Journal of Intelligent Systems and Internet of Things
Vol. 16, No. 01, PP. 132-141, 2025
DOI: https://doi.org/10.54216/JISIoT.160111

Implement Intelligent YOLOv8 for Car Crowd Detection and Counting in the Roads

Noor Abdul Khaleq Zghair[1], Rand A. Atta[2], Hussein M. Hasan[2], Asmaa S. Zamil[2], Saja B. Attallah[2]
[1]Computer Engineering Department, University of Technology- Iraq, Baghdad, Iraq
[2]Biomedical Engineering Department, University of Technology- Iraq, Baghdad, Iraq
Emails: Noor.A.Zghair@uotechnology.edu.iq; rand.a.ata@uotechnology.edu.iq; Hussein.m.Hasan@uotechnology.edu.iq; Asma.s.zamil@uotechnology.edu.iq; 060191@uotechnology.edu.iq'

Abstract

Car crowd management refers to the process of efficiently and safely managing the movement and flow of cars in crowded areas, such as parking lots, traffic intersections, event venues, and busy streets. Effective car crowd management is essential to ensure smooth traffic flow, prevent accidents, reduce congestion, and optimize the utilization of available parking spaces. It is a critical aspect of urban planning and traffic management to enhance the overall transportation experience and safety for both drivers and pedestrians. Deep learning methods are used to create an artificial system that is shown in this study. Proposed in detecting cars in streets and traffic intersections, in addition to determining the quantity of cars based on the YOLOv8 algorithm. Where the proposed system was trained on three types of datasets for the purpose of testing the algorithm used to determine the number of cars in each direction of the traffic intersection and then give priority to the most crowded direction with cars and then less and less. Where the system reached a high accuracy in detecting cars, reaching 98%, and through it conclude that the YOLOv8 algorithm used was suitable to be employed in solving the problem of determining the priority of traffic by identifying places of congestion with high accuracy.

Received: November 25, 2024 Revised: January 16, 2025 Accepted: February 17, 2025

Keywords: Car Crowd; Deep Learning; YOLOv8; Traffic Light Priority

1. Introduction

Because of the worldwide economy's accelerated expansion, the standard of living has increased [1]. Urban road traffic congestion is the most prevalent problem due to traffic density and population growth [2]. Congestion at traffic intersections is one of the important problems that must be drawn attention to, and the reason for this is that the traffic lights operate steadily without the use of any smart technology, meaning that the time allocated to all directions is equal in the event that one direction may be empty and the other full. Analysing traffic information is a crucial undertaking in a transportation system that is intelligent (ITS) [3], By providing effective information for the administration and control of traffic, we can ensure efficient operations Approximating the quantity of vehicles in motion during video sequences is a crucial procedure utilized in a variety of applications that aid governments in making decisions regarding the construction of new highways and the expansion of the traffic system, regulating traffic lights, and selecting the most efficient routes. These figures denote the traffic situation, including the intensity of road traffic, lane occupancy, and congestion level [4]. Road congestion mitigation, automated route planning, and early incident detection can all benefit from this type of information. Dedicated sensors often carry out the process of counting vehicles in conventional intelligent transportation systems. Nevertheless, the high installation cost and the uncomplicated structure pose certain limitations for these sensors [5].

Hence, vehicle identification and monitoring are critical traditional for the purpose of forecasting the amount of traffic on the Internet, Deep learning and machine learning algorithms are implemented [6]. Deep learning algorithms, conversely, have replaced conventional approaches to object recognition in light of the substantial progress that has been achieved in these domains. A departure from conventional approaches to object recognition has occurred within the computer vision and deep learning domains, wherein deep learning algorithms have essentially supplanted them. This category includes the detection of objects in real time. Deep learning techniques are superior in terms of effectiveness and accuracy, and they outperform other methods in detecting objects in complicated situations. Nevertheless, the object detection system still faces some of the common detection issues, such as demanding illumination [7]. Therefore, attention turned to the use of effective methods to implement this, traffic surveillance applications have made extensive use of video cameras, and these applications efficiently provide useful information regarding traffic flow. Incorporating modern technology into traffic-surveillance video cameras is another step forward in the field, such as computer graphics, high-end computing capabilities, state-of-the-art cameras, and automatic video analysis [8]. The primary responsibility of a transport monitoring system is vehicle detection, as this paves the path for more efficient use of other aspects. Many efforts are being made to develop automatic vehicle detection and categorization technologies. Image processing is used in conjunction with various algorithms (based on shape, texture, color, size, etc.) to classify vehicles in videos, with the results being sent to a centralized controlling and monitoring system [9]. One of the most useful tools for managing and controlling traffic is counting vehicles. Synchronous vehicle detection and tracking is necessary for vehicle counting. Target identification is just one of many real-world domains where artificial intelligence AI and deep learning (DL) have recently demonstrated promising applications [10].

