sciendo

Journal of Smart Internet of Things (JSIoT)
VOL 2023, No.02 | 65-80 | 2023
DOI: 10.2478/jsiot-2022-0005

Article

Exploration of Vehicle Target Detection and Classification Method Based on Sea Lion Optimization with Deep Convolutional Neural Network

Zain Jaffer Ahmed[1], Mohammed Ali Mustfai[1]

[1] Department of Computer Engineering, College of Engineering, Amman Arab University, Amman, Jordan

*Corresponding author: (email: zain.jafferahmed@gmail.com)

Received 20 August 2023; Accepted 21 October 2023; Issue Published 15 December 2023

Abstract: Presently, urban environments over the globe are highly employed to obtain solutions for enhancing the quality of the livers and enhance the usage of city infrastructure and resources with minimal operational cost. Urban remote sensing acts as a significant part in the ability of mapping, monitoring, and controlling infrastructure. High-resolution remote sensing data renders worldwide images faster than traditional data collection strategies. Hence, small objects such as cars are easily detected. Vehicle recognition on aerial remote sensing images (RSIs) in the complicated background of urban zones has always gained a lot of interest in the remote sensing field. The automatic vehicles enumeration research domain had a significant contribution in several applications, including traffic management and monitoring. Target detection technology will be a crucial part of computer vision (CV) technology, and target detection techniques were enforced in several domains. Therefore, this study develops a new Vehicle Recognition and Classification using Sea Lion Optimization with Deep Learning (VRC-SLODL) model on RSI. In the presented VRC-SLODL technique, the major intention lies in recognising and classifying vehicles present in the images. The bilateral filtering (BF) technique can initially improve the RSI quality. The VRC-SLODL technique employs a modified residual network (ResNet) model to produce a collection of feature vectors. Eventually, the SLO approach with long short-term memory (LSTM) technique was exploited for vehicle classification, where the SLO algorithm acts as a hyperparameter optimizer. The experiments were performed on a benchmark dataset to examine the better performance of the VRC-SLODL method. The obtained values reported the improved classification performance of the VRC-SLODL technique over other models.

Keywords: Urban remote sensing; Remote sensing images; Sealion optimization; Deep learning; Vehicle classification; Target detection.

Copyright © 2023 Journal of Smart Internet of Things (JSIoT) published by Future Science for Digital Publishing and Sciendo . This is an open access article license CC BY (https://creativecommons.org/licenses/by/4.0/)

1. Introduction

With the progression of remote sensing technologies, remote sensing images (RSI) have seen exponential growth [1]. Compared with aerospace RSI, aerial RSI has the merits of accurate geometric correction, large imaging scale, and high resolution. Thus, aerial remote sensing is considered a significant remote sensing way [2, 3]; it generally leverages balloons or aeroplanes as working platforms, and flying altitude lies from meters to kilometres. In aerial RSI, vehicle recognition was a crucial technology in military and civil surveillance [4], like urban planning and traffic management. Still, the technique of manual analysis for vehicle detection has low data usage and poor data timeliness, which can be affected easily by subjective consciousness, physical conditions, and mentality [5, 6]. Thus, it is mostly required for executing automated vehicle recognition on RSI accurately and efficiently.

Remote sensing target recognition was to spot the object of interest in RSI after predicting the location and type of this target [7]. In the conventional recognition database, the target was focused. In contrast, the aviation data was not, and object strength in the aviation image generally appeared in arbitrary orientation that relies on the perception of the Earth vision platform [8, 9]. Object recognition refers to identifying samples of semantic objects of some classes (like birds, humans, or aeroplanes) in video and digital images. Minor target recognition was problematic in target recognition tasks [10]. The RSI study has vital applications in environmental management, military, transport planning, and disaster management. In addition, vehicles in RSI as a special category, whether civilian transportation or military, have a significant meaning and are simultaneously very difficult [11].

Though the obtainability of high-resolution satellite images hastens the object recognition task and automates such applications, vehicle identification from satellite images is difficult [12]. The reason behind this is that even in high spatial resolution images, vehicles were visible as minute spots unidentifiable from foreground areas to identify [13, 14]. Categorizing the identified vehicles was very serious since it sometimes could not differentiate small and large vehicles utilizing the eye itself from high-resolution images. Over the years, DL approaches were broadly employed in several research domains, and the evolving advancement of convolutional neural networks (CNN) made certain enhancements [15]. Kim et al. [16] examine integrating more predictive layers as to typical Yolov3 utilizing spatial pyramid pooling to complement the recognition accuracy of vehicles to huge scale change or make occlusion by another object. In [17], an innovative nearby ground vehicle recognition method in aerial infrared images dependent upon a CNN was presented. The UAV and infrared sensor utilized during this application can be initially established. Afterwards, a new aerial moving environment was created, and an aerial infrared vehicle database was unprecedentedly created. The authors released this database (NPU_CS_UAV_IR_DATA) utilized in the subsequent study under this domain. Afterwards, an end-to-end CNN was constructed. A real-time ground vehicle method was generated with huge counts of detection features being iteratively learned.

