Research & Reviews: Journal of
Embedded System & Applications
https://journals.stmjournals.com/rrjoesa

ISSN: 2395-6712 (Online)
ISSN: 2321-8533 (Print)
Volume 12, Issue 1, 2024
January—April
DOI (Journal): 10.37591/RRJoESA

Review RRJoESA

Enhancing Road Safety with the Latest Breakthrough: Real-time Vehicle Classification, Counting, and Speed Estimation Using YOLOv8n and Deep SORT Algorithm

K. Jeevitha*

Abstract

The real-time vehicle classification, counting, and speed estimation system based on YOLOv8n is an important tool for monitoring traffic flow on highways. However, because they are distinct objects from their surroundings, it is still difficult to detect them, which has an impact on how accurate vehicle counts are. To tackle this concern, this paper suggests the implementation of a vision-centric system for real-time vehicle monitoring and identification. The approach involves utilizing a You Only Look Once (YOLOv8n)-based Deep SORT model to detect, count, and estimate the speed of vehicles from video clips. Additionally, the integration of a Deep SORT algorithm, based on deep learning, enhances the accuracy of tracking the actual presence of vehicles in video frames predicted by YOLOv8n. Two processes are involved in vehicle counting: first, the recorded video is fed into a deep learning framework based on YOLO to identify, count, and classify each vehicle. Multi-vehicular tracking is adopted using Deep SORT to enhance the tracking performance by associating detected vehicles across frames, ensuring reliable tracking in crowded and dynamic environments.

Keywords: Vehicle counting, object detection, tracking, YOLOv8n, Deep SORT

INTRODUCTION

Real-time vehicle classification, counting, and speed estimation systems are crucial tools that help transportation authorities and law enforcement agencies monitor traffic conditions and optimize traffic flow [1–3]. Government agencies and private institutions may seek to comprehend the vehicular movement in a particular area to enhance infrastructure for the overall convenience of the public. Examples of such initiatives include road expansion projects, optimizing traffic signal timings, and creating parking facilities. Historically, traffic analysis has relied on manual methods, where an individual would be stationed at a location to manually record the number and types of vehicles. While more recently, sensors have been employed to address counting issues, they still fall short in identifying the specific vehicle types.

In the swiftly changing contemporary environment, efficient traffic control and monitoring have gained utmost significance. Real-time vehicle classification, counting, and speed estimation systems are crucial tools that help transportation authorities and law enforcement agencies monitor traffic conditions and optimize traffic flow [1–3]. Despite their limitations, video surveillance systems play a crucial role in monitoring street situations. However, their effectiveness is largely centered around capturing stable recordings through fixed cameras under simple conditions. These methods rely on

***Author for Correspondence**
K. Jeevitha
E-mail: jeevitha.rajsekaran@gmail.com

PG Student, Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Madagadipet, Puducherry, India

Received Date: February 16, 2024
Accepted Date: March 21, 2024
Published Date: April 04, 2024

Citation: K. Jeevitha. Enhancing Road Safety with the Latest Breakthrough: Real-time Vehicle Classification, Counting, and Speed Estimation Using YOLOv8n and Deep SORT Algorithm. Research & Reviews: Journal of Embedded System & Applications. 2024; 12(1): 10–18p.

background deduction to identify the stationary elements of the scene, allowing for mass analysis and tracking of moving vehicles by detecting bounding boxes. However, these techniques often fail to address various obstacles such as shadows, unstable cameras, inaccurate camera placement, and complex backgrounds. Smart cities will greatly benefit from the implementation of this concept [4]. The counting of vehicles in parking lots can also be implemented using the vehicle counting system. In the context of traffic management systems, the utilization of the YOLO (You Only Look Once) algorithm for vehicle counting is of great significance, given its many applications [5, 6]. The current traffic management system has identified several issues that could be addressed through an analysis utilizing the YOLO algorithm and deep sorting for vehicle counting [7].

