AIMS Mathematics, 9(5): 11352–11371.
DOI: 10.3934/math.2024557
Received: 13 December 2023
Revised: 08 February 2024
Accepted: 27 February 2024
Published: 25 March 2024

http://www.aimspress.com/journal/Math

Research article

An intelligent water drop algorithm with deep learning driven vehicle detection and classification

Thavavel Vaiyapuri[1], M. Sivakumar[2], Shridevi S[3], Velmurugan Subbiah Parvathy[4], Janjhyam Venkata Naga Ramesh[5], Khasim Syed[6,*] and Sachi Nandan Mohanty[7]

[1] College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Kingdom of Saudi Arabia
[2] Department of Networking and Communications, College of engineering and Technology, SRM institute of science and technology, Kattankulathur 603203, India
[3] Centre for Advanced Data Science, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India
[4] Department of Electronics and Communication Engineering, Kalasalingam Academy of Research and Education, Tamilnadu, India
[5] Dept. of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur Dist., Andhra Pradesh – 522502, India
[6] School of Computer Science & Engineering (SCOPE), VIT AP University, Amaravati, Andhra Pradesh, India
[7] School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

* **Correspondence:** Email: profkhasim@gmail.com.

Abstract: Vehicle detection in Remote Sensing Images (RSI) is a specific application of object recognition like satellite or aerial imagery. This application is highly beneficial in different fields like defense, traffic monitoring, and urban planning. However, complex particulars about the vehicles and the surrounding background, delivered by the RSIs, need sophisticated investigation techniques depending on large data models. This is crucial though the amount of reliable and labelled training datasets is still a constraint. The challenges involved in vehicle detection from the RSIs include variations in vehicle orientations, appearances, and sizes due to dissimilar imaging conditions, weather, and terrain. Both specific architecture and hyperparameters of the Deep Learning (DL) algorithm must be tailored to the features of RS data and the nature of vehicle detection tasks. Therefore, the current study proposes the Intelligent Water Drop Algorithm with Deep Learning-Driven Vehicle Detection

and Classification (IWDADL-VDC) methodology to be applied upon the Remote Sensing Images. The IWDADL-VDC technique exploits a hyperparameter-tuned DL model for both recognition and classification of the vehicles. In order to accomplish this, the IWDADL-VDC technique follows two major stages, namely vehicle detection and classification. For vehicle detection process, the IWDADL-VDC method uses the improved YOLO-v7 model. After the vehicles are detected, the next stage of classification is performed with the help of Deep Long Short-Term Memory (DLSTM) approach. In order to enhance the classification outcomes of the DLSTM model, the IWDA-based hyperparameter tuning process has been employed in this study. The experimental validation of the model was conducted using a benchmark dataset and the results attained by the IWDADL-VDC technique were promising over other recent approaches.

Keywords: vehicle detection; remote sensing images; unmanned aerial vehicle; YOLO-v7; water drop algorithm
Mathematics Subject Classification: 11Y40

1. Introduction

In the present scenario, the application of Unmanned Aerial Vehicle (UAV) has become a popular area in Remote Sensing (RS) domain, determined by its academic and commercial achievements [1]. However, these practices are highly dissimilar for the same application, primarily attributed to the fact that the data acquisition using sensors is highlighted to be more flexible than the existing traditional methods [2]. In optical Remote Sensing Images (RSIs), Object Detection (OD) corresponds to the action in which an assumed satellite or aerial image holds a single object or multiple objects that suit the period of interest and the location is identified for every forecast object in the image [3]. The term 'object', employed in this study, denotes a general structure of the artificial objects such as vehicles, buildings, ships, etc. with sharp borders. In this terminology, the background location and landscape substances such as land use/land cover (LULC) parcels are not considered since it contains unclear borders and measures of the background location [4]. In order to overcome the essential difficulty in aerial and satellite image analysis, the OD process plays a vital part in optical RSIs for a huge sort of uses such as the Geographic Information System (GIS) upgrades, LULC mapping, environmental hazard recognition, environmental observation, precision agriculture and urban development [5]. In optical RSIs, the OD process frequently encompasses numerous growing tasks including huge differences in the visual appearance of objects affected by perspective variation, background clutter, brightness, shadow, and much more [6].

The OD process further plays an essential role in military as well as civilian uses using the RSIs [7]. However, this process faces multiple hindrances in the form of changing visual form of the objects affected by illumination, obstruction, shadow, resolution, viewpoint variation, polarization, speckle sound, and much more [8]. Moreover, the explosive development of the RSIs in both quality as well as quantity incurs heavy computational cost, which in turn complicates the real-time application of OD. Robust and quick vehicle recognition, from the RSIs, has prospective applications in emergency management, traffic surveillance and financial analysis. Besides, the location and density data of the vehicles act as a vital information for creating Intelligent Transport Systems [9]. However, precise and strong vehicle classification from the RSIs remains one of the most challenging tasks to

accomplish. Traditional vehicle recognition techniques are mainly based on hand-crafted features that are removed from the sliding windows with dissimilar measures. However, these models heavily depend on physically-intended features and it cannot effectively handle huge differences in both backgrounds as well as the targets [10]. Currently, the Convolutional Neural Networks (CNNs) have been employed in aerial image OD and it attained promising outcomes. For example, You Only Look Twice (YOLT) and R-CNNs (Region-CNN) have been applied as a benchmark in various studies.

Numerous tasks are involved in vehicle recognition and classification process from the RSIs utilizing DL technique. One important difficulty is the restricted robustness of the existing techniques to meet the needs for different environmental conditions such as changing lighting, climate, and terrain. This drawback impacts the models' flexibility across dissimilar remote sensing states. In addition to these, it is also challenging to achieve both scalability and simplification of the existing methods to handle large-scale and high-resolution datasets. The need for modification of the hyperparameters to accommodate the remote sensing data parameters frequently demand human intervention. Furthermore, the explainability and the interpretability of the classification choices in DL techniques for remote sensing make use of important yet under-addressed challenges. In order to overcome these hindrances, there is a pressing need exists to develop a method that not only combines the innovative DL architectures but also joins the real hyperparameter tuning plans, thus offering a holistic solution for strong and precise vehicle recognition and classification under various remote sensing atmospheres.

In this background, the current study developed the Intelligent Water Drop Algorithm with a Deep Learning-Driven Vehicle Detection and Classification (IWDADL-VDC) technique to be used upon the RSIs. The goal of the projected technique is to modernize the vehicle recognition and classification processes conducted upon the Remote Sensing Images (RSI) using DL technique. By incorporating innovative DL architectures and executing real hyperparameter tuning plans, the technique seeks to improve adaptability, robustness, and scalability across different environmental states. The IWDADL-VDC model exploits a hyperparameter-tuned DL approach for both recognition and classification of the vehicles. For the vehicle detection process, the IWDADL-VDC technique uses the improved YOLO-v7 model. After identifying the vehicles, the next classification stage is performed using the Deep Long Short-Term Memory (DLSTM) model. In order to improve the classification outcomes of the DLSTM model, the IWDA-based hyperparameter tuning process has been employed in this study. The proposed IWDADL-VDC model was experimentally validated using a benchmark dataset.