In [18], a new end-to-end DL network, such as orientation-aware feature fusion single-stage detection (OAFF-SSD), was designed to precise recognition of dense several creating vehicles utilizing images in UAV. The projected OAFF-SSD contains 3 important elements as a multi-level extracting feature, new feature fusion, and novel orientation-aware bounding box (OABB) offer and regression. In the meantime, detailed approaches can be selected for the fast convergence of trained loss. In [19], a productive and effective moving vehicle recognition system in aerial infrared image orders was presented using fast image registration and the YOLOv3 network. Because of insufficiency of infrared vehicle instances, transfer learning (TL) was established for training the

enhanced YOLOv3 network. Koga et al. [20] examine an unsupervised domain adaptation (DA) approach, which could not need to be labelled trained data and, therefore, could preserve recognition efficiency in targeted domain at a minimal cost. The authors executed Correlation alignment (CORAL) DA and adversarial DA to region-related vehicle detectors and enhanced the recognition precision by over 10% from the objective domain.

This study develops a new Vehicle Recognition and Classification using Sea Lion Optimization with Deep Learning (VRC-SLODL) model on RSI. In the presented VRC-SLODL technique, the RSI quality can be initially improved by the bilateral filtering (BF) technique. The VRC-SLODL technique employs a modified residual network (ResNet) model to generate a set of feature vectors. Finally, the SLO method with long short-term memory (LSTM) model was exploited for vehicle classification, where the SLO algorithm acts as a hyperparameter optimizer. The experiments were performed on a benchmark dataset to examine the better performance of the VRC-SLODL method.

2. The Proposed VRC-SLODL Model

In this study, we have introduced a new VRC-SLODL algorithm developed for vehicle classification on RSI. In the presented VRC-SLODL technique, the major intention lies in recognising and classifying vehicles present in the images. It encompasses four major processes: BF-based preprocessing, modified ResNet feature extraction, LSTM classification, and SLO-based hyperparameter tuning. Fig. 1 illustrates the overall procedure of the VRC-SLODL system.

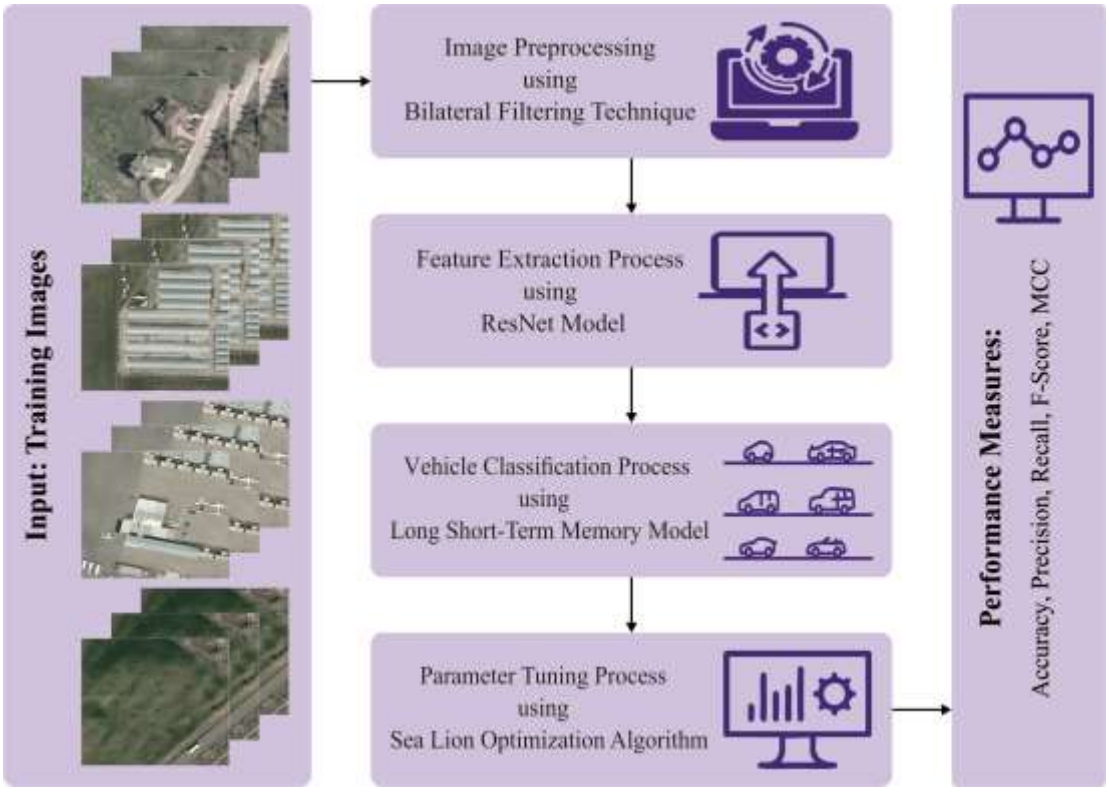

Figure 1: Overall process of VRC-SLODL system.

2.1. BF-based Preprocessing

At the primary level, the quality of the RSI can be enhanced by the BF technique. A BF can be exploited for de-noising the input images. The BF used spatial weighted averaging without the help of smoothing edges [21]. With the combination of two Gaussian filters, this could be accomplished in the spatial domain, one operating and in the intensity domain, the other functioning. The intensity

and spatial distance were applied for weight, and it is expressed below:

$$\overline{F}(p) = \frac{1}{N}\sum_{z\epsilon S(p)} e^{\frac{-\|q-p\|^2}{2\varepsilon_e^2}} \frac{-F(q)-F(p)\ \ ^2}{2E_S^2} F(q) \quad (1)$$

From the expression, $S(p)$ represents the pixel spatial neighbourhood (p), ε_e and ε_r symbolize the variable directs weight in intensity domain and spatial begin to reduce and N specifies the normalization constant.

$$e^{\frac{-\|q-p\|^2}{2\varepsilon_e^2}} e^{\frac{-F(q)-F(p)\ \ ^2}{2\varepsilon_e^2}} \quad (2)$$

BF was exploited in tone mapping, volumetric de-noising, and other applications like denoising images. They may generate basic conditions for down-sampling the crucial procedure and attain acceleration by formulating 2 modest non-linearity and augmented space. The BF can be applied as linear convolution.