There have been many advances in the past decade toward efficient vehicle detection and tracking. Gaussian mixed models (GMM) and other conventional vehicle detection algorithms give good results, but they do not work as well as they should when illumination changes or when there is background clutter. Researchers prefer deep learning approaches over more conventional ones because of their built-in feature extraction capacity, as it reduces the number of errors that happen in classification tasks when features are extracted by hand [11]. The primary contributions derived from your paper:

- Deployment of YOLOv8 for Vehicle Detection — Employing the YOLOv8 model to precisely identify and quantify cars in real-time from aerial traffic imagery.
- Traffic Prioritization According to Vehicle Density — creating a system that dynamically priorities traffic flow depending on congestion levels, enhancing traffic management at junctions.
- Utilizing Multiple Datasets – Employing two distinct datasets (Arial-Car-Dataset and Aerial View Car Detection) to improve the model's generalization and robustness.
- Exceptional identification Precision — Attaining up to 98% accuracy in automobile identification, illustrating the efficacy of YOLOv8 in diverse situations.
- Overcoming Common Detection Challenges – Addressing issues such as low image quality, varying lighting conditions, and vehicle occlusion, ensuring reliable performance.
- Comparison with Traditional Methods – Discussing how YOLOv8 outperforms older detection techniques in terms of accuracy and efficiency.

The rest of this paper is organized as follows: Section (2) presents a literature review of existing models alongside the proposed model. Section (3) delineates the characteristics of the proposed system. Section (4) details the results achieved and discussion with performance metrics of models; lastly, the conclusions of the paper are discussed in Section (5).

2. Related Work

In recent years, the scientific community has seen a proliferation of vehicle detection algorithms. Various vehicle detection strategies affect our exploratory investigation. Vehicle detection has emerged as a significant domain of interest in autonomous driving systems, capitalizing on the swift progress in computer vision and deep learning technology. Traffic management videos have implemented algorithms for vehicle detection and monitoring to forecast traffic flow. Therefore, counting vehicles is a prevalent method of managing traffic flow that employs object detection and monitoring. The following section examines recent developments in the region of the vehicle detection in addition to monitoring, object detection, and object tracking.

In 2020 [10], this article focuses on the creation of an approach that makes use of photographs taken by unmanned aerial vehicles (UAVs) in order to perform the rapid and automatic identification of motion vehicles. The gathering of a vehicle dataset for the purpose of target recognition is done before moving further with the process. Because of this, a unique YOLOv3 framework for vehicle identification is proposed. This framework possesses the following characteristics: In the image captured by the unmanned aerial vehicle (UAV), the vehicle targets are concentrated and relatively small. The average accuracy (AP) is now 97.49 percent, which is a 5.48 percent improvement over the previous level of 92.01 percent, which is still a relatively quick rate of processing for the YOLO network. The proposed framework is subsequently evaluated utilizing the COWC, VEDAI, and CAR datasets.

In 2021[11], a method is proposed for detecting monitoring vehicles when they are on the road in different weather situations. The method employs CNN uses multiple region proposal networks (RPNs) to accelerate the object detection process in computer vision. In this specific field of study, the investigation of quicker R-CNN using multiple region proposal networks (RPNs) is largely uncharted territory. The traditional Faster R-CNN algorithm produces areas of interest (ROIs) by employing a solitary fixed-size Region Proposal Network (RPN). This, in turn, compromises its ability to identify vehicles of varying dimensions. This study, on the other hand, presents an end-to-end methodology for Vehicle detection on the road. Multiple RPNs of diverse sizes generate ROIs, thereby facilitating Vehicle detection over a range of sizes. The proposed approach introduces an innovation by integrating numerous RPNs of diverse dimensions into the traditional Faster R-CNN is a computer vision algorithm. The autos have been acknowledged. In diverse elements situations. We evaluated the proposed system's performance using three distinct public datasets: DAWN, CDNet 2014, and LISA are the datasets in question. In terms of average precision rates, the system obtained 89.48% and 91.20% respectively, and 95.16% on the relevant datasets.