2. Literature works

In literature [11], the improved Chimp optimization algorithm with DL-based vehicle detection and classification (ICOA-DLVDC) algorithm under RSIs was proposed. This algorithm had two different stages such as object detection and classification. The EfficientNet architecture was used in this study for vehicle detection. For classification, the sparse autoencoder (SAE) network was used. The study introduced ICOA that simplifies the parameter tuning process which in turn effectively enhanced the hyperparameters of the SAE. Javadi et al. [12] introduced a DNN using YOLOv3 with different baseline networks such as DenseNet-201, DarkNet53, MobileNetv2, and SqueezeNet. In this study, 3D depth maps were produced by parallax displacement and a pair of aerial images. Then, the FCNN model was trained under 3D feature mappings of trucks, trailers, and semi-trailers. At last, the network force was introduced for detecting the vehicles in aerial images.

Gao et al. [13] presented a vehicle recognition module named binary full convolution one-step

object detection (FCOS) and a dataset termed 4MVD for vehicle recognition in the RSIs. During the RPN step, the FCOS module was utilized to generate the candidate box in different images. The dual-step positive and negative example models were implemented to improve the positive plan sampling effect. During the RCNN step, the two-stage classification models were introduced. Ragab et al. [14] devised the Improved DL-assisted Vehicle Detection for Urban Applications employing the RSI (IDLVD-UARSI) method. This algorithm exploited the enhanced RefineDet framework for vehicle detection. The model of classification was incorporated by employing the CAE module. At last, the Quantum-enabled Dwarf Mongoose Optimizer (QDMO) method was exploited in the optimum hyperparameter tuning model.

The foremost aim of the study conducted earlier [15] was two-fold in which the first was the construction of a novel space object image collection. The second aim was to propose a unique RSIC model to achieve high performance classification utilizing the recently-generated data. In general, feature extraction can be performed using the pre-trained MobileNet_V2 and discrete wavelet transform (DWT) methods. In this study, an integration of the DWT and MobileNet_V2 models was used to generate several features. Next, Iterative Neighborhood Component Analysis (INCA) selected the optimum features. Finally, the features selected were fed into SVM for automatic classification. Ahmed et al. [16] developed an IoT-aided smart surveillance model for object recognition through segmentation. Especially, this research encompasses the concepts of IoT, cooperative drones, and DL to improve the surveillance initiatives in smart cities. During segmentation, an AI-based system was introduced in this study by employing the DL-based Pyramid Scene Parsing Network (PSPNet).

Alotaibi and Nagappan [17] presented an automatic vehicle detection and classification method based on chaotic equilibrium optimizer algorithm with a DL (VDTC-CEOADL) model. This technique exploited the YOLO-HR object sensor in which ResNet was used as a support method. Further, the CEOA-based hyperparameter optimizer was introduced for hyperparameter tuning of the ResNet architecture. This technique used the attention-based LSTM (ALSTM) model for the purpose of classification. Gadamsetty et al. [18] presented a technique in which a supervised image detection method was exploited for the categorization of the images, tracked by object recognition, through the YOLOv3 model so as to extract the DCNN features. In general, semantic segmentation and image segmentation can be performed to detect the object's category of all the pixels using class labels. Next, the idea of hashing with SHA-256 was used together with the shipping amount as well as the position of the bounding box for satellite images.

Xie et al. [19] presented a Dense Sequential Fusion (DSF) structure, specifically intended to fuse the camera and LiDAR device data. The main intention of the study was to improve the robustness and accuracy of 3-D object recognition, mainly for distant objects. Alshahrani et al. [20] presented the Artificial Ecosystem Optimizer with DCNN for Vehicle Detection (AEODCNN-VD) technique for the RSIs. The projected AEODCNN-VD method concentrated on the classification of transports in a rapid and precise manner. Ahmed et al. [21] projected the Chicken Swarm Optimizer with Transfer Learning-Driven Vehicle Detection and Classification on the RSI (CSOTL-VDCRS) method to resolve these problems. This method made use of mask-region-based CNN (Mask RCNN) system for vehicle recognition. After observing the transportation in the RSIs, it was then categorized by following the Fuzzy Wavelet Neural Network (FWNN) method. Aljebreen et al. [22] developed the new honey badger optimizer algorithm with an ensemble learning-based vehicle detection and classification (HBOAEL-VDC) system. The main purpose of the research was to project the ensemble DL techniques for precise vehicle classification process.

The research gap in vehicle recognition and identification on the RSIs by employing the DL lies in the restricted survey of hyperparameter tuning plans within the current methods. Though the DL technique has established its importance in terms of precise object detection in the RSI, the optimum outline of hyperparameters that are vital for model performance remains under-addressed. The different and difficult nature of the RSIs need a nuanced method of hyperparameter tuning to safeguard the flexibility and robust generalization. So, it is important to investigate and develop effective hyperparameter tuning methods that achieve precise vehicle recognition from the RSIs. This is crucial in terms of improving the classification accuracy, addressing the tasks like changing environmental states and image resolutions, and finally increasing the reliability and applicability of DL techniques in the area of RSI for vehicle recognition and identification.

3. Proposed model

In the current research work, the authors present an innovative IWDADL-VDC methodology to be applied on the RSIs. The IWDADL-VDC model exploits the hyperparameter-tuned DL model for detection and classification of the vehicles. To accomplish this, the IWDADL-VDC model follows two major stages such as vehicle detection and classification. Figure 1 depicts the workflow of the IWDADL-VDC technique.

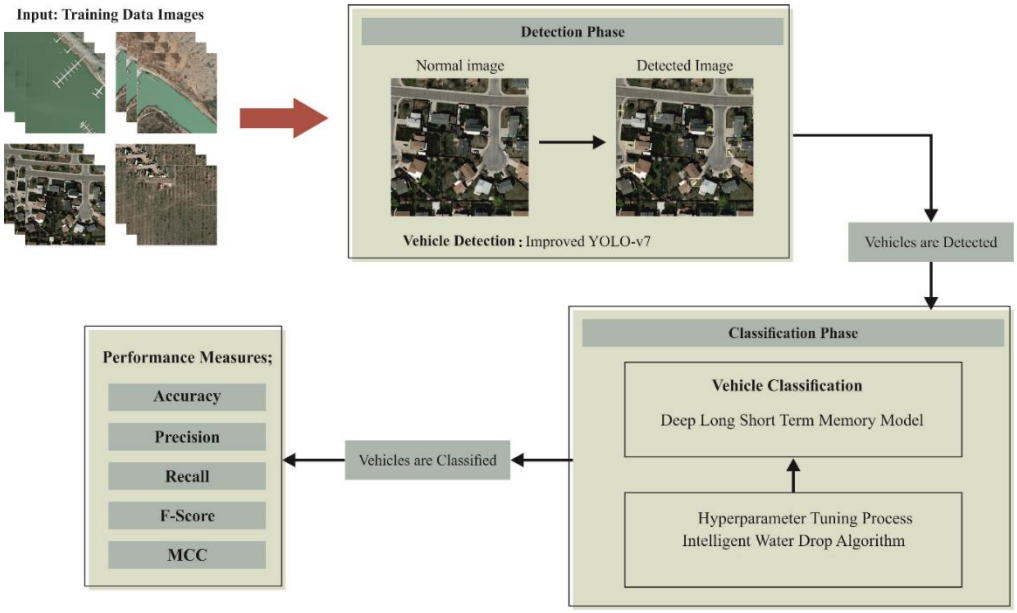

Figure 1. The Workflow of the IWDADL-VDC algorithm.