2.2. Feature Extraction using modified ResNet Model

The VRC-SLODL technique employed the modified ResNet model for producing a set of feature vectors. The representation capability of the model becomes strong once the network is strengthened [22]. In other words, the model extracts additional features from the image. However, the network is complex to train. The initial problem encountered during training is gradient exploding or vanishing. The amount of calculation increases rapidly due to the surge in the number of layers. The gradient becomes unstable in the backpropagation model. For the gradient vanishing problems, researcher workers developed various solutions, namely batch normalization, initialization of MSRABN, etc. A further challenge is network degradation, viz., with the rise of network depth, the model efficiency becomes poor. Especially, the training accuracy would be decreased. In the case of overfitting, the training accuracy must be higher. ResNet resolves these obstacles using residual learning. Note that ResNet gives two choices, residual mapping and identity mapping. Let us assume that the network's input has obtained the optimum level. In other words, is optimum. The residual mapping will be set to once they continue to deepen the network. Currently, the network has identity mapping. The crucial thing is that network deepening won't reduce the model performance.

Formally, indicates the original mapping, and then the stacked non-linear layer fits another mapping of . Now, represents the residual mapping. The original mapping transformed into

This study uses Modified ResNet-152, a variation of ResNet-152, as a recognizer to categorize facial expressions. In the Modified ResNet-152 framework, the initial layer is 77 convolutions. Next, there are four major components with 9 layers, layers, layers, and 9 layers correspondingly. Lastly, implement the softmax, global average pooling, and a 7-way FC layer.

2.3. LSTM-based Classification

Here, the LSTM model is used for vehicle classification. Once the feature is made, they are inputted into LSTM classification for allotting the appropriate class. LSTM is an exclusive RNN model with RNN features, while the memory cell sequence was exploited to improve the learning procedure of time series and randomly handling input datasets [23]. Furthermore, the input dataset's long-term dependence is captured to prevent the gradient disappearance of data communication, a significant enhancement to capture the active modification of time sequence. It is the most productive RNN structure and is widely applied. This approach incorporated the input gate, output gate, output unit, input unit, and memory cells.

Input gate:

$$g_i = \sigma(W_{ix} \cdot X_z + W_{ih} \cdot h_{z-1} + b_i) \quad (3)$$

Forget gate:

$$g_f = \sigma\left(W_{fx} \cdot X_z + W_{fh} \cdot h_{z-1} + b_f\right) \quad (4)$$

Output gate:

$$g_o = \sigma(W_{ox} \cdot X_z + W_{oh} \cdot h_{z-1} + b_o) \quad (5)$$

Input unit:

$$C_{-\in\beta(W_{cx} \cdot X_z + W_{ch} \cdot h_{z-1} + b_c)} \quad (6)$$

Memory cell:

$$C_z = g_f \cdot C_{z-1} + g_i \cdot C_{-\in} \quad (7)$$

Output unit:

$$h_z = g_0 \cdot \theta(C_t) \quad (8)$$

$$y = \phi\left(W_y \cdot h_z + b_y\right) \quad (9)$$

From the expression, , , and correspondingly characterize the nonlinear activation function; signifies the resulting unit function; and represent the weight and bias coefficients that are implemented for establishing the connection amongst the units.

2.4. Hyperparameter Tuning using SLO Algorithm

In this work, the SLO algorithm acts as a hyperparameter optimizer. SLO has been proposed to resolve global-scale optimization [24]. It stimulates the hunting behavior of sea lions, including how they use their tail and whiskers or capture and encircle prey. SLO provides more competitive outcomes than other PSO techniques. SLO is given in the following. Initially, SLO creates (the size of population) -dimension solution through standard distribution in the searching space. Next, in the swarming of sea lions, they recognize the position of prey and gathers others who join subgroups for organizing the net, ensuring encircling model. The prey will be regarded as the solution nearby or the current optimum solution to the optimum solution. This behaviour is given as follows.

$$X_{ij}^{init} = X_{ij}^{min} + rand_{i,j}\left(X_{ij}^{max} - X_{i,j}^{min}\right) \quad (10)$$

Whereas:

$i = 1,2, \ldots, N$

$j = 1,2, \ldots, D$

$X_{i,j}^{init}$indicates the initial location vector of $thei\text{-}th$ solution;

$X_{i,j}^{min}$represent the minimal value for $thej\text{-}th$ parameter of $i\text{-}th$ solutions;

X_{ij}^{max}symbolizes the maximum value for a $j\text{-}th$ parameter of $ii\text{-}th$ solutions;

$rand$indicates a random integer within $[0,1]$.

The solution is estimated for fitness based on the objective function.

$$X^{g+1} = X_{best} - C|2rX_{best} - X^g| \quad (11)$$

$$C = 2\left(1 - \frac{g}{g_{max}}\right) \quad (12)$$

Whereas:

X_{best}^{g}indicates the location vector of the optimum solution;

X^{g}characterize sea lions in iteration g;

$g$signifies the existing iteration of generations;

g_{max}embodies the maximal amount of generations;

$r$symbolizes a arbitrary number withinand 1;

that is multiplied by 2 to improve search range;

X^{g+1}signify the novel location of the searching agent afterwards upgrading;

Cembodies a parameter with a linearly declined value from two to zero, which indicates the encircling model once they move to the prey and encircle them.

