International Journal of Electrical and Computer Engineering (IJECE)
Vol. 13, No. 4, August 2023, pp. 4612~4622
ISSN: 2088-8708, DOI: 10.11591/ijece.v13i4.pp4612-4622

Application of improved you only look once model in road traffic monitoring system

Shridevi Jeevan Kamble, Manjunath R Kounte
School of Electronics and Communication Engineering, REVA University, Bangalore, India

Article Info

Article history:

Received Jun 28, 2022
Revised Nov 4, 2022
Accepted Nov 9, 2022

Keywords:

Computer vision
Intelligent traffic
Neural network
Traffic analysis
YOLOv3

ABSTRACT

The present research focuses on developing an intelligent traffic management solution for tracking the vehicles on roads. Our proposed work focuses on a much better you only look once (YOLOv4) traffic monitoring system that uses the CSPDarknet53 architecture as its foundation. Deep-sort learning methodology for vehicle multi-target detection from traffic video is also part of our research study. We have included features like the Kalman filter, which estimates unknown objects and can track moving targets. Hungarian techniques identify the correct frame for the object. We are using enhanced object detection network design and new data augmentation techniques with YOLOv4, which ultimately aids in traffic monitoring. Until recently, object identification models could either perform quickly or draw conclusions quickly. This was a big improvement, as YOLOv4 has an astoundingly good performance for a very high frames per second (FPS). The current study is focused on developing an intelligent video surveillance-based vehicle tracking system that tracks the vehicles using a neural network, image-based tracking, and YOLOv4. Real video sequences of road traffic are used to test the effectiveness of the method that has been suggested in the research. Through simulations, it is demonstrated that the suggested technique significantly increases graphics processing unit (GPU) speed and FSP as compared to baseline algorithms.

Corresponding Author:

Shridevi Jeevan Kamble
School of Electronics and Communication, REVA University
Rukmini Knowledge Park, Kattigenahalli, SH 104, Srinivasa Nagar, Bengaluru, Karnataka 560064, India
Email: shridevi.havanoor@gmail.com

1. INTRODUCTION

In the present times, the use of advanced technology such as artificial intelligence (AI) has increased in the different work spheres to reduce the frequency of repetitive work and make use of machine features for simplifying human tasks. While focusing on the present road traffic condition, congestion and irregular movement could be recorded regularly. Due to the advancement of AI, there is a requirement for the development of a regulated environment to exercise control over expanding traffic. Over the last few years, there has been a high increase in the volume of vehicles globally which creates major problems like traffic congestion, safety, contamination, and rapid requirement for movement. These new highways have additional lanes to manage the traffic; however, they do not seem feasible because of being too expensive. As a result, it is necessary to implement the traffic analysis system to gather insights related to the current traffic condition on the road. Traffic analysis system is processing the video stream and developing a system that will help to determine the numeral value of the vehicles that are passing on the lane. It will also help in determining in which direction the vehicles are moving on the lane. This system helps in categorizing the traffic flow into different segments for accurate prediction and optimizing the coordination plan. It is also due to the concentration of vehicles on a specific route so that it is easy to introduce a better traffic control mechanism [1]. For instance, a detection algorithm is used for the précised regulation of road traffic. It also entails the use of image processing to characterize the object and pinpoint its location in the image with great accuracy and speed. Registration detection plates that search the process of vehicles, it is regarded as one of the easiest methods because it provides contrast to the background. As a result, the characters and their number are visible on the screen for estimation purposes. However, this approach is not useful in identifying the vehicle in a situation where there are no licensed plates. Here, the main drawbacks of both methods are their inability to modify the perspective because they use blurry frames and their decreased ability to classify one large frame into a little frame [2]. There are different techniques for tracking the object. It may be in the form of a moving object or the dynamic component that is to be recognized from the static backdrop of the illustration. Apart from this, particle filter dependent Tracker is useful in tracking the movement of vehicles [3]. To focus on the background of the subtraction method, it has a drawback in processing data when there is congested traffic order because there is the creation of vehicle fusion.

It further leads to partial occlusion that is observed in the processed image data leading to predicting the incorrect bounding box. To resolve this issue gaussian mixture modelling (GMM) is used in combination with expectation optimization so that there is developing the background model [4]. The occultation was also developed by using colored histograms so that it helps to propose and identify the number of vehicles in dark videos by removing headlights [5]. The offline training method also helps to detect the vehicle by using a support vector machine so that it estimates the Haar wavelet function [6]. Furthermore, the multilayered neural network approach is also used to perform classification and feature extraction with the application. It includes using a new algorithm by the network so that there is the creation of an attribute of space. It includes the pattern that is accompanied by enviable statistical distribution [7]. AdaBoost classification and gradient analysis help in detecting the vehicle by identifying its back by using cascade of boosted classified presented [8].

Apart from this, an algorithm for teaching real-time tracking is used for the tracking of the vehicle by including a video sequence. It is based on the use of an advanced feature selection technique that also helps in the tracking of the object [9]. The technologies such as machine learning and cloud computing also play a very important role in reducing the congestion of traffic on the road. It helps to predict the latest traffic situation and provide guidance respectively. For example, to determine a free route or a shorter route for an ambulance, the city traffic control system can use real-time analysis techniques to determine the road network density. Figure 1. identifying and classifying the vehicle on video with the help of a surveillance camera. This technique can be used to classify security cameras in order to solve traffic analysis problems.

