Journal of Information Systems Engineering and Management
2025, 10(18s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Real-Time Vehicle Detection and Classification Using Deep Learning based Approach

Riddhi Mehta[1], Dr. Ankit Shah[2]
[1]*Research Scholar, Faculty of Engineering and Technology, Sankalchand Patel University, Visnagar, Gujarat, India*
[2]*Professor, Adani Institute of Digital Technology Management, Gandhinagar, Gujarat, India*

ARTICLE INFO

Received: 20 Dec 2024

Revised: 29 Jan 2025

Accepted: 14 Feb 2025

ABSTRACT

Computer vision is a key component of many technologies, including Automation in industry, robotics, character recognition, human-machine interfaces, text analytics, and motion detection technologies. One significant area within this field is the identification of moving objects, crucial for applications like security cameras and intelligent transportation systems. This paper explores vehicle detection and classification using YOLO base algorithm, focusing on YOLOv5 due to its efficiency and speed. The study evaluates the proposed method using two datasets: Highway Traffic Videos and Vehicle Detection Image Dataset. The YOLOv5 model demonstrated high accuracy in detecting and classifying vehicles, with Confidence scores for vehicles, specifically cars and trucks, fall within the range of 0.41 to 0.86. Results indicate that YOLOv5 achieves effective vehicle detection in both low and high traffic scenarios, addressing challenges such as occlusion and low-resolution footage. The proposed model significantly enhances real-time traffic monitoring and management, highlighting its potential for intelligent transportation systems. The findings underscore YOLOv5's applicability in improving vehicle detection accuracy, performance and efficiency.

Keywords: Vehicle Detection, YOLO, Machine Learning.

INTRODUCTION

Technologies including Automation in industry, robotics, character recognition, human machine interfaces, text analytics and motion detection technologies heavily rely on computer vision as a field of study. A particularly fascinating field of study within computer visionand image processing applications is the identification of moving objects. For this reason, recognizing a moving object is essential for many different applications, such as security cameras. The intelligent transportation system grows with the population. Significance and application are therefore required [1].

Vehicle detection is the foundation of the entire process, and tracking and detection of vehicles within the transportation system are critical to the realization of an intelligent transportation system. The video image processing-based detection method acquires video images by means of camera capture and employs computer vision and image processing techniques to process them [2]. Vehicle detection and tracking represent a vital technology for vehicular networks that are enabled by cognitive radio, which gathers road traffic videos and uses image processing techniques to direct traffic operations [3]. The implementation of an effective detection and tracking model for vehicles has demonstrated significant benefits in the real time of intelligent traffic management. Its benefits include immediate, cost-effective, and high-efficiency.

There is still a great deal of room for advancement in this field, despite its many applications and extensive research. In practice, the existing methods perform poorly and have imperfect accuracy. One of the main causes is the abundance of barriers in this field. Building an automated vehicle classification system presents a number of challenges include a lot of objects, scale variation, occlusion, and changing lighting [4]. Dense traffic has always been a problem on real-world urban roads. When a model is trained to detect and classify vehicles in such a crowded and obscured scenario, the problem becomes more significant. The most significant obstacle that has to be overcome is occlusion. Severe vehicle congestion in dense traffic leads to serious occlusion issues. The image quality from the security footage presents another difficulty. Obtaining high resolution surveillance data is not always feasible. The majority of the time, surveillance cameras capture low-quality footage, which makes it challenging to detect and classify such data [5]. Therefore, there are two issues with real-world surveillance footage. These surveillance cameras record scenes of extremely clogged traffic in the first place, and they also record video of low quality. Real-world surveillance footage has these two traits, which make detection and classification extremely challenging.

In order to investigate the viability of YOLO-based techniques, we focus on the recognition and categorization of cars in photos using deep learning techniques. The YOLO algorithm family is a first-order object detection technique that integrates different object localization using an anchor box. The YOLO family of algorithms has been released in seven versions thus far. The YOLOv3 represents a significant advancement in the YOLO algorithm family's speed and performance. We selected the YOLOv5 detection model because, in comparison to its predecessors in the model family, it has a smaller architecture and significantly faster detection speed. The YOLOv5 algorithm stands out as an excellent choice for vehicle detection, as researchers across various fields have recently enhanced the original YOLOv5 model to better suit the specific characteristics of their detection targets [6-8].

OBJECTIVES

The surveillance system will serve as a vital component in improving the efficiency and effectiveness of traffic monitoring. By continuously monitoring traffic conditions, the system can help in identifying and preventing frequent traffic jams, ensuring smoother flow of vehicles. It will also play a significant role in detecting traffic rule violations, such as speeding, running red lights, and illegal parking, leading to better enforcement and adherence to traffic laws.

