INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Vehicle Detection in SAR Satellite Images Using Yolov8 Oriented Bounding Box Detection Algorithm

Sudheer Reddy Bandi[1]*, **G Merlin Linda**[2], **S Nagarjuna Chary**[3], **K Spurthy**[4]

1 Department of Artificial Intelligence and Data Science, J.B. Institute of Engineering and Technology, Telangana, Hyderabad, 500075, India
2 Department of Computer Science and Engineering, Vidya Jyothi Institute of Technology, Hyderabad, 500085, Telangana, India
3 Department of Electronics and Instrumentation Engineering, Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering and Technology, Hyderabad, 500090, Telangana, India
4 Department of Computer Science and Engineering, Swarna Bharathi Institute of Science and Technology, Khammam, 507001, Telangana, India

OPEN ACCESS

Received: 08-08-2024

Accepted: 08-05-2025

Published: 30-05-2025

Editor: Special Issue Editors: Dr. N. Pothanna, Dr.T. Jayashree, Dr. V. Ganesh Kumar

Citation: Bandi SR, Linda GM, Chary SN, Spurthy K (2025) Vehicle Detection in SAR Satellite Images Using Yolov8 Oriented Bounding Box Detection Algorithm. Indian Journal of Science and Technology 18(SP1): 62-71. https://doi.org/10.17485/IJST/v18si1.icamada38

* **Corresponding author**.

sudheer653@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To detect object localization in visualizing Synthetic Aperture Radar, known as SAR Satellite Images, by introducing an additional angle to rotate and fit the orientation of the object placement, which is more precise, essential for investigating vehicle targets in a wide area of remote sensing technology. **Methods**: The orientation of the object detection is carried out on Ka, Ku, and X bands of data by exhibiting richness, stability, and challenge as fundamental properties. The SAR Image data for the VEhicle Detection (SIVED) dataset involves complex background information with 1044 chips and 12,013 vehicle instances of SAR images. The collected dataset demonstrates the usage of detecting buildings, vehicles, ships, oil tanks, security, and damage control assessment. The pre-trained deep learning bounding box algorithm is used to predict and detect vehicle targets with high speed and efficiency. **Findings**: The achieved results for detecting the object localization evaluation score for YOLOv8 Oriented Bounding Box (OBB) one-stage detection are 98.837 and 98.52 for recall and precision, respectively. The proposed YOLOv8 one-stage detection algorithm surpassed all other traditional models, achieving an outstanding mean Average Precision (mAP) of 99.83. Further, the losses obtained are minimal with box loss and class loss of 0.5015 and 0.29082, respectively. **Novelty**: Enhancing object localization in vehicle targets helps to systematically estimate the detection rate, computational complexity, and resource utilization by optimizing traffic planning and reducing environmental effects. Orientation performance and features play a key role in measuring the relative importance of SAR images, leading to a better decision-making process. The earlier research efforts considered the architecture of axis-aligned bounding boxes to optimally enclose an object. This prevents customized axis-angle representations. The proposed work is enhanced by a more optimized architecture with specific transformations for handling rotated objects. Moreover, the proposed work is used to fine-tune anchor boxes for better bounding box representation in SAR images. Further, the geometry of rotating boxes is crucial in improving the convergence rate between the predicted region and the ground truth angle. To be more precise, the speckle noise augmentation, refined anchor-free mechanism, and multi-scale feature aggregation attributes are effective in detecting SAR vehicles as depicted in the experimental results. The code and data are available at the GitHub repository given as SudheerReddyBandi/ICAMADA2024VNRVJIET (github.com).

Keywords: OBB; YOLOv8; Remote Sensing; Object Localization; Rotating Angle; Vehicle Detection; Sentinel Data

1 Introduction

In recent times, the usage of high-resolution SAR satellite images has been widely adopted to automatically detect moving targets in infrastructure maintenance, resource inspection, disaster prevention, urban agricultural planning, and other global economic assessments. Due to complex angle alignments and arbitrary background indifference in satellite imagery, there is a need to analyze semantic information while expanding vehicle target detection. Incorporating potential enhancements in target detection focuses on adapting anchor boxes to different objects and aspect ratios by explicitly mapping relevant image regions for improved feature representation. The existing works are limited to tracing rectangular axis areas; on the other hand, not capable of locating rotated bounding box angles in detecting densely positioned vehicle targets. With advancements in Deep Learning (DL) algorithms gradually integrated with the SAR image dataset to provide a comprehensive understanding of robust target detection in traffic surveillance scenarios. Although reducing the quality of the image and making object detection difficult due to increasing noise level, hard to identify the heavy vehicles due to poor contrast, misrepresentations of object shapes due to geometry occlusions and the need for arbitrary OBB detection for analyzing the similarity of target objects are the unique challenges to justify the review in more depth. Considering the huge quantity of data in monitoring the visuals of traffic situations, the proposed work emphasizes core components by building a highly significant architecture to detect the vehicles' movements precisely in real-time applications. Moreover, deep pre-trained network models help to automatically abstract the high-level features in target detection and tune their object shapes with rotated anchor bounding boxes in complex backgrounds. Due to dominant performance in various fields, many researchers practically use deep networks to detect SAR vehicles for long-term groundwork improvements and global transportation.