The primary objective of the proposed work is to identify vehicles or objects and monitor their movements within a specific area of interest. The primary objective is to precisely count the total number of vehicles on the road. To achieve this, an algorithm called You Only Look Once or YOLO is put forward as a means of detecting and counting all the objects observed in the video frames [7–10]. The output of the YOLO algorithm is subsequently inputted into the Deep SORT (simple online and real time) tracker, enabling the tracking of objects found within the video frames.

YOLOv8: Revolutionizing Object Detection

One of the most important and often used tasks in computer vision is object detection. The YOLO algorithm is an extremely potent method for real-time object detection. YOLOv8n, the latest version, has taken this technology to new heights. It provides exceptional accuracy and performance by partitioning images into a grid and concurrently predicting bounding boxes and class probabilities within each grid cell.

Efficient and Accurate Classification

One significant advantage of YOLOv8n is its ability to classify vehicles accurately. It has the capability to recognize and classify a diverse range of vehicles, including cars, trucks, motorcycles, bicycles, and pedestrians, with remarkable precision. This reliable classification provides vital insights into traffic density, vehicle composition, and other crucial parameters for traffic management.

Real-Time Processing

The real-time processing capability of YOLOv8n surpasses traditional methods. With its speed and efficiency, it allows for instant monitoring of multiple vehicles in a scene simultaneously. This enhanced processing speed is particularly beneficial for situations.

Deep SORT Algorithm Tracking and Counting Vehicles

While YOLOv8n excels in accurately detecting and classifying vehicles, the Deep SORT algorithm takes it a step further by introducing vehicle tracking and counting capabilities. Deep SORT combines the power of deep learning and multi-object tracking to create a robust system that can handle complex scenarios efficiently.

Accurate Tracking and ID Maintenance

Deep SORT employs sophisticated techniques to associate the detected vehicles over consecutive frames, ensuring precise tracking. By assigning unique IDs to each vehicle, it becomes possible to monitor individual vehicles' movements, speeds, and behaviors accurately.

Real-Time Speed Estimation

Another significant aspect of the YOLOv8n and Deep SORT based system is real-time speed estimation. By utilizing the tracked vehicles' positions over time, this system can estimate the speed at which each vehicle is traveling with impressive accuracy.

RELATED WORK

The method for determining how congested a route is for intelligent transportation systems (ITS) is the main emphasis of the work by Kejriwal et al. [4]. The primary issues addressed by the current traffic systems are speeding and signal infractions, but they frequently overlook the volume of traffic. The suggested method counts and tracks automobiles using footage from mobile and CCTV cameras in smart cities. A YOLO-based deep learning model is used to identify, count, and classify automobiles in the video. Then, several vehicles are tracked in the video frames using the Deep SORT algorithm. The model has been carefully trained for the conditions of Indian roads. According to the review, the model's categorization and counting accuracy is mediocre. In order to improve infrastructure, this model can assist in determining how congested the roadways are. It plays a crucial role in augmenting the intelligence and efficiency of traffic surveillance in smart cities.

The increased use of autos and the corresponding growth in activity mishaps necessitate strong vehicle recognition [11]. Specifically, it focuses on using deep learning, with an emphasis on the SSD (single shot multi-box locator) computation. Finding a reliable calculation with a high acknowledgment rate is the goal. The SSD computation and the central convolutional neural network (CNN) theory are examined in detail in this research. At that point, it tests the method by determining whether the vehicles are moving or stationary based on a video and images of stationary vehicles. Picture recognition is used to assess the accuracy and practicality of the suggested method, which provides a practical solution for vehicle recognition with a high degree of solidity and precision.

The importance of a YOLOv7-based real-time vehicle monitoring system for highway applications is highlighted in the study by Rouf et al. [12]. The system's primary goal is to reduce traffic congestion, violations, and road incidents by improving traffic management, planning, and accident prevention. It provides real-time traffic flow statistics by recognizing, tracking, and keeping an eye on a variety of vehicle types on urban roads and highways. The system's essential phases include classification, tracking, monitoring, and vehicle detection. A revised version of YOLOv7 is proposed to overcome some of its drawbacks, including low feature extraction and high missed detection rates. The goal of this upgraded system is to provide real-time vehicle tracking, providing information on traffic dynamics and assisting in well-informed traffic management decision-making. The study essentially presents a YOLOv7-based architecture for accurate real-time vehicle identification, tracking, and counting, enhancing our comprehension of traffic patterns and aiding in traffic management decisions.