3.1. Vehicle detection using YOLOv7 object detection network

For the vehicle detection process, the IWDADL-VDC technique uses the improved YOLO-v7 model. Being the baseline method in YOLO series, the YOLO_v7 model accepts plans as protracted effective long-range attention network (E-ELAN), a scaling method that depends on the convolution reparameterization and concatenation-related methods. This model accomplishes a better balance between detection accuracy and efficiency [23]. The detection idea followed by the YOLO_v7 model

remains the same alike YOLO_v4 and YOLO_v5 models in YOLO series. There exist four modules in YOLO_v7 network namely head, backbone, input, and prediction. The head module comprises of Path Aggregation Feature Pyramid Network (PAFPN) model. The input module is used to scale the input images into even pixel dimensions so as to meet the network requirements. The module comprises of MPConv convolution layer, Bconv convolution layer E, and -ELAN convolution layer while BConv has LeakyReLU activation function, convolution, and Batch Normalization (BN) layers, which are used in the extraction of image data of varying scales. With the emergence of bottom-up path, it is easy to transmit the fundamental data to a higher level, thereby realizing an effective incorporation of the features at dissimilar levels. A forecast element is used to adjust the image networks for P3, P4, and P5 features of dissimilar scales' output through PAFPN via repeg block (REP) and the outcome is passed over 1×1 convolutions, which predicts the anchor frame, confidence, and category. In this scenario, the field vehicle detection module must meet the requirements in terms of accuracy and real-time detection. YOLOv7 model has been chosen as the basic model in this study since it strikes a better balance between speed and accuracy of detection.

Initially, many attention mechanism modules of SimAM are embedded into the YOLOv7 network architecture. In general, the Attention module provides dissimilar weights to the input part of the network. The technique disregards the unrelated data and focuses only on the relevant data that could efficiently enhance the feature extraction capability in a complicated background. SimAM is an attention model that does not enhance the parameter count of the network. It is embedded in all the positions of the model and has plug-and-play features. The fundamental objective of SimAM lies in the computation of attention weight using its energy function. SimAM reduces the intervention of complicated background on vehicle recognition by producing a spatial inhibition on the neighboring neurons of the vehicle. This phenomenon highlights the basic features of the vehicle and also improves the capability of extracting those features as given below.

$$\hat{X} = sigmoid\left(\frac{1}{E}\right) \otimes X, \tag{1}$$

$$E = \frac{4(\sigma^2+\lambda)}{(t-\mu)^2+2\sigma^2+2\lambda}, \tag{2}$$

$$\mu = \frac{1}{Q}\sum_{i=1}^{Q} x_i, \tag{3}$$

$$\sigma^2 = \frac{1}{Q}\sum_{i=1}^{Q}(x_i - \mu)^2. \tag{4}$$

The sigmoid function limits the values of E so as to avoid its value from getting larger. Here, the improved mapping feature of the vehicle is $\hat{X}$, E denotes the energy function on the channel. There is a lesser and greater distinction between the adjacent and vehicle neurons. X indicates the input vehicle feature map. $\otimes$ shows dot product operation, σ^2 refers to the variance of all the networks in the input vehicle feature map, t denotes the target vehicle neuron, μ represents the mean value of the network in the input vehicle and λ implies a super-parameter.

Next, downsampling is the major role of MPConv in which the feature sizes are to be reduced using specific feature loss method. It is to be noted that the dual branches of MPConv model in YOLOv7 exploit the convolution of 3×3 kernels for ConV process. During this procedure, some of the features might get lost while an ineffective feature learning might take place in the system, once

the step size is 2. The convolution of 3×3 kernels in the branch, below the MPConv layer, is substituted by the focus module based on the focus module in YOLOv5. With the help of a halved feature map, the learning efficacy of the features and accuracy of vehicle detection under complicated backgrounds are enhanced while the loss of features gets reduced.

3.2. Classification using a DLSTM model

In this stage, the classification is performed with the help of the DLSTM model. LSTM is a specific Recurrent Neural Network (RNN) that can solve the issue of long-term reliance upon time sequence, gradient withdrawal and an explosion of RNN in the extended series training procedures [24]. The storage part of the LSTM strengthens the links between long- and short-term-time sequences. This method is capable of upgrading, preserving and removing the data in the storage unit via three gates namely forget, input, and output. The gate plan holds the data for a long time while it also manages the information movement. When compared with usual RNNs, the LSTM technique accomplishes superior performance in lengthy orders. In this study, the LSTM technique was executed by employing MATLAB $2022b$ version deep learn- $Q \approx \sigma$ ing toolbox. The detailed steps of the LSTM technique are mentioned below:

Step 1. The forget gate reads the data of h_{t-1} and x_t, and selects either to preserve the data of earlier time via sigmoid function (σ):

$$f_t = \sigma\big(a_f \cdot [h_{t-1}, x_t] + b_f\big). \tag{5}$$

Here, f_t refers to forget gate function, a_f signifies the weight, x_t denotes the input at time t, h_{t-1} is the earlier intended output and b_f denotes the bias of the forget gate.

Step 2. The input gate adopts the data that is kept in a cell state. Primarily, the upgrade value i_t is defined by sigmoid function whereas a novel candidate value $\tilde{C}_t$ is produced by $\tanh$ function:

$$i_t = \sigma(a_i \cdot [h_{t-1}, x_t] + b_i), \tag{6}$$

$$\tilde{C}_t = \tanh(a_c \cdot [h_{t-1}, x_t] + b_c). \tag{7}$$

Here, a_i and a_c denote the weights while b_i and b_c refer to biases.

Step 3. Upgrade the old cell state via forget and input gates so as to produce an upgraded value:

$$C_t = f_t x C_{t-1} + i_t x \tilde{C}_t. \tag{8}$$

Step 4. The output gate consequence depends on cell state. Primarily, the output value 0_t of cell state is defined using the sigmoid function. An upgrade value C_t refers to normalization by $\tan h$ function and then multiplied with o_t, and to attain result value (h_t) at time t:

$$o_t = \sigma(a_o \cdot [h_{t-1}, x_t] + b_o), \tag{9}$$

$$h_t = o_t \times \tanh(C_t). \tag{10}$$

Here, a_o and b_o correspond to weight and bias respectively. Figure 2 demonstrates the architecture of the DLSTM model.

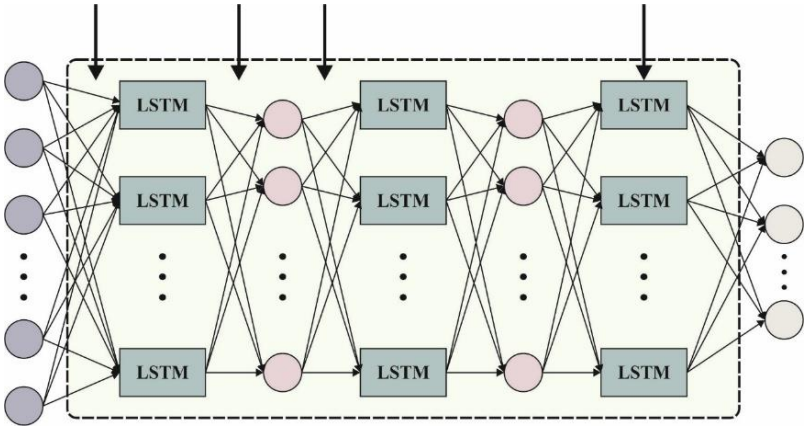

Figure 2. Architecture of the DLSTM model.

3.3. IWDA-based hyperparameter-tuning process

In order to enhance the detection performance of the DLSTM technique, the IWDA-based hyperparameter tuning process has been employed in this study. The range of the IWDA for hyperparameter tuning is accepted based on its exclusive optimization features and suitability for the exact desires of the hyperparameter optimization task. The IWDA is stimulated by the foraging behavior of water drops in nature and is mainly suitable for resolving difficult and non-linear optimizer issues. It displays a fine balance between the exploration and exploitation phases, which is vital for identifying the optimum hyperparameter configurations. It travels the hyperparameter space while it also exploits some of the promising areas that help in stopping the system from getting fixed in local goals.