Once the sea lion identifies a prey, it calls other members to create and gather a net for capturing prey. Sea lions are regarded as leaders, and it leads a group of sea lions towards them and decides the group's behaviour. The mathematical expression of this behaviour can be given in the following:

$SP_{leader} =$ (13)

$V_1 = sin(\theta)$ (14)

$V_2 = sin(\phi)$ (15)

Whereas:

SP_{leader}characterize value that demonstrates decisions of leader ensued by others;

θcharacterize the angle of voice reflection in the water;

ϕsignifies the angle of voice refraction in the water;

In this study, $\theta = 2\pi r$ and $\varphi = 2\pi(1 - r)$, whereas r denotes the random value within$[0,1]$.

The hunting activity of sea lions led by the leader is defined below:

Dwindling encircling method: These behaviours depend on the value of C, and it declined linearly from 2 to 0. Hence it enables searching space round the present optimum location to force and shrink other searching agents for updating the searching space. Consequently, a recently upgraded location of the sea lion is positioned wherever in searching space between its existing location and position of optimum agent.

Circling updating location: Sea lion chases bait ball of fish and hunts them beginning from the edge, with m a random integer within [−1, 1].

$X^{g+1} = X_{best} + cos(2\pi m)|X_{best} - X^{g}|$

Searching for prey (Exploration stage)

In this stage, the searching agent upgrades the location according to the randomly chosen sea lion. The condition which enables the exploitation stage to perform was once the value of C is higher than 1, and procedure of identifying novel agent can be formulated as follow:

$X^{g+1} = X_{rand}^{g} - C\left|2rX_{rand}^{g} - X^{g}\right|$ (17)

In Eq. (17), X_{rand}^{t} represents the sea lions randomly chosen from the existing population. r represents random integer range from zero to one.

The work results show that SLO faces noticeable issues with nature-inspired techniques, including slow convergence and being trapped in local optimal. In the presented method, the exploration and exploitation stage for ISLO is enhanced than the original SLO algorithm.

Algorithm 1: Pseudocode of SLO Algorithm

Input: Population size N, the maximal amount of generations g_{max}

Output: The better solution X_{best}

Randomly initialize the Sea Lion population $X_i(i = 1,2, \dots, n)$

Arrange the population by the fitness values and find the better global solution X_{best}

$$g = 1$$

While $g < g_{max}$ do

Compute the value of C using Eq. (12)

Compute SP_{leader} by Eq. (11)

for $i < N$ do

if $SP_{leader} < 0.25$ then

if $C \vee 1$ then

Upgrade the position of the existing searching agent by Eq. (11)

Else

Choose a random agent X_{rand} from the existing population

Upgrade the position of the current individual using Eq. (17)

end if

else

Upgrade present searching agent position by Eq. (16)

end if

Check the bound and compute the fitness of the novel solution.

Replace the older solution with the newer one if the fitness value is better 20:

end for

Sort the population and Upgrade the global optimal solution X_{best}

$$g = g + 1$$

end while

Return: X_{best}

3. Results and Discussion

In this section, the experimental result analysis of the VRC-SLODL method conducted on the VEDAI dataset [25]. The dataset comprises 3687 instances with nine class labels, as shown in Table 1. Fig. 2 showcases the sample images.

Table 1: Details of VEDAI Dataset

Class	**No. of Instances**
Car	1340
Truck	300
Van	100
Pickup Car	950
Boat	170
Camping Car	390
Other	200
Plane	47
Tractor	190
Total Number of Instances	**3687**

Figure 2: Sample images

Fig. 3 provides the confusion matrices of the VRC-SLODL method on vehicle classification results. The figure pointed out that the VRC-SLODL scheme has properly classified different types of vehicles under all aspects.

