_l$ of 89.93%, F_{score} of 91.35%, and MCC of 90.24%. Also, on 20% of the TSST the WSODL-ODCRSI method achieves an average $accu_y$ of 98.14%, $prec_n$ of 93.74%, $reca_l$ of 91.22%, F_{score} of 92.39%, and MCC of 91.29%.

Table 2: Classifier outcomes of WSODL-ODCRSI algorithm on 80:20 of TRST/TSST

Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC
Training Phase (80%)					
Maize	97.87	95.68	97.76	96.71	95.15
Banana	98.31	95.24	98.43	96.81	95.69
Forest	96.76	91.47	92.10	91.79	89.77
Other	98.49	95.07	91.94	93.48	92.64
Legume	98.04	89.96	73.63	80.98	80.40
Structure	98.90	91.14	85.71	88.34	87.81
Average	98.06	93.09	89.93	91.35	90.24
Testing Phase (20%)					
Maize	97.83	95.59	97.86	96.71	95.11
Banana	98.45	94.38	99.69	96.96	95.99
Forest	96.82	93.90	89.88	91.85	89.91
Other	98.29	93.48	90.85	92.14	91.20
Legume	98.37	87.88	81.69	84.67	83.88
Structure	99.07	97.18	87.34	92.00	91.65
Average	98.14	93.74	91.22	92.39	91.29