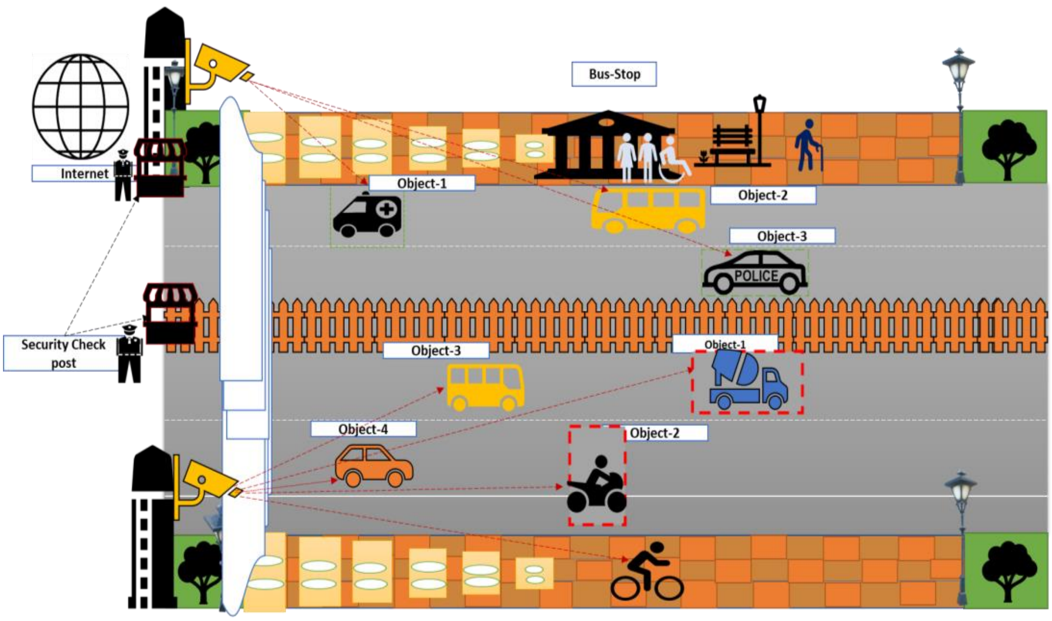

Figure 1. Shows the clear picture of the system used for identifying and classifying the vehicle on video with the help of a surveillance camera

This paper, describes the use of an intelligent vehicle tracking that is based on the use of the improved you only look once (YOLOv4) algorithm. This system is used for identifying and classifying the vehicle on video with the help of a surveillance camera. This system can be used for the categorization of the surveillance cameras so that there is resolving traffic analysis issues. It focuses on the examination of the test data that plays a major role in detecting the vehicles. It also helps in analyzing the intensity and density of the traffic flow in all directions so that there is the development of a system for regulated traffic flow. It includes effective use of the angles and camera position to evaluate the traffic flow. The present research also examines the efficiency of a YOLOv4-based traffic monitoring system that can be used for traffic management. The accompanying paper is structured as follows: section 1 describes AI-based technology on traffic analysis system to gather information about the present traffic situation on the road, coupled with an explanation of YOLO based on intelligent vehicle tracking. The major purpose of the paper is outlined in section 2 of the text, we explain about the main contribution of the paper. Section 3 is all about literature reviews about different YOLO techniques used for different purpose within vehicular communication. Section 4 explains you the methodology about what techniques and tools are used to carry out the research work. In section 5, we spoke about the outcomes and how our YOLOv4 is superior to previous YOLOs. Conclusion and closure are covered in section 6.

2. MAIN OBJECTIVE

To examine a new improved YOLOv4 based traffic monitoring system. This is based on the CSPDarknet53 architecture, introduces with significant number of novel concepts, including CIoU loss, cross-minibatch normalization, weighted residual connections, cross-stage partial connections, and mish activation. Our main objective is traffic monitoring by detecting vehicles on the road with high accuracy and DeepSORT algorithm is used to track vehicles in real-time. In order to detect vehicles with many targets from traffic video, we additionally apply the DeepSORT learning approach. YOLOv4 detects vehicles with an accuracy of 85%. We have also added extras like the Kalman filter, which estimates unidentified objects. To include moving items in the image, they also model several states. The estimate model development method includes motion prediction and feature generation. Our research also uses the Hungarian approach, which can tell if an object in the current frame and one in the previous frame are the same. It will also be used for id attribution and association. An algorithm known as a Kalman filter can predict where things will be in the future based on their current positions.

3. LITERATURE REVIEW

3.1. Enhanced YOLO-based traffic monitoring system for roads

According to Al-qaness *et al.* [10] improved YOLOv3 algorithm can be used to detect and categorize the vehicles that are moving on roads. It is based on acquiring vehicle data from the surveillance cameras that are installed in different locations on the road so that problems related to traffic analysis are resolved. YOLOv3 method is based on marking the visibility percentage to identify the vehicles of interest and evaluate traffic flow intensity during peak hours. It also helps in determining the traffic flow intensity in different directions by ensuring resistance to camera position during changing directions.

Ren *et al.* [11] analyzed that the YOLOv3 model is based on the convolutional neural network so that there is the detection of different vehicles that have been covered in the fame. YOLOv3 has open-source code and license that allows its implementation to be free and increases its usability for complete retraining purposes in the neural network. As a result, there is efficient utilization of the Darknet-53 network to predict the coordinates of objects along with their sizes from three different imaging scales.

Gustafsson *et al.* [12] analyzed that probabilistic methods can also be included to increase the efficacy of the YOLOv3 method so that there is the detection of objects. Each object is known to be present in its internal state which can be estimated in each frame by using Kalman filter and particle filter. In the Kalman filter, it is considered as the object is present in the form of a random variable associated with normal distribution.

Sun *et al.* [13] asserted that neural networks are to be included in the YOLOv3 network so that its efficiency increases for the image recognition process. It is based on the use of a fine-tuning approach that is used along with an already-trained network so that there is performing of tasks in the core network working system. Additionally, a neural network is also trained for the selection of the selected dataset that helps in determining the training speed at a lower level.