By integrating digital traffic monitoring technology into an intelligent transportation system, the initiative by Wu et al. [13] seeks to improve urban transportation. Traffic data is collected by elevated surveillance cameras, and various vehicle kinds are identified through the use of YOLO object identification technology. The study focuses on particular urban settings, including T-shaped intersections, cross-type intersections, and straight-line bidirectional roads and uses deep convolutional neural networks and visual processing for object detection training. In order to accurately count the number of vehicles, the system uses object tracking to draw a line in the direction of the pathways taken by the vehicles and tracks the center coordinate as objects cross this line.

The Ministry of Public Works and Public Housing (PUPR) in Indonesia developed a web application utilizing the deep learning model YOLOv4 in order to increase the efficiency of their manual traffic surveys [3]. Based on Bina Marga classifications, this model can recognize and classify cars automatically. Using video data from traffic records, the app generates a car count report. During testing, a small collection of 1000 images were used for training on the Google Colab platform, and the YOLOv4 model, along with the Deep SORT technique and a bespoke dataset, obtained a respectable accuracy.

PROPOSED SYSTEM

In response to the growing significance of highways and the increasing role of intelligent vehicle detection, counting, and speed estimation, this paper tackles the persistent challenge posed by the diverse sizes of vehicles. Precisely identifying vehicles becomes notably difficult when confronted with variations in size. To overcome this obstacle, the paper proposes a vision-based approach for vehicle detection, counting, and speed estimation. By leveraging advanced computer vision techniques, this methodology aims to enhance the accuracy of these crucial tasks, offering a promising solution for the evolving landscape of intelligent transportation systems on highways.

Figure 1 shows the block diagram of proposed method. Deep CNNs have proven highly successful in vehicle object detection by demonstrating remarkable proficiency in learning intricate image features and executing various tasks, including classification and bounding box regression. Detection methods generally belong to two main categories. The first, a two-stage method, involves generating a candidate box for the object using diverse algorithms and subsequently employing a CNN for object classification. In contrast, the second, a one-stage method, simplifies the process by directly converting the object's bounding box positioning challenge into a regression problem, bypassing the need to generate a candidate box. CNNs are capable of addressing complex tasks in vehicle detection through both approaches [10].

This project is aimed at implementing the YOLO-V8-based Deep SORT model for detecting and tracking vehicles in video sequences in real-time. The addition of the Deep SORT algorithm enhances the tracking of actual vehicles identified by YOLOv8 in video frames. The video is segmented into frames, serving as input for YOLOv8 to detect vehicles. Frames that are identified are subsequently examined by the Deep SORT algorithm for the tracking of vehicles. If a vehicle is tracked, Deep SORT overlays a bounding box and increments the tracking counts. The model is trained using COCO (Common Objects in Context) datasets.

Dataset

The COCO dataset comprises images featuring cars, trucks, motorbikes, and people. All 80 images within the dataset are stored in the "img/" folder. In addition, there are three separate folders, named "train/," "val/," and "test/," that are designated for training, validation, and testing purposes.

Preprocessing and Frame Separation

Digital image processing involves altering digital images through the use of computer algorithms. Compared to analog image processing, it has various advantages.

It is within the category of digital signal processing. The capacity to apply a wider variety of algorithms to the input data, increasing the adaptability and versatility of image alteration, is one important benefit.