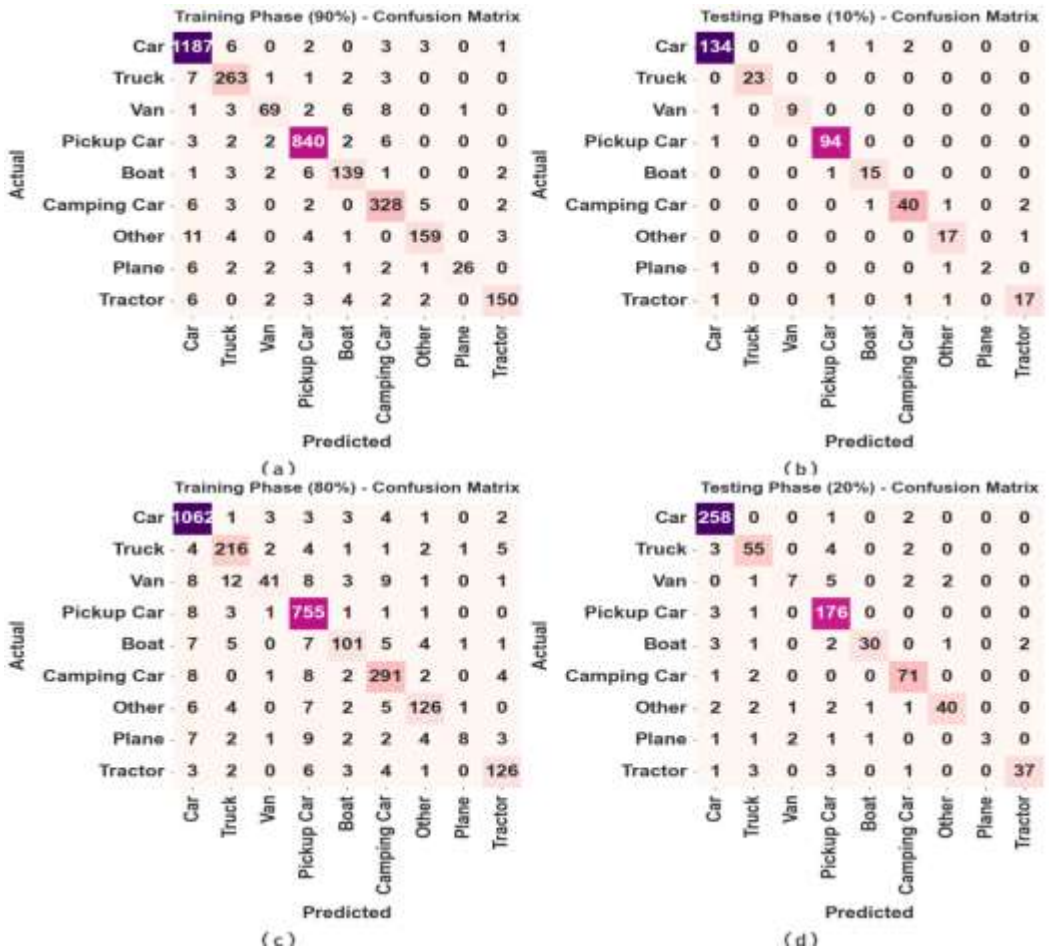

Figure 3: Confusion matrices of VRC-SLODL system (a-b) TR and TS databases of 90:10 and (c-d) TR and TS databases of 80:20

Table 2 shows detailed vehicle classification results of the VRC-SLODL model under 90% of TR and 10% of TS datasets. Fig. 4 illustrates a brief vehicle classification outcome of the VRC-SLODL approach on 90% of the TR datasets. On car class, the VRC-SLODL model has obtained an accu_y of 98.31%, prec_n of 96.66%, reca_l of 98.75%, F_score of 97.70%, and MCC of 96.38%. Simultaneously, on van class, the VRC-SLODL method has gained accu_y of 99.10%, prec_n of 88.46%, reca_l of 76.67%, F_score of 82.14%, and MCC of 81.90%.Parallelly, on boat class, the VRC-SLODL technique has attained anaccu_y of 99.07%, prec_n of 89.68%, reca_l of 90.26%, F_score of 89.97%, and MCC of 89.48%. Finally, ontractor class, theVRC-SLODL approach has gained accu_y of 99.19%, prec_n of 94.94%, reca_l of 88.76%, F_score of 91.74%, and MCC of 91.37%.

Table 2: Vehicle classification outcome of VRC-SLODL method under 90:10 of TR/TS databases

Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	**MCC**
Training Phase (90%)					
Car	98.31	96.66	98.75	97.70	96.38

Truck	98.88	91.96	94.95	93.43	92.83
Van	99.10	88.46	76.67	82.14	81.90
Pickup Car	98.85	97.33	98.25	97.79	97.02
Boat	99.07	89.68	90.26	89.97	89.48
Camping Car	98.70	92.92	94.80	93.85	93.13
Other	98.98	93.53	87.36	90.34	89.86
Plane	99.46	96.30	60.47	74.29	76.08
Tractor	99.19	94.94	88.76	91.74	91.37
Average	**98.95**	**93.53**	**87.81**	**90.14**	**89.78**
Training Phase (10%)					
Car	97.83	97.10	97.10	97.10	95.37
Truck	100.00	100.00	100.00	100.00	100.00
Van	99.73	100.00	90.00	94.74	94.74
Pickup Car	98.92	96.91	98.95	97.92	97.19
Boat	99.19	88.24	93.75	90.91	90.53
Camping Car	98.10	93.02	90.91	91.95	90.89
Other	98.92	85.00	94.44	89.47	89.04
Plane	99.46	100.00	50.00	66.67	70.52
Tractor	98.10	85.00	80.95	82.93	81.95
Average	**98.92**	**93.92**	**88.46**	**90.19**	**90.03**

Fig. 5 demonstrates a brief vehicle classification outcome of the VRC-SLODL approach on 10% of the TS database. On car class, the VRC-SLODL algorithm has attained an accu_y of 97.83%, prec_n of 97.10%, reca_l of 97.10%, F_score of 97.10%, and MCC of 95.37%. In parallel, on van class, the VRC-SLODL algorithm has anaccu_y of 99.73%, prec_n of 100%, reca_l of 90%, F_score of 94.74%, and MCC of 94.74%. Parallelly, in boat class, the VRC-SLODL technique has achieved anaccu_y of 99.19%, prec_n of 88.24%, reca_l of 93.75%, F_score of 90.91%, and MCC of 90.53%. Lastly, on tractor class, the VRC-SLODL methodology has obtained an accu_y of

98.10%, prec_n of 85%, reca_l of 80.95%, F_score of 82.93%, and MCC of 81.95%.