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

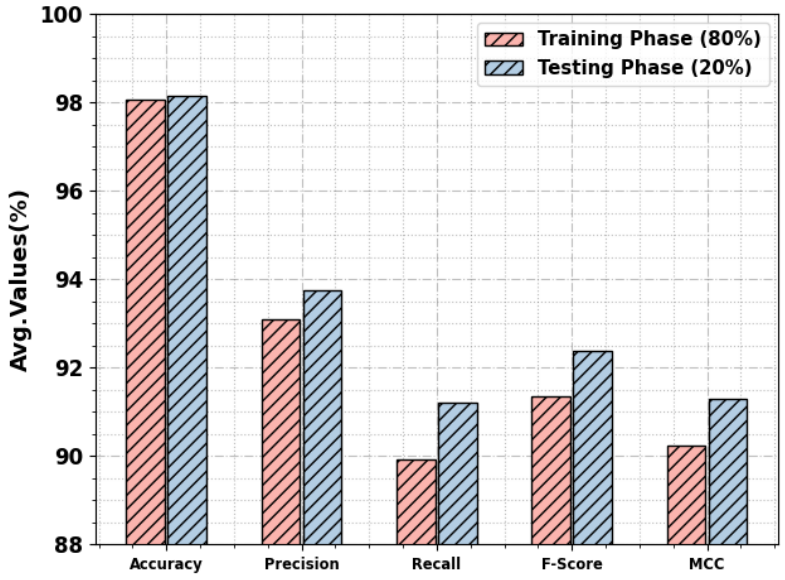

Figure 4. Average outcome of WSODL-ODCRSI algorithm on 80:20 of TRST/TSST

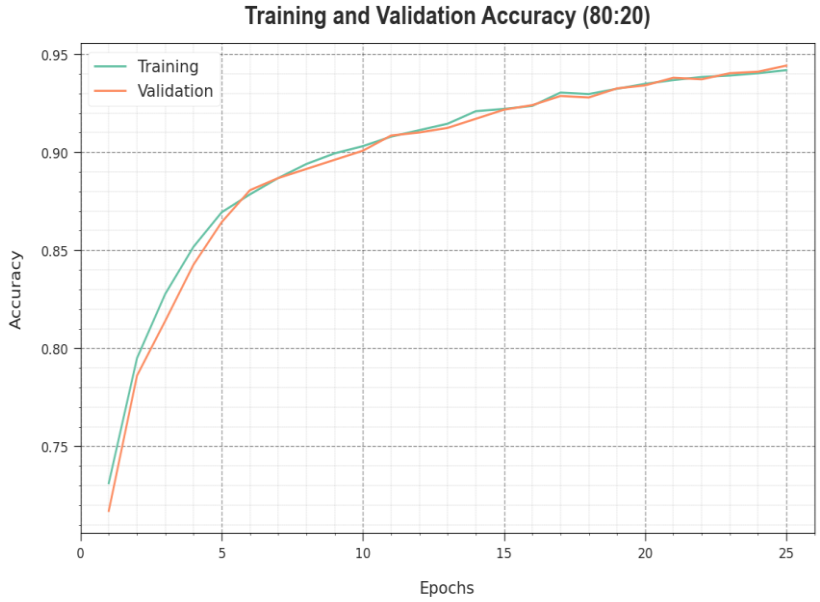

Figure 5. $Accu_y$ Curve of WSODL-ODCRSI method on 80:20 of TRST/TSST

Fig. 5 demonstrates the training $accu_y$ TR_accu_y and VL_accu_y of the WSODL-ODCRSI methodology on 80:20 of the TRST/TSST. The TL_accu_y has been described through the approximation of the WSODL-ODCRSI technique on TR datasets while the VL_accu_y was designed by approximating the outcomes on individual testing datasets. The results show that TR_accu_y and VL_accu_y raise through an increase in epoch counts. Thus, the outcomes of the WSODL-ODCRSI algorithm acquires to enhance both dataset using an increased epoch counts.

In Fig. 6, the TR_loss and VR_loss curve of the WSODL-ODCRSI technique on 80:20 of the TRST/TSST is proven. The TR_loss evaluates the error amongst the original values and predicted values by the TR data. The VR_loss demonstrates the compute of the outcomes of the WSODL-ODCRSI model on individual validation data's. The results indicate that the TR_loss and VR_loss have a tendency to decrease with better epoch counts. It represented the superior values of the WSODL-ODCRSI methodology and its abilities to make precise classification. The reduced values of TR_loss and VR_loss proves the better outcomes of the WSODL-ODCRSI technique in taking pattern and relationship.

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

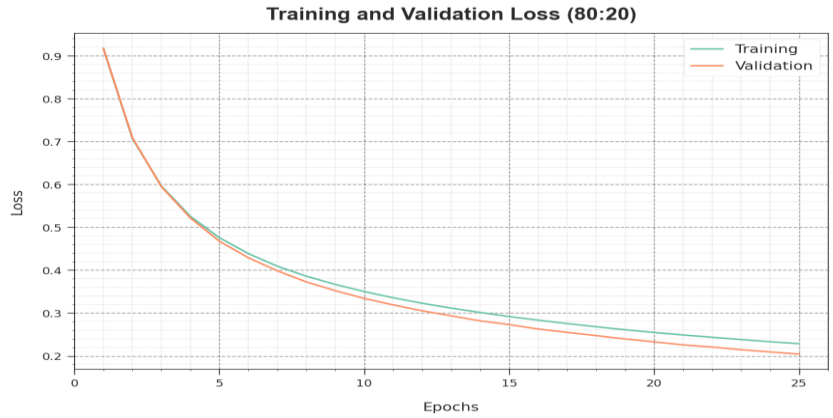

Figure 6. Loss curve of WSODL-ODCRSI technique on 80:20 of TRST/TSST

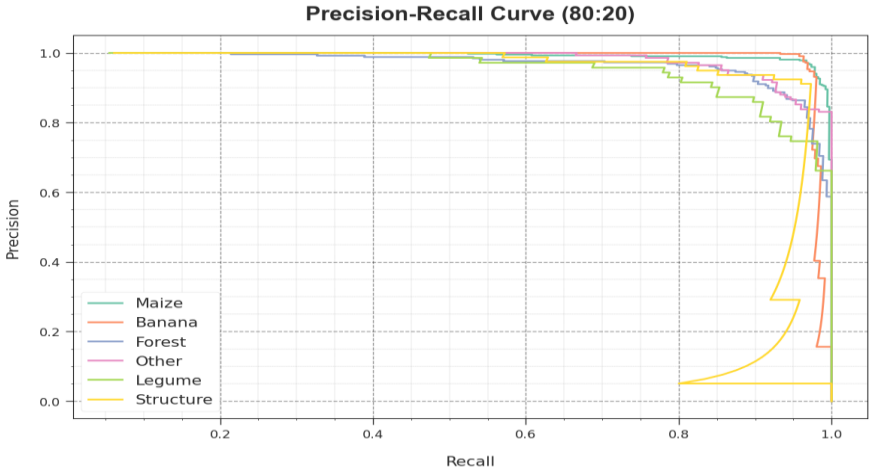

Figure 7. PR curve of WSODL-ODCRSI method on 80:20 of TRST/TSST

A complete investigation of PR of the WSODL-ODCRSI methodology has been proved on 80:20 of the TRST/TSST in Fig. 7. This outcomes indicated that the WSODL-ODCRSI approach resulted in increasing PR values. Additionally, the WSODL-ODCRSI method might gain to better value of PR on every classes.