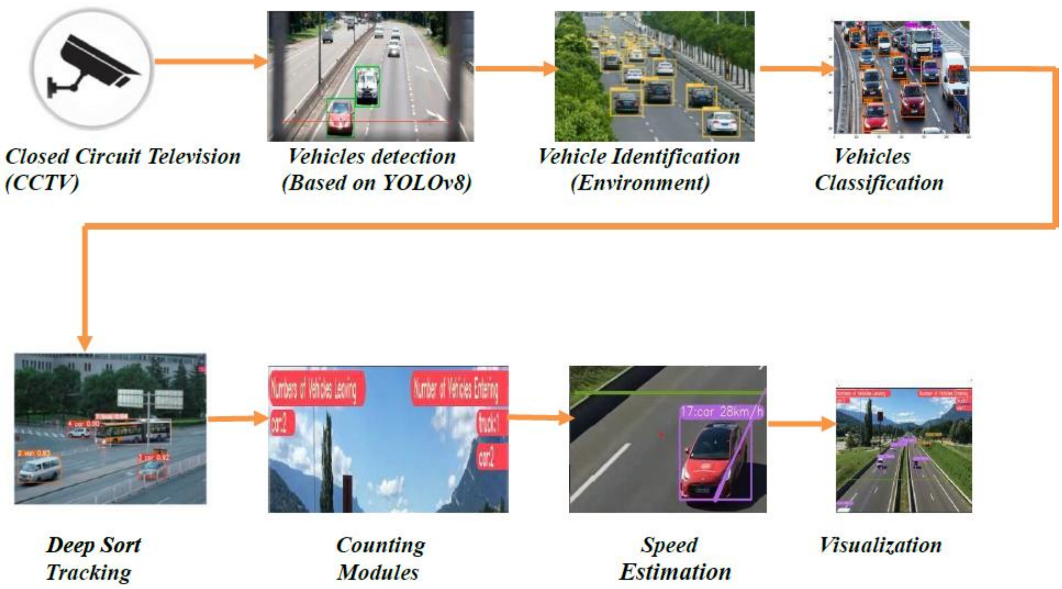

Figure 1. Block diagram of vehicle detection and counting.

Digital image processing seeks to enhance the data (features) within pictures by improving crucial aspects and/or reducing unwanted distortions. This ensures that our artificial intelligence (AI)–computer vision models can effectively operate with enhanced data [10]. For the training and adaptation of the model to new data, it is essential that our photos conform to the network's input size. If adjustments are necessary to align the image size with the network's requirements, resizing or cropping the data to the specified size can be performed.

With the use of randomized augmentation, they can efficiently enhance the quantity of training data. Training networks to be invariant to aberrations in visual data is another benefit of augmentation [10]. An effective strategy to guarantee that a network remains unaffected by the occurrence of rotation in input images is to introduce randomized rotations into the images. Applying a restricted set of augmentations to two-dimensional (2D) images for classification challenges is made easier using an enhanced image data store. For training, prediction, and classification, one option is to utilize an augmented image data store or a resized 4D array. Alternatively, for prediction and classification purposes only, a resized 3D array can be employed.

Resizing image data to fit a network's input size can be done in two different ways. The image's height and width are multiplied by a scaling factor for rescaling. Rescaling modifies the aspect ratio and the spatial extents of the pixels if the scaling factor is not the same in the vertical and horizontal axes.

By cropping an image, you can extract a sub-region while maintaining the spatial extent of each pixel. Images can be cropped from random locations throughout the image or from the center. An image can be described as a 2D array of numerical values (or pixels), with values ranging from 0 to 255. Mathematically, it is represented by the function f (x, y), where x and y denote the horizontal and vertical coordinates, respectively [10].

Resize Image

During the pre-processing stage, resizing a picture is necessary since not all camera captured photos that are given into our AI algorithms are the same size. As a result, all images should have a baseline size as shown in Figure 2.

YOLOv8n

To implement the algorithms, a test video is selected from the datasets and divided into multiple frames. These frames are then processed by YOLOV4 for vehicle detection and Deep SORT for vehicle tracking. The final output video is generated, showcasing both vehicle detection and tracking.

In the realm of transportation research, vehicle tracking is a significant aspect. Deep SORT is a modern tracking algorithm, an extension of the SORT tracking algorithm. Initially designed for the multiple object tracking (MOT) task, the original algorithm prioritizes online and real-time applications. This implies that the tracker associates detected objects solely between previous and current frames.