IWDA is a group intelligence algorithm based on the principles of water drops that interrelate with residue to create a water flow track, once they move [25]. The erosion of the riverbed by water movement results in ravines on the riverbed. Water flow is a collection of unit droplets while every individual drop consists of sediment and velocity attributes. Under gravity, once the water droplet chooses the direction with comparatively small resistance, viz., the direction with low sediment, the water droplet takes away more deposition and attains a great speed. Once, two water drops (a) and (b) correspondingly, with similar assets pass over the region, the water droplet carries more sediment and attains a high-velocity increment.

Water drops tend to move thus conferring to discretization in the abstract model and conveying double significant possessions such as the sediment-carrying property soil (IWD) and motion attribute, $vel(IWD)$. Both the possessions change with the flow of water drops. Assume the present state of water drop during movement to the following location so that the water droplet undergoes subsequent changes.

Initially, the water droplet tends to select the path with a lesser number of sediments. Here, probability $p(i,j)$ denotes that the water droplet chooses j^{th} particle at i^{th} location as the following location, which is inversely relative to the volume of sediment in (i,j) path, and the selection probability is defined as given below.

$$p(i,j) = \frac{f(soil(i,j))}{\sum f(soil(i,k))}, \tag{11}$$

$$f\big(soil(i,j)\big) = \frac{1}{\varepsilon + g\big(soil(i,j)\big)}. \tag{12}$$

Here, the small positive value is $\mathcal{E}$ and

$$g\big(soil(i,j)\big) = \begin{cases} soil(i,j), & \min\big(soil(i,k)\big) \geq 0 \\ soil(i,j) - \min\big(soil(i,k)\big), & else \end{cases}. \tag{13}$$

Once the water droplets travel from i^{th} to j^{th} position, its velocity changes and this speed increment $\Delta vel(IWD)$ is inversely proportional to the sediment content soil (i,j) on the running path.

$$\Delta vel(IWD) = \frac{a_v}{b_v + c_v(soil(i,j))^2}. \tag{14}$$

In Eq (14), $a_{1Z\backslash}\ b_{1Z}$ and c_v denote the predetermined parameters.

The quantity of sediment transported by water drop flow is equivalent to the quantity of sediment reduction $\Delta soil(i,j)$ in the path (i,j), as mentioned below:

$$\Delta soil(IWD) = \Delta soil(i,j). \tag{15}$$

After the water droplet operation, the sediment reduction in the path is inversely proportional to the time required for the water droplet to be passed over path$(s\ i,j)$.

$$\Delta soil(i,j) = \frac{a_s}{b_s + c_s(time(i,j))^2}. \tag{16}$$

Here, a_s, b_s and c_s correspond to predetermined parameters and time taken for a water drop to transfer from i^{th} position to j^{th} position as given below:

$$time(i,j) = \frac{HUD(i,j)}{vel(IWD)}. \tag{17}$$

In Eq (17), the heuristic concerning the road segment (i,j) is denoted by $HUD(i,j)$.

Once the water droplet arrives at j^{th} position from i^{th} location, the sediment in (i,j) is updated to provide the feedback on course planning of other water droplets.

$$soil(i,j) = (1-p) \bullet soil(i,j) - p \bullet \Delta soil(i,j). \tag{18}$$

Here, the coefficient lies in the range of [0 and 1].

Every drop finishes its path planning from the beginning to the endpoint through the steps (1) to (4); consider that the course of k^{th} droplet is T_k, and it uses evaluation function $q(T)$ to choose the optimum way in the droplet group T^{IB}, as,

$$T^{IB} = \max q(T_k). \tag{19}$$

To make the optimum path, the planning process have controlling effects on succeeding course planning. This further enhances the capability of water droplets to find the optimum path and is essential to form a feedback model for updating the global sediment volume on optimum path.

$$soil(i,j) = (1-\rho)soil(i,j) + \rho\frac{2soil(IWD)}{N_{lB}(N_{lB}-1)}. \tag{20}$$

In Eq (20), the updated parameter within [0,1] is ρ and the node count for the path is N_{IR}.

Fitness selection is an extensive factor that influences the performance in IWDA methodology. The hyperparameter selection method contains a solution encoding model to estimate the efficiency of candidate solutions. In this study, the IWDA model reflects 'correctness' as the chief measure to project the fitness function formulated below.

$$Fitness = \max(P), \tag{21}$$

$$P = \frac{TP}{TP+FP}. \tag{22}$$

In the expressions given above, TP denotes True Positive value and FP represents false positive values.

4. Experimental validation

The current section details about the performance validation of the IWDADL-VDC technique using two datasets such as the VEDAI dataset [26] and ISPRS Postdam dataset [27]. The VEDAI dataset includes 3,687 samples under nine classes. Similarly, the ISPRS Postdam dataset includes 2244 samples under four classes. Figure 3 represents some of the sample detected images.

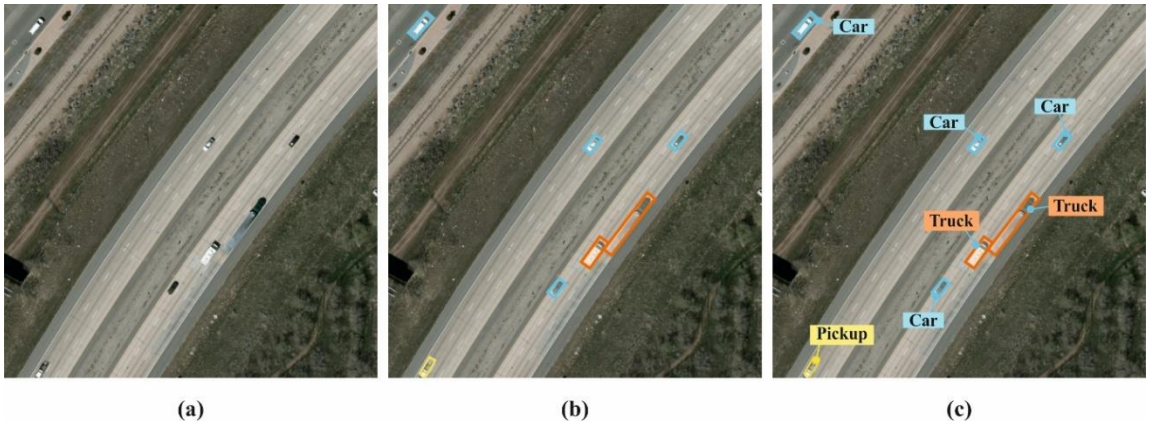

Figure 3. Sample detected images.

Figure 4 shows the classifier analysis outcomes of the IWDADL-VDC system upon the VEDAI dataset. Figures 4a and b reveal the confusion matrices generated by the IWDADL-VDC model with 70:30 of training phase (TRPH)/testing phase (TSPH). The figure infers that the IWDADL-VDC methodology can exactly categorize and recognize all the nine classes. Furthermore, Figure 4c displays the PR analysis outcomes of the IWDADL-VDC method. The figure exhibits that the IWDADL-VDC methodology achieved an excellent PR performance under each class. Also, Figure 4d demonstrates the ROC analysis outcomes achieved by the IWDADL-VDC method. This figure indicates that the IWDADL-VDC model prompts successful outcomes with higher ROC values on diverse classes.