3.2. Investigation of a promoted you only look once algorithm and its application in traffic flow monitoring

Cao *et al.* [14] examined that the rising number of vehicles on the road has increased the incidence of traffic accidents that make a negative impact on the transportation experience of people and overall their life. The intelligent transportation system has been developed for the proper management of traffic by the traffic management department. This department makes decisions for the management of traffic control based on the data acquired for the road traffic flow. The data provided information on the road to the traffic department by determining the particular road that has high flow so that better traffic information is acquired by the driver. It is important as it manages the traffic flow statistic.

Cao *et al.* [14] examined the importance of vehicle classifier in that it performs the protection of vehicles on a set area of a video. It is a simple method that is based on a few factors for the implementation of the vehicle classifier. Apart from this, the virtual call method is also used because it includes a set candidate area in a video and also focuses on the detection of the traffic flow by changing the call state and protecting the surface of roads. However, the use of this system gets restricted because of a lack of effectiveness in the identification of the multi-lane traffic flow.

Liu *et al.* [15] examined video surveillance systems also used on roads that have been constructed in the urban regions. The system helps in obtaining video images from the traffic video systems that are simple and convenient to use. It also helps to detect the amount of data that is very large and it will not affect the normal travel of the vehicle. Apart from this, a Gaussian model is also considered to be an easy way for developing the background images that are acquired from traffic video surveillance systems. In this method, there is targeting of the image by using a background difference method and also focus on the elimination of shadow and morphological processing by reducing the illumination intensity and noise.

Sarikan *et al.* [16] examined in the present time deep learning and artificial intelligence are used for traffic semantic segmentation image categorization and other detection actions in different installations. Apart from this, the region-based method is also considered as a common method for targeting the detection when used with the faster R-convolutional neural network (CNN), SPP-net, Region-CNN (R-CNN), and fast R-CNN algorithms. R-CNN is considered as a discerning investigating technique to remove the image from the surrounding region. It also focuses on training and test speed because it is considered slow. Additionally, lightweight PeleeNet models and DenseNet models are used to enhance the speed and accuracy of detection mobile devices. These models are also used for highest targeting the categorization precision.

Ren *et al.* [17] examined that YOLO is a single convolutional network that is used for predicting class probabilities and multiplicity of vehicles with a clear view. YOLO can be used for the training of full images that play a major role in optimizing detection efficacy. It includes using a one-stage algorithm and an end-to-end detection process that is well efficient in combining the target location and recognition of the target. It must be ascertained that most YOLO network is trained with datasets that help in acquiring accurate information and ensuring flow statistics based on model tests. It showcases the outcomes that are related to the monitoring of the traffic flow.

3.3. A deep learning approach of vehicle multitarget detection from traffic video

Li *et al.* [18] intelligent transport system is considered as an efficient way to resolve the traffic-related problem in the urban area in future. Apart from this, vehicle detection is also considered to be important for the attainment of data from the transportation system because it provides valid data from different traffic intelligent control applications such as congested traffic data. It is based on machine learning that extracted data from 4G features such as histogram-based gradients. It also includes the features of scale-invariant transform that play a major role in identifying the local binary pattern the next step includes the extracted feature that is to be fed for pretend classifier that is in the form of a support vector machine for cortical identification of the task it includes the data files to extract feature. However, the major restriction is experienced while using this process because there is the problem of generalization that creates the difficulty of result for particular identification among the practical problem.

Li *et al.* [18] examined that machine learning has to supervised and unsupervised method to extract and convert the feature the planning also considered as a sub-area of machine learning because it is an algorithm of the model that is defined the complex relationship with between the data with the help of multilevel representation it also transfers the original data into higher-level and more simplified expression with the help of simple and nonlinear model by showing a particular wood result for detecting the target. For instance, the mean average precision (mAP) is known to reach the level of 30% in the segment of the VOC2007 dataset. In the VOC2007 dataset, the R-CNN technique, on the other hand, a combination of standard machine learning and deep learning plays a big role in improving the mAP to 48%. Though RCNN provides benefits, its use gets restricted because of the training of several classifiers that usually result in the creation of complicated computational issues. Though there has been a modification in the network structure, there has been an augmentation to 66% in the year 2014. Later on, there was the use of other mechanisms such as fast RCNN, SPP-Net, faster RCNN, and YOLO to bring improvements in the transfer of the original data. It is followed by the use of the YOLO v2 target detection algorithm (YOLO-vocR) to acquire précised outcomes. In this method, there is the conversion of the problem in binary classification to acquire accurate datasets. It leads to the creation of the YOLO-vocRV network that helps in the attainment of more than 20,000 iterations resulting in an increase in the accuracy of the system for the detection of different traffic densities that are larger than 90%.

Based on the above facts, it can be said that in comparison to the traditional machine learning method, the process that has been proposed in the research is more efficient. A major reason behind it is that it can be used for the recalling value which is a limitation in the previous models. Therefore, it can be said that YOLO9000, YOLO-voc, and YOLO v3 models, the YOLO-vocRV is considered to be the most efficient one because it has less recall time and ensures more precision. The final research analysis indicated that the error rate of free flow was recorded to be 1.4%, while the false detection rate was recorded to be 3.7%, and the accuracy rate was recorded to be 96.3%. It includes focusing on light or low light conditions for the estimation of facts in the research.