Intelligent vehicle detection system visually optimizes the movements of traffic patterns in urban areas using innovative technologies and big data analytics to measure the location of weak lane lines robustly, identify road areas during climatic conditions, and environmental illuminations circumstances, thus improving the community livelihood[1]. Cloud servers are equipped to store the captured images using sensors, cameras, drones, satellites, and other complex algorithms to process and generalize data efficiently. This technological revolution helps to use advanced data transmission to save time and cost while extracting the needed information in a notable time[2]. The multi-scale rotational dense network will enhance ship detection using rotational-invariance features across various axes in SAR images. This model does not generalize well for other types of detection because of the complex background noise[3]. The concept of the YOLO algorithm is introduced to optimize target detection by filtering the noise ratio and identifying objects with arbitrary orientations with improved accuracy. Its performance is limited due to variations in the background poses, axis anchor alignment, and data imbalances[4]. The advanced anchor-based one-stage model detects oriented bounding boxes in recognizing the overlapped vehicle movements, leading to average performance. The features, such as horizontal and vertical alignments, are marked easily in satellite images from a wide resolution angle. It helps to visualize the highlighted information easily by making manual adjustments in axis alignments. The problem with this model is that additional parameters are required to make the bounding box fit the target concept[5]. To reduce the parameters, deep network models integrate the spatio-temporal imbalances and address the system performance with improved speed and accuracy. The automatic localization of objects is not handled well in the YOLOV7 model[6]. The fusion of algorithms using YOLOV5 and YOLOV8 measures the conjunction rate and automatic rotational scales through the deep features of network functions[7].

Further, various studies suggest that the YOLO deep network algorithm is utilized for most important object detection systems. YOLOV8 will significantly handle complex data representation and enhanced feature similarity by including a Deformable Attention Module (DAM). The challenge of upholding the densely allocated vehicles needs to be considered for higher evaluation results[8]. Further, in the literature[8], there are a few limitations such as data quality, data availability, high resolution challenges, training complexity, and oriented boundary boxes, which are the main objective of the proposed work. These limitations shall be overcome with the proposed YOLOv8-OBB technique. In later stages, improvements in major YOLOV8 algorithms make the target detection converge quickly with fewer threshold variations. Thus, introducing the hyperparameter in the network model provides a change of loss function, adapting to minimize the error rate in analyzing the target object[9]. Also, the new customized deep network uses a fusion of model features to prioritize the two-stage multi-scale detection of targets effectively with less error and an increase in accuracy[10]. Data variability brought on by varying weather conditions, the likelihood of false positives and false negatives, and imprecise tracking are some of the limitations related to data, tracking algorithms, and modeling[11]. Further, the literature works on SAR ships. The YOLOv8-OBB object detection framework overcomes these limitations through a dynamic anchor-free design to develop data in various scales and sizes. In addition, DeepSORT is constructed for improved consistency and precision, and learning complex feature representations to recognize subtle object properties, reducing false negatives. The technique in[12] faces issues like detecting small objects, false positives, false negatives, and image quality sensitivity. YOLOv8-OBB, designed for varying object orientations with FPN and PAN, improves detection accuracy by reducing false negatives and positives by focusing on relevant features and reducing background noise. Also, the unique feature of Bi-directional Feature Pyramid Network (BiFPN) supports in detecting small objects. On the other hand, the methodology in[13] is limited in detecting small shapes and optimizing the detection head with various challenges, such as rotated objects detection, computational complexity, and generalization. However, the OBB nature and large architecture of YOLOv8 address the issues. Based on the above-mentioned methods, there is a research gap concerning the architecture's ability to manage data at various scales, rotations, and sizes. Furthermore, crucial problems remain unanswered about the speckle noise available in SAR images. Further, there are challenges with occlusions and clutter. The aforementioned challenges are addressed with the contributions described in a five-fold work as follows:

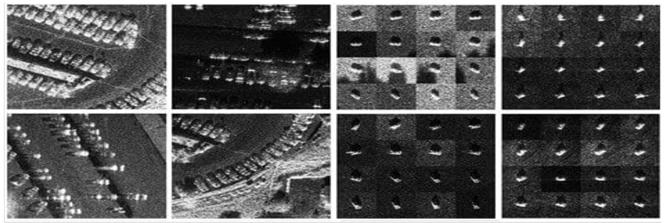
- The implementation of novel YOLOv8-OBB is oriented boundary box detection for better accuracy, noise reduction, and detection of small-sized objects. Further, the important attributes are data augmentation, multi-scale feature fusion, and attention mechanisms.
- Working with the variable SAR dataset with less spectral resolution and integrating different polarizations in adverse environmental conditions and geometric distortions.
- This will be one of the best works in the SIVED dataset to detect the vehicles and show the path to future work for diverse object detection methods.
- Implement the DeepSORT mechanism to improve the performance of false positives and negatives.