In Fig. 8, the ROC inspection of the WSODL-ODCRSI approach has been exposed under 80:20 of the TRST/TSST. The figure designated that the WSODL-ODCRSI system leading to superior values of ROC. Furthermore, it's obvious that the WSODL-ODCRSI methodology may spread superior value of ROC on every class labels.

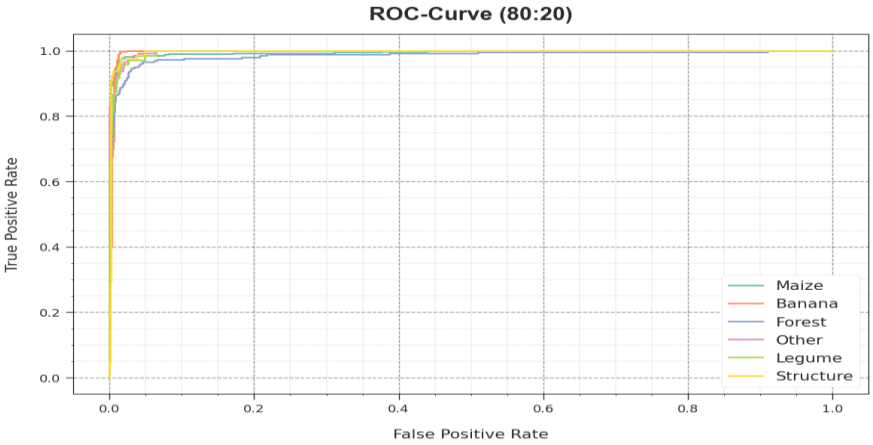

Figure 8. ROC curve of WSODL-ODCRSI approach on 80:20 of TRST/TSST

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

In Table 3 and Fig. 9, the complete classifier results of the WSODL-ODCRSI system are demonstrated. The outcomes showed the effective performance of the WSODL-ODCRSI system. On 70% of the TRST, the WSODL-ODCRSI algorithm gets an average $accu_y$ of 97%, $prec_n$ of 88.40%, $reca_l$ of 84.13%, F_{score} of 85.73%, and MCC of 84.17%. Also, on 30% of the TSST the WSODL-ODCRSI model gains an average $accu_y$ of 97.54%, $prec_n$ of 91.19%, $reca_l$ of 86.22%, F_{score} of 88.18%, and MCC of 86.97%.

Table 3: Classifier outcome of WSODL-ODCRSI algorithm on 70:30 of TRST/TSST

Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC
Training Phase (70%)					
Maize	96.32	92.83	95.81	94.30	91.61
Banana	96.79	92.41	95.61	93.99	91.82
Forest	96.63	90.54	92.58	91.55	89.45
Other	97.43	90.57	87.15	88.82	87.39
Legume	97.17	87.12	57.03	68.93	69.18
Structure	97.63	76.96	76.62	76.79	75.54
Average	97.00	88.40	84.13	85.73	84.17
Testing Phase (30%)					
Maize	97.47	94.20	98.44	96.27	94.41
Banana	97.26	91.72	97.69	94.61	92.86
Forest	97.16	94.05	91.34	92.68	90.93
Other	97.78	91.59	88.69	90.11	88.88
Legume	97.57	93.51	63.16	75.39	75.75
Structure	97.98	82.11	78.00	80.00	78.97
Average	97.54	91.19	86.22	88.18	86.97