3.4. Optimized YOLOv2 based vehicle classification and tracking for intelligent transportation system

Sivaraman and Trivedi [19] automated vehicle classification is used for tracking the intelligent transportation system. It is executed in three essential systems such as data collection, feature extraction, and training. While focusing on the data collection aspect, it is based on collecting data from two different locations. The obtained data is categorized into four lanes in which every lane has its own set of enforcement cameras an installation overview. As a result, the presence of equipment in each lane helps in continuous sampling from the light curtain to identify edge detection, shadow removal, and execute multifaceted filtering operations. To recreate the image, feature extraction, filtering, and image morphing algorithms are used. Using a 3×3 frame, the filtering phase is related with median filtering [20]. It will aid in the skeletonization of processes by decreasing foreground regions and connecting pixels. On the other hand, in image morphing, operations are carried out to determine joints and endpoints so that there is a highlighting of the selected features. In the training phase, there is the use of heat maps for the identification of face and vehicle. It includes the use of the expression.

$$HM_c = \frac{1}{n^n} \sum Xcii \tag{1}$$

In this formula, c stands for the vehicle class name, n for the training set's element count, and x for the associated training data set item. It helps in classifying the images and testing them based on the corresponding heat maps outcomes. This method of image identification is highly beneficial for vehicles such as motorcycles, cars, SUVs, trucks, and buses that help to identify the age and vehicle data by using classification techniques. It supports the implementation of the intelligent transport system by accurately executing image processing and machine learning techniques.

Zhu *et al.* [21] examined that a similar approach has also been used for image classification by using the data that has been acquired from cameras that are placed on top of the road. It helped in overcoming the issues that were faced on classifying vehicles by using training and neural network that is based on the use of machine learning application. It also helps in differentiating motorcycles from other vehicles such as cars.

4. METHOD

Trekking of vehicle plays an important role and considered as the big part in intelligent transportation and intelligent traffic incident. Detection based tracking and deduction free tracking are considered as two major objects that tracking framework currently available [22]. In order to make focus on object trekking; it is used to analyses the number of the vehicle that pass through a pole so that person can develop a plan to analyses the type of road. Detection mainly used the presence of things that are imaginary and using a bounding box. As a result, it can be said that input to a system is considered as the image of one and more act it may be in the form of photograph video. In order to make focus on image of input with high resolution; it is used with the help of YOLOv4 network that help to detect the diversified items especially the smaller size of object. It also includes 110 convolutional layers with 21 root layer 3 layer of Max pooling 23 layer of shortcut and 3 layers of YOLO [23]. It has to avoid the conjunction by providing negative values and help greater gradient and contrast in a slight manner. We use YOLOv4 at the backend trained using the darknet framework. The model will be converted to tensor flow format and then the detections will be further utilized to track the object. This can be used for any type of object tracking. Deep sort framework which is an extension of Sort with a deep association matrix will be used to track the objects and reduce false positives. Deep sort internally uses Kalman filters to approximate the path and then reduce the errors cascading through the best approximation as specified in Figure 2(a). The main focus of our research work is on object tracking mainly and the major applications of object tracking are counting number of vehicles passing through a poll so that we can plan the type of road that needs to be made and theft detection, ambulance identification and traffic management. Figure 2(b). Show the flow the vehicle detection method, which is in flow with multiple object tracking [24].

4.1. Deep sort algorithm

It is useful to track based on velocity. It helps to add the feature by computing the feature which is considered as a bounding box and highly used to make similarity between the features. It also includes Kalman filter for tracking the objective and object. This mainly related with the new detection and protection and especially for the tracking that performed on image space by using metric and compute bounding box and its related overlapping.

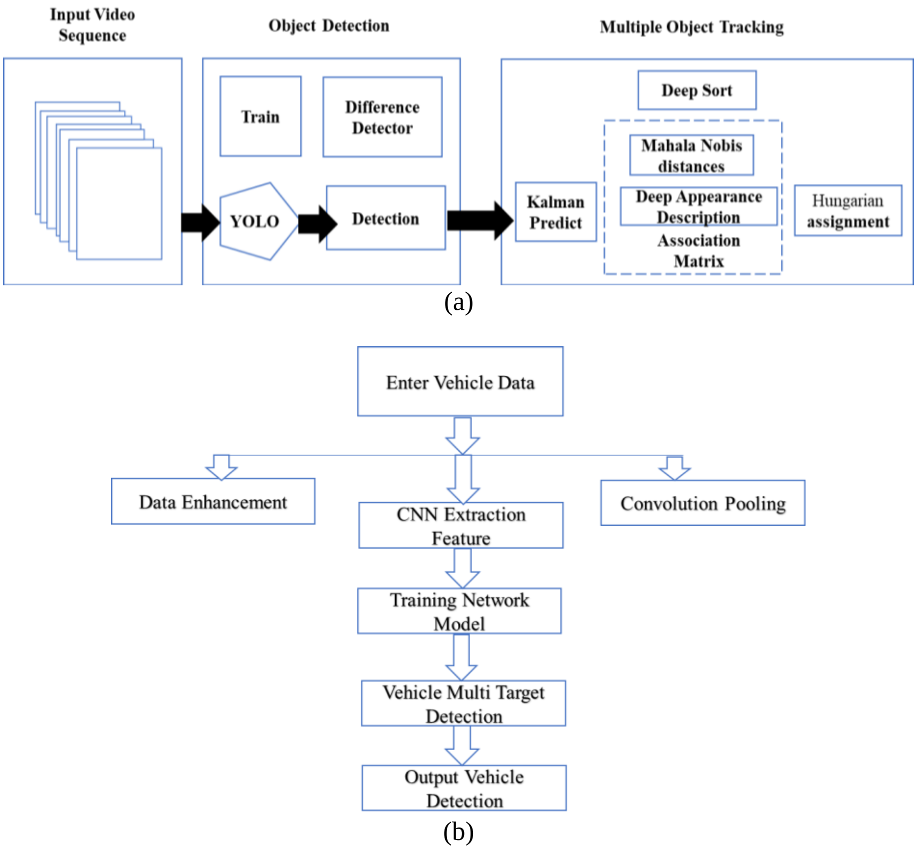

Figure 2. Multi-object tracking and vehicle detection (a) multi-object tracking and analysis method and (b) vehicle detection using YOLO4 algorithm