The huge research gap allows for evolving strong models capable of competently analyzing and combining data from diverse vehicle detection technologies. The proposed work will include potential adaptation via transfer learning to lower noise levels, anchor box orientation, and data augmentation in noisy situations. Moreover, tweaking model parameters increases precision and model performance. This research will help to control SAR vehicle detection in road safety and dependability, which is an important part of urban civilization planning and supervision, as well as traffic management.

2 Methodology

2.1 Dataset preparation

The dataset used in the proposed work is SIVED: A SAR Image Dataset for Vehicle Detection Based on Rotatable Bounding Box in X, Ka, Ku bands, as demonstrated in[14] and given in the downloadable link GitHub – CAESAR-Radi/SIVED. The dataset information is revealed in Table 1. The dataset is a combination of Moving and Stationary Target Acquisition and Recognition (MSTAR)[15], MiniSAR[16], and FARAD[16] images and depicted in Figure 1. The first two columns denote the images from MiniSAR, FARAD, and the last two columns indicate the images of MSTAR. The proposed SIVED dataset has a complex background with trees, plantations, road networks, and buildings.



(a) (b)

Fig 1. (a) MiniSAR and FARAD images. (b) MSTAR images

Table 1. Demonstration of the SIVED dataset used in the experiments

Data	Band	Polarization	Resolution	Total	Train	Test	Valid
FARAD	Ka/X	VV/HH	0.1m x 0.1m	621	498	61	62
miniSAR	Ku	-	0.1m x 0.1m	100	80	10	10
MSTAR	X	HH	0.3m x 0.3m	323	259	32	32

2.2 Methodology and Architecture of YOLOv8 1 OBB

The YOLOv8.1 obb was built on the YOLOv8 version with some modifications, and it is depicted in Figure 2. The strong arrival of object detection in computer vision brought good application scenarios like target classification, self-driving cars, face detection, manual feature extraction, and so on. With the development of neural networks, as discussed in the previous sections, there are OSDs and TSDs. In OSD, the YOLOv8.1 version gives maximum frames per second (FPS) and floating point operations (FLOPs) of 130 and 3.9G, respectively, that minimizes the computational complexity compared to the state-of-the-art approaches. The name indicates that from a single-shot algorithm, detection and classification probabilities of an object are found. Therefore, the time complexity is lower and is used for autonomous applications. On the other hand, TSDs are more complex, with the two steps involving the feature extraction of objects using deep learning and the identification of objects from the features in the second stage. Hence, these are not trained for end-to-end applications. The basic components of YOLOv8.1 obb are detailed below:

Backbone Network Module (BNM): This is in the early stage, where the convolutional blocks are used along with the Darknet to extract the image features from low-level to high-level at different resolutions. The activation function Sigmoid Linear Unit (SiLU) is implemented to avoid vanishing gradient problems and benefit from smooth gradients.

Neck Network Module (NNM): This is also known as Path Aggregation Network (PANet). The features extracted from the previous networks are merged to make good predictions of objects of varying sizes and in adversarial conditions. However, this is possible with the help of the Spatial Pyramid Pooling Fast (SPPF) block. Both the Backbone and Neck Network modules are custom-designed for real-time performance.

Head Network Module (HNM): In this module, the features from the PANet are considered for predictions of class bounding boxes. Further, there are numerous results such as confidence scores indicating the object class, class probabilities, and the possibility of an object being present. The HNM is designed to overcome the limitations of the Yolov5 OBB version in terms of OBB parameter predictions, geometric distortions, and image contrast enhancements.

2.3 Innovations in YOLOv8 1 Oriented Bounding Box (OBB)

In the earlier versions of Yolo, the bounding boxes are rectangular grids without any angles. The v8.1 version breaks through the earlier versions to find the objects with various rotations and angles under different environmental and occlusion conditions. Ahead of digging into YOLOv8, it's important to understand the particular challenges of SAR image vehicle detection. The most trivial challenges are speckle noise (dense noise), geometric distortions (terrain, sensor geometry, and motion), vehicle orientation, and image contrast (sensor characteristics). The architecture of YOLOv8.1 OBB is designed over other approaches to overcome these shortcomings. For instance, the dense noise will impact the detection rate if we use YOLOv5 OBB or other