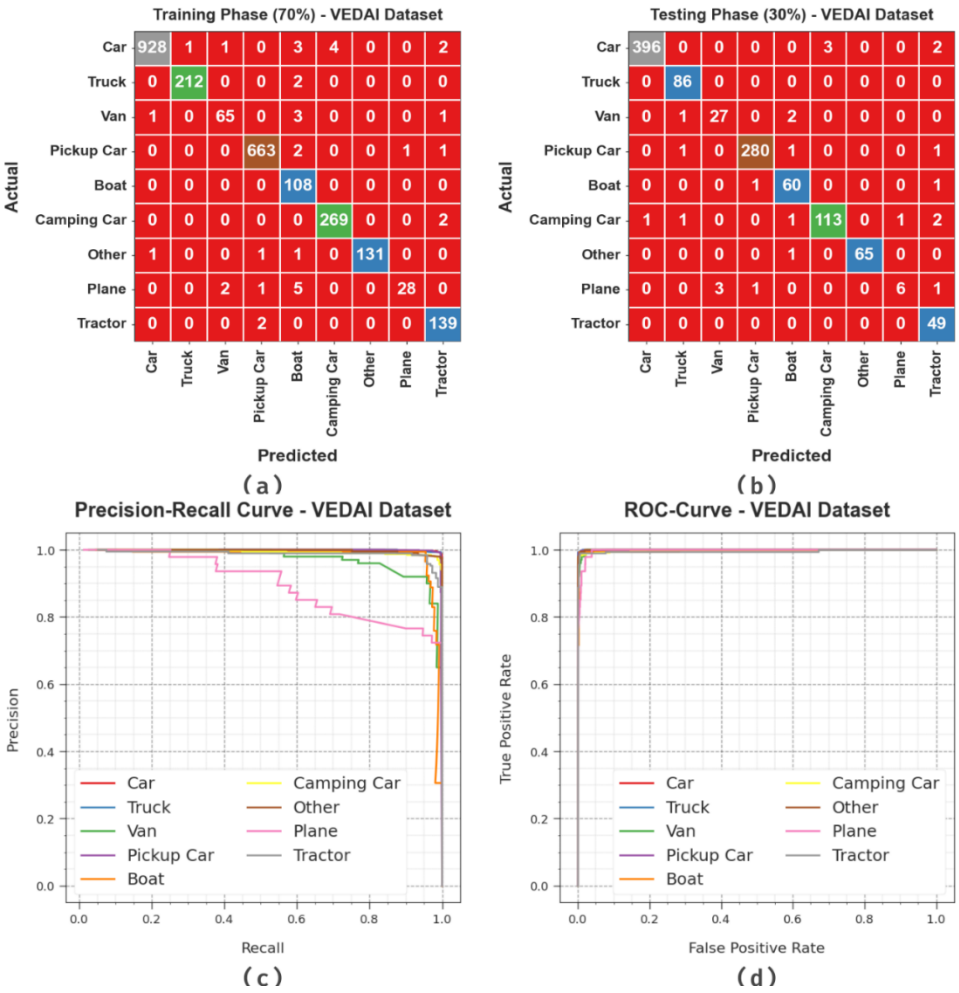

Figure 4. VEDAI dataset (a and b) Confusion matrices (c) PR-curve and (d) ROC-curve.

The vehicle classification results accomplished by the IWDADL-VDC system under the VEDAI dataset are shown in Table 1 and Figure 5. The achieved outcomes display that the IWDADL-VDC system achieved an increase in its performance at each class. Upon 70% of TRPH, the IWDADL-VDC technique achieved an average $accu_y$ of 99.68%, $prec_n$ of 96.93%, $reca_l$ of 95.95%, F_{score} of 96.27%, and an MCC of 96.17%. Also, based on 30% of TSPH, the IWDADL-VDC technique provided an average $accu_y$ of 99.50%, $prec_n$ of 94.29%, $reca_l$ of 92.50%, F_{score} of 92.95%, and an MCC of 92.89%, respectively.

Table 1. Vehicle classification analysis outcomes of the IWDADL-VDC system under the VEDAI dataset.

Labels	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC
TRPH (70%)					
Car	99.50	99.78	98.83	99.30	98.91
Truck	99.88	99.53	99.07	99.30	99.23
Van	99.69	95.59	92.86	94.20	94.05
Pickup Car	99.69	99.40	99.40	99.40	99.19
Boat	99.38	87.10	100.00	93.10	93.02
Camping Car	99.77	98.53	99.26	98.90	98.77
Other	99.88	100.00	97.76	98.87	98.81
Plane	99.65	96.55	77.78	86.15	86.50
Tractor	99.69	95.86	98.58	97.20	97.05
Average	99.68	96.93	95.95	96.27	96.17
TSPH (30%)					
Car	99.46	99.75	98.75	99.25	98.83
Truck	99.73	96.63	100.00	98.29	98.16
Van	99.46	90.00	90.00	90.00	89.72
Pickup Car	99.55	99.29	98.94	99.12	98.81
Boat	99.37	92.31	96.77	94.49	94.18
Camping Car	99.19	97.41	94.96	96.17	95.73
Other	99.91	100.00	98.48	99.24	99.19
Plane	99.46	85.71	54.55	66.67	68.14
Tractor	99.37	87.50	100.00	93.33	93.23
Average	99.50	94.29	92.50	92.95	92.89

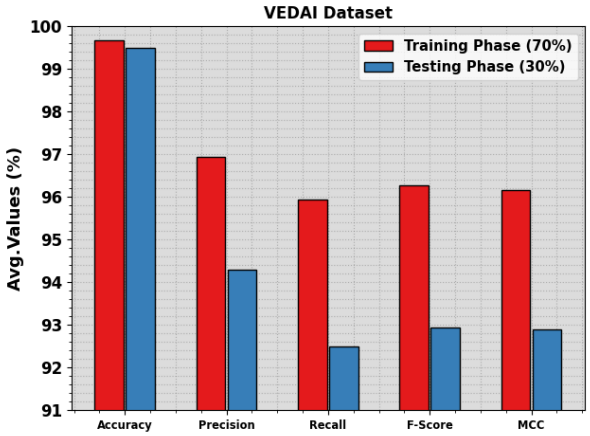

Figure 5. Vehicle classification analysis outcomes of the IWDADL-VDC model under the VEDAI dataset.

Figure 6 shows the training and validation $accu_y$ analysis curves achieved by the IWDADL-VDC system upon the VEDAI dataset. The figure offers knowledgeable insights about the effectiveness of the IWDADL-VDC algorithm over the increasing number of epochs. Two curves show essential insights about the capabilities and learning evolution of the model for generalization.

Additionally, it is made clear that the proposed model consistently enhanced its performance in terms of training and testing $accu_y$ over increasing number of epochs. This outcome displays the capacity of the model to identify and learn patterns within training and testing databases. The augmentation $accu_y$ analysis suggests that the model not only varies to the training data but also exceeds to make $accu_y$ predictive on prior unnoticed data, thus emphasizing its powerful generalization abilities.