4.2. Kalman filter

Kalman filtering is an algorithm that is useful to estimate unknown variables over time. It is a two-stage algorithm prediction and updation also known as propagation and correction respectively. It also includes linear model with additive Gaussian noises, The Kalman filter offers the most accurate results. Figure 3 shows the Kalman filter algorithm estimation over time. Based on the behaviors of the front moving object and the previous state matrix, model system equations are used to generate the predicted state in order to estimate the object parameter for the upcoming state. The Kalman filter uses the measurement from the detector to update the prediction state for locating the position of the item when it is time for updating. Where X_K is the moving object's state vector and k is the discrete time divided into frame interval. Estimating optimal X_K, which is updated by measurementY_k, is the main goal of this tracking.

Kalman filter prediction equations are as:

$$\text{Prediction state: } \hat{X}_{k|k-1} = \widehat{F_k} x_{k-1|k-1} + W_k \quad (2)$$

$$\text{Process covariance: } \hat{P}_{k|k-1} = \widehat{F_k} P_{k-1|k-1} + F^T{}_k + Q_k \quad (3)$$

F_k is the transition matrix applying the effect of every state parameter, W_k is zero mean which predicts noise along with covariance Q_k. P_k is process noise covariance between terms in the state vector.

Kalman filter correlation equations are:

Kalman gain $K_k = \hat{P}_{k-1|k-1} H^T{}_k (H_k \hat{P}_{k|k-1} H^T{}_k + R_k)^{-1}$ (4)

Measurement $Y_k = H_k X_k + v_k$ (5)

Correction state $\hat{X}_{k_|k} = \hat{X}_{k_|k-1} + K_k (Z_k - H_k \hat{X}_{k|k-1})$ (6)

Update covariance $\hat{P}_{k|k-1} = (1 - K_k H_k) \hat{P}_{k|k-1}$ (7)

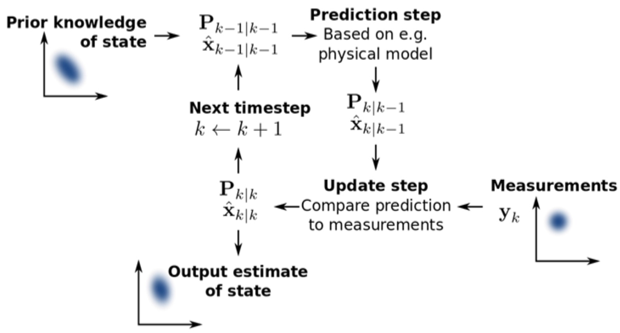

Figure 3. Shows the Kalman filter algorithm estimation over time

4.3. Hungarian algorithm

The Hungarian algorithm is helpful to find out the minimum costs that are related to the problem in assignment. It also includes a finding the people to the activity and organize the data into matrix with vehicles and also include rows and activities as a column. We follow each vehicle tracks using a Hungarian algorithm, and then we calculate the tracked vehicles' velocities using our understanding of camera calibration. Velocity of vehicles are calculated as the vehicle gradually approach towards the camera, the video plays. So, we use the method of a simple to install camera calibration for observing traffic. Calculating velocity:

$$i_x = f_w = \frac{w_{xc\theta} - w_{ys\theta}}{xC_r\ S_\theta + yC_r\ C_\theta + h|\ S_r} \quad (8)$$

$$w_x = f_w = \frac{-(i_y\, S_\theta + i_x\, C_\theta\, S_T\ )h}{S_r\,(i_y\, C_r\ + f\, S_T} \quad (9)$$

$$i_y = f_w = \frac{-w_x\ S_T S_\theta\ - w_y S_r C_\theta}{xC_r\ S_\theta + WyC_r\ C_\theta + h|\ S_r} \quad (10)$$

$$w_y = \frac{-(i_y\, c_\theta + i_x\, C_\theta\, S_T\ )h}{S_r\,(i_y\, C_r\ + f\, S_T} \quad (11)$$

4.4. Tensor flow

It helped to develop data flow graph structure and also analyses how the data move with the help of graph and series of processing nodes and also include application which are convenient for the target because it includes local machine cluster in the cloud iOS and android devices, CPU and GPU. a platform for creating machine learning applications that is open source and comprehensive is called TensorFlow. The primary application of this symbolic math library was to create differentiate programming and perform different type of data that totally related to the training and deep neural network. It also helps the developer to create machine learning language with the help of various tools library community and resources.

4.5. Dataset

The dataset consists of 500 images of buses, trucks, bike, bicycle and cars. The images of the vehicles were personally taken from Mumbay-Pune express highway state of Maharashtra-MH 48. The graphics card is NVIDIA GeForce RTX 2070 super (8192 mb). The system environment for simulation is: experiment is implemented in Windows system, CPU-Intel core i-5 2.5 GHz CPU and 16 GB memory, 10400F CPU @2.90 GHz. Google Colab serves as the test environment for vehicle detection. A training set and a test set are created from the data set. The experiment dataset has 250 and 105 images, respectively, whereas the vehicle's training dataset has two items from distinct categories with a ratio of 7:3. Figure 4 shows data creation for vehicle detection accuracy using YOLOv4. Figure 5 show represents the YOLOv4 network building and creations using dataset creation.