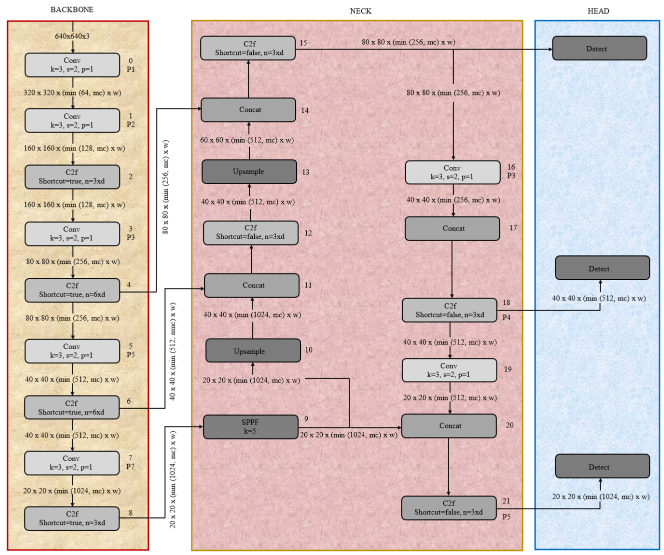


Fig 2. The three stages of Yolov8 architecture

standard approaches, but this shall be reduced with the non-max suppression nature and the architectures of the neck and backbone, respectively. Geometric modifications included in data augmentation can enhance the model's ability to withstand distortion. Rotation, scaling, shearing, and perspective transformation are methods that can replicate real-life situations and enhance the model's adaptability. The primary component of YOLOv8.1, probably a convolutional neural network, is designed to extract significant features from images. These characteristics can remain unchanged during certain geometric transformations, enabling the model to handle distorted objects. However, the orientation shall be controlled by the OBB nature of the proposed work. Finally, the image contrast can be effectively handled by the special segmentation capabilities and classification qualities of YOLOV8.1 OBB incorporated into the pipeline. Enhancing the training database with images with varying contrast levels may improve the model's robustness. This includes raising or lowering contrast and maybe generating synthetic contrast swings. Using histogram equalization on photos during training can help standardize contrast and improve overall performance. YOLOv8.1 most likely makes use of a strong framework capable of extracting distinguishing features even in low-light settings and high contrast settings.

Hence, the model is particularly used in satellite remote sensing to discover numerous objects such as ships, buildings, and vehicles. Further, the OBB technique excels in object outlining through image segmentation, uncovering various application scenarios, and object analysis. The proposed model is created on the previous techniques to improve the general correctness and robustness. Therefore, various benefits of using the OBB version include efficiency, speed, background noise reduction, improved accuracy, and so on. YOLOv8.1 OBB modifies the head to predict OBB parameters instead. These parameters typically include:

a) Center Coordinates (x, y): Axis-aligned bounding boxes that represent the object's centre point.

b) Width (w): Symbolizes the width of the OBB.

c) Height (h): Symbolizes the height of the OBB.

d) Rotation Angle (θ): This additional parameter specifies the rotation of the OBB relative to the image axes. This allows for more accurate bounding of angled objects.

2.4 Framework of the complete workflow

The main objective is to localize the objects and detect the bounding boxes with certain orientations, hence the name OBB. This feature of YOLOv8.1 obb intensifies the performance of segmentation, classification, and performance optimizations.

1. As usual, the CNN takes the input from the user for feature extraction from the input data in the backbone network.

2. In the second step, the features from the first step are considered across different scales and predict the class probabilities with bounding box coordinates and orientation for information fusion in PANet. The grid cells with angles are formed around the image for predicting a set of bounding boxes. This step is for finding the orientation and generating the horizontal object with no background.

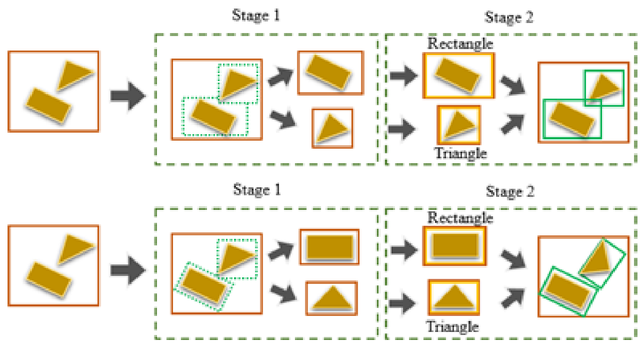


Fig 3. Stages at the top indicate the demonstration of normal YOLO. The stages at the bottom depict the YOLO OBB working principle

3. Now, the general approach of bounding box detection is applied with orientation offset, where each cell predicts the bounding box with class probabilities and generates the oriented bounding box. This is achieved with a modified prediction head network.

4. In this stage, the boxes that network with the best probability are considered. During training, the model is optimized based on the loss calculated between the predicted OBBs and the ground truth OBBs in the training dataset. This refines the model's ability to accurately detect and localize rotated objects. The difference is depicted in Figure 3.

2.5 Training and Testing Process

The total number of images is 1044, with 837 images for training, 103 images for validation, and testing with 104 samples. The division and properties of FARAD, MiniSAR, and MSTAR are given in Table 1. The training/testing/validation process involves the Ultralytics framework and supervision library with deep learning model support for object detection, segmentation, enhancement, and semantic segmentation. The various files used for the SAR vehicle detection are the data file, the YAML file, the labels, and the images. YAML represents the path of the training, testing, and validation images for the reading process. The values of the MiniSAR_61.txt file contain the following information, with variation values. Generally, the format given in SIVED is x1 y1 x2 y2 x3 y3 x4 y4 class_index class_id.