In 2021 [12], the work presents a cutting-edge implementation of the technology of the Mask R-CNN programming language. This approach utilizes transfer learning to identify vehicles by utilizing instance-wise segmentation; it is possible to simultaneously generate a bounding box and an object mask. In order to ensure accurate and immaculate identification of vehicles, autonomous systems employ a segmentation-based approach. The model also exhibits satisfactory performance when applied to occluded and small-sized objects. This research is conducted utilizing a Google Colab-provided online GPU and cloud services in conjunction with the TensorFlow and Keras frameworks. A mAP of 90.27% and a mAR of 92.38% are attained through the utilization of a composite set of benchmark datasets.

In 2022 [13], In this article, an effective real-time method for identifying and quantifying moving vehicles is described, utilizing YOLOv2 and point motion analysis features. In order to produce precise counting results, the effort is predicated regarding synchronous surveillance and detection of vehicle features. The suggested approach operates in two distinct phases: vehicle detection occurs first, followed by the enumeration of moving vehicles. They are exploring various convolutional neural networks, including regression networks and Networks that classify images on a pixel-by-pixel basis, to improve their detection and counting abilities. They used the cutting-edge YOLOv2 deep learning object in its current state identification method for initial object detection. Subsequently, they enhanced the findings by applying The K-means clustering algorithm with the KLT (Kanade-Lucas-Tomasi) tracker. Then present an effective method that utilizes the temporal information of tracking and detection framesets that are separated by feature points to accurately tally every label on the vehicle by its consistent trajectory. The experimental outcomes pertaining to twelve arduous videos have demonstrated that, on average, the suggested methodology surpasses currently available approaches. Moreover, the YOLOv2-based suggested methodology improves the mean execution time for the twelve sequences by 93.4%. The suggested method outperforms the BS-CNN (background subtraction-based CNN) method, which achieves a frame rate of 0.19 frames per second, as well as F Rapid Convolutional Neural Network is abbreviated as F R-CNN., which gets 18.7 frames per second.

In 2022 [14], One way to fix an adaptive clipping approach that use identification of objects using the You Only Look Once (YOLO) algorithm system as its foundation technique should be employed throughout the data preparation and detection phase in order to compensate for the loss of features that is brought about through the process of compressing high-resolution photographs while the normalization phase is continuing. The process begins by augmenting a high-resolution training dataset using an adaptive cropping algorithm. A subsequent training set is produced in order to maintain the complex features that the object detection network has acquired must acquire knowledge of. During network detection, the adaptive clipping algorithm segments the image and uses position mapping to combine the coordinates correspond to the detected results. The segmented detection outcomes, situated beside the overall detection outcomes, make up the output. Experiments are conducted using the enhanced YOLO algorithm to gauge its performance in comparison to the original algorithm in detecting vehicles in the test set. The experimental findings indicate that Compared to the initial YOLO object detection system, the mAP@0.5 increases from 47.9% to 89.6%. This represents a significant increase, the recall increases the precision goes from 79.5% to 91.9%, while the percentage of accuracy goes from 44.2% to 82.5%. The process of putting the algorithm into action used is the adaptive algorithm for clipping. Significantly enhances the performance concerning the conventional object detection algorithm when it comes to vehicle detection.

In 2024 [15], the findings of this study imply that an enhanced feature extraction network may be developed by adding an efficient long-range aggregation network for vehicle detection (ELAN-VD) onto the backbone layer. This would allow for the creation of a more advanced network. Increasing the accuracy with which small vehicles may be identified in aerial imagery is the purpose of an architectural enhancement that has been proposed for the YOLOv7-tiny model. Furthermore, in order to enhance performance, the image dimensions produced by the subsequent two predictions boxes are enlarged. The findings of this investigation suggest that the mean average precision (mAP) generated by the suggested approach is 57.94%, which exceeds the mAP of the traditional YOLOv7-tiny.