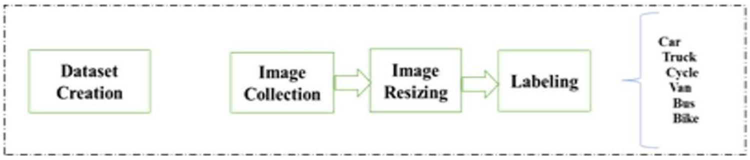

Figure 4. Data creation for vehicle detection

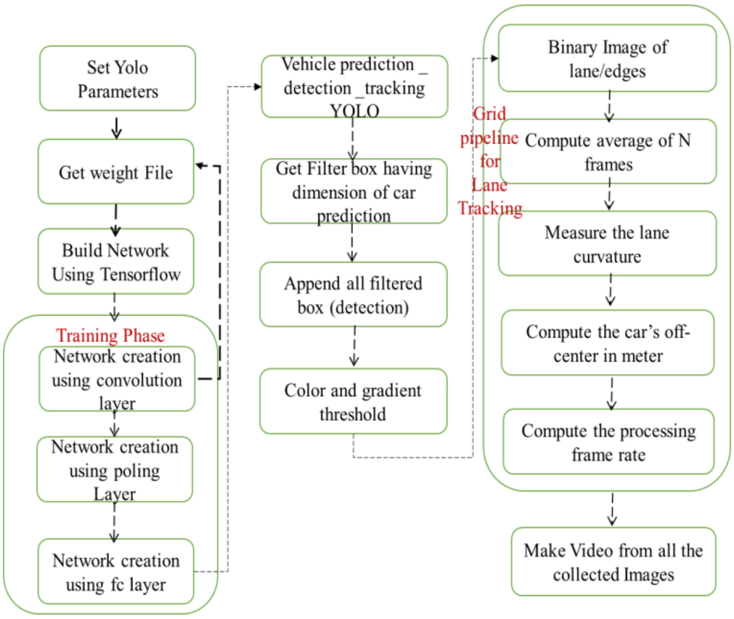

Figure 5. Represents the YOLOv4 network building and creations

The following algorithm is included in the proposed system's programming:

Step 1 : Collection of data from various sources like real time video

Step 2 : Vehicle localization using the optimization YOLOv4 object detector.

Step 3 : Predicting the position in the next frame using Kalman filter

Step 4 : Computing the intersection-over-union (IOU) between the detection boxes, prediction bounding boxes of vehicles and the associated results are obtained through Hungarian algorithms.

Step 5 : The distance between the confirmed tracklets and detections and their appearance feature similarity are calculated. Following that, the Hungarian method is used to determine the associations between the confirmed tracklets and detections.

Step 6 : Assigning detection on vehicles by Kalman filter and the motion prediction model. Then we initialize new tracklets for unassociated detection.

5. RESULT AND DISCUSSION

Based on the above-discussed facts, it can be said that advanced technologies such as AI can be used in daily human working to reduce redundancies of tasks. The technology includes the use of new frameworks and systems such as YOLOv4 so that there is the implementation of a traffic analysis system for collecting related to the current traffic situation on the road. It focused that the machine learning method is also to be included for the supervision of extracted and converted features that are essential for planning of road maintenance and safety. Figure 6. Shows the simulations, it is demonstrated that the suggested technique significantly increases C as compared to baseline algorithms. The picture images with bounded box tagged in Figures 7(a) to (d) shows the output of our proposed objection detection through a video tracking system using an improvised YOLOv4 vehicle detection algorithm. YOLOv4 is extremely fast and accurate however no retraining of the model is required. Earlier detection techniques use localizers or classifiers as detectors. Various locations and scales are used to apply the model to an image. Detections are the areas of the image that scored highly. We employ a special strategy. One neural network is used to process the entire image. The image is divided into regions by this network, which forecasts bounding boxes and probabilities for each region. Projected probabilities are used to weigh these bounding boxes [25].

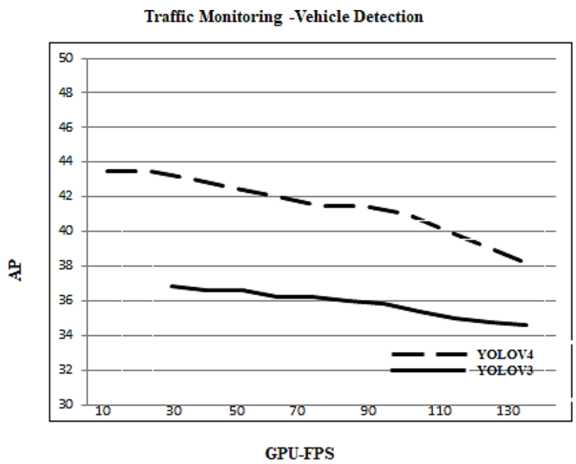

Figure 6. Traffic Monitoring: YOLOv4 is both quick and effective

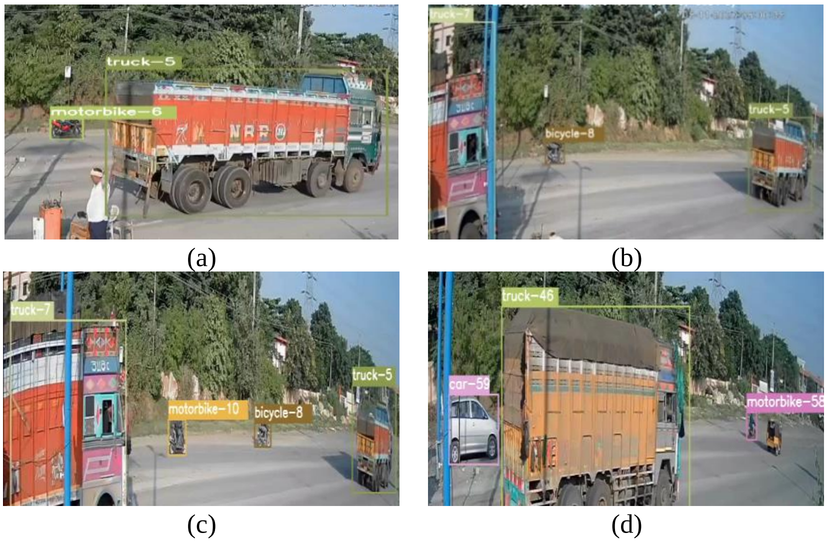

Figure 7. Shows vehicle detection using YOLOv4 (a) image samples with bounding box, (b) bounded box to detect vehicles, (c) bounded box and class label, and (d) bounded box with class label outputs from object detection