YOLO, on the other hand, is a CNN-based object detection system. It approaches object detection as a regression problem, mapping image pixels to bounding boxes along with their associated class probabilities. Since it trains on whole images directly, its performance is far superior to other conventional methods of object detection. Twenty-four convolutional layers, two fully linked layers, and a final detection layer make up the 27 CNN layers that make up YOLO.

After dividing the input photographs into N-by-N grid cells, YOLO predicts multiple bounding boxes for each image during processing, hence increasing the likelihood that an item will be discovered. Hence, it is necessary to compute a loss function. In the case of YOLO, the initial step involves calculating the intersection over union (IoU) for each bounding box as shown in Figure 3.

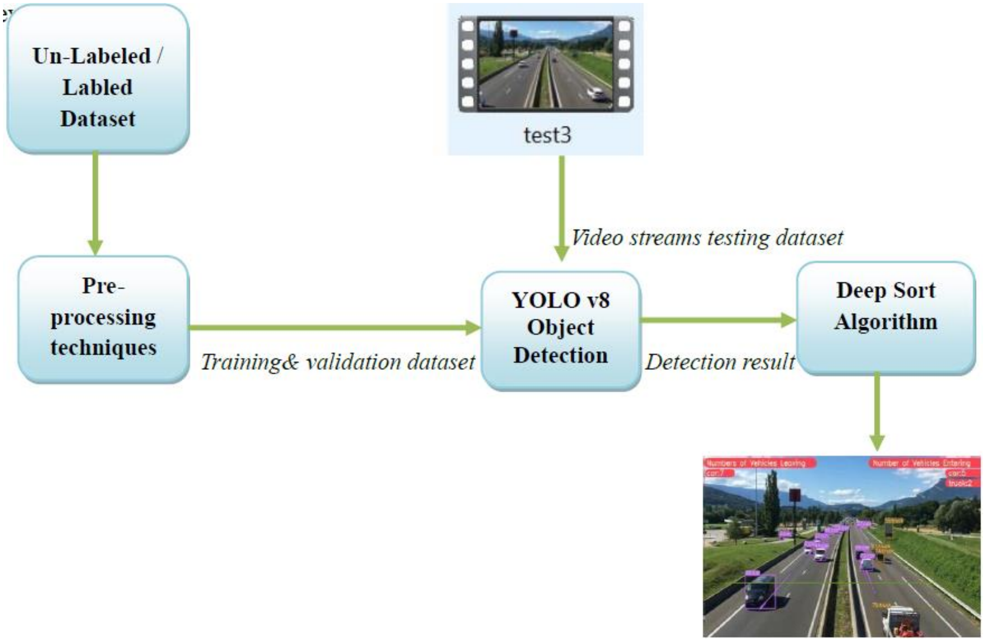

Figure 2. The methodology of the proposed system.