Through previous research, notice many different experiments using different techniques for the purpose of contributing to the system's increased precision in detecting vehicles, and the decision to detect vehicles may be affected depending on Many circumstances, including the blurring of the captured images, the number of vehicles in one image, or small detected objects, in addition, the results from the previous study can be summed as shown in Table 1.

Table 1: Summary of previous works

Ref.	Year	Dataset Used	Technique	Accuracy
[10]	2020	COWC VEDAI CAR	YOLOv2	97.49%
[11]	2021	DAWN CDNet 2014 LISA	Faster R-CNN	89.48% 91.20% 95.16%
[12]	2021	KITTI Dataset-BDD 100K	Mask R-CNN	90.27%
[13]	2022	CDNet 2014	YOLOv2	93.4%
[14]	2022	MS COCO	Improved YOLO	91.9% 82.5% 89.6%
[15]	2024	Images capture by UAVs	YOLOv7	57.94%

3. Materials and Methods

In the proposed system, two types of datasets were relied upon for testing the proposed system. Two datasets downloaded from the Internet were used. Downloaded dataset-containing images with top view. The datasets are as a following.

3.1 Arial-Car-Dataset

This dataset contains 160 top view images for traffic intersection, each images contain different numbers and types of cars. This dataset was obtained from the Internet by clicking on the link (https://github.com/jekhor/aerial-cars-dataset). Fig. 1 illustrate samples of Arial-car-dataset.

Figure 1. Samples of Arial-Car-Dataset

3.2 Aerial View Car Detection

This dataset contains 300 photos of parking places. Fig. 2 illustrate samples of second dataset. This dataset has been downloaded from the link that is provided below (https://www.kaggle.com/datasets/braunge/aerial-view-car detectionforyolov5?resource=download).

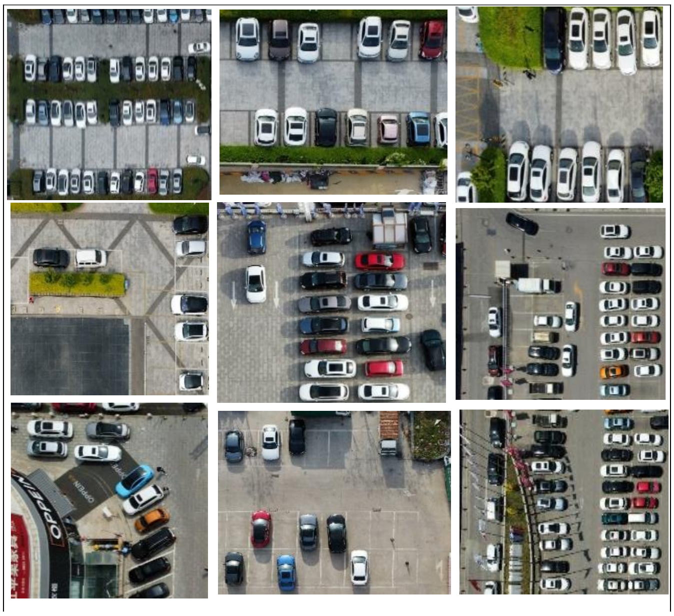

Figure 2. Samples of Arial View Car Detection

The YOLOv8 algorithm is one of the modern algorithms that have proven its high accuracy in segmenting objects by training it on a number of images of the object to be revealed. The YOLOv8 model consists of three main parts Fig. 3:

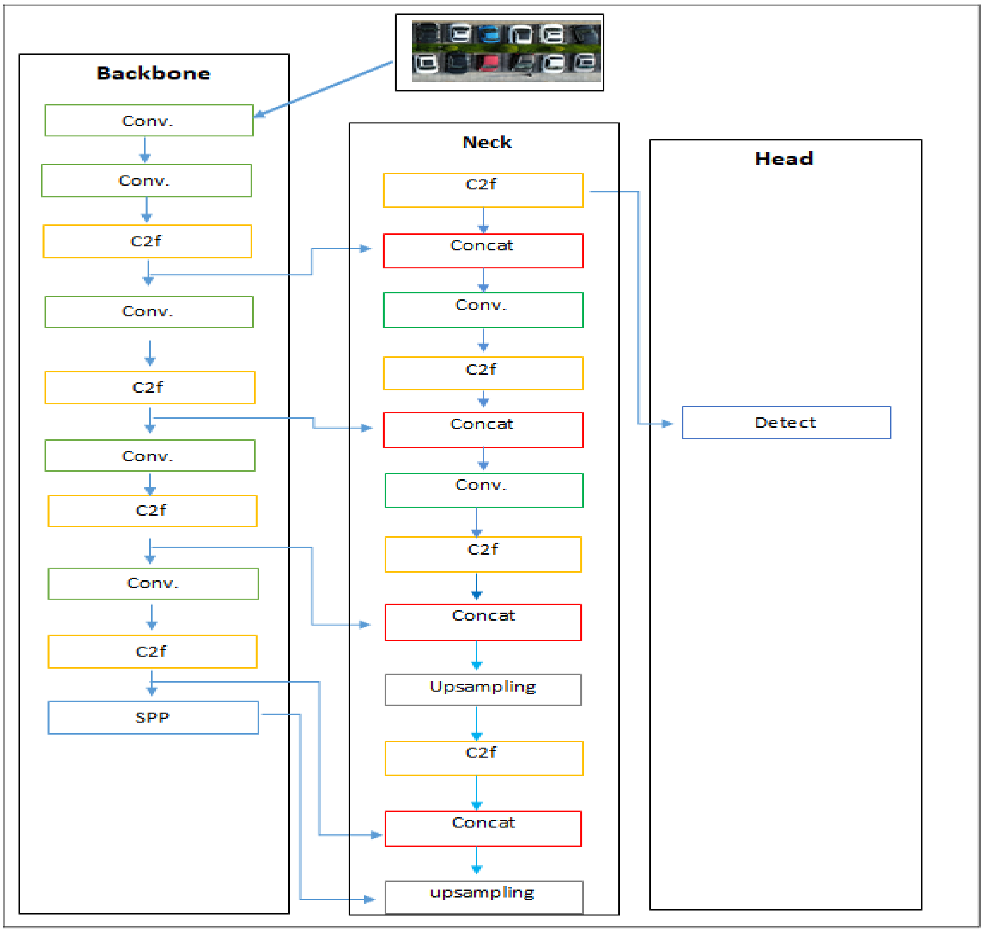

Figure 3. YOLOv8 segmentation model

a. Backbone: this part consists of 10 modules.
b. Neck: this part consists of 12 modules.
c. Head: this part consists of 1 module.

Where the YOLOv8 model consists of 23 modules and the specifications of each module have been accurately determined to suit the process of car segmentation. Table 2 illustrate specification of YOLOv8 segmentation model. YOLOv8 segmentation model consists of 331 layers and 27240227 parameters.

Table 2: YOLOv8 segmentation model specification

Part	**Params**	**Module Name**	**Arguments**
Backbone	1392	Conv.	[3, 48, 3, 2]
	41664	Conv.	[48, 96, 3, 2]
	111360	C2f	[96, 96, 2, True]
	166272	Conv.	[96, 192, 3, 2]
	813312	C2f	[192, 192, 4, True]
	664320	Conv.	[192, 384, 3, 2]
	3248640	C2f	[384, 384, 4, True]
	1991808	Conv.	[384, 576, 3, 2]
	3985920	C2f	[576, 576, 2, True]
	831168	SPPF	[576, 576, 5]
Neck	0	Upsample	[None, 2, 'nearest']
	0	Concat	[1]
	1993728	C2f	[960, 384, 2]
	0	Upsample	[None, 2, 'nearest']
	0	Concat	[1]
	51763	C2f	[576, 192, 2]
	332160	Conv.	[192, 192, 3, 2]
	0	Concat	[1]
	1846272	C2f	[576, 384, 2]
	1327872	Conv.	[384, 384, 3, 2]
	0	Concat	[1]
	4207104	C2f	[960, 576, 2]
Head	5159603	Segment	[1, 32, 192, [192, 384, 576]]