As compared with the previous YOLOv3, YOLOv4 has the following advantages: YOLOv3 improved upon earlier models by incorporating an objectness score into bounding box prediction and adding connections to the backbone network layers. In addition to improving performance on smaller objects, we used CSPDarknet53 as the framework for YOLOv4 Vehicle detection. Typically, ImageNet categorization is used as pretraining for the backbone network of an object detector. Pretraining means that the network's weights have already been modified to captured important features in an image, even if they will be modified for the new task of object detection. YOLOv4 is a powerful object detection model that help a person to achieve accurate and superpower object detector with the help of 1080 Ti or 2080 Ti GPU. By using state-of-the-art "Bag-of-Freebies" and "Bag-of-Specials", it helps to detect the object during detector training in a verified manner. Cross interaction batch normalization and path aggregation network help to gain single GPU training. DenseNet serves as the foundation for CSPDarknet53. DenseNet was created to connect layers in convolutional neural networks with the following goals in mind: to improve feature propagation, stimulate feature reuse, and decrease the number of network parameters. Table 1 shows parameters for neural network for image classification and Object detectors on the MS COCO dataset, in terms of speed and precision. In comparison to its predecessor, YOLOv3, YOLOv4 achieves measurements with 10% greater average precision and 12% better frames per second (FPS).

Table 1. Parameter of neural network for image classification and object detection in terms of speed and precision

Methods	Backbone	Input Network Resolution	FPS	AP
YOLOv4	CPSDarknet53	512	40%	39%
YOLOv3	Darknet53	360	23%	34%

6. CONCLUSION AND RECOMMENDATION

The YOLOv3 model had a number of shortcomings, including the inability to accurately predict results when several vehicles in a video were close to one another. It also discovered positioning issues as a result of various leakage statistics, which have a detrimental effect on the accuracy of statistical data. However, in the present research, it was analyzed that when the YOLO algorithm was optimized by generalized intersection over union (GIOU), it provided precise information about traffic flow data by making use of the adjusting threshold parameters. It helped in achieving rightful outcomes related to the estimation of the actual number of vehicles. In the optimized algorithm, there was the use of statistical analysis that helped in providing reliable traffic flow monitoring information for multiple sensing. It included analyzing different weather conditions for the attainment of the reference value. As a result, it was estimated that the YOLOv3 model was not appropriate for acquiring reliable vehicle insights on a rainy day because of its inability to have a recall rate. Hence, in such conditions, the YOLO network must be replaced with the deep residual network. There may also be the implementation of batch normalization at the time of training to bring improvements in the adaptability of the system during different weather conditions. It will also help in resolving the issues related to category imbalances and focal loss. YOLOv4 is a powerful object detection model that help a person to achieve accurate and superpower object detector with the help of 1080 Ti or 2080 Ti GPU. In comparison to its predecessor, YOLOv3, YOLOv4 achieves measurements with 10% greater average precision and 12% better frames per second.

ACKNOWLEDGEMENTS

The authors would like to thank Dr P Shyamaraju, Chancellor, REVA University for all the facilities provided to carry out the research activities.

REFERENCES

[1] N. Rachburee and W. Punlumjeak, "An assistive model of obstacle detection based on deep learning: YOLOv3 for visually impaired people," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 4, pp. 3434–3442, Aug. 2021, doi: 10.11591/ijece.v11i4.pp3434-3442.

[2] M. B. W. de Oliveira and A. de Almeida Neto, "Optimization of traffic lights timing based on artificial neural networks," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Oct. 2014, pp. 1921–1922, doi: 10.1109/ITSC.2014.6957986.

[3] S. J. Kamble and M. R. Kounte, "SG-TSE: Segment-based geographic routing and traffic light scheduling for EV preemption based negative impact reduction on normal traffic," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 12, 2021, doi: 10.14569/IJACSA.2021.0121236.

[4] T. Long, W. Jiao, G. He, and W. Wang, "Automatic line segment registration using gaussian mixture model and expectation-maximization algorithm," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 5, pp. 1688–1699, May 2014, doi: 10.1109/JSTARS.2013.2273871.

[5] Y.-L. Chen, B.-F. Wu, C.-T. Lin, C.-J. Fan, and C.-M. Hsieh, "Real-time vision-based vehicle detection and tracking on a moving vehicle for nighttime driver assistance," *International Journal of Robotics and Automation*, vol. 24, no. 2, 2009, doi: 10.2316/Journal.206.2009.2.206-3089.

[6] K. Wang, Y. Liang, X. Xing, and R. Zhang, "Target detection algorithm based on Gaussian mixture," in *Proceedings of the 2015 Chinese Intelligent Automation Conference*, 2015, pp. 439–447, doi: 10.1007/978-3-662-46469-4_47.

[7] O. L. Junior and U. Nunes, "Improving the generalization properties of neural networks: an application to vehicle detection," in *2008 11th International IEEE Conference on Intelligent Transportation Systems*, Oct. 2008, pp. 310–315, doi: 10.1109/ITSC.2008.4732681.