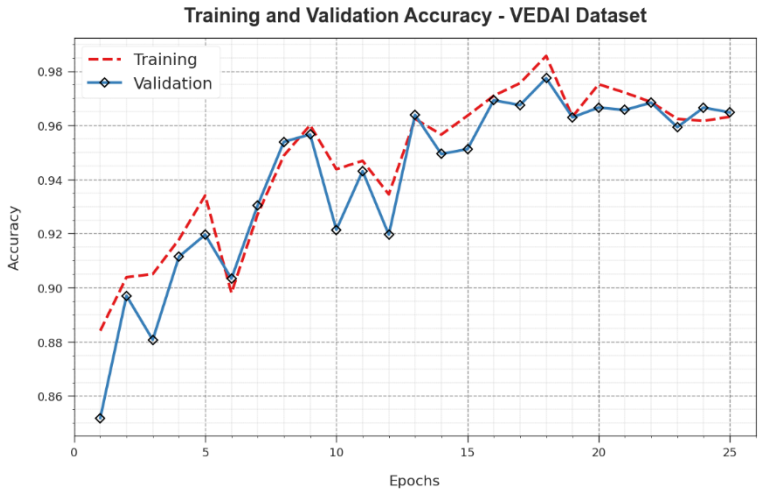

Figure 6. $Accu_y$ curve of the IWDADL-VDC model under the VEDAI dataset.

Figure 7 shows an extensive view of the IWDADL-VDC approach with the VEDAI dataset in terms of training and testing loss values for the IWDADL-VDC system at multiple number of epochs. The training loss progressively diminishes as the model increases its weights to lessen the classification errors at both testing and training databases. These loss curves perfectly represent the appropriate alignment of the model with that of the training data. Further, it also highlights the model's capability to adept at holding patterns on these databases. This is a valuable inference to note that the IWDADL-VDC method frequently improves its parameters to reduce the discrepancies between the actual and the predicted training labels.

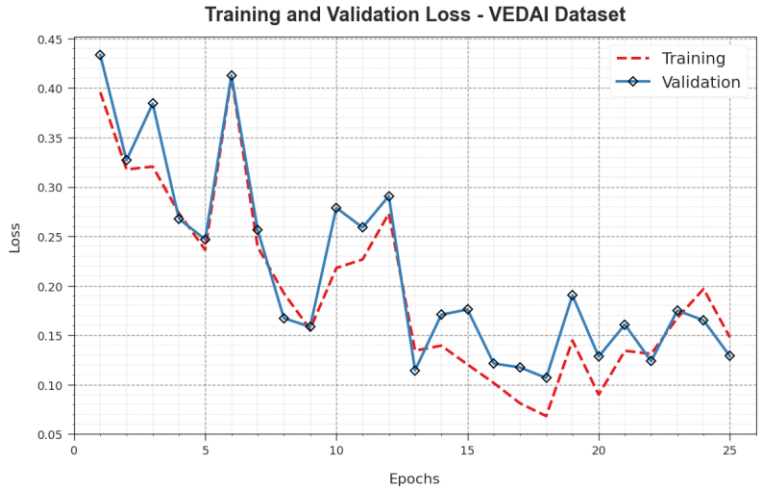

Figure 7. Loss curve of the IWDADL-VDC model under VEDAI dataset.

Figure 8 shows the outcomes of a brief comparison study conducted upon the IWDADL-VDC technique using the VEDAI dataset [11]. The attained results highlight that the IWDADL-VDC technique achieved superior performance with a maximum $accu_y$ of 99.68%. Conversely, the existing models such as the ICOA-DLVDC, CSOTL-VDCRS, LeNet, AlexNet, and the VGG-16 algorithms achieved the least $accu_y$ values such as 99.50%, 98.07%, 79.74%, 88.98%, and 94.46%, respectively.

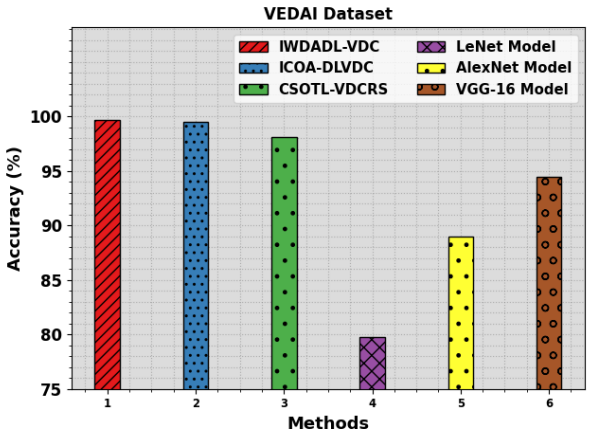

Figure 8. $Accu_y$ analysis outcomes of the IWDADL-VDC model under the VEDAI dataset.

Figure 9 displays the classifier analysis outcomes of the IWDADL-VDC system when using the ISPRS Postdam dataset. Figures 9a and b show the confusion matrices generated by the IWDADL-VDC technique with 70:30 of TRPH/TSPH. The figure infers that the IWDADL-VDC methodology can accurately categorize and recognize all the four classes. Moreover, Figure 9c represents the PR analysis outcomes of the IWDADL-VDC algorithm. The figure reveals that the IWDADL-VDC methodology achieved a remarkable PR performance. Further, Figure 9d shows the ROC analysis results of the IWDADL-VDC model. The figure indicates that the IWDADL-VDC model accomplished efficacious outcomes with higher ROC values under different classes.

The outcomes from the vehicle classification analysis conducted upon the IWDADL-VDC system using the ISPRS Postdam dataset are shown in Table 2 and Figure 10. The accomplished values infer that the IWDADL-VDC system gets boosted achieved supreme performance under all the classes. Upon 70% of TRPH, the IWDADL-VDC system achieved an average $accu_y$ of 99.87%, $prec_n$ of 97.27%, $reca_l$ of 99.75%, F_{score} of 98.47%, and an MCC of 98.18%. Besides, with 30% of TSPH, the IWDADL-VDC technique provided an average $accu_y$ of 99.70%, $prec_n$ of 97.95%, $reca_l$ of 96.50%, F_{score} of 97.11%, and an MCC of 96.41%, correspondingly.

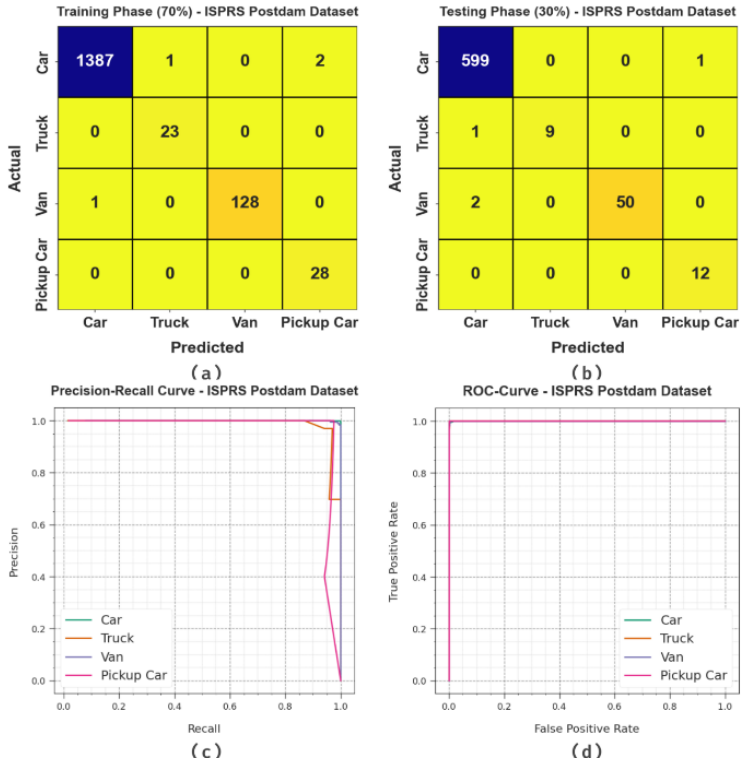

Figure 9. ISPRS Postdam Dataset (a-b) Confusion matrices (c) PR-curve and (d) ROC-curve.