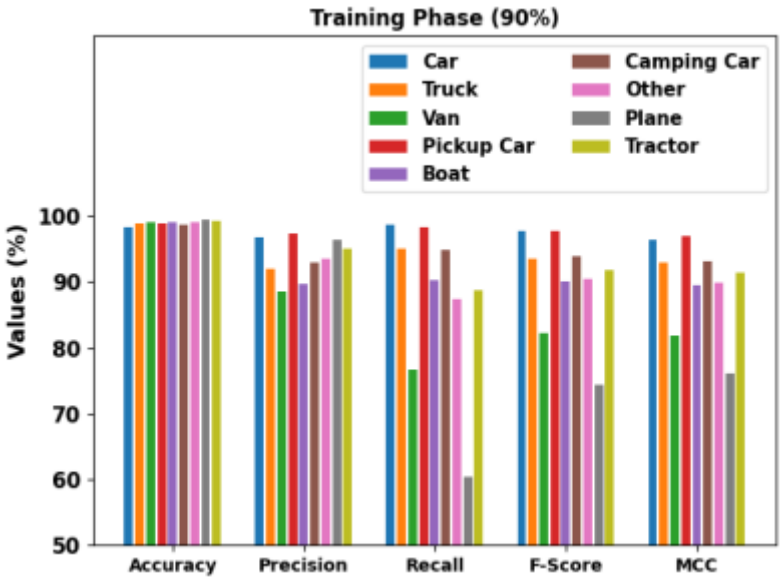

Figure 4: Result analysis of VRC-SLODL approach on 90% of TR database

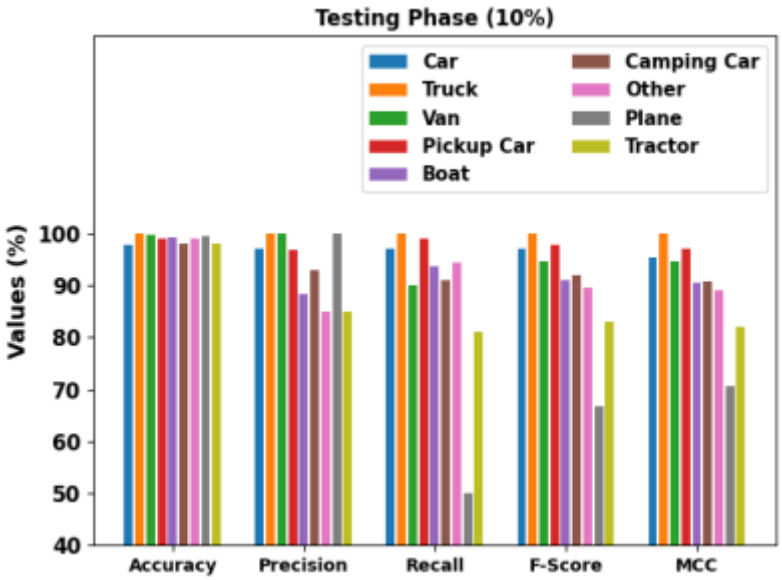

Figure 5: Result analysis of VRC-SLODL approach on 10% of TS database

Table 3 shows a brief vehicle classification outcome of the VRC-SLODL technique under 80% of TR and 20% of TS databases.

Fig. 6 portrays a brief vehicle classification outcome of the VRC-SLODL algorithm on 80% of the TR datasets. On car class, the VRC-SLODL method has attained anaccu_y of 97.69%, prec_n of 95.42%, reca_l of 98.42%, F_score of 96.90%, and MCC of 95.09%. In parallel, on van class, the VRC-SLODL technique has attained an accu_y of 98.30%, prec_n of 83.67%, reca_l of 49.40%, F_score of 62.12%, and MCC of 63.55%. Concomitantly, on boat class, the VRC-SLODL approach has gained accu_y of 98.41%, prec_n of 85.59%, reca_l of 77.10%, F_score of 81.12%, and MCC

of 80.41%. Finally, on tractor class, the VRC-SLODL approach has gained accu_y of 98.91%, prec_n of 88.73%, reca_l of 86.90%, F_score of 87.80%, and MCC of 87.19%.

Table 3: Vehicle classification outcome of VRC-SLODL method under 80:20 of TR/TS databases

Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	**MCC**
Training Phase (80%)					
Car	97.69	95.42	98.42	96.90	95.09
Truck	98.34	88.16	91.53	89.81	88.93
Van	98.30	83.67	49.40	62.12	63.55
Pickup Car	97.73	93.56	98.05	95.75	94.25
Boat	98.41	85.59	77.10	81.12	80.41
Camping Car	98.10	90.37	92.09	91.22	90.16
Other	98.61	88.73	83.44	86.01	85.32
Plane	98.88	72.73	21.05	32.65	38.76
Tractor	98.81	88.73	86.90	87.80	87.19
Average	**98.32**	**87.44**	**77.55**	**80.38**	**80.41**
Training Phase (20%)					
Car	97.70	94.85	98.85	96.81	95.06
Truck	97.29	83.33	85.94	84.62	83.14
Van	98.24	70.00	41.18	51.85	52.89
Pickup Car	97.02	90.72	97.78	94.12	92.24
Boat	98.51	93.75	76.92	84.51	84.18
Camping Car	98.51	89.87	95.95	92.81	92.04
Other	98.37	93.02	81.63	86.96	86.30
Plane	99.19	100.00	33.33	50.00	57.50

Tractor	98.64	94.87	82.22	88.10	87.63
Average	**98.16**	**90.05**	**77.09**	**81.08**	**81.22**

Fig. 7 illustrates a brief vehicle classification outcomes of the VRC-SLODL model on 20% of the TS database. On car class, the VRC-SLODL model has obtained an accu_y of 97.70%, prec_n of 94.85%, reca_l of 98.85%, F_score of 96.81%, and MCC of 95.06%. Simultaneously, on van class, the VRC-SLODL approach has obtained accu_y of 98.24%, prec_n of 70%, reca_l of 41.18%, F_score of 51.85%, and MCC of 52.89%. Simultaneously, on boat class, the VRC-SLODL approach has achieved accu_y of 98.51%, prec_n of 93.75%, reca_l of 76.92%, F_score of 84.51%, and MCC of 84.18%. Finally, on tractor class, the VRC-SLODL method has gained accu_y of 98.64%, prec_n of 94.87%, reca_l of 82.22%, F_score of 88.10%, and MCC of 87.63%.