The above information is converted into the format class_index x1 y1 x2 y2 x3 y3 x4 y4, where the values of the coordinates are normalized between 0 and 1. The .pt files observed are the pre-trained PyTorch files. After the training process completion, two PyTorch files are available in the train folder: (i) last.pt for resuming the training process, (ii) best.pt for the validation and prediction process. The parameters are discussed in section 3.1 and tabulated in Tab. 3. The evaluation metrics used in the experiments are box_loss, cls_loss, dfl_loss, precision, recall, mAP50, and mAP50-95. The mAP50 means the IoU (Intersection over Union) average precision measured at a threshold of 0.5. The training, validation, and prediction process is executed with the following command line interface (CLI).

!yolo task=obb mode=train model=yolov8x-obb.pt data=/content/SIVED/data.yaml epochs=150 imgsz=512 batch= 8

!yolo task=obb mode=val model=/content/runs/obb/train/weights/best.pt source="/content/SIVED/ImageSets/val/images" save=true

!yolo task=obb mode=predict model=/content/runs/obb/train/weights/best.pt conf=0.4 source="/content/SIVED/ImageSets/test/images" save=true

2.6 Loss function

The loss function for the proposed model is given as the sum of Classification loss (CL_loss), Binary Cross Entropy (BCE), Distribution Focused Loss (DFL), and Complete Intersection over Union (CIoU) loss. The segmentation loss is the overall loss of all the losses to give pixel-wise classification.

Classification loss: This function handles the correct object prediction and reduces the probabilities for incorrect object detection. This is also known as class loss, which is categorical cross-entropy loss.

BCE loss: This is also known as confidence score loss, which predicts the presence of an object in a grid cell. The loss calculated is the Binary cross-entropy loss. This loss is considered a high score for the actual object and a lower confidence score for the incorrect object or background noises. This is also known as Bounding Box Loss (BBL).

CIoU loss: This loss is known as the bounding box loss that differentiates the predicted OBB and the actual OBB, with a higher IoU value specifying a higher loss. This value is detected with the amount of region overlapped between the ground truth and the predicted region.

DFL loss: As the name suggests, this loss is used to maintain the balance among classes in the learning process. The loss of the easily predicted class is overburdened and improves the bounding box regression.

total_loss = BCE_loss + DFL_loss + CIoU_loss + CL_loss + seg_loss

3 Results and Discussion

3.1 Environmental setup

The experiments are conducted with different models of YOLOv8.1 OBB, and the results for 150 epochs and a batch size of 8 are tabulated in Table 2, and other parametric values are given in Table 3. Also, as the OBB was introduced in YOLOv5, the results are evaluated with the same. The proposed work is executed in Google Colaboratory with a Graphical Processing Unit (GPU) and Python 3.9. The parameters inside the table indicate the class loss (cls_loss) and distribution focal loss (dfl) to balance the imbalanced data classes. In the table, the first box_loss, cls_loss, and dfl_loss relate to the training data, secondly, the metrics are tabulated, finally, the validation data for the losses are given respectively. The predicted results are given in Figure 4. The results in bold indicate better results when compared to other versions.

Table 2. Results were evaluated from the YOLOv8.1 versions of vehicle detection experiments

	Train			Metrics				Validation		
Model	**box_loss**	**cls_loss**	**dfl_loss**	**precision**	**recall**	**mAP50**	**mAP50-95**	**box_loss**	**cls_loss**	**dfl_loss**
YOLOv5	0.75369	0.46957	1.22307	0.9527	0.94481	0.98150	0.79989	0.75057	0.45380	1.32541
YOLOv8m	0.51116	0.29994	1.0976	0.982	0.97528	0.9896	0.84118	0.65565	0.36496	1.2288
YOLOv8n	0.6689	0.394	1.1593	0.97541	0.96948	0.98746	0.81012	0.74652	0.43302	1.2858
YOLOv8l	0.503	0.29436	1.0971	0.97832	0.97791	0.98985	0.85236	0.62891	0.35756	1.2141
YOLOv8s	0.56557	0.33203	1.1131	0.9724	0.96727	0.98822	0.83088	0.68015	0.39955	1.2578
YOLOV8x	**0.5015**	**0.29082**	**1.0178**	**0.9852**	**0.98837**	**0.9983**	**0.86723**	**0.61955**	**0.35086**	**1.2031**

As the oriented boundary box detection was introduced in YOLOv5 and YOLOv8.1 obb, the experiments are performed through these techniques. There is a good progression in the YOLO techniques, as is observed from the results and the normal boundary box detection to the OBB detection. The values in the pre-process, inference, post-process of an image, and the training process indicate the time taken in milliseconds (ms). On the other hand, the parameter values are in the millions. Further, from the literature and Table 3, it is clear that the Nano version of YOLO takes less time. The training time is longer in yolov5 and yolov8x, and moderate in yolov8s.