YOLOv8 model start with 2 convolution layers then C2f module, where C2f module used in YOLOv8 instead of C2, this enables the model to acquire Additional Extensive Data on the gradient flow. Connected ConvModules and n DarknetBottleNecks are the components that make up the C3 module, whereas the C2f module is composed of two ConvModules and n DarknetBottleNecks. Moreover, YOLOv8 integrates the Spatial Pyramid Pooling Fusion (SPPF) module, a feature that YOLOv5 and other architectures also share. The ability of the module to extract contextual information from images representing a variety of dimensions significantly improves the model's power to generalize that knowledge. Using a picture as an input, for contributing to the system has increased precision. Once the neck merges these feature maps through path aggregation blocks, the head uses these features to segment vehicles from the testing image. This study used datasets of automobiles, dividing them into two categories: 80% of the data was designated for use in training, while the 40% was designated for testing reasons.

4. Results and Discussion

The YOLOv8 algorithm is one of the algorithms that have a high quality in detecting cars of different shapes and sizes, as may notice that dealing with cars that look small while taking pictures may be difficult for some algorithms, and this does not apply to the YOLOv8 algorithm, as it has been noted that it is not significantly affected in the following cases, which are the lack of image quality, different lighting, and also the large number of cars in one shot. The YOLOv8 algorithm was tested on two datasets; it turned out to be highly accurate in detecting cars. Figure 4 illustrate results of training YOLOv8 on Arial-Car-Dataset, Figure 5 illustrate results of training YOLOv8 on Arial View Car Detection.

Several measures have been used by which system performance can be evaluated, including precision, recall, and mean average precision (mAP). These measures depend on confusion matrix, when the ground truth and the predictions generated by the neural networks are combined, the samples can be categorized as true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). These classifications are based on the combination of the ground truth and the forecasts. Based on the findings of the neural networks, these classifications have been developed.

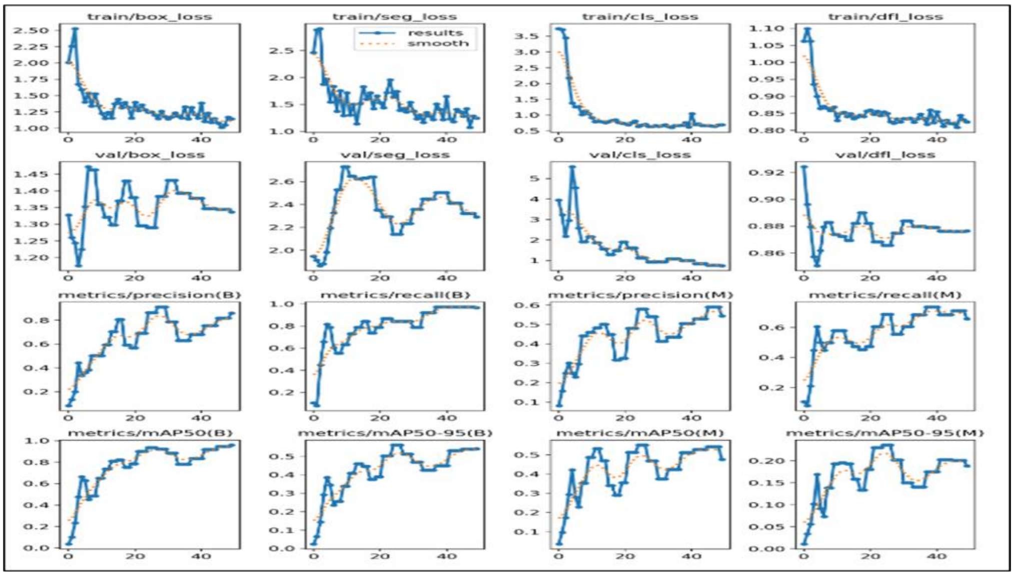

Figure 4. Results of YOLOv8 training on Arial-Car-Dataset.