Moreover, the surveillance system will contribute to the reduction of fatal road accidents by providing real-time data that can help authorities respond swiftly to accidents, hazardous driving behaviors, or unsafe road conditions.

In addition, the system will assist in combating vehicle thefts by tracking stolen vehicles and alerting law enforcement agencies in real time. It will also help identify vehicles involved in hit-and-run incidents, aiding in the timely apprehension of offenders and improving public safety. Ultimately, this comprehensive surveillance approach will lead to safer and more organized road networks.

METHODS

This section evaluates the effectiveness of the proposed YOLO v5 for Vehicle Detection and classification in experiments. First, the dataset descriptions are presented. Then, the algorithm explained with network architecture. Finally, result of vehicle detection and classification is presented.

Dataset

For our experiment, we are used two Datasets: Highway Traffic Videos Dataset and Vehicle Detection Image Dataset. The Highway Traffic Videos Dataset was used to collect the highway traffic videos for the presented research model. The three categories of traffic in the video: light, medium and heavy were manually assigned, and they represent, respectively, free-flowing, reduced-speed, and stopped or extremely slow-moving traffic. In addition, vehicle detection image dataset was used class "Vehicle" which includes a broad range of vehicles like cars, trucks, and buses. It contains 626 images. The dataset is standardized by resizing each image to 640 x 640 pixels, which is the same resolution across the board. The training data, which consists of 536 images, underwent augmentations to strengthen the model's capacity for generalization. To maintain the integrity of the performance evaluation, the validation set, which consists of ninety images, has not been expanded.

YOLO v5 Algorithm

The foremost algorithm in the realm of single-stage object detection is referred to as YOLO [9]. This approach integrates the bounding box and classification into a single regression problem, thereby removing the need for candidate box extraction found in two-stage algorithms. The following is how the YOLO algorithm works: First, $S \times S$ meshes are created from the image. The task assigned to each grid is to forecast the location in the grid center where the actual box will land. The total number of bounding boxes generated by these meshes amounts to $S \times S \times B$.A bounding box is defined by five key parameters: the center point coordinates of the target, the width and height dimensions (x, y, w, h), and the confidence level indicating whether the target is included to determine the category score for every prediction box, multiply the category probability and prediction bounding box confidence. To get the final prediction results, these prediction boxes are filtered using non-maximum suppression (NMS) [10].

YOLO v5 algorithm achieve high accuracy on Highway Traffic Videos Dataset. The YOLO v5 version served as the benchmark network model in this chapter, focusing on the speed and accuracy requirements essential for vehicle detection algorithms in the context of highway monitoring to increase the detection algorithm's accuracy even more, the network model was enhanced for highway scenarios.

YOLO v5 is recognized as the most refined detection network in the YOLO object detection algorithm to increase the detection speed, arithmetic set innovation was used, this is built upon the YOLO v3 and YOLO v4 algorithms. Anchor boxes were a concept that YOLO v5 adopted in order to speed up the R-CNN algorithm, manually chosen anchor boxes were dropped. YOLOv5 distinguishes itself from earlier versions by employing Prefer PyTorch over Darknet. The backbone of this system is based on CSPDarknet53. A Path Aggregation Network (PANet) is integrated into the model's neck, facilitating better information flow. PANet presents an innovative feature pyramid network (FPN) that includes several bottom-up and top-down layers. This architecture enhances the transmission of low-level features throughout the model. Consequently, by enhancing localization in the lower layers, PANet significantly boosts the accuracy of object localization within the model [11]. Additionally, the head architecture in YOLOv5 resembles that of YOLOv4 and YOLOv3, producing 3 distinct feature map outcomes to enable multi-scale prediction. YOLO v5 network architecture shown in fig.1.