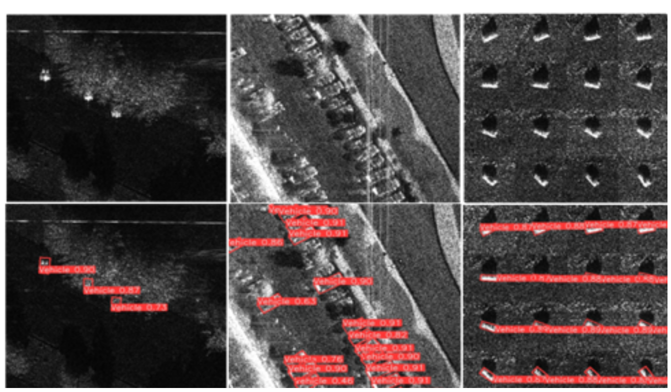


Fig 4. Predicted results using YOLOv8.1 for FARAD, MiniSAR, and MSTAR images, respectively

3.2 Comparison with existing approaches

The proposed methodology performance is enhanced by comparing the proposed work with various SOTA techniques such as Region of Interest Transfer (RoI-T) [17], Gliding Vertex (GV) [18], Single Shot Alignment Network (SSA-Net) [19], Kullback-Leibler Divergence (KLD) [20], Rotated RetinaNet (RR-Net) [21], Rotated Fast Region-Based Convolutional Neural Network (RFR-CNN) [22], Rotated Fully Conventional One Stage detection (R-FCOS) [23], and the object detection using oriented representation of points (ORP) [24]. As the proposed network is a single-shot object detector, the related techniques considered are also single-stage object detection techniques. As the proposed work is focused on Yolov8.1 obb, the discussion of other techniques is not in our scope, but a brief description is given.

Table 3. Parameters observed from the YOLO vehicle detection experiments

Model	No. of layers	Parameters	GFLOPs	Pre-process	Inference	Post-process	Training
YOLOv5	290	21.6	50.6	0.9	22.6	20	6.12e+6
YOLOv8m	237	26.4	80.8	0.3	21.2	18.7	2667600
YOLOv8n	**187**	30.7	**8.3**	0.3	**15**	17.8	**2376000**
YOLOv8l	287	44.4	168.5	0.3	25	**16.3**	2466000
YOLOv8s	**187**	**11.4**	29.4	**0.2**	16.1	**16.3**	3700800
YOLOv8x	390	69.4	263.9	0.7	25.8	16.6	5522400

The RoI-T performs better on low-resolution images and has low accuracy on high-resolution images, and the self-attention mechanism adds computational complexity, which shall be overcome by the YOLOv8 version through the advanced head design. Whereas the GV provides the target's precise location as stated in the findings, however the methodology is sensitive to noisy objects and works less efficiently for longer objects. The anchor design technique and robust backbone network with Squeeze-and-Excitation Modules (SAEM) in the proposed technique help in handling the occlusion of vehicles. The KLD technique uses the difference between two or more distributions in the bounding box alignment, and hence, more loss is generated. On the other hand, the proposed work has a well-designed loss function (e.g., L1, L2, IoU, angle-specific) and is necessary for OBB detection. RR-Net is a restructured version of RetinaNet to form oriented boxes that cannot deal with anchor boxes and generates novel regression loss, increasing the training and inference time complexity. Nonetheless, the current work handles the arbitrary anchor boxes through the flexible and direct approach of OBB. RFR-CNN is a two-stage detector that involves more complexity, and Yolo versions overcome the complexity with streamlining feature fusion, model depth reduction, and efficient backbone design. Further, R-FCOS uses the pixel-wise technique to build the bounding boxes with the assistance of a prediction head and needs to depend on varying scales. The FPN network of the proposed work represents the instinctive nature to address this drawback. Finally, the ORP technique [24] proposed in the literature, as the name suggests, focuses on the random prediction of oriented points. This challenge shall be addressed with the instinctive nature to address this limitation. In addition, there are limitations on the data quality, data scarcity, background clutter, and sensitivity to noise and occlusions. These drawbacks are overcome with the features of Decoupled Head, Oriented points, and C2f available in the proposed YOLOv8-OBB work. Therefore, the proposed technique overcomes some of the general drawbacks of the existing techniques, such as reduced loss in the training and inference process with the combination of losses. Also, there are various benefits such as higher accuracy (advanced head design and loss function engineering), faster inference speed, better generalization (handling different scales, aspect ratios, and orientations), and simpler pipeline (end-to-end training). Furthermore, the difference in the results is validated in Figure 5 . The graphical description of the metrics evaluation of various approaches is demonstrated in Figure 6.