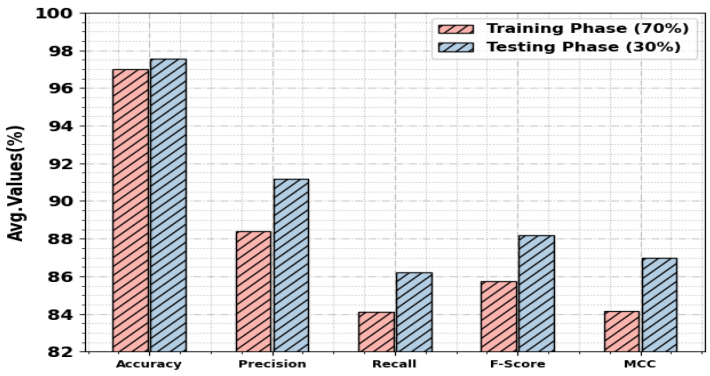

Figure 9. Average outcome of WSODL-ODCRSI algorithm on 70:30 of TRST/TSST

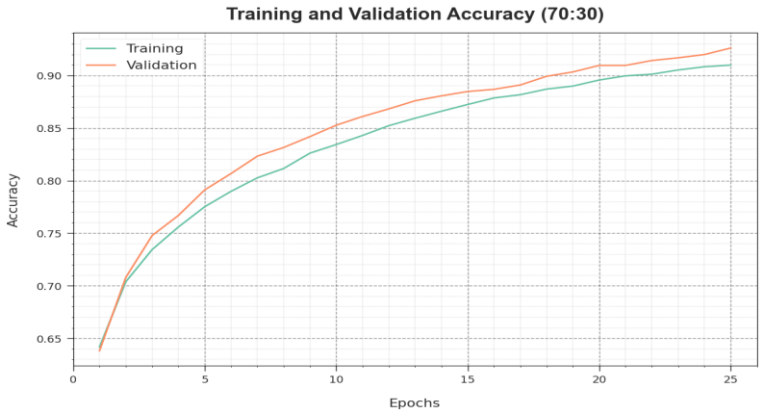

Figure 10. $Accu_y$ curve WSODL-ODCRSI method on 70:30 of TRST/TSST

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

Fig. 10 shows the training $accu_y$ TR_accu_y and VL_accu_y of the WSODL-ODCRSI technique at 70:30 of the TRST/TSST. The TL_accu_y are identified through the valuation of the WSODL-ODCRSI system on TR datasets however the VL_accu_y has been measured by approximating the outcomes on seperate testing datasets. The results showed that TR_accu_y and VL_accu_y raise through an improvement in epoch counts. Thus, the performance of the WSODL-ODCRSI technique was obtained to enhance both datasets with an increased epochs.

In Fig. 11, the TR_loss and VR_loss investigation of the WSODL-ODCRSI technique on 70:30 of the TRST/TSST can be revealed. The TR_loss identifies the error among the original and predicted values by the TR dataset. The VR_loss indicates the calculation of the performance of the WSODL-ODCRSI methodology on individual authentication data. The results indicate that the TR_loss and VR_loss have a tendency to lessen through growing epoch counts. It establishes the superior outcomes of the WSODL-ODCRSI system and its capabilities to make precise classification. The diminished value of TR_loss and VR_loss establishes the superior outcomes of the WSODL-ODCRSI method in taking relationship and pattern.