Table 2. Vehicle classification analysis outcomes of the IWDADL-VDC model under the ISPRS Postdam dataset.

Labels	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC
TRPH (70%)					
Car	99.75	99.93	99.78	99.86	98.75
Truck	99.94	95.83	100.00	97.87	97.86
Van	99.94	100.00	99.22	99.61	99.58
Pickup Car	99.87	93.33	100.00	96.55	96.55
Average	99.87	97.27	99.75	98.47	98.18
TSPH (30%)					
Car	99.41	99.50	99.83	99.67	96.94
Truck	99.85	100.00	90.00	94.74	94.80
Van	99.70	100.00	96.15	98.04	97.90
Pickup Car	99.85	92.31	100.00	96.00	96.00
Average	99.70	97.95	96.50	97.11	96.41

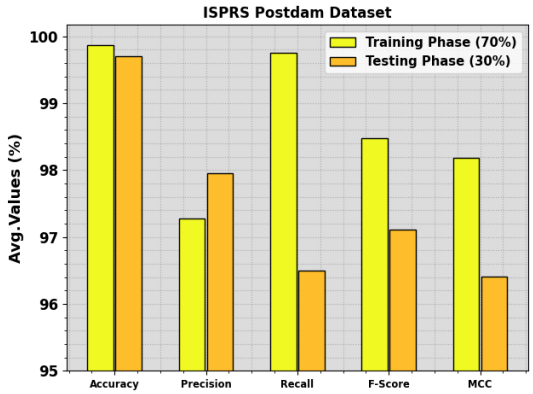

Figure 10. Average analysis outcomes of the IWDADL-VDC algorithm under the ISPRS Postdam dataset**.**

Figure 11 displays the training and validation $accu_y$ curves of the IWDADL-VDC technique when using the ISPRS Postdam dataset. The figure provides highly commendable insights about the effectiveness of the IWDADL-VDC system over multiple epochs. The two curves correspond to some crucial insights into the capabilities as well as the learning progression of the model for generalization. Moreover, it is apparent that the model is a consistent performer with enhanced training and testing $accu_y$ values over increasing number of epochs. This further represents the capacity of the model to learn and identify the patterns within training and testing databases. The enhancement testing $accu_y$ suggests that the model not only modifies the training data but also surpasses to make $accu_y$ predictions on the previously disregarded data, thus emphasizing its powerful generalization abilities.

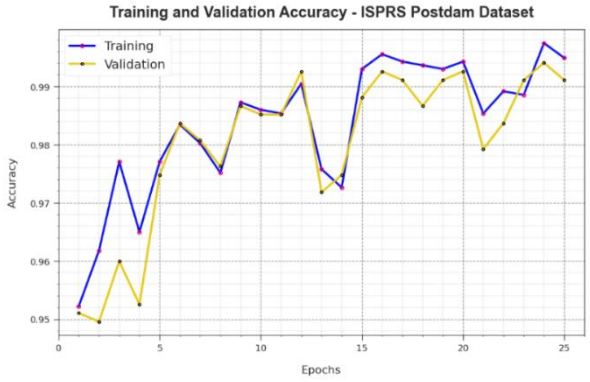

Figure 11. $Accu_y$ curve of the IWDADL-VDC system under the ISPRS Postdam dataset.

Figure 12 demonstrates an extensive view of the results achieved by the IWDADL-VDC methodology under the ISPRS Postdam dataset, training and testing loss values at different number of epochs. There was a gradual decline observed in the training loss as the model enhanced its weights to lessen the classification errors for testing and training datasets. These loss curves offer a better representation of the model's alignment with that of the training data and also emphasize its capacity for competently holding patterns on these databases. This assessment outcomes infer that the IWDADL-VDC algorithm frequently improves its parameters to reduce the discrepancies between the prediction and training labels.

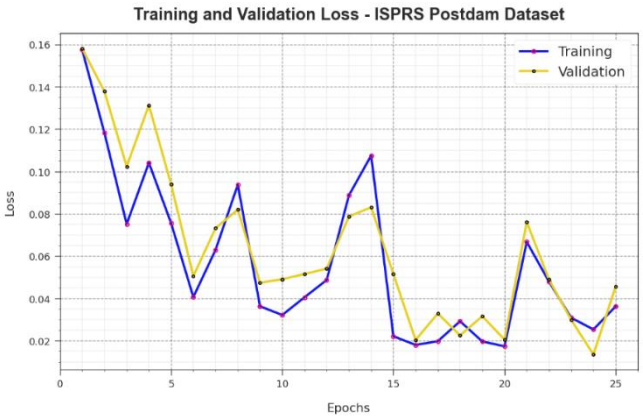

Figure 12. Loss curve of the IWDADL-VDC approach under the ISPRS Postdam dataset.

Figure 13 shows the extensive comparison analysis outcomes achieved by the IWDADL-VDC system upon the ISPRS Postdam dataset. The attained outcomes display that the IWDADL-VDC technique achieved an increased performance with a maximum $accu_y$ of 99.87%. Alternatively, the ICOA-DLVDC, CSOTL-VDCRS, LeNet, AlexNet, and the VGG-16 methodologies achieved the least $accu_y$ values such as 99.70%, 98.67%, 94.54%, 95.86%, and 89.54%, correspondingly. Thus, it can be inferred that the IWDADL-VDC technique can be utilized for an accurate and automated vehicle detection process.

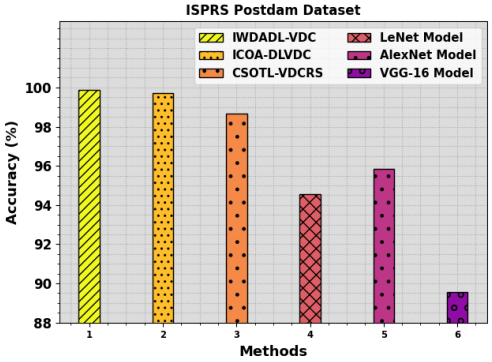

Figure 13. $Accu_y$ analysis outcomes of the IWDADL-VDC algorithm under the ISPRS Postdam dataset.

5. Conclusions

In the current research work, the authors have developed a new vehicle detection and classification model named IWDADL-VDC technique to be applied on the RSIs. The IWDADL-VDC methodology exploits the improved YOLO-v7 object detection with DLSTM classifier and IWDA-based hyperparameter tuning process. The IWDADL-VDC model was experimentally validated using two benchmark datasets and the results attained by the IWDADL-VDC technique were promising compared to other recent approaches. The proposed IWDADL-VDC technique achieved the maximum accuracy values such as 99.68% and 99.87% under VEDAI and the ISPRS Postdam datasets,

respectively. The future works in vehicle recognition and identification can concentrate on the incorporation of innovative DL designs, like transformer-based methods, to improve the extraction of features and classification accuracy. In addition to this, when the application of multi-modal sensor fusion by uniting the data from cameras, LiDAR, and radar is explored further, then it may additionally enhance the robustness of vehicle recognition methods, especially in challenging environmental states. Further research could also examine real-time execution for dynamic traffic states and explore the possibility of edge computing to reduce the latency in decision-making methods.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare that they have no conflicts of interest. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