[8] P. Negri, X. Clady, S. M. Hanif, and L. Prevost, "A cascade of boosted generative and discriminative classifiers for vehicle

detection," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, Dec. 2008, doi: 10.1155/2008/782432.
[9] G. Zhang, R. P. Avery, and Y. Wang, "Video-based vehicle detection and classification system for real-time traffic data collection using uncalibrated video cameras," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1993, no. 1, pp. 138–147, Jan. 2007, doi: 10.3141/1993-19.
[10] M. A. A. Al-qaness, A. A. Abbasi, H. Fan, R. A. Ibrahim, S. H. Alsamhi, and A. Hawbani, "An improved YOLO-based road traffic monitoring system," *Computing*, vol. 103, no. 2, pp. 211–230, Feb. 2021, doi: 10.1007/s00607-020-00869-8.
[11] P. Ren, L. Wang, W. Fang, S. Song, and S. Djahel, "A novel squeeze YOLO-based real-time people counting approach," *International Journal of Bio-Inspired Computation*, vol. 16, no. 2, pp. 94–101, 2020, doi: 10.1504/IJBIC.2020.109674.
[12] F. Gustafsson *et al.*, "Particle filters for positioning, navigation, and tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 425–437, 2002, doi: 10.1109/78.978396.
[13] Z. Sun, G. Bebis, and R. Miller, "Monocular precrash vehicle detection: features and classifiers," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 2019–2034, Jul. 2006, doi: 10.1109/TIP.2006.877062.
[14] C.-Y. Cao, J.-C. Zheng, Y.-Q. Huang, J. Liu, and C.-F. Yang, "Investigation of a promoted you only look once algorithm and its application in traffic flow monitoring," *Applied Sciences*, vol. 9, no. 17, Sep. 2019, doi: 10.3390/app9173619.
[15] Y. Liu, O. Kosut, and A. S. Willsky, "Sampling from Gaussian Markov random fields using stationary and non-stationary subgraph perturbations," *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 576–589, Feb. 2015, doi: 10.1109/TSP.2014.2375134.
[16] S. S. Sarikan, A. M. Ozbayoglu, and O. Zilci, "Automated vehicle classification with image processing and computational intelligence," *Procedia Computer Science*, vol. 114, pp. 515–522, 2017, doi: 10.1016/j.procs.2017.09.022.
[17] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.
[18] X. Li, Y. Liu, Z. Zhao, Y. Zhang, and L. He, "A deep learning approach of vehicle multitarget detection from traffic video," *Journal of Advanced Transportation*, vol. 2018, pp. 1–11, Nov. 2018, doi: 10.1155/2018/7075814.
[19] S. Sivaraman and M. M. Trivedi, "A general active-learning framework for on-road vehicle recognition and tracking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 267–276, Jun. 2010, doi: 10.1109/TITS.2010.2040177.
[20] C. Tu and S. Du, "A hierarchical RCNN for vehicle and vehicle license plate detection and recognition," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 1, pp. 731–737, Feb. 2022, doi: 10.11591/ijece.v12i1.pp731-737.
[21] Y. Zhu, J. Wang, and H. Lu, "A study on urban traffic congestion dynamic predicts method based on advanced fuzzy clustering model," in *2008 International Conference on Computational Intelligence and Security*, Dec. 2008, pp. 96–100, doi: 10.1109/CIS.2008.194.
[22] D. Withopf and B. Jahne, "Learning algorithm for real-time vehicle tracking," in *2006 IEEE Intelligent Transportation Systems Conference*, 2006, pp. 516–521, doi: 10.1109/ITSC.2006.1706793.
[23] M. Karthikeyan and T. S. Subashini, "Automated object detection of mechanical fasteners using faster region based convolutional neural networks," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 6, pp. 5430–5437, Dec. 2021, doi: 10.11591/ijece.v11i6.pp5430-5437.
[24] A. Kausar, A. Jamil, N. Nida, and M. H. Yousaf, "Two-wheeled vehicle detection using two-step and single-step deep learning models," *Arabian Journal for Science and Engineering*, vol. 45, no. 12, pp. 10755–10773, Dec. 2020, doi: 10.1007/s13369-020-04837-4.
[25] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *Prepr. arXiv1804.02767*, pp. 1–6, 2018.

BIOGRAPHIES OF AUTHORS

Shridevi Jeevan Kamble is an Assistant Professor and research scholar at REVA University. She holds a Bachelor-ECE and M.Tech in Digital Electronics from VTU Belgaum, Karnataka. She is an IEEE student member. She has over nine years of Teaching, Research, and administrative experience. Her research area includes the internet of vehicles, cyber-physical system, machine learning, and deep learning VANETs. She has also published 2 books "Electronics Product Design" for SPPU and "Control System-II" for Mumbai University for final year Engineering students. Published papers at international conference paper at Elsevier, Springer and IEEE explore. She has also published journal paper, Lecture Notes on Data Engineering and Communications Technologies, vol 31. Springer, Cham. 2018. She can be contacted at: shridevi.jeevank@reva.edu.in.

Manjunath R Kounte has completed Bachelors of Engineering in Electronics and Communication Engineering during 2007, Masters of Technology in Computer Network Engineering during 2010 from VTU, Belagavi, Karnataka, India, PhD from JAIN University during 2017. He is currently working as Associate Professor and Head, Electronics and Computer Engineering, School of ECE, REVA University, Bangalore, India. He has over 15+ years of Teaching, Research and Administrative experience. Dr. Manjunath Kounte is Senior Member, IEEE and member of various international professional bodies such as IETE, ISTE. He has published over 12+ Indexed Journal papers, 25+ International Conference papers,1 book and 3 book chapters. His research interests include machine learning and IoT Edge computing, internet of vehicles, signal processing, fog computing, blockchain technologies. He is also heading Machine Learning and Blockchain (MLB) research lab in REVA University. More details about him can be found in his website https://www.kountesir.com/ He can be contacted at: kounte@reva.edu.in.