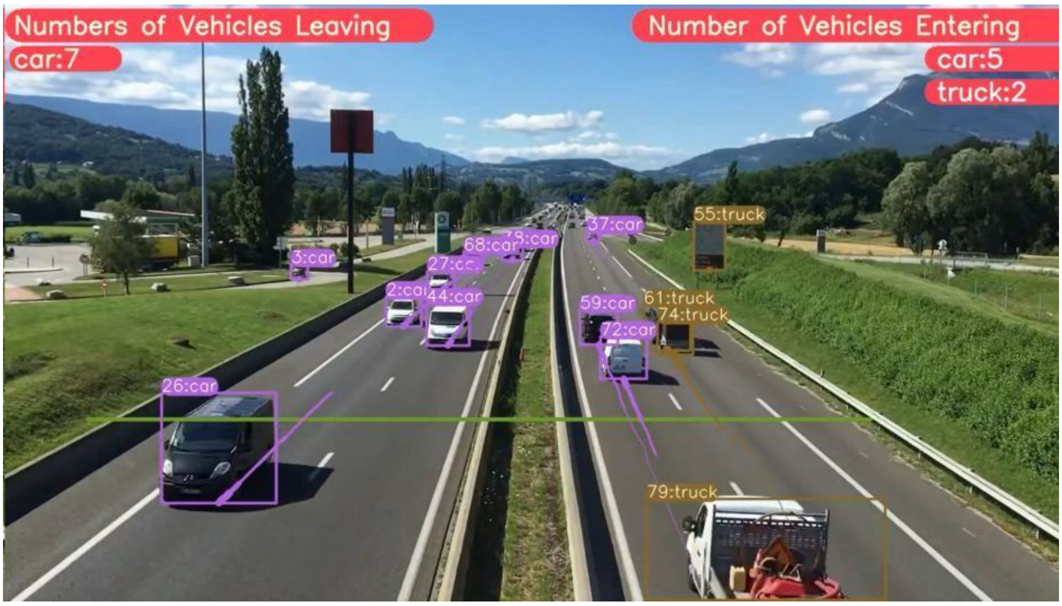

Figure 3. Detection of vehicles using YOLOV8.

Deep SORT

The tracking functionality is executed through the utilization of the Deep SORT algorithm. The improved system uses a convolutional neural network to integrate motion and appearance data, replacing the association metric with an informed metric. Using the detection outputs from the previous step, the algorithm tracks each object that is discovered. The quality of detection findings determines tracking accuracy when using a detection technique [4].

The Deep SORT algorithm is used to implement the tracking function. The updated system replaces the association metric with an informed metric by integrating motion and appearance data using a convolutional neural network. Every object that is found is tracked by the algorithm using the detection outputs from the previous stage. When employing a detection approach, tracking accuracy is determined by the caliber of detection findings.

A number of cutting-edge object detection and tracking algorithms, such as SORT and Deep SORT, were used to identify and monitor various vehicle classes within their area of interest. However, it has been reported that the trackers' performance in predicting vehicles [10].

Speed Estimation

A real-time speed estimation system leveraging YOLOv8 and Deep SORT combines the speed and accuracy of YOLO's vehicles detection with the robust vehicles tracking capabilities of Deep SORT. In this integrated framework, YOLOv8 is employed to detect and identify vehicles in real-time video frames, while Deep SORT is utilized for continuous tracking of the identified vehicles across consecutive frames. By measuring the distance traveled by each vehicle and recording the corresponding time intervals between frames, the system calculates real-time speed estimates. This approach provides a comprehensive solution for accurate and efficient speed estimation in dynamic environments, making it applicable for traffic monitoring, intelligent transportation systems, and other scenarios requiring precise real-time speed data as Figures 4 and 5.

RESULTS AND DISCUSSION

The YOLOv8n and Deep SORT algorithms have been used for vehicle classification and counting using the COCO dataset and video clips. Based on a predetermined boundary line, the system counts the number of vehicles entering and exiting the frame and classifies them using speed estimate.

```
line = [(100, 500), (1050, 500)]
speed_line_queue = {}
1 usage  new *
def estimatespeed(Location1, Location2):
    #Euclidean Distance Formula
    d_pixel = math.sqrt(math.pow(Location2[0] - Location1[0], y: 2) + math.pow(Location2[1] - Location1[1], y: 2))
    # defining thr pixels per meter
    ppm = 8
    d_meters = d_pixel/ppm
    time_constant = 15*3.6
    #distance = speed/time
    speed = d_meters * time_constant

    return int(speed)
```

Figure 4. Speed estimation code.