References

1. F. Safarov, K. Temurbek, D. Jamoljon, O. Temur, J. C. Chedjou, A. B. Abdusalomov, et al, Improved agricultural field segmentation in satellite imagery using TL-ResUNet architecture, *Sensors*, **22** (2022), 9784. https://doi.org/10.3390/s22249784
2. M. A. Momin, M. H. Junos, A. S. M. Khairuddin, M. S. A. Talip, Lightweight CNN model: Automated vehicle detection in aerial images. *Signal Image Video P.*, **17** (2023), 1209–1217. https://doi.org/10.1007/s11760-022-02328-7
3. Y. Wang, F. Peng, M. Lu, M. A. Ikbal, Information extraction of the vehicle from high-resolution remote sensing image based on convolution neural network, *Recent Adv. Electr. El.*, **16** (2023), 168–177. https://doi.org/10.2174/2352096515666220820174654
4. L. Wang, Y. Shoulin, H. Alyami, A. A. Laghari, M. Rashid, J. Almotiri, et al., A novel deep learning—based single shot multibox detector model for object detection in optical remote sensing images, *Geosci. Data J.*, 2022, 1–15. https://doi.org/10.1002/gdj3.162
5. R. Ghali, M. A. Akhloufi, Deep learning approaches for wildland fires remote sensing: Classification, detection, and segmentation, *Remote Sens.*, **15** (2023), 1821. https://doi.org/10.3390/rs15071821
6. C. Anusha, C. Rupa, G. Samhitha, Region-based detection of ships from remote sensing satellite imagery using deep learning. In: *2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM)*, 2022, https://doi.org/10.1109/ICIPTM54933.2022.9754168
7. Y. Chen, R. Qin, G. Zhang, H. Albanwan, Spatial-temporal analysis of traffic patterns during the COVID-19 epidemic by vehicle detection using planet remote-sensing satellite images, *Remote Sens.*, **13** (2021), 208. https://doi.org/10.3390/rs13020208

8. L. K, S. Karnick, M. R. Ghalib, A. Shankar, S. Khapre, I. A. Tayubi, A novel method for vehicle detection in high-resolution aerial remote sensing images using YOLT approach, *Multimed. Tools Appl.*, **81** (2022), 23551–23566. https://doi.org/10.1007/s11042-022-12613-9
9. B. Wang, B. Xu, A feature fusion deep-projection convolution neural network for vehicle detection in aerial images, *PLoS One*, **16** (2021), e0250782. https://doi.org/10.1371/journal.pone.0250782
10. J. Wang, X. Teng, Z. Li, Q. Yu, Y. Bian, J. Wei, VSAI: A multi-view dataset for vehicle detection in complex scenarios using aerial images, *Drones*, **6** (2022), 161. https://doi.org/10.3390/drones6070161
11. M. Alajmi, H. Alamro, F. Al-Mutiri, M. Aljebreen, K. M. Othman, A. Sayed, Exploiting remote sensing imagery for vehicle detection and classification using an artificial intelligence technique, *Remote Sens.*, **15** (2023), 4600. https://doi.org/10.3390/rs15184600
12. S. Javadi, M. Dahl, M. I. Pettersson, Vehicle detection in aerial images based on 3D depth maps and deep neural networks, *IEEE Access*, **9** (2021), 8381–8391. https://doi.org/10.1109/ACCESS.2021.3049741
13. P. Gao, T. Tian, T. Zhao, L. Li, N. Zhang, J. Tian, Double FCOS: A two-stage model utilizing FCOS for vehicle detection in various remote sensing scenes, *IEEE J. STARS*, **15** (2022), 4730–4743. https://doi.org/10.1109/JSTARS.2022.3181594
14. M. Ragab, H. A. Abdushkour, A. O. Khadidos, A. M. Alshareef, K. H. Alyoubi, A. O. Khadidos, Improved deep learning-based vehicle detection for urban applications using remote sensing imagery, *Remote Sens.*, **15** (2023), 4747. https://doi.org/10.3390/rs15194747
15. C. H. Karadal, M. C. Kaya, T. Tuncer, S. Dogan, U. R. Acharya, Automated classification of remote sensing images using multileveled MobileNetV2 and DWT technique, *Expert Syst. Appl.*, **185** (2021), 115659. https://doi.org/10.1016/j.eswa.2021.115659
16. I. Ahmed, M. Ahmad, A. Chehri, M. M. Hassan, G. Jeon, IoT enabled deep learning based framework for multiple object detection in remote sensing images, *Remote Sens.*, **14** (2022), 4107. https://doi.org/10.3390/rs14164107
17. Y. Alotaibi, K. Nagappan, G. Rani, S. Rajendran, Vehicle detection and classification using optimal deep learning on high-resolution remote sensing imagery for urban traffic monitoring, 2023. Preprint. https://doi.org/10.21203/rs.3.rs-3272891/v1
18. S. Gadamsetty, R. Ch, A. Ch, C. Iwendi, T. R. Gadekallu, Hash-based deep learning approach for remote sensing satellite imagery detection, *Water*, **14** (2022), 707. https://doi.org/10.3390/w14050707
19. C. Xie, C. Lin, X. Zheng, B. Gong, H. Liu, Dense sequential fusion: Point cloud enhancement using foreground mask guidance for multimodal 3D object detection, *IEEE T. Instrum. Meas.*, **73** (2024), 9501015, https://doi.org/10.1109/TIM.2023.3332935
20. S. M. Alshahrani, S. S. Alotaibi, S. Al-Otaibi, M. Mousa, A. M. Hilal, A. A. Abdelmageed, et al., Optimal deep convolutional neural network for vehicle detection in remote sensing images, *CMC Comput. Mater. Con*, **74** (2023), 3117–3131. https://doi.org/10.32604/cmc.2023.033038
21. M. A. Ahmed, S. A. Althubiti, V. H. C. de Albuquerque, M. C. dos Reis, C. Shashidhar, T. S. Murthy, et al., Fuzzy wavelet neural network driven vehicle detection on remote sensing imagery, *Comput. Electr. Eng.*, **109** (2023), 108765. https://doi.org/10.1016/j.compeleceng.2023.108765

22. M. Aljebreen, B. Alabduallah, H. Mahgoub, R. Allafi, M. A. Hamza, S. S. Ibrahim, et al., Integrating IoT and honey badger algorithm based ensemble learning for accurate vehicle detection and classification, *Ain Shams Eng. J.*, **14** (2023), 102547. https://doi.org/10.1016/j.asej.2023.102547
23. Y. Lai, R. Ma, Y. Chen, T. Wan, R. Jiao, H. He, A pineapple target detection method in a field environment based on improved YOLOv7, *Appl. Sci.*, **13** (2023), 2691. https://doi.org/10.3390/app13042691
24. Y. F. Shi, C. Yang, J. Wang, Y. Zheng, F. Y. Meng, L. F. Chernogor, A hybrid deep learning-based forecasting model for the peak height of ionospheric F2 layer, *Space Weather*, **21** (2023), e2023SW003581. https://doi.org/10.1029/2023SW003581
25. B. O. Alijla, C. P. Lim, L. P. Wong, A. T. Khader, M. A. Al-Betar, An ensemble of intelligent water drop algorithm for feature selection optimization problem, *Appl. Soft Comput.*, 65 (2018), 531–541. https://doi.org/10.1016/j.asoc.2018.02.003
26. S. Razakarivony, F. Jurie, Vehicle detection in aerial imagery: A small target detection benchmark, *J. Vis. Commun. Image R.*, **34** (2016), 187–203. https://doi.org/10.1016/j.jvcir.2015.11.002
27. F. Rottensteiner, G. Sohn, J. Jung, M. Gerke, C. Baillard, S. Benitez, U. Breitkopf, The ISPRS benchmark on urban object classification and 3D building reconstruction, *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, **1–3** (2012), 293–298.