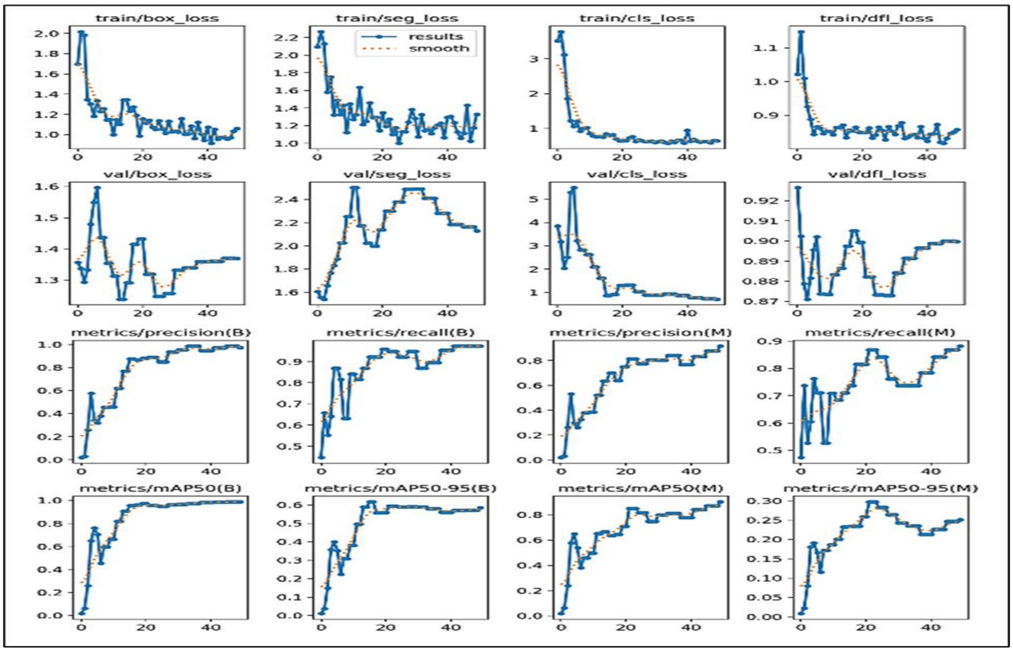

Figure 5. Results of YOLOv8 training on Arial View Car Detection

This is the formula of computing these measures is in Equations (1), (2), and (3).

$$Precision = \frac{TP}{TP + FP} \times 100\% \qquad (1)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \qquad (2)$$

$$mAP = \frac{1}{n}\sum_{k=1}^{n} J(Precision, Recall)_k \qquad (3)$$

Where n is categories number, k is current category number, J(Precision,Recall) is the average precision function. The YOLOv8 algorithm was trained within 50 epochs, In Table 3, you will get a summary of the findings that were obtained from the training.

Table 3: Results of training YOLOv8 on two datasets

Measure	**Value of Arial-Car-Dataset**	**Value of Arial View Car Detection**
train/box_loss	1.0604	1.1426
train/seg_loss	1.3299	1.2434
train/cls_loss	0.64798	0.69004
train/dfl_loss	0.8578	0.82416
metrics/precision(B)	0.97235	0.85954
metrics/recall(B)	0.97368	0.96632
metrics/mAP50(B)	0.9881	0.9794
metrics/mAP50-95(B)	0.58505	0.57106
metrics/precision(M)	0.91788	0.54453
metrics/recall(M)	0.88264	0.65789
metrics/mAP50(M)	0.90205	0.47606
metrics/mAP50-95(M)	0.25142	0.18845
val/box_loss	1.3697	1.3364
val/seg_loss	2.1292	2.2916
val/cls_loss	0.71592	0.7437
val/dfl_loss	0.89966	0.87637

A network uses the mAP performance metric to assess its object detection capabilities. This refers to the region that lies beneath the Precision-Recall curve of the network when the threshold for the detection ratio is set to mAP@0.5. The accuracy of the system reached 98.81% in the Arial-Car-Dataset and 97.94% in the Arial View Car Detection. The system was tested on 64 images from the Arial-Car-Dataset and 120 images from the Arial View Car Detection dataset, and the Fig. 6 shows the results of detecting cars in each image.