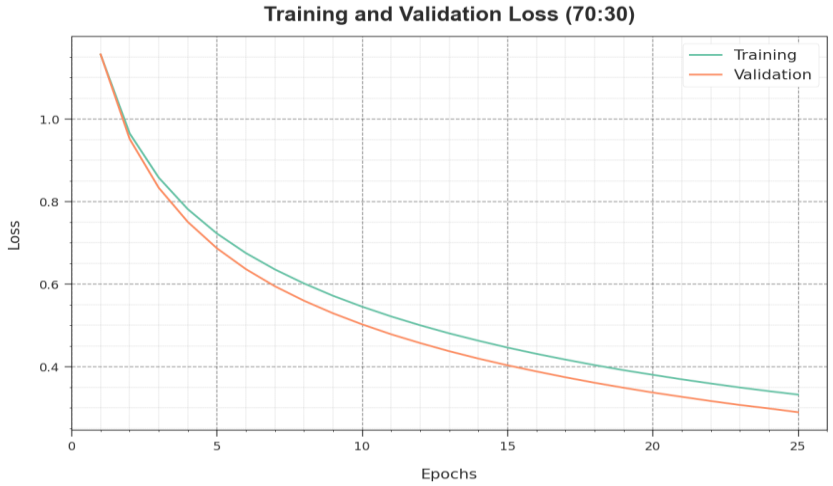

Figure 11. Loss curve of WSODL-ODCRSI method on 70:30 of TRST/TSST

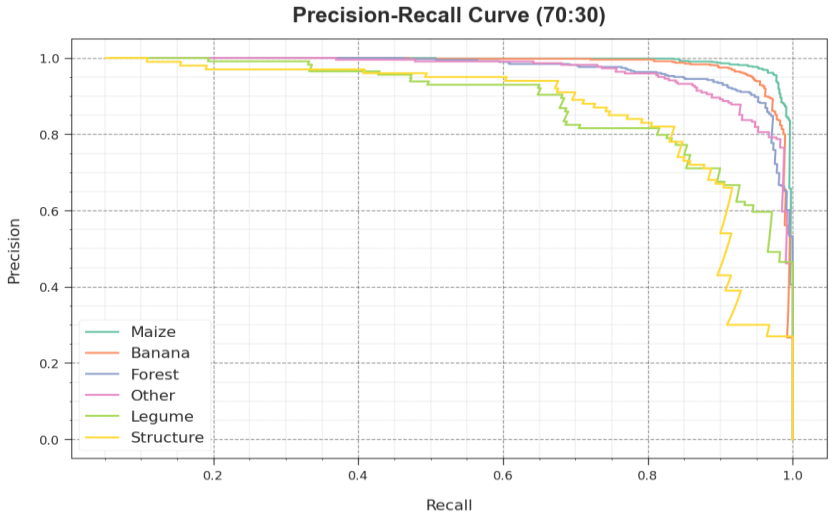

Figure 12. PR curve of WSODL-ODCRSI method on 70:30 of TRST/TSST

A complete PR inspection of the WSODL-ODCRSI approach is presented at 70:30 of the TRST/TSST in Fig. 12. The outcomes indicated that the WSODL-ODCRSI method resulted in maximum PR value. Additionally, the WSODL-ODCRSI system can reach superior values of PR on every classes.

In Fig. 13, the ROC curve of the WSODL-ODCRSI technique is presented at 70:30 of the TRST/TSST. The figure identified that the WSODL-ODCRSI system leading to better ROC values. Moreover, the WSODL-ODCRSI methodology obtained greater values of ROC on every classes.

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

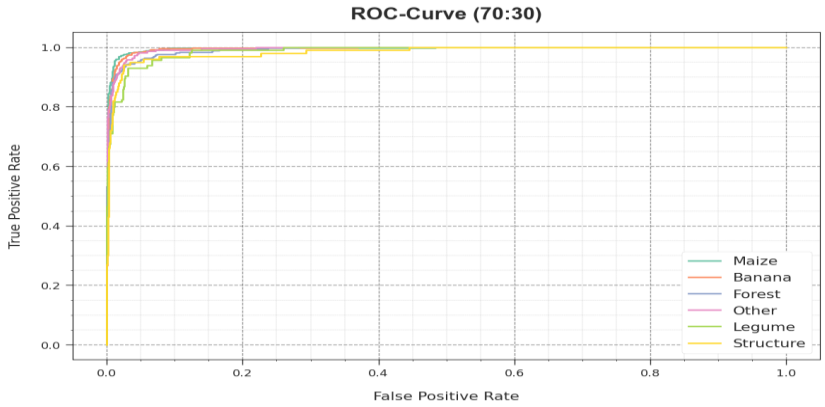

Figure 13. ROC curve of WSODL-ODCRSI method on 70:30 of TRST/TSST

Table 4: Comparative outcome of WSODL-ODCRSI method with existing approaches

Methods	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
WSODL-ODCRSI	98.14	93.74	91.22	92.39
SBODL-FCC	97.43	89.02	85.03	86.74
DNN Model	86.23	86.11	84.39	86.29
AlexNet Model	90.49	87.68	81.7	83.36
VGG-16 Model	90.35	85.28	81.35	85.7
ResNet Algorithm	87.7	86.42	81.18	83.02
SVM Model	86.69	87.99	83.61	84.21