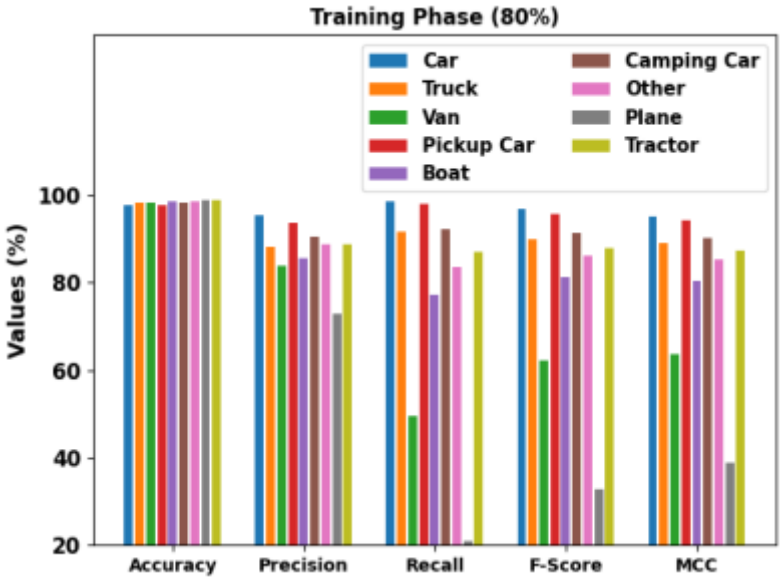

Figure 6: Result analysis of VRC-SLODL method on 80% of TR dataset

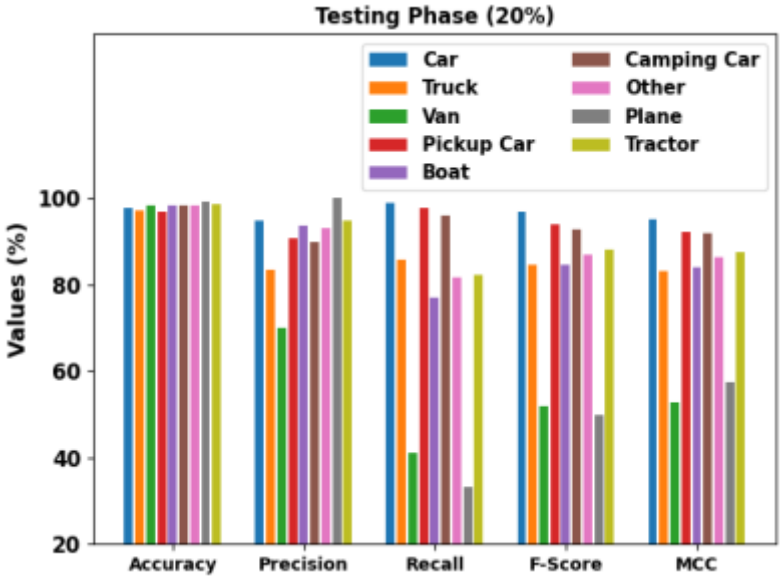

Figure 7: Result analysis of VRC-SLODL method on 20% of TS dataset

For exhibiting the enhanced performance of the VRC-SLODL approach, a comparative $accu_y$examination with existing models takes place in Table 4. The obtained values inferred that the LeNet model had shown decreased performance with an $accu_y$ of 80.02%. In contrast, the AlexNet model has reported a slightly improved $accu_y$ of 88.92%. Moreover, the VGG-16 model has tried to accomplish a moderately reasonable $accu_y$ of 94.38%. But the VRC-SLODL model has shown superior performance with an $accu_y$ of 98.95%. These results signified the betterment of the VRC-SLODL method over other DL methods.

Table 4: Accuracy analysis of VRC-SLODL technique with recent systems

Methods	$Accu_y$(%)
LeNet	80.02
AlexNet	88.92
VGG-16	94.38
VRC-SLODL	98.95

4. Conclusion

In this study, we have introduced a new VRC-SLODL method developed for vehicle classification on RSI. In the presented VRC-SLODL technique, the major intention lies in recognising and classifying vehicles present in the images. Firstly, the quality of the RSI can be enhanced by the BF technique. The VRC-SLODL technique employed by the modified ResNet model is used to produce a set of feature vectors. Eventually, the SLO algorithm with the LSTM approach is used for vehicle classification, where the SLO algorithm acts as a hyperparameter optimizer. The experiments were conducted on a benchmark dataset to examine the better performance of the VRC-SLODL method. The obtained values reported the improved classification performance of the VRC-SLODL technique over other models. In the upcoming years, the performance of the VRC-SLODL method can be improvized by fusion-based DL models.

References

[1] X. Wu, W. Li, D. Hong, J. Tian, R. Tao et al., "Vehicle detection of multi-source remote sensing data using active fine-tuning network," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 167, pp. 39-53, 2020, doi: 10.1016/j.isprsjprs.2020.06.016

[2] M. Mandal, M. Shah, P. Meena, S. Devi and S.K. Vipparthi, "AVDNet: A small-sized vehicle detection network for aerial visual data," IEEE Geoscience and Remote Sensing Letters, vol. 17, no. 3, pp.494-498, 2019.