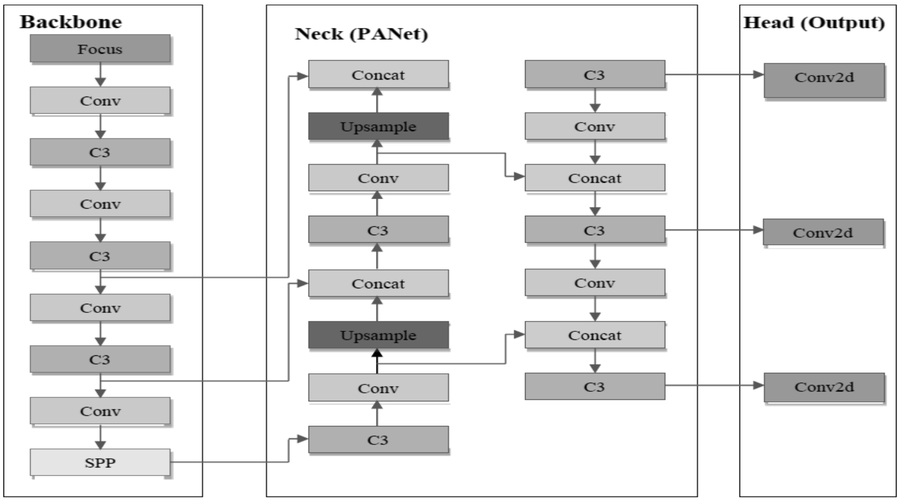

Fig. 1.Network architecture for YOLO v5 [10]

RESULTS

For the experiment, we used two Datasets, which contains vehicle videos as well as images and it is detected by our proposed algorithm. The experimental result shown in fig. 2. and fig. 3. Here we used YOLO v5 algorithm for vehicle detection and classification from both dataset video and image.

Fig. 2. Vehicle Detection and Classification from Highway Traffic Videos Dataset

In Fig. 2., an advanced object detection system was used to detect and label multiple vehicles on a section of highway in the image. All thetype of vehicles, including cars, trucks, and buses, arelabeled and color-coded within a bounding box. The trucks are indicated by the green bounding boxes, which have confidence scores of 0.453, which indicates how certain the model is about its detection. Cars and buses are surrounded by orange boxes; the scores for cars range from 0.731 to 0.834, while the scores for buses identify by 0.430. These confidence scores show how likely it is that the objects are recognized by the system correctly; a score that is closer to 1 indicates a higher level of confidence. For traffic analysis, automated vehicle detection is crucial as it aids in flow monitoring, congestion management, and increased road safety.Vehicle detection image dataset result shown in fig. 4.Vehicles have been identified and labeled by an object detection system on a highway in the image. cars with confidence scores ranging from 0.41 to 0.82 and trucks with scores of 0.86 have been identified. Orange bounding boxes outline the cars, and green bounding boxes indicate the trucks. This detection system is essential to traffic monitoring because it offers precise data that can be used to assess vehicle density and flow.

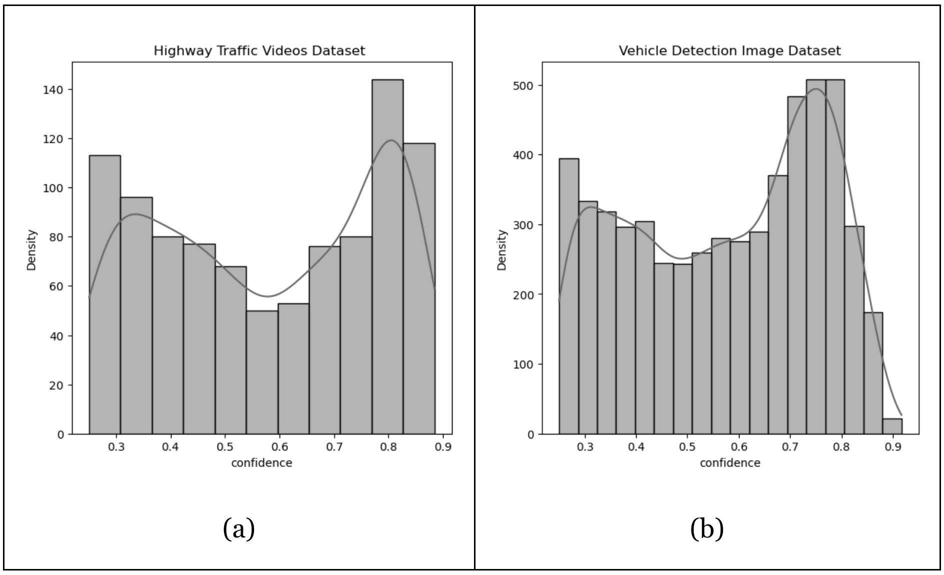

Fig. 3. Confidence Graph: (a) Highway Traffic Videos Dataset, (b) Image Dataset

The graph (a) displays the confidence value distribution in the "Highway Traffic Videos Dataset," where the y-axis denotes frequency and the x-axis represents confidence levels. The histogram shows two prominent peaks around confidence values of 0.3 and 0.8, indicating that these levels occur most frequently, and a noticeable dip around 0.6, indicating that these values occur less frequently. The overlaid density plot highlights this bimodal distribution further, emphasizing the clustering of data points at particular confidence levels.The graph (b) displays the distribution of vehicle detection confidence scores from 0.3 to 0.9, with the y-axis representing density. The highest concentration of detections is between 0.7 and 0.8, peaking at around 500 detections near a confidence score of 0.75.