The comparison investigation of the WSODL-ODCRSI algorithm with present classifiers are completed in Table 4 and Fig. 14 [23]. The stimulated outcomes highlighted the supremacy of the WSODL-ODCRSI methodology over another one. According to $accu_y$, the WSODL-ODCRSI model provides a better $accu_y$ of 98.14% however the SBODL-FCC, DNN, AlexNet, VGG-16, ResNet, and SVM system offer decreasing $accu_y$ of 97.43%, 86.23%, 90.49%, 90.35%, 87.7%, and 86.69% respectively. Concurrently, based on $prec_n$, the WSODL-ODCRSI method provides a raising $prec_n$ of 93.74% however the SBODL-FCC, DNN, AlexNet, VGG-16, ResNet, and SVM model give a $diminishingprec_n$ of 89.02%, 86.11%, 87.68%, 85.28%, 86.42%, and 87.99% correspondingly. Concurrently, based on $reca_l$, the WSODL-ODCRSI approach offers raising $reca_l$ of 91.22% but the SBODL-FCC, DNN, AlexNet, VGG-16, ResNet, and SVM methodology provide decreasing $reca_l$ of 85.03%, 84.39%, 81.7%, 81.35%, 81.18%, and 83.61% correspondingly.

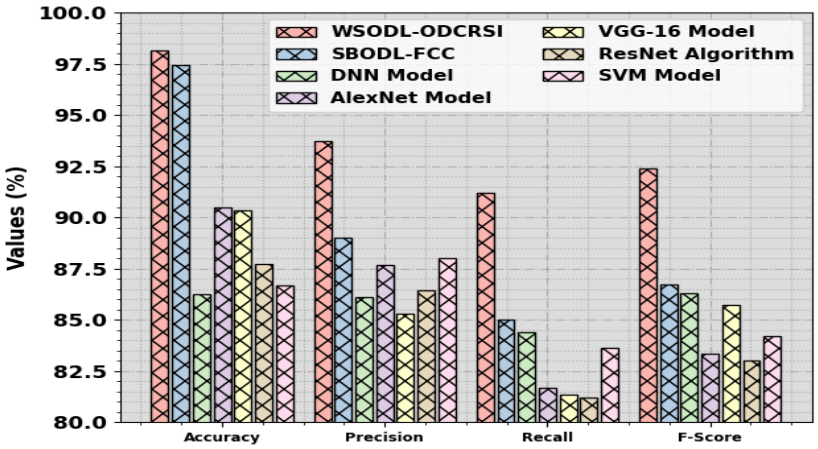

Figure 14. Comparative outcome of WSODL-ODCRSI approach with recent methods

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

These experimental results presented the higher performance of the WSODL-ODCRSI methodology over additional techniques.

5. Conclusion

In this paper, an automatic WSODL-ODCRSI algorithm has been determined for object detection and classification on RSI. The aim of the WSODL-ODCRSI model is to classify and detect the existence of the objects in the RSI. To achieve this, the WSODL-ODCRSI method follows dual phases of operations such as MSSD-based object detector and WSO with ENN-based object classification. Primarily, the MSSD model properly recognizes the occurrence of the objects in the RSI. Next, in the second stage, the recognized objects are categorized using the ENN model and the WSO algorithm can choose its parameters. The experimental study of the WSODL-ODCRSI technique has been investigated on a benchmark dataset and the outcomes highlighted the encouraging performance of the WSODL-ODCRSI algorithm on the object classification technique.

Funding: "The author gratefully acknowledges technical support provided by the Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] Wen, L., Cheng, Y., Fang, Y. and Li, X., 2023. A comprehensive survey of oriented object detection in remote sensing images. Expert Systems with Applications, p.119960.

[2] Sun, H., Chen, Y., Lu, X. and Xiong, S., 2023. Decoupled Feature Pyramid Learning for Multi-Scale Object Detection in Low-Altitude Remote Sensing Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

[3] Ahmed, I., Ahmad, M., Chehri, A., Hassan, M.M. and Jeon, G., 2022. IoT Enabled Deep Learning Based Framework for Multiple Object Detection in Remote Sensing Images. Remote Sensing, 14(16), p.4107.

[4] Fu, S., He, Y., Du, X. and Zhu, Y., 2023. Anchor-free object detection in remote sensing images using a variable receptive field network. EURASIP Journal on Advances in Signal Processing, 2023(1), pp.1-19.

[5] Liu, L., Liu, Y., Yan, J., Liu, H., Li, M., Wang, J. and Zhou, K., 2022. Object Detection in Large-Scale Remote Sensing Images with a Distributed Deep Learning Framework. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, pp.8142-8154.

[6] Zhang, J., Lei, J., Xie, W., Li, Y., Yang, G. and Jia, X., 2023. Guided Hybrid Quantization for Object Detection in Remote Sensing Imagery via One-to-one Self-teaching. IEEE Transactions on Geoscience and Remote Sensing.