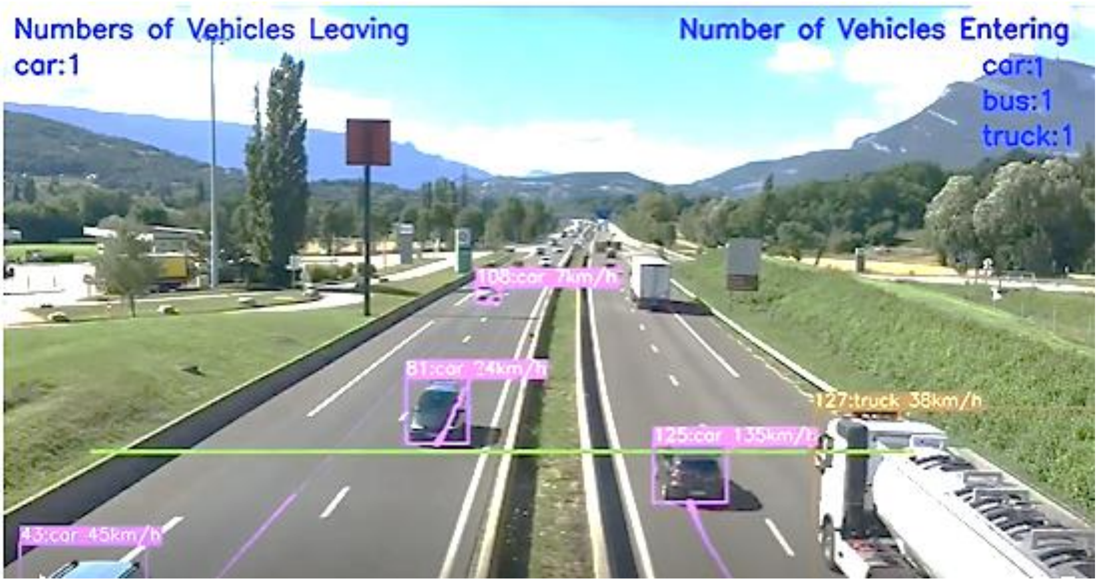

Figure 5. Vehicle speed estimation.

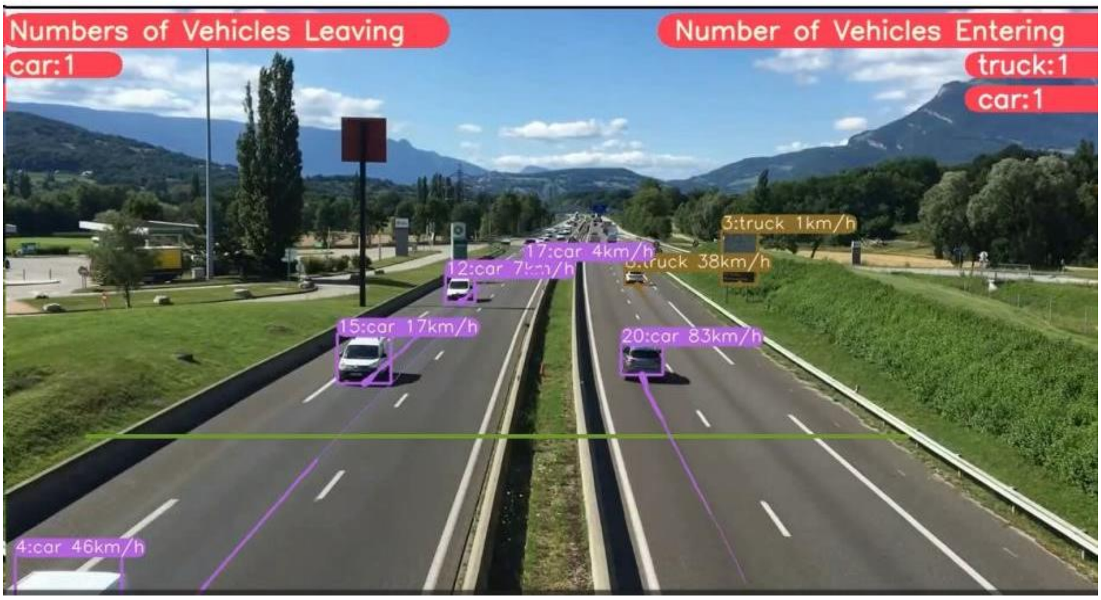

Figure 6. Output visualization.