[3] S.A. Ahmadi, A. Ghorbanian and A. Mohammadzadeh, "Moving vehicle detection, tracking and traffic parameter estimation from a satellite video: A perspective on a smarter city," International journal of remote sensing, vol. 40, no. 22, pp. 8379-8394, 2019.

[4] Y. Koga, H. Miyazaki and R. Shibasaki, "A CNN-based method of vehicle detection from aerial images using hard example mining," remote sensing, vol. 10, no. 1, p. 124, 2018.

[5] B. Ma, Z. Liu, F. Jiang, Y. Yan, J. Yuan et al., "Vehicle detection in aerial images using rotation-invariant cascaded forest," IEEE Access, vol. 7, pp. 59613-59623, 2019, doi: 10.1109/ACCESS.2019.2915368.

[6] H. Song, H. Liang, H. Li, Z. Dai and X. Yun, "Vision-based vehicle detection and counting system using

deep learning in highway scenes," European Transport Research Review, vol. 11, no. 1, pp. 1-16, 2019.
[7] F. Zhang, C. Li and F. Yang, "Vehicle detection in urban traffic surveillance images based on convolutional neural networks with feature concatenation," Sensors, vol. 19, no. 3, p. 594, 2019.
[8] S. Zhang, G. He, H.B. Chen, N. Jing and Q. Wang, "Scale adaptive proposal network for object detection in remote sensing images," IEEE Geoscience and Remote Sensing Letters, vol. 16, no. 6, 2019, pp.864-868.
[9] K. Li, G. Wan, G. Cheng, L. Meng and J. Han, "Object detection in optical remote sensing images: A survey and a new benchmark," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 159, pp. 296-307, 2020, doi: 10.48550/arXiv.1909.00133.
[10] H. Yao, R. Qin and X. Chen, "Unmanned aerial vehicle for remote sensing applications—A review," Remote Sensing, vol. 11, no. 12, p. 1443, 2019.
[11] R. Bibi, Y. Saeed, A. Zeb, T.M. Ghazal, T. Rahman et al., "Edge AI-based automated detection and classification of road anomalies in VANET using deep learning," Computational intelligence and neuroscience, 2021, doi: 10.1155/2021/6262194.
[12] S.M. Azimi, E. Vig, R. Bahmanyar, M. Körner and P. Reinartz, "Towards multi-class object detection in unconstrained remote sensing imagery," In Asian conference on computer vision, pp. 150-165, 2018, doi: 10.1007/978-3-030-20893-6_10.
[13] J. Rabbi, N. Ray, M. Schubert, S. Chowdhury and D. Chao, "Small-object detection in remote sensing images with end-to-end edge-enhanced GAN and object detector network," Remote Sensing, vol. 12, no. 9, p. 1432, 2020.
[14] X. Zhang, L. Han, L. Han and L. Zhu, "How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery?," Remote Sensing, vol. 12, no. 3, p. 417, 2020.
[15] S.A. Ahmadi, A. Ghorbanian and A. Mohammadzadeh, "Moving vehicle detection, tracking and traffic parameter estimation from a satellite video: A perspective on a smarter city," International journal of remote sensing, vol. 40, no. 22, pp. 8379-8394, 2019.
[16] K.J. Kim, P.K. Kim, Y.S. Chung and D.H. Choi, "Multi-scale detector for accurate vehicle detection in traffic surveillance data," IEEE Access, vol. 7, pp. 78311-78319, 2019, doi: 10.1109/ACCESS.2019.2922479.
[17] X. Liu, T. Yang and J. Li, "Real-time ground vehicle detection in aerial infrared imagery based on convolutional neural network," Electronics, vol. 7, no. 6, p. 78, 2018.
[18] Y. Guo, Y. Xu and S. Li, "Dense construction vehicle detection based on orientation-aware feature fusion convolutional neural network," Automation in Construction, vol. 112, p. 103124, 2020, doi: 10.1016/j.autcon.2020.103124
[19] X.X. Zhang and X. Zhu, "Moving vehicle detection in aerial infrared image sequences via fast image registration and improved YOLOv3 network," International Journal of Remote Sensing, vol. 41, no. 11, pp. 4312-4335, 2020.
[20] Y. Koga, H. Miyazaki and R. Shibasaki, "A method for vehicle detection in high-resolution satellite images that uses a region-based object detector and unsupervised domain adaptation," Remote Sensing, vol. 12, no. 3, p. 575, 2020.
[21] M. Elhoseny and K. Shankar, "Optimal bilateral filter and convolutional neural network based denoising method of medical image measurements," Measurement, vol. 143, pp. 125-135, 2019, doi: 10.1016/j.measurement.2019.04.072
[22] W. Xu and R.S. Cloutier, "A facial expression recognizer using modified ResNet-152," EAI Endorsed Transactions on Internet of Things, vol. 7, no. 28, p. e5, 2022.
A. Girma, X. Yan and A. Homaifar, "Driver identification based on vehicle telematics data using LSTM-recurrent neural network," In EEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), pp. 894-902, 2019, doi: 10.1109/ICTAI.2019.00127.
[23] B.M. Nguyen, T. Tran, T. Nguyen and G. Nguyen, "An Improved Sea Lion Optimization for Workload Elasticity Prediction with Neural Networks," International Journal of Computational Intelligence Systems, vol. 15, no. 1, pp. 1-26, 2022.
[24] S. Razakarivony and F. Jurie, “Vehicle detection in aerial imagery: A small target detection benchmark,” Journal of Visual Communication and Image Representation, vol. 34, pp. 187–203, 2016, doi:10.1016/j.jvcir.2015.11.002.