[7] Li, Z., Wang, Y., Zhang, N., Zhang, Y., Zhao, Z., Xu, D., Ben, G. and Gao, Y., 2022. Deep learning-based object detection techniques for remote sensing images: A survey. Remote Sensing, 14(10), p.2385.

[8] Su, H., Wei, S., Yan, M., Wang, C., Shi, J. and Zhang, X., 2019, July. Object detection and instance segmentation in remote sensing imagery based on precise mask R-CNN. In IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium (pp. 1454-1457). IEEE.

[9] Yang, L., Yuan, G., Zhou, H., Liu, H., Chen, J. and Wu, H., 2022. RS-Yolox: A high-precision detector for object detection in satellite remote sensing images. Applied Sciences, 12(17), p.8707.

[10] Li, X., Deng, J. and Fang, Y., 2021. Few-shot object detection on remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 60, pp.1-14.

[11] Devi, N.B., Kavida, A.C. and Murugan, R., 2022. Feature extraction and object detection using fast-convolutional neural network for remote sensing satellite images. Journal of the Indian Society of Remote Sensing, 50(6), pp.961-973.

[12] Chen, J., Sun, J., Li, Y. and Hou, C., 2022. Object detection in remote sensing images based on deep transfer learning. Multimedia Tools and Applications, pp.1-17.

[13] Gong, Y., Xiao, Z., Tan, X., Sui, H., Xu, C., Duan, H. and Li, D., 2019. Context-aware convolutional neural network for object detection in VHR remote sensing imagery. IEEE Transactions on Geoscience and Remote Sensing, 58(1), pp.34-44.

[14] Zhao, Z., Tang, P., Zhao, L. and Zhang, Z., 2021. Few-shot object detection of remote sensing images via two-stage fine-tuning. IEEE Geoscience and Remote Sensing Letters, 19, pp.1-5.

[15] Jiang, B., Li, X., Yin, L., Yue, W. and Wang, S., 2019, March. Object recognition in remote sensing images using combined deep features. In 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) (pp. 606-610). IEEE.

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024

[16] Zhang, X., Zhu, K., Chen, G., Tan, X., Zhang, L., Dai, F., Liao, P. and Gong, Y., 2019. Geospatial object detection on high-resolution remote sensing imagery based on a double multi-scale feature pyramid network. Remote Sensing, 11(7), p.755.

[17] Zhang, S., He, G., Chen, H.B., Jing, N. and Wang, Q., 2019. Scale adaptive proposal network for object detection in remote sensing images. IEEE Geoscience and Remote Sensing Letters, 16(6), pp.864-868.

[18] Cheng, G., Si, Y., Hong, H., Yao, X. and Guo, L., 2020. Cross-scale feature fusion for object detection in optical remote sensing images. IEEE Geoscience and Remote Sensing Letters, 18(3), pp.431-435.

[19] Wang, L., Shoulin, Y., Alyami, H., Laghari, A.A., Rashid, M., Almotiri, J., Alyamani, H.J. and Alturise, F., 2022. A novel deep learning-based single shot multibox detector model for object detection in optical remote sensing images.

[20] Yang, M. and Liu, Y., 2023. Research on the potential for China to achieve carbon neutrality: A hybrid prediction model integrated with elman neural network and sparrow search algorithm. Journal of Environmental Management, 329, p.117081.

[21] Parveen, N., Chakrabarti, P., Hung, B.T. and Shaik, A., 2023. Twitter sentiment analysis using hybrid gated attention recurrent network. Journal of Big Data, 10(1), pp.1-29.

[22] J. Rineer, R. Beach, D. Lapidus, M. O'Neil, D. Temple, N. Ujeneza, J. Cajka, and R. Chew, ''Drone imagery classification training dataset for crop types in Rwanda,'' Version 1.0, Radiant MLHub, 2021. [Online]. Available: https://mlhub.earth/data/rti_rwanda_crop_type, doi: 10.34911/rdnt.r4p1fr.

[23] Ahmed, M.A., Aloufi, J. and Alnatheer, S., 2023. Satin Bowerbird Optimization with Convolutional LSTM for Food Crop Classification on UAV Imagery. IEEE Access.

DOI: https://doi.org/10.54216/FPA.170211
Received: January 28, 2024 Revised: April 24, 2024 Accepted: September 26, 2024