It can be demonstrated that the utilization of YOLOv8n is not only acceptable but also faster compared to its predecessors. This model can be applied effectively in real-time surveillance cameras along highways or during video recording to assess the volume of passing vehicles over a specified time period. The collected data can then be employed for traffic management, offering insights into congestion levels at a particular location. Opting for the YOLOv8n model is deemed optimal for achieving the highest accuracy and speed in the system. For precise vehicle tracking, an algorithm inspired by the Deep SORT framework is proposed, capable of accurately tracking the nonlinear motion of vehicles. This algorithm incorporates YOLOv8n with Darknet, an open-source neural network framework, for vehicle localization and identification, ensuring a high level of tracking accuracy as Figure 6.

CONCLUSION

In this research, we have suggested a YOLOv8n-based real-time vehicle tracking, detection, and classification system. The suggested system can be used to monitor traffic flow on highways and is capable of reliably detecting, tracking, classifying, and speed estimating various sorts of vehicles in real-time. Additionally, we have created the predict.py model, which is capable of object detection, tracking, object counting, and real-time video. Transportation authorities can make better judgments to increase the effectiveness and safety of the transportation system by using the system's insightful data on highway traffic flow. Future development will focus on improving accuracy, which will include features such as lane recognition for over-speed alarms.

REFERENCES

1. Mohammadnazar A, Arvin R, Khattak AJ. Classifying travelers' driving style using basic safety messages generated by connected vehicles: application of unsupervised machine learning. Transport Res Part C: Emerg Technol. 2021; 122: Article 102917. doi: 10.1016/j.trc.2020.102917.
2. Wang Y, Ban X, Wang H, Wu D, Wang H, Yang S, Liu S, Lai J. Detection and classification of moving vehicle from video using multiple spatio-temporal features. IEEE Access. 2019; 7: 80287–80299. doi: 10.1109/ACCESS.2019.2923199.
3. Kusumah AP, Djayusman D, Setiadi GR, Nugraha AC, Hidayatullah P. Counting various vehicles using YOLOv4 and DeepSORT. J Integr Adv Eng. 2023; 3 (1): 1–6. doi: 10.51662/jiae.v3i1.68.
4. Kejriwal R, Ritika HJ, Arora A, Mohana. Vehicle detection and counting using deep learning based YOLO and Deep SORT algorithm for urban traffic management system. In: 2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichy, India, February 16–18, 2022. pp. 1–6. doi: 10.1109/ICEEICT53079.2022.9768653.

5. Gomaa A, Abdelwahab MM, Abo-Zahhad M, Minematsu T, Taniguchi R. Robust vehicle detection and counting algorithm employing a convolution neural network and optical flow. Sensors. 2019; 19 (20): 4588.
6. Sonnleitner E, Barth O, Palmanshofer A, Kurz M. Traffic measurement and congestion detection based on real-time highway video data. Appl Sci. 2020; 10 (18): 6270.
7. Mandal V, Adu-Gyamfi Y. Object detection and tracking algorithms for vehicle counting: a comparative analysis. J Big Data Anal Transport. 2020; 2: 251–261.
8. Meghana RK, Chitkara Y, Apoorva S. Mohana. Background-modelling techniques for foreground detection and tracking using Gaussian mixture model. In: 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, March 27–29, 2019. pp. 1129–1134.
9. Chethan Kumar B, Punitha R, Mohana. Performance analysis of object detection algorithm for intelligent traffic surveillance system. In: Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, July 15–17, 2020. pp. 573–579.
10. Rishika AL, Aishwarya Ch, Sahithi A, Premchender M. Real-time vehicle detection and tracking using YOLO-based Deep SORT model: a computer vision application for traffic surveillance. Turk J Computer Math Educ. 2023; 14 (1): 255–264.
11. Zhou Z, Li H. Vehicle object detection based on deep learning. Acad J Sci Technol. 2023; 5 (1): 38–45.
12. Rouf MA, Wu Q, Yu X, Iwahori Y, Wu H, Wang A. Real-time vehicle detection, tracking and counting system based on YOLOv7. Embedded Self-Organ Syst. 2023; 10 (7): 4–8. doi: 10.14464/ess.v10i7.598.
13. Wu JD, Chen BY, Shyr WJ, Shih FY. Vehicle classification and counting system using YOLO object detection technology. Traitement du Signal. 2021; 38 (4): 1087–1093.