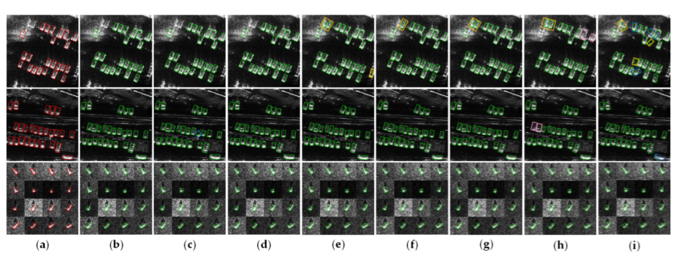


Fig 5. The first column represents (a) Ground truth and Predicted results using (b) ROI-T, (c) GV, (d) SSA-Net, (e) KLD, (f) RR-Net (g) RFR-CNN (h) R-FCOS (i) Proposed method of FARAD, mini-SAR, and MSTAR images respectively

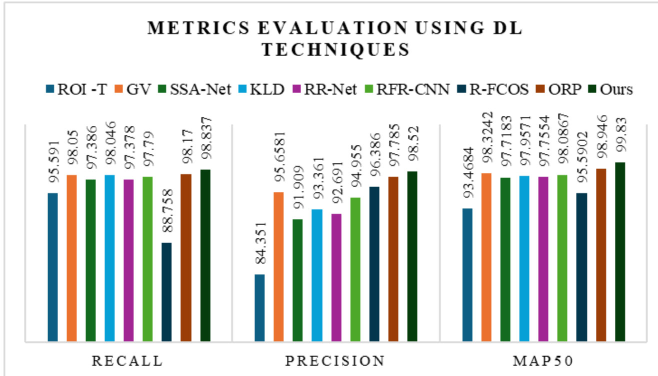


Fig 6. Metrics comparison of various DL approaches

4 Conclusion

In the proposed methodology, the OBB versions of YOLOv8.1 were used for object detection. The unique ideas presented in the proposed work were orientation agnostic detection (arbitrary angle detection), enhanced noise robustness (insensitive to occlusions), anchor-free detection that works for the generalized SAR dataset, and optimized speed. These features were not observed in the existing approaches. The drawbacks of the proposed work were that the layer structure is difficult to understand and has more computational complexity. Additionally, in dense environments, the bounding boxes of vehicles may overlap, leading to the limitation of OBB versions. Furthermore, other challenges are small vehicle detection, proper data annotation, false positives, and data scarcity.

Moreover, the proposed work is open to many future enhancements, such as comparison to deep learning techniques with diverse datasets (ships, buildings, and oil tanks). Also, there is a good research scope to implement the work with less energy hardware platforms that will be useful in applications with fewer resources. Consequently, the employment of ensemble approaches with one-stage detection and two-stage detection techniques achieves better benchmark results. Furthermore, the model refinement should be continued on Yolov8 OBB approaches to achieve higher accuracy.

Acknowledgements

This paper has been presented in “International Conference on Applied Mathematics and Advanced Data Analytics for Industry 5.0 (ICAMADA-2024)” held during 25th to 27th April 2024 at VNR VJIET, in collaboration with APTSMS and CCOE. The authors express their sincere gratitude to the guest editors for their tireless efforts and dedication by means of comments, editing and overseeing the manuscripts to bring in this final form.

The APC is deferred partially by Indian Society for Education and Environment.

The authors would like to extend their gratitude to the authors of the SIVED data for providing vehicle datasets as training, testing, and validation data with proper labels and XML files.

References

1) Han Y, Wang B, Guan T, Tian D, Yang G, Wei W, et al. Research on Road Environmental Sense Method of Intelligent Vehicle Based on Tracking Check. *IEEE Transactions on Intelligent Transportation Systems*. 2023;24:1261–1275. Available from: https://dx.doi.org/10.1109/tits.2022.3183893.
2) Shokri D, Larouche C, Homayouni S. A Comparative Analysis of Multi-Label Deep Learning Classifiers for Real-Time Vehicle Detection to Support Intelligent Transportation Systems. *Smart Cities*. 2023;6(5):2982–3004. Available from: https://dx.doi.org/10.3390/smartcities6050134.
3) Tang L, Tang W, Qu X, Han Y, Wang W, Zhao B. A Scale-Aware Pyramid Network for Multi-Scale Object Detection in SAR Images. *Remote Sensing*. 2022;14(4). Available from: https://dx.doi.org/10.3390/rs14040973.
4) Betti A, Tucci M. YOLO-S: A Lightweight and Accurate YOLO-like Network for Small Target Detection in Aerial Imagery. *Sensors*. 2023;23(4). Available from: https://dx.doi.org/10.3390/s23041865.
5) Li Z, Hou B, Wu Z, Ren B, Yang C. FCOSR: A Simple Anchor-Free Rotated Detector for Aerial Object Detection. *Remote Sensing*. 2023;15(23):549. Available from: https://dx.doi.org/10.3390/rs15235499.
6) Tang H, Gao S, Li S, Wang P, Liu J, Wang S, et al. A Lightweight SAR Image Ship Detection Method Based on Improved Convolution and YOLOv7. *Remote Sensing*. 2024;16(3). Available from: https://dx.doi.org/10.3390/rs16030486.
7) Kılıçkaya FN, Taşyürek M, Öztürk C. Performance evaluation of YOLOv5 and YOLOv8 models in car detection. *Imaging and Radiation Research*. 2024;6(2). Available from: https://dx.doi.org/10.24294/irr.v6i2.5757.

8) Zhao C, Guo D, Shao C, Zhao K, Sun M, Shuai H. SatDetX-YOLO: A More Accurate Method for Vehicle Target Detection in Satellite Remote Sensing Imagery. *IEEE Access*. 2024;12:46024–46041. Available from: https://dx.doi.org/10.1109/access.2024.3382245.
9) Yang Y, Ma C, Li G, Wang Y, Zhang L, Song S. ADERLNet: Adaptive Denoising Enhancement Representation Learning for Low-Latency and High-Accurate Target Detection on SAR Sensors. *IEEE Sensors Journal*. 2024;24(5):6430–6450. Available from: https://dx.doi.org/10.1109/jsen.2023.3348141.
10) Zhao S, Tao R, Jia F. DML-YOLOv8-SAR image object detection algorithm. *Signal, Image, and Video Processing*. 2024;18:6911–6923. Available from: https://link.springer.com/article/10.1007/s11760-024-03361-4.
11) Yasir M, Shanwei L, Mingming X, Jianhua W, Hui S, Nazir S, et al. YOLOv8-BYTE: Ship tracking algorithm using short-time sequence SAR images for disaster response leveraging GeoAI. *International Journal of Applied Earth Observation and Geoinformation*. 2024;128. Available from: https://dx.doi.org/10.1016/j.jag.2024.103771.
12) Song Y, Wang S, Li Q, Mu H, Feng R, Tian T, et al. Vehicle Target Detection Method for Wide-Area SAR Images Based on Coarse-Grained Judgment and Fine-Grained Detection. *Remote Sensing*. 2023;15(13). Available from: https://dx.doi.org/10.3390/rs15133242.
13) Xu Z, Yang Y, Wei Y, Magnagna OL. Improvement of YOLOv8 for Vehicle Small Object Detection Research. *International Conference on Artificial Intelligence, Automation and High Performance Computing*. 2024;2024:84–89. Available from: https://doi.org/10.1145/3690931.3690946.
14) Lin X, Zhang B, Wu F, Wang C, Yang Y, Chen H. SIVED: A SAR Image Dataset for Vehicle Detection Based on Rotatable Bounding Box. *Remote Sensing*. 2023;15(11):282. Available from: https://dx.doi.org/10.3390/rs15112825.
15) Chapman J, Chen B, Tan Z, Calder J, Miller K, Bertozzi AL. Novel batch active learning approach and its application on the synthetic aperture radar datasets. *Algorithms for Synthetic Aperture Radar Imagery XXX*. 2023;12520:13–13. Available from: https://doi.org/10.48550/arXiv.2307.10495.
16) Complex SAR Data. . Available from: https://www.sandia.gov/radar/pathfinder-radar-isr-and-synthetic-aperture-radar-sar-systems/complex-data/.
17) Wu M, Yue H, Wang J, Huang Y, Liu M, Jiang Y, et al. Object detection based on RGC mask R-CNN. *IET Image Processing*. 2020;14(8):1502–1508. Available from: https://dx.doi.org/10.1049/iet-ipr.2019.0057.
18) Chen Z, Luo R, Li J, Du J, Wang C. U-Net Based Road Area Guidance for Crosswalks Detection from Remote Sensing Images. *Canadian Journal of Remote Sensing*. 2021;47(1):83–99. Available from: https://dx.doi.org/10.1080/07038992.2021.1894915.
19) Alom MZ, Yakopcic C, Hasan M, Taha TM, Asari VK. Recurrent residual U-Net for medical image segmentation. *Journal of Medical Imaging*. 2019;6(1). Available from: https://dx.doi.org/10.1117/1.jmi.6.1.014006.
20) Bagheri F, Tarokh M, Ziaratban M. Semantic Segmentation of Lesions from Dermoscopic Images using Yolo-DeepLab Networks. *International Journal of Engineering*. 2021;34(1):458–469.
21) Aly GH, Marey M, El-Sayed SA, Tolba MF. YOLO Based Breast Masses Detection and Classification in Full-Field Digital Mammograms. *Computer Methods and Programs in Biomedicine*. 2021;200. Available from: https://dx.doi.org/10.1016/j.cmpb.2020.105823.
22) Huang H, Zhou H, Yang X, Zhang L, Qi L, Zang AY. Faster R-CNN for marine organisms detection and recognition using data augmentation. *Neurocomputing*. 2019;337:372–384. Available from: https://dx.doi.org/10.1016/j.neucom.2019.01.084.
23) Zhang Y, Yao Y, Wan Y, Liu W, Yang W, Zheng Z, et al. Histogram of the orientation of the weighted phase descriptor for multi-modal remote sensing image matching. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2023;196(1):1–15. Available from: https://dx.doi.org/10.1016/j.isprsjprs.2022.12.018.
24) Li W, Chen Y, Hu K, Zhu J. Oriented reppoints for aerial object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022;p. 1829–1838. Available from: https://openaccess.thecvf.com/content/CVPR2022/papers/Li_Oriented_RepPoints_for_Aerial_Object_Detection_CVPR_2022_paper.pdf.