Figure 6. Test Samples of cars detection

5. Conclusion

This paper succeeded in presenting the implementation of the YOLOv8 algorithm to detect cars in various conditions, which includes images of any resolution, the number of different cars and different car sizes, and it was found that the algorithm has high accuracy in detecting cars as well as their speed. The accuracy of the algorithm after applying it to two datasets reached approximately 98%, and this accuracy is excellent. The proposed system can be applied in various fields to reduce vehicle traffic congestion by counting cars at traffic intersections in all directions and giving preference to traffic in the most crowded direction with cars.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] D. K. Dewangan and S. P. Sahu, "RCNet: road classification convolutional neural networks for intelligent vehicle system," *Intelligent Service Robotics*, vol. 14, pp. 199–214, 2021.

[2] S. Yassine and A. Stanulov, "A comparative analysis of machine learning algorithms for the purpose of predicting Norwegian air passenger traffic," *Int. J. Math., Stat. Comput. Sci.*, vol. 2, pp. 28–43, 2024. doi: 10.59543/ijmscs.v2i.7851.

[3] M. A. B. Zuraimi and F. H. K. Zaman, "Vehicle detection and tracking using YOLO and DeepSORT," in *Proc. 2021 IEEE 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, 2021.

[4] W. Farag, "A comprehensive vehicle-detection-and-tracking technique for autonomous driving," *Int. J. Comput. Digit. Syst.*, vol. 9, no. 4, pp. 567–580, 2020.

[5] J. Azimjonov and A. Özmen, "A real-time vehicle detection and a novel vehicle tracking system for estimating and monitoring traffic flow on highways," *Adv. Eng. Informatics*, vol. 50, no. 11, pp. 101393, 2021.

[6] M. Contreras and E. Gamess, "Real-time counting of vehicles stopped at a traffic light using vehicular network technology," *IEEE Access*, vol. 8, pp. 135244–135263, 2020.

[7] V. Mandal and Y. Adu-Gyamfi, "Object detection and tracking algorithms for vehicle counting: a comparative analysis," *J. Big Data Anal. Transp.*, vol. 2, pp. 251–261, 2020.

[8] A. Appathurai, R. Sundarasekar, C. Raja, E. J. Alex, C. A. Palagan, and A. Nithya, "An efficient optimal neural network-based moving vehicle detection in traffic video surveillance system," *Circuits Syst. Signal Process.*, vol. 39, pp. 734–756, 2020.

[9] N. Balamuralidhar, S. Tilon, and F. Nex, "MultEYE: Monitoring system for real-time vehicle detection, tracking, and speed estimation from UAV imagery on edge-computing platforms," *Remote Sensing*, vol. 13, no. 4, pp. 573, 2021.

[10] X. Luo, X. Tian, H. Zhang, W. Hou, G. Leng, W. Xu, H. Jia, X. He, M. Wang, and J. Zhang, "Fast automatic vehicle detection in UAV images using convolutional neural networks," *Remote Sensing*, vol. 12, no. 12, pp. 1994, 2020.

[11] R. Ghosh, "On-road vehicle detection in varying weather conditions using faster R-CNN with several region proposal networks," *Multimedia Tools Appl.*, vol. 80, no. 17, pp. 25985–25999, 2021.

[12] A. Ojha, S. P. Sahu, and D. K. Dewangan, "Vehicle detection through instance segmentation using mask R-CNN for intelligent vehicle system," in *Proc. 2021 5th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, IEEE, 2021.

[13] A. Gomaa, T. Minematsu, M. M. Abdelwahab, M. Abo-Zahhad, R. Taniguchi, "Faster CNN-based vehicle detection and counting strategy for fixed camera scenes," *Multimedia Tools Appl.*, vol. 81, no. 18, pp. 25443–25471, 2022.

[14] Z. Chen, L. Cao, and Q. Wang, "YOLOv5-based vehicle detection method for high-resolution UAV images," *Mobile Inf. Syst.*, vol. 2022, no. 6, 2022.

[15] J. L. Andika, A. S. M. Khairuddin, H. Ramiah, and J. Kanesan, "Improved feature extraction network in lightweight YOLOv7 model for real-time vehicle detection on low-cost hardware," *J. Real-Time Image Process.*, vol. 21, no. 3, pp. 1–12, 2024.