There are fewer detections at the extremes, with the density dropping significantly above 0.8. This indicates that most detections are made with moderate to high confidence, particularly around 0.75. The confidence graph shown in fig. 3.

Fig. 4. Vehicle Detection and Classification from Image Dataset

Table 1. Vehicle Detection Confidence Scores

Dataset	Vehicle Type	Confidence Score
Highway Traffic Videos	Trucks	0.45
	Cars	0.83
	Buses	0.43
Image Dataset	Cars	0.82
	Trucks	0.86

The "Image Dataset" and "Highway Traffic Videos" Dataset are vehicle detection confidence scores are displayed in the Table 1. For trucks, cars, and buses, the confidence scores in the "Highway Traffic Videos" dataset are 0.45, 0.83, and 0.43, respectively. Vehicles, on the "Image Dataset," score 0.82, while trucks score 0.86. When comparing the datasets, it can be seen that the "Image Dataset" has higher confidence scores for trucks, while the "Highway Traffic Videos" dataset includes bus detection and has a slightly higher score for cars.

DISCUSSION

The study concludes by highlighting the vital role that computer vision plays in technologies like intelligent transportation systems and industrial automation. An effective transportation system must have reliable vehicle detection and classification, and the YOLOv5 algorithm has proven to be highly accurate and effective in this area. Even with these notable advancements, there are still challenges to overcome, particularly in terms of managing occlusion and dealing with poor-quality video in real-world contexts. By employing efficient segmentation and classification methods, the suggested YOLOv5-based model effectively tackles these problems and demonstrates significant gains in vehicle detection precision. The model's ability to handle a range of traffic conditions is validated by the experimental results obtained from various datasets. All things considered, this research makes a significant contribution to the understanding and development of more intelligent and dependable vehicle detection systems.

REFERENCES

[1] Wieczorek, G., Tahir, S. B. U. D., Akhter, I., &Kurek, J. (2023). Vehicle detection and recognition approach in multi-scale traffic monitoring system via graph-based data optimization. Sensors, 23(3), 1731.

[2] Ge, D. Y., Yao, X. F., Xiang, W. J., & Chen, Y. P. (2023). Vehicle detection and tracking based on video image processing in intelligent transportation system. Neural Computing and Applications, 35(3), 2197-2209.

[3] Wang, J., Dong, Y., Zhao, S., & Zhang, Z. (2023). A high-precision vehicle detection and tracking method based on the attention mechanism. Sensors, 23(2), 724.

[4] Ali, S., Jalal, A., Alatiyyah, M. H., Alnowaiser, K., & Park, J. (2023, June). Vehicle detection and tracking in UAV imagery via YOLOv3 and Kalman filter. In CMC.

[5] Basak, S., & Suresh, S. (2024). Vehicle detection and type classification in low resolution congested traffic scenes using image super resolution. Multimedia Tools and Applications, 83(8), 21825-21847.

[6] Farid, A., Hussain, F., Khan, K., Shahzad, M., Khan, U., & Mahmood, Z. (2023). A fast and accurate real-time vehicle detection method using deep learning for unconstrained environments. Applied Sciences, 13(5), 3059.

[7] Li, S., Chen, J., Peng, W., Shi, X., & Bu, W. (2023). A vehicle detection method based on disparity segmentation. Multimedia Tools and Applications, 82(13), 19643-19655.

[8] Basak, S., & Suresh, S. (2024). Vehicle detection and type classification in low resolution congested traffic scenes using image super resolution. Multimedia Tools and Applications, 83(8), 21825-21847.

[9] Mehta, R., & Shah, A. (2024). An Insight into Real Time Vehicle Detection and Classification Methods using ML/DL based Approach. Procedia Computer Science, 235, 598-605.

[10] Zhang, Y., Guo, Z., Wu, J., Tian, Y., Tang, H., & Guo, X. (2022). Real-time vehicle detection based on improved yolo v5. Sustainability, 14(19), 12274.

[11] Olorunshola, O. E., Irhebhude, M. E., &Evwiekpaefe, A. E. (2023). A comparative study of YOLOv5 and YOLOv7 object detection algorithms. Journal of Computing and Social Informatics, 2(1